NBER WORKING PAPER SERIES

WHAT'S MY EMPLOYEE WORTH? THE EFFECTS OF SALARY BENCHMARKING

Zoe B. Cullen
Shengwu Li
Ricardo Perez-Truglia

Working Paper 30570
http://www.nber.org/papers/w30570

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2022, Revised January 2023

Special thanks to Brent Weiss and Ben Hanowell for all of their help and feedback. We are also thankful for comments by Sydnee Caldwell, Laura Giuliano, Matthew Grennan, Simon Jäger, Asim Khwaja, Pat Kline, Ray Kluender, Felix Koenig, Claudio Labanca, Alex MacKay, Alex Mas, Filip Matejka, Enrico Moretti, Bobby Pakzad-Hurson, Simon Quinn, Ben Roth, Benjamin Schoefer, Jesse Shapiro, Isaac Sorkin, Shoshana Vasserman and other colleagues and seminar discussants at NBER Summer Institute (Labor Studies), Harvard University, Columbia University (Econ), Columbia University (GSB), University of Chicago (Booth), U.S. Census, CEPR Labor Studies, Essex University, Universita della Svizzera Italiana, Norwegian School of Economics, UC-Berkeley (PF), UC-Berkeley (Labor), Amazon, University of Delaware, University of Copenhagen, University of Cologne, Goethe Univeristy, Firms and Labor Workshop, and the Texas A&M Labor and Public Economics Workshop. This project was reviewed and approved in advance by the Institutional Review Board at Harvard Business School (IRB #20-1779). We thank the collaborating institution for granting access to their data and for all of their help. The collaborating institution did not provide any financial support for the research being conducted. Julia Gilman, Dylan Balla-Elliott, Romina Quagliotti and Xinmei Yang provided excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

What's My Employee Worth? The Effects of Salary Benchmarking
Zoe B. Cullen, Shengwu Li, and Ricardo Perez-Truglia
NBER Working Paper No. 30570
October 2022, Revised January 2023
JEL No. D83,J31,J38,M52

ABSTRACT

While U.S. legislation prohibits employers from sharing information about their employees' compensation with each other, companies are still allowed to acquire and use more aggregated data provided by third parties. Most medium and large firms report using this type of data to set salaries, a practice that is known as salary benchmarking. Despite their widespread use across occupations, there is no evidence on the effects of salary benchmarking. We provide a model that explains why firms are interested in salary benchmarking and makes predictions regarding the effects of the tool. Next, we measure the actual effects of these tools using administrative data from one of the leading providers of payroll services and salary benchmarks. The evidence suggests that salary benchmarking has a significant effect on pay setting and in a manner that is consistent with the predictions of the model. Our findings have implications for the study of labor markets and for ongoing policy debates.

Zoe B. Cullen
Rock Center 210
Harvard Business School
60 N. Harvard
Boston, MA 02163
and NBER
zcullen@hbs.edu

Ricardo Perez-Truglia
Haas School of Business
University of California, Berkeley
545 Student Services Building #1900
Berkeley, CA 94720-1900
and NBER
ricardotruglia@berkeley.edu

Shengwu Li
Harvard University
Department of Economics
Littauer Center
1805 Cambridge Street
Cambridge, MA 02138
shengwu_li@fas.harvard.edu

A data appendix is available at http://www.nber.org/data-appendix/w30570

1 Introduction

Employee compensation is the largest source of expenditures for firms. Setting the right salaries is of first order importance. How do firms find out what their employees are worth?

While U.S. legislation, in an effort to hinder collusive practices, prohibits employers from sharing compensation information with each other, employers are still allowed to acquire and use more aggregated data provided by third parties. This practice of using market pay data to identify the typical market salaries for an internal position is known as *salary benchmarking*. According to historical accounts, salary benchmarking has long been central to pay setting strategies (Adler, 2020a). In our own survey of members of the U.S. Society for Human Resource Managers, 87.6% report using salary benchmarks to set pay. Interviews with HR executives also indicate that salary benchmarking plays a prominent role in their pay-setting practices (Adler, 2020b). Even the Human Resources textbooks dedicate entire chapters on how to use salary benchmarking tools (e.g., Berger and Berger, 2008; Zeuch, 2016).

Despite their ubiquity, salary benchmarking tools rarely make their way into public view, and their broad application has not been studied by economists. Understanding how these tools affect pay-setting can shed light on how labor markets operate in practice. Furthermore, the effects of these tools are of direct interest to policy-makers, who are seeking to determine whether "the information exchange may have an anticompetitive effect that outweighs any procompetitive justification for the exchange" (White House, 2021).

Our analysis focuses on the compensation of new hires. We provide a simple theoretical framework based on a standard model of competitive bidding (Milgrom and Weber, 1982b). Our model provides an economic rationale for why firms care about salary benchmarks, and it generates testable predictions. More precisely, we model the market for new hires as a first-price private values auction in which firms bid for employees. This captures two key aspects of our setting. First, making a higher offer raises the probability it is accepted, but at the cost of a higher salary. Second, firms make offers without knowing what the competing offers are, creating a role for salary benchmarks.

In our model, each firm j observes its own marginal revenue of hiring worker i to fill position X, a random variable that we denote V_{ji}^X (the "value"). Each firm chooses a bid b_{ji} for worker i, the worker accepts the highest bid, and the highest bidder receives profit $V_{ji}^X - b_{ji}$. We assume that the marginal revenues within each position are **affiliated**. In essence, this means that if worker i is valuable to firm j, then it is more likely that workers eligible for the same position are valuable to other firms. Affiliation is a standard technical condition in auction theory, that ensures that equilibria are tractable and well-behaved even when values are correlated. This formulation allows that the joint distribution of values might be different across positions—for instance, that the marginal revenue generated by distinct bank tellers is highly correlated, but the marginal revenue generated by distinct software developers is not.

Suppose that earlier auctions have been conducted for other workers eligible for the same position, with disjoint sets of firms bidding for each worker. Let S denote the salary benchmark, which we model as the median accepted offer in the earlier auctions. To generate testable predictions, we study the *direct* effect of the benchmark. That is, one firm covertly learns S while the other firms' bidding strategies are held constant. We prove that, in expectation, the access to the benchmark information must reduce the offers at the top end of the distribution. Intuitively, if the firm was going to make a high offer even without the benchmark, then raising their offer cannot increase the probability of hiring by much. Hence, their use for the benchmark is to safely lower their offer when they were already likely to win. On the other hand, the effects at the lower end of the distribution can be positive, negative, or zero, depending on the distribution of firm values.

Additionally, we use the model to study the *equilibrium* effect of the benchmark. More precisely, we consider the thought experiment in which we move from the equilibrium with no benchmarks to the equilibrium when the benchmark is common knowledge. While we cannot test the equilibrium effects with our data, this thought experiment can be quite informative for the policy discussion. Consistent with what policy-makers may refer to as procompetitive effects (White House, 2021), we find that the equilibrium effect of salary benchmarking is to raise salaries. This result builds on a canonical result of Milgrom and Weber (1982b). Intuitively, in a first-price auction, firms exploit their private information by shading their bids below their value. The salary benchmark helps to inform firm j that firm j' has a high value, so that firm j makes higher offers, and it is less safe for firm j' to shade its bid. Thus, in equilibrium the benchmark leads to less bid-shading and hence higher salaries.

In a first-price auction, firms each firm must choose its offer without knowing its competitors' offers. If instead each firm could perfectly observe and match competing offers, as in Postel-Vinay and Robin (2002) and other popular models of the wage determination process, then it is a dominant strategy to do so whenever the competing offer is below that firms' value for the worker. Thus such a model predicts salary benchmarks have no effect on the worker's final salary. As we will see, that prediction is at odds with the data.

Next, we provide empirical evidence on the effects of the benchmark tool on pay-setting. We collaborated with the largest U.S. payroll processing company serving 20 million Americans and approximately 650,000 firms. In addition to the payroll services, the company aggregates the salary data from their payroll records in the form of salary benchmarks. Clients can access these tools online, through a website. This online search tool allows firms to search for any job title they want in a user-friendly way. Currently, this benchmark tool is among the most advanced tools of their kind and is being used by many prominent firms.

Our analysis is made possible thanks to the combination of three sources of administrative data. The first dataset corresponds to the payroll records, which include detailed information such as the hire date, position and compensation. The second dataset contains information about the usage of the benchmark tool, allowing us to reconstruct which firms looked up which positions and when. Third, we have the historical data on the salary benchmarks, allowing us to observe the salary benchmarks that a firm saw (or would have seen) in the compensation explorer when searching for a specific position at a particular point in time.

Our data covers the roll-out of the benchmark tool when it was first introduced to the market. Our sample includes 586 "treatment" firms that gained access to the tool and 1,419 "control" firms that did not gain access to the tool but were selected to match treatment firms along observable characteristics. We focus on new hires that took place between January 2017 and March 2020, and during a narrow window of 10 quarters around the firm's onboarding date.

Our identification strategy is based on a differences-in-differences design. We leverage three sources of plausible exogenous variation. First, while some firms gain access to the tool, some other firms do not. Among the firms who gain access to the tool, some gain access earlier than others. And even within firms with access to the tool, some positions are searched and others are not. According to the provider of the benchmark tool, which firms end up gaining access to the tools, and when they gain access, is largely arbitrary. For example, when the benchmark tool was introduced to the market, its adoption relied heavily on direct contact from the sales representative of the payroll firm to its clients. As a result, some firms adopted earlier than others, to a great extent, due to the arbitrary order in which they were approached by the sales team. Rather than taking these sources of exogenous variation as granted, we conduct a series of empirical tests (e.g., event-study) to test their validity.

We assign each new hire into one of three categories. *Searched positions* correspond to the 5,266 unique hires in positions that are (eventually) searched in treatment firms. *Non-Searched positions* correspond to the 39,686 hires in positions that are not searched by treatment firms. *Non-Searchable positions* correspond to the 156,865 hires in control firms, who by construction could not be searched in the tool. For treatment firms, we analyze how the salaries in Searched and Non-Searched positions evolved around the date when the firm gained access to the benchmark tool. For control firms, we analyze how the salaries in Non-Searchable positions evolved around the date when the firm could have gained access to the benchmark tool: for each "control" firm we assign a "hypothetical" onboarding date, equal to the actual onboarding date of the treatment firm that is most similar in observables.

To assess whether the results were surprising or predictable, we conduct a forecast survey using a sample of 97 experts, most of whom are professors doing research on these topics. After receiving a brief explanation of the context, the experts are asked to make forecast about some of the potential effects of salary benchmarking (or lack thereof).

We start by measuring the effects of salary benchmarking on the distribution of salaries. According to the theoretical framework, there should be compression from above: firms who would have otherwise paid above the market benchmark should reduce salaries, thus moving towards the benchmark. On the other hand, the model predicts that there may be compression from below too: in some cases, but not always, firms who would have otherwise paid below the market benchmark will increase salaries, thereby moving towards the benchmark. Notably, this ambiguous prediction is present in the expert forecasts too. Some respondents predict compression from above, others from below, others from both above and below – and many others predict something entirely different from compression. Moreover, experts show low confidence in their own predictions.

Our evidence suggests that salaries get compressed towards the benchmark, both from above and below. Among Searched positions, and after gaining access to the tool, the distribution of salaries gets more compressed towards the median market benchmark. To quantify the compression effect more parametrically, we construct a dependent variable equal to the absolute %-difference between the employee's starting salary and the corresponding market benchmark. This formula is closely related to a common measure of dispersion in statistics and economics: the Mean Absolute Percentage Error.[1] Among Searched positions, the dispersion to the benchmark was on average 19.8 pp before the firms gained access to the tool. After gaining access to the tool, the dispersion dropped from 19.8 pp to 14.9 pp. This drop is not only highly statistically significant (p-value<0.001), but also large in magnitude, corresponding to a 25% decline. Moreover, our event-study analysis indicates that these effects on salary compression coincide precisely with the timing of access to the benchmark: the compression was stable in the quarters before the firm gained access to the tool, dropped sharply in the quarter after the firm gained access, and remained stable at the lower level afterwards.

Next, we use the Non-Searched and Non-Searchable positions as two alternative control groups, in a differences-in-differences fashion. Because the firms never see the relevant benchmark, we should not expect compression towards the benchmark for Non-Searched positions. We show that, indeed, the compression around the benchmark is stable before the firm gains access to the tool, and remains stable at the same level after the firm gains access to the tool. Next, we use Non-Searchable positions as an alternative control group. Because firms cannot see the benchmarks for the Non-Searchable positions, we should not expect compression towards the benchmark either. We show that, indeed, the compression around the benchmark is stable before the (hypothetical) onboarding date, and remains stable at that same level afterwards. Comparing the evolution of Searched positions to each of these control groups yields treatment effects that are similar in magnitude and statistically indistinguishable from each other. The fact that the results are consistent across these two identification strategies is reassuring. Moreover, these results are robust to a host of additional validation checks.

While our estimated effects on compression are economically significant, they are probably a lower bound on the *true* effects due to various sources of attenuation bias. We also note that this average effect masks substantial heterogeneity. We categorize positions by skill levels. We define low-skill

[1]More precisely, the relevant "error" in our context is be the difference between the employee's starting salary and the corresponding benchmark (i.e., the median salary for that position).

positions as those that typically require no more than a High-School diploma, that typically employ younger employees and with modest pay. Around 42% of the sample is classified as low-skill, and the remaining 58% as high-skill. Some examples of low-skill positions are Bank Teller and Receptionist, and some examples of high-skill positions are Ophthalmic Technician and Software Developer.

When we break down the effects on salary compression by skill levels, we observe large and statistically significant differences, with stronger effects in the low-skill positions. In low-skill positions, dispersion around the benchmark drops from 14.5 pp to 8.7 pp (p-value<0.001), equivalent to a 40.0% decline. By comparison, for high-skill positions the change in dispersion is smaller, dropping from 21.9 pp to 18.9 pp (p-value=0.021), a 14.6% decline. This finding is in sharp contrast to the expert forecasts, which predicted that the effects would be concentrated on high-education positions. However, this finding is largely consistent with the anecdotal accounts in interviews with compensation managers, according to which low-skill positions are treated as commodities and thus should be paid the market rate (Adler, 2020b). This heterogeneity is also predicted by our model. Salary benchmarks may be more informative for low-skill positions because there is less heterogeneity across workers in the marginal revenue they generate for competing firms. If marginal revenue is less heterogeneous across workers, then the salaries of past workers are more informative about the offers for present workers, resulting in larger reactions to the benchmark.

The above evidence suggests that the use of salary benchmarks has a significant effect on the wage determination process. The natural next question is what the average effects of this practice may be. Is salary benchmarking having a negative effect on the average salary? Is salary benchmarking helping companies to retain their newly hired employees? These average effects can be quite relevant for employers, as it may indicate whether it is in their best interest to use salary benchmarks. These average effects can be particularly relevant to policy-makers too, as it may provide hints on who are the winners and losers from salary benchmarking. The empirical evidence is particularly valuable, as the model provides ambiguous predictions. This ambiguity is present in the expert forecasts too: only a minority of experts feel confident about the effects on average salary. And the expert forecasts vary widely, with some predicting negative effects and others positive effects.

To estimate the average effects of salary benchmarking, we use the same identification strategy from the analysis of compression described above. The key difference is that, instead of using salary compression as dependent variable, we use other outcomes, such as the salary level or retention. Our evidence suggests that, for the average employee, and regardless of the specification, salary benchmarking does not have a negative effect on the average salary. For the whole sample, the effect on the average salary is positive, but small in magnitude and statistically insignificant. When considering the low-skill positions, the evidence points to a modest increase in the average salary. Depending on whether the Non-Searched or Non-Searchable positions are used as control, the gains in average salary are estimated at 5.0% (p-value=0.014) and 6.7% (p-value=0.001), respectively. We also find evidence suggesting that, among low-skill positions, the gains in average salary were followed by an increase in retention rates, measured as the probability that the employee is still working at the firm 12 months after the hiring date. Depending on whether the Non-Searched or Non-Searchable positions are used as control, the gains in retention probability are estimated at 6.6 pp (p-value=0.101) and 6.8 pp (p-value=0.029), respectively. The ratio between the effects on average salary and retention imply a retention elasticity that is consistent with the average estimates in the literature (e.g. Sokolova and Sorensen, 2021). This evidence suggests that firms may be using salary benchmarking to raise some salaries in an effort to improve, among other things, the retention of their employees.

This study contributes to various strands of literature. First and foremost, we contribute to the fields of labor economics, personnel economics and management by measuring the effects of salary benchmarking tools. In spite of their widespread use, there is no evidence on their effects. We fill that gap by providing the first casual estimates. Moreover, to the best of our knowledge, ours is the first study to analyze the effects of business analytic tools more generally. The existing literature is either theoretical (Blankmeyer et al., 2011; Duffie et al., 2017) or descriptive (Schiemann et al., 2018).[2]

This study is related to a recent but growing body of literature on pay transparency. Evidence from field experiments and natural experiments indicate that *employees'* perceptions about others' salaries affects a variety of employee outcomes such as satisfaction, effort, turnover and pay (Card et al., 2012; Mas, 2016; Cullen and Pakzad-Hurson, 2016; Mas, 2017; Breza et al., 2018; Dube et al., 2019; Perez-Truglia, 2020; Cullen and Perez-Truglia, 2022; Bennedsen et al., 2022; Duchini et al., 2022; Baker et al., 2023). Relatedly, there is work documenting significant misperceptions of employees about salaries, even the salaries of coworkers at the same firm (Cullen and Perez-Truglia, 2022, 2018; Caldwell and Harmon, 2018; Caldwell and Danieli, 2021; Jäger et al., 2021; Roussille, 2021). This literature is, however, entirely focused on the information frictions on the employee's side. The whole literature implicitly assumes that the transparency policies operate by affecting the information that employees can see. We contribute to this literature by showing that firms too, even the large ones, face significant information frictions. Our evidence suggests that some of the documented effects of transparency policies may be driven by the beliefs and decisions of firms, not just employees.

This project is also related to a small but growing literature on "behavioral firms" (DellaVigna and Gentzkow, 2019), more specifically on a series of biases in setting wages such as rounding (Dube et al., 2018), wage anchoring (Hjort et al., 2020; Hazell et al., 2021) and downward wage rigidities (Kaur, 2019; Grigsby et al., 2021). While the existing evidence is focused on *optimization* frictions, we contribute to this literature by showing direct evidence that firms face *information* frictions.

Our study is also related to a literature on dispersion in wages for similar workers, and more specifically studies attributing variation in wages to firm wage setting policies (Abowd et al.,

[2] One notable exemption is Grennan and Swanson (2020), which is discussed below. More broadly, our findings are related to the effects of information technology (e.g., Jensen, 2007).

1999; Mortensen, 2005). Canonical models in this literature start from the premise that workers have limited information about the wages that firms are offering, and as a consequence employers engage in a wage setting game that results in differentiated offers. Recent empirical advances in this literature focus on measuring firm-specific premiums and rent sharing elasticities (Card et al., 2018). Our evidence suggests that firm-level pay setting decisions do in fact impact the extent of wage dispersion among observably similar workers; however, we highlight information frictions on the firm side as a novel factor contributing to this dispersion.

Finally, our study is related to a literature on auction theory and industrial organization. On the theoretical side, Milgrom and Weber (1982a) and Milgrom and Weber (1982b) study what happens when bidders in an auction can observe private and public signals. On the empirical side, Tadelis and Zettelmeyer (2015) conducted a field experiment in wholesale automobile auctions and show that disclosing quality information about the goods being auctioned leads to higher revenues. Luco (2019) provides evidence that, in the context of retail gasoline industry, a policy of online price disclosure increased the average margins. And Grennan and Swanson (2020) provides evidence that, in the context of U.S. hospitals, access to a web-based benchmarking database has a significant effect on price negotiations for health services.

The rest of the paper proceeds as follows. Section 2 presents the theoretical predictions. Section 3 describes the institutional context, data and research design. Sections 4 and 5 present the empirical results. The last section concludes with implications for researchers and policy-makers.

2 The Model

We study the wage offer process as a first-price auction with affiliated private values, a canonical model due to Milgrom and Weber (1982b). This modeling approach captures two key features of our setting. First, modeling the process as a first-price auction implies that higher offers increase the probability it is accepted, but at the cost of raising their salary.[3] Second, the affiliated values assumption allows wage benchmarks to matter, in the sense that they can convey information about the distribution of competing offers that a firm faces.[4][5]

To start with, consider one worker and $n \geq 2$ firms. Each firm j has a value for the worker that is a real-valued random variable V_j, known to that firm. This captures the marginal revenue that the worker would generate at firm j. The salary benchmark S is a real-valued random variable.

[3]Alternative auction formats such as second-price auctions and English auctions do not exhibit this trade-off, since raising the winning firm's offer, holding all other offers fixed, has no effect on the worker's salary. In such formats, information about other firms' offers is strategically irrelevant.

[4]If firms' values are drawn independently across workers according to a known distribution, then past salaries convey no further information about current offers.

[5]Appendix C.5 provides empirical support for affiliated values.

We interpret this as capturing past offers made by other firms for similar workers.

We assume that V_j has support on interval $[\underline{v}, \overline{v}]$ with $0 \leq \underline{v} < \overline{v} < \infty$, and that S has support on some arbitrary interval $\mathcal{S}$. Let $f(s, v_1, ..., v_n)$ denote the joint density of the benchmark and the firm values. We assume that the density f is symmetric in its last n arguments and uniformly continuous with respect to each v_j.

We assume that the random variables $(S, V_1, ..., V_n)$ are **affiliated**, as we now define. Let $z, z' \in \mathbb{R}^k$ for some integer k. $z \vee z'$ denotes the component-wise maximum, and $z \wedge z'$ the component-wise minimum. Random variables are *affiliated* if for all z and z', their joint density f satisfies

$$f(z \vee z') f(z \wedge z') \geq f(z) f(z'). \tag{1}$$

Example 2.1. *There are two firms, and the marginal revenue of worker i to firm j is*

$$V_j = f(Q + Q_i + Q_{ij}) \tag{2}$$

where $f: \mathbb{R} \to \mathbb{R}_{\geq 0}$ is a continuous increasing function, and Q, Q_i, and Q_{ij} are random variables. Q is a position-specific component, Q_i is a worker-specific component, and Q_{ij} is a match-specific component, all independent and normally distributed. Each firm observes its marginal revenue V_j, but not the individual components. Then the variables (Q, V_1, V_2) are affiliated.

We define another random variable $Y_1 \equiv \max_{j \neq 1} V_j$. Let $f_{Y_1}(x | v, s)$ denote the density of Y_1 conditional on $V_1 = v$ and $S = s$, with cumulative distribution $F_{Y_1}(x | v, s)$.

By a standard argument[6], affiliation implies that $\frac{f_{Y_1}(x|v,s)}{F_{Y_1}(x|v,s)}$ is non-decreasing in s. We use $f_S(s|v)$ to denote the density of S conditional on $V_1 = v$, with cumulative distribution $F_S(s|v)$.

We start by studying the **no-benchmark equilibrium** of the first-price auction. Each firm j observes V_j and then chooses a bid b_j. Firm j's payoff is equal to $(V_j - b_j)$ if $b_j > \max_{k \neq j} b_k$ and 0 otherwise.[7] A standard argument, adapting the proof of Theorem 14 of Milgrom and Weber (1982b), yields this characterization of the equilibrium:

Theorem 2.2. *There exists a symmetric no-benchmark equilibrium of the first-price auction. The equilibrium strategy $b^*: [\underline{v}, \overline{v}] \to \mathbb{R}$ is strictly increasing and satisfies the first-order linear differential equation defined by*

$$b^*(\underline{v}) = \underline{v}, \tag{3}$$

$$b^{*\prime}(v) = (v - b^*(v)) \frac{E[f_{Y_1}(v | v, S) | V_1 = v]}{E[F_{Y_1}(v | v, S) | V_1 = v]}. \tag{4}$$

We assume that the benchmark is **locally relevant**, meaning that for all v, there exists s such that $0 < F_S(s|v)$ and

$$\frac{f_{Y_1}(v | v, s)}{F_{Y_1}(v | v, s)} < \frac{E[f_{Y_1}(v | v, S) | V_1 = v]}{E[F_{Y_1}(v | v, S) | V_1 = v]}. \tag{5}$$

[6]Lemma 1 of Milgrom and Weber (1982b).

[7]Ties are zero-probability events, so the analysis does not depend on the tie-breaking condition.

This condition essentially requires that the benchmark is informative about the ratio $\frac{f_{Y_1}(v|v,s)}{F_{Y_1}(v|v,s)}$. If firm 1 were to slightly reduce its bid from $b^*(v)$, the cost is a reduced probability of winning, proportional to $f_{Y_1}(v\,|\,v,s)$. The benefit is that firm 1 pays less if it wins, which is proportional to $F_{Y_1}(v\,|\,v,s)$. So local relevance implies that the benchmark is informative about the expected profits from slightly changing 1's bid.

We now study the direct effect of the benchmark. That is, suppose that firm 1 covertly observes S before placing its bid, while believing that the other firms continue to bid according to b^*. Consider the informed firm's best-response correspondence, which depends on its value V_1 and the benchmark S,

$$\operatorname*{argmax}_{b\geq 0} E[(V_1-b)\mathbb{1}_{\{b^*(Y_1)<b\}}\,|\,V_1=v,S=s]. \tag{6}$$

The correspondence (6) is monotone non-decreasing in v, by Topkis's theorem. Let $\tilde{b}(v,s)$ be an arbitrary selection from (6).

The next theorem states that if the firm's value is high enough, then covertly observing the benchmark strictly reduces its expected bid as well as its expected payment.

Theorem 2.3. *There exists $\tilde{v}<\overline{v}$ such that for all $v>\tilde{v}$, we have that*

$$E\Big[\tilde{b}(v,S)\Big|V_1=v\Big]<b^*(v), \tag{7}$$

and also that

$$E\Big[\tilde{b}(v,S)\mathbb{1}_{\tilde{b}(v,S)\geq b^*(Y_1)}\Big|V_1=v\Big]<b^*(v). \tag{8}$$

The proof is in Appendix A.1.

For lower quantiles of the distribution, one can show using the boundary condition (3) that the direct effect is non-negative as V_1 approaches $\underline{v}$. Formally, there exists a selection from the informed firm's best-response correspondence that satisfies

$$\lim_{v\downarrow\underline{v}}\Big(E\Big[\tilde{b}(v,S)\Big|V_1=v\Big]-b^*(v)\Big)\geq 0. \tag{9}$$

However, without further assumptions on the distribution, one cannot sign the direct effect at low quantiles, in the sense that we now state.

Remark 2.4. *There exist affiliated joint distributions of $(S,V_1,...,V_n)$ such that the direct effect is positive on an open interval $(\underline{v},\tilde{v})$ for $\underline{v}<\tilde{v}$. There also exist affiliated joint distributions of $(S,V_1,...,V_n)$ such that the direct effect is negative on an open interval $(\underline{v},\tilde{v})$ for $\underline{v}<\tilde{v}$.*

These results motivate the following prediction:

Prediction 2.5 (Direct effect)**.** *Gaining access to the benchmark will reduce salaries at high quantiles of the distribution, but not necessarily at low quantiles of the distribution.*

The next theorem states that if the firm's value is high enough, then covertly observing the benchmark strictly reduces its expected probability of hiring the worker.

Theorem 2.6. *Suppose that the joint density f is strictly positive everywhere on $[\underline{v},\overline{v}]^N \times \mathcal{S}$. There exists $\tilde{v} < \overline{v}$ such that for all $v > \tilde{v}$, we have that*

$$P\Big(\tilde{b}(v,S) \geq b^*(Y_1)\Big| V_1 = v\Big) < P(b^*(v) \geq b^*(Y_1)|V_1 = v). \tag{10}$$

The proof is in Appendix A.2. Suppose that a firm was making high wage offers before observing the benchmark. Theorem 2.6 predicts that such a firm will become less likely to hire the worker after observing the benchmark. However, it is possible that firms that were initially making low wage offers will become more likely to hire, and the sign of the overall effect on hiring probability is an open question.

In our data, individual firms gain access to the salary benchmark, which does not tell us what would occur if many firms gained access to the benchmark, and if it was common knowledge that they used the benchmark to set salaries. To speak to this question, we now examine the model's predictions for the new equilibrium that arises when all firms observe the benchmark and best-respond to each others' bidding strategies.

Let $b^{**} : [\underline{v},\overline{v}] \times \mathcal{S} \to \mathbb{R}$ be the symmetric equilibrium strategy in a first-price auction after all bidders observe the benchmark. This is characterized by the first-order linear differential equation

$$b^{**}(\underline{v},s) = \underline{v}, \tag{11}$$

$$b^{**\prime}(v,s) = (v - b^{**}(v)) \frac{f_{Y_1}(v\,|\,v,s)}{F_{Y_1}(v\,|\,v,s)}. \tag{12}$$

Theorem 2.7. *The equilibrium with the benchmark yields higher expected salaries than the no-benchmark equilibrium, that is*

$$E\Big[\max_i b^{**}(V_i,S)\Big] \geq E\Big[\max_i b^*(V_i)\Big]. \tag{13}$$

Theorem 2.7 is a special case of Theorem 16 of Milgrom and Weber (1982b). Both the benchmark equilibrium and the no-benchmark equilibrium lead to the same winner, namely the firm with the highest value V_i. But in the benchmark equilibrium, each firm's bid b^{**} is increasing in the benchmark S, which is affiliated with that firm's private information V_i. In this way, the benchmark strengthens the statistical linkage between the bid $b^{**}(V_i,S)$ and the firm's private information V_i, reducing the firm's information rents and raising salaries.

2.1 Extensions to the Model

In Appendix B, we provide a number of extensions, which we summarize below. So far we have assumed that there is only one signal S, and examined comparative statics from allowing one firm to observe S, and allowing all firms to observe S. But the firms in our data already had access to other, arguably less accurate, salary benchmarks before gaining access to the one that we study. In Appendix B.1, we extend the model to allow for multiple signals, some of which the firms already observe, and find that the same comparative statics hold for the effects of observing an *additional* signal.

The baseline model treats the benchmark as an arbitrary signal S such that $(S,V_1,...,V_n)$ are affiliated; the results do not require further structure on S. However, the benchmarks in our data are not arbitrary—in particular, we focus on the median of past salaries for each position-title. Why should this be affiliated with the marginal revenue each firm has for the current worker? In Appendix B.2, we provide a simple foundation for affiliated signals. We consider a sequence of auctions, imposing that firms' values for the current worker are affiliated with other firms' values for past workers. In equilibrium, the median of the winning bids in past auctions is affiliated with firms' values in the present auction. Hence, rather than an exogenous signal S, we can regard S as being determined by the equilibrium offers made by other firms to similar workers.

Our model can also make predictions about heterogeneous effects by position type, such as low-skill vs. high-skill. Intuitively, low-skill positions are easier to standardize and monitor, so any two workers in that position can provide similar productivity. Appendix B.3 shows that, if we model low-skill positions as having less individual productivity variation, then the prediction is that the benchmark will have a stronger effect on this group.

We modeled firms making simultaneous offers to each worker. This was an intentional design choice, as a dynamic model would have significantly complicated the setting. In reality, firms may be motivated to use benchmarks, among other things, because of employee retention. In Appendix B.4 we show that the theoretical predictions hold in a stylized model of retention concerns.

In our baseline model, each firm treats the benchmark as exogenous. In practice, some employers may be large enough to influence the market benchmarks. In essence, their wages "pass-through" to the benchmark and hence affect its competitors' offers. While these cases are rare, we observe them in our data.[8] In Appendix B.5 we study a model of benchmark pass-through, finding that the benchmark has ambiguous effects on the large firm's equilibrium offers, but consistently reduces the large firm's equilibrium profits.

[8] See for example Appendix J and Appendix K.

3 Institutional Context and Data Sources

3.1 Background on Salary Benchmarking

The use of salary benchmarks dates back to 1980s (Adler, 2020a).[9] This practice can be found in the private as well as public sectors (Faulkender and Yang, 2010; Thom and Reilly, 2015), and is used for all echelons of the organization, even executive pay.[10] Many Human Resources handbooks dedicate entire chapters to the practice of salary benchmarking.[11] The following excerpt from one of those textbooks provides an illustration of the type of trade-offs that HR professionals have in mind when using salary benchmarks:

> *"Using surveys to benchmark compensation levels ensures that the pay levels determined by the organization are not extraordinarily misaligned with market practice – i.e., pay is not too low or too high. Determining the appropriate amount of compensation is a balancing act. No organization wants to waste their financial resources by paying too high relative to the market; and those who pay too low risk unwanted turnover from employees looking for a better deal elsewhere."* – Berger and Berger (2008), p. 125.

The earliest forms of salary benchmarks were compensation surveys administered by consulting firms. To meet these demands, some personnel management consultants grew specialized in providing market data through compensation surveys, with some notable examples being Abbott, Langer and Associates, Korn Ferry, Hayes Group, Mercer, Radford, and Willis Towers Watson. In the last decade, some tech companies started offering online tools that allow employees and employers to find information about the market salaries in specific positions. Some of these websites, such as Glassdoor, Comparably, and LinkedIn, have become popular because they allow anyone to conduct searches for free. These websites rely primarily on crowdsourcing: i.e., employees who visit the website can fill out a quick survey reporting their pay at their current or past companies. These data are probably not the highest quality, among other things, because of biases in who decides to self-report their salary, whether they self-report it truthfully, and also the limited number of observations.[12]

---

[9] Adler (2020a) argues that the use of external benchmarks was, at least in part, motivated by a need to reduce the firms' liability for discrimination lawsuits.

[10] In 2006, the Securities and Exchange Commission issued a new disclosure requirement, requiring companies to state whether they engaged in "any benchmarking of total compensation, or any material element of compensation, identifying the benchmark and, if applicable, its components (including component companies)" (Securities and Exchange Commission, 2006). In fiscal year 2015, over 95% of the S&P 500 companies disclosed a peer group of firms that they used to benchmark executive salaries against (Larcker et al., 2019).

[11] For example, Chapter 48 from Zeuch (2016) is dedicated to the "Essentials of Benchmarking." And Chapters 9 and 10 of Berger and Berger (2008) are dedicated to "Salary Surveys" and "Benchmarking".

[12] There are other online tools that require a paid subscription, such as Salary.com and Payscale.com. These other tools are based mainly on data from traditional salary surveys.

More recently, the largest U.S. provider of payroll services started to offer data analytics tool to their clients, including but not limited to salary benchmarking tools. Payroll data is arguably the highest-quality data one could think of to construct salary benchmarks. Any error in payroll is immediately corrected as it impacts someone's day to day life. The most comparable data is probably tax records, but tax records fall short of payroll records in terms of frequency, accuracy and detail. For example, payroll records include information about the position title of the employee, which is missing from tax records. And while tax records include the gross taxable income of the employee, it does not show the critical break down by base salary, commissions, bonuses, etc. The payroll data has even bigger advantages over salary surveys and crowd-sourced data, which raise flags about the smaller sample sizes, measurement error and biases due to selection into the survey. Moreover, due to the massive sample sizes of payroll, covering several millions of employees at any point, salary benchmarks are much more precisely estimated. And due to the high-frequency nature of the payroll data, the benchmarks can be updated more frequently.

Salary benchmarking is part of the broader phenomenon of people analytics, brought about by growth in business data capacity. HR functions at leading companies leverage data to attract and retain talent, predict employee turnover, identify talent shortages, and other aspects of workforce planning (Davenport and Shapiro, 2010). In a survey of more than 10,000 HR and business leaders across 140 countries implemented by Deloitte in 2017, 71% of companies saw people analytics as a high priority in their organizations, and recruiting came up as the highest-priority area of focus within that (Collins et al., 2017). HR has come to be one of the most data-driven functions at major companies (Davenport, 2019).

3.2 Survey on Uses of Salary Benchmarking

To provide evidence on how firms use salary benchmark tools, and to assess the validity of assumptions of the model, we conducted a survey of U.S. Human Resources professionals in collaboration with the Society for Human Resource Management (hereafter referred to as the SHRM survey). We collected responses from July 15, 2022 to July 20, 2022, using SHRM's Voice of Work Research Panel. From a sample of 9,537 panelist, we had 2,696 responses, for a response rate of 28.3%. As a filter to access the main module of the survey, we asked them if they participate in setting salaries for employees (2,085 respond affirmatively) and, then, if they use salary benchmarks (1,827, or 87.6%, respond affirmatively).[13] From these, 1,350 complete the entire survey and constitute our main sample – all of the results reported below are based on this sample. The sample broadly covers firms across industries and size in both the public and private sector. More details on

[13]The magnitude of this estimate is consistent with the results from an industry survey of 5,003 U.S. firms: 96.3% of them reported that they use some form of salary benchmarking to inform their compensation strategy and structure (PayScale, 2021).

the implementation of the survey, sample characteristics and results are provided in Appendix C.[14]

The results from our survey provide some useful facts to guide our analysis and identification strategy. A majority of employers (72.3%) use multiple sources to obtain market data on salaries. The most popular sources are industry surveys and free online data sources (68.0% and 58.1% of participants, respectively, indicate they use these). Other popular options are government data (37.1%), paid online data sources (34.4%), compensation consultants (26.3%) and payroll data services (23.2%).[15]

Our survey also elicited how employers use the benchmarks, and how frequently. A vast majority (97.4%) of respondents use the salary benchmarks for setting the pay of new hires.This fraction is calculated for respondents that participate in setting pay for new hires and in organizations that use salary benchmarking. There is a lot of variation in how frequently they use the benchmark information, though. Only 36.6% of respondents report using benchmarks to set salaries for all of their new hires, with the rest using it for some, but not all, of the new hires. Using an open-ended question, we asked respondents why they use the salary benchmark in some cases but not others. There is a wide range of answers that vary substantially across employers. For example, some respondents said that they consult the benchmark for positions in which they have less hiring experience.[16] Setting the salary of new hires is by no means the only use of salary benchmarks. A vast majority of the employers report using the salary benchmark for their existing employees too, and again they typically use it for some employees but not for all of them. And they are other uses for the benchmark tools, such as financial planning for headcount.

3.3 The Compensation Explorer Tool

The study builds on an ongoing collaboration with the largest payroll processing firm in America, a publicly-traded firm with a current market cap of $72.5 billion. This company provides payroll services for 650,000 firms, including many prominent ones, for a total of 20 million employees. In addition to providing payroll services, this firm uses the massive payroll data from its clients to provide business analytic tools as a subscription service. In this study, we are interested in the *Compensation Benchmark Tool*, consisting of a search engine to view detailed compensation statistics. To better illustrate how the compensation explorer works, Figure 1 provides a screenshot of this online tool.[17] The online tool allows the user to browse the benchmarks in different ways. Most prominently, there is a search bar at the top of the screen.

One challenge for the creators of this tool was to aggregate data across different job titles. For

[14] The full survey instrument is attached as Appendix L.

[15] Among our survey respondents, 9.5% use the compensation explorer offered by our partner organization.

[16] For more details, see Appendix C.

[17] This is a screenshot of how the tool looked like in 2020. There have been some changes to the tool during the period of study, but the overall look and functionality remained similar.

example, one company might call a job "warehouse handler," another might call the same job "inventory handler" or "material handler." The firm converts the raw position titles into a standardized taxonomy with the use of machine learning tools for probabilistic matching, and the firm directly seeks approval of matches from clients, creating new inputs for the algorithm. Our data includes a match score that reflects the quality of the match between the firm-specific job title and the title in the taxonomy.[18] Until August 2020, which covers the vast majority of our sample (95.7%), the company used a taxonomy that spanned 2,236 distinct position titles.[19] To understand the granularity of this taxonomy, the taxonomy includes 31 position titles for "teacher" that distinguish between preschool, primary, secondary, middle school, substitute, and special education teachers. There are on average 3.84 unique position titles per each 6-digit Occupational Information Network (O*NET) code.

Users can search by the position names in the proprietary taxonomy using auto-complete functionality. Because this is the default option, the vast majority of the search results originate through the proprietary taxonomy, comprising around 70.9% of the searches. Additionally, a drop-down menu allows users to search using alternative taxonomies, such as the client company's own position titles (22.6% of searches) or O*NET taxonomy (6.5% of searches).[20]

Once the user selects a position title, the tool provides a job description. For illustrative purposes, we will use the position of "Accountant," which is the same example featured in Figure 1. The tool describes the "Accountant" position with the following tasks: "(i) Maintains the accounting operations for a department within the organization; (ii) Checks and verifies records, prepares invoices, vouchers, and filings; (iii) Posts ledgers and general journal entries and balances all records related to accounts receivables and payable; (iv) Assists the financial services manager with accounting and administrative duties; (v) Undertakes responsibility for financial analysis and administration or overseeing the projects occasionally." The job description also includes information about the typical qualifications of the candidate, which in the case of an accountant are: "Requires an undergraduate degree or equivalent experience. For some jobs this may also require a graduate degree or additional certification. This is typically a knowledge worker who applies information and judgment in a specific area to achieve results and solve problems."

Once a position has been selected, the benchmark tool provides rich data on compensation statistics for that position. The most salient figure is the median base salary, the first estimate shown in the screen and highlighted (e.g., highlighted in purple in the bottom panel of Figure 1). The screen also shows the sample size upon which statistics are based, namely the number of

[18] We restrict our main sample to observations with match scores above the 20th percentile match score in each quarter. The results are similar without this restriction (for details, see Table F.2 and Table G.2).

[19] Starting September 2020, the company switched to a new taxonomy that expanded the number of position titles. Since our main sample stops at March 2020, our baseline results are not affected by this change.

[20] O*NET searches are excluded from our analysis because we lack historical O*NET benchmarks.

organizations and the number of employees.[21] The fact that the tool highlights the median base salary is no coincidence, as conversations with the product team indicate that this is the metric that their clients are most interested in, and also the type of information highlighted in handbooks on Human Resources (e.g., Berger and Berger, 2008; Zeuch, 2016).[22]

The compensation tool defines base salary clearly, and in a manner consistent with other compensation studies (Grigsby et al., 2021). For salaried employees, the base pay is the yearly base salary (i.e., before commissions or bonuses). For the hourly employees, the annual base salary is defined as the annual equivalent of hourly pay: i.e., the hourly wage times 40 hours times 52 weeks.[23] The vast majority of the total cash compensation comes from base salary.[24] While the median base salary is the most salient piece of information, the tool offers more comprehensive information about pay. In addition to the median, the tool shows a chart with additional information about the distribution of base salary (see the bottom of Figure 1): the 10th, 25th, 75th and 90th percentiles, as well as the average. Likewise, in addition to base salary, the tool allows the user to learn about bonuses, overtime and total cash compensation.

The tool also allows the user to apply filters to the set of employers and employees included in the benchmark. For instance, users can click on drag-and-drop menus to zoom into a specific industry, or they can use a map to filter by geography, for example by clicking on their own state. However, these filters are only available to the extent that there is enough data, more precisely at least 5 other firms collectively hiring at least 10 employees in the position of interest – for instance, if you tried to zoom in by industry and state, and that left you with an insufficient sample size, you would not be able to see the statistics.

When using this and other benchmarking tools, the user has many pieces of information to pick from. For example, the user can look at the median salary or the average salary, and the user can apply filters (e.g., by state) or use the unfiltered results. Our survey data gives us a glimpse of the typical usage patterns. When asked about what pieces of information they typically care about (e.g., median, 25th percentile), the most popular choice was the median (ranked in first place by 56.73% of the respondents), with the second most popular piece of information being the average salary (ranked first by 32.59% of respondents). And regarding the filters applied when

[21] The tool also indicates the specific date to which the statistics refer, and it even shows some information about the change of the median salary during the past 12 months. The benchmark is generally stable on quarter-over-quarter basis. For example, the median absolute quarter-over-quarter change in the benchmark is 1.12%.

[22] There is also some evidence that employees and labor platforms, not just employers, pay special attention to median salaries (Roussille, 2021).

[23] In our sample, 81.1% of the employees are hourly and the rest are salaried.

[24] In addition to base salary, employees may receive other forms of compensation such as bonuses and commissions, observed in payroll and reported in the benchmark. 93.2% of the total cash compensation comes in the form of base salary. A negligible fraction of positions (<1%) receive less than 60% of total cash compensation as base salary. However, our data does not include equity compensation, which may be a significant part of compensation for some employees, especially at the executive level.

searching for a salary benchmark, the most popular choices are to filter by industry and by state (87.33% and 84.15% of participants indicate they typically apply these filters, respectively).

3.4 Data Sources

We have access to the following datasets:

Payroll Database: this is the key dataset covering all employees in a firm, including the new hires, and with a monthly frequency, from January 2017 through July 2021. It includes detailed information about the position of the employee, exact hire date and compensation details. Our main focus of interest is the base salary, but we also have additional information such as on bonuses. The data include employee characteristics such as gender and age.

Tool Usage Database: this is the key dataset that indicates which positions were searched for and which were not. These data track the web navigation of clients using the benchmark tool. The data include a timestamp for each search, and the position searched. Due to the firm's pre-existing data storage policy, we have access to data starting on September 2019 and until August 2021.[25]

Benchmark Database: this is the database that allows us to reconstruct the search result for each search that we observe in the tool usage dataset, and is available from the first quarter of 2017 through the second quarter of 2021.[26] This database contains the compensation benchmarks, at each point in time and for all positions.[27] As explained in Section 3.3 above, users can apply filters for their search results. The usage data does not indicate which filters the user applied, or whether they applied any filters at all. In our baseline specification, we assume that, whenever there is enough data, subjects applied filters by state and industry, which were by far the most popular filters according to the SHRM survey.[28] Likewise, we do not know whether the user was looking at the median salary, the average salary or some of the other statistics. In the baseline specification, we focus on the median salary, again because that was the most popular choice in the SHRM survey. In any case, we show that the results are robust under alternative specifications.

There are some additional details about the data that deserve mention. To prevent the influence of outliers, we winsorize all dependent variables in the analysis. For example, in the baseline

[25]Due to the default setting in the tool, the company would automatically delete the usage data older than 6 months. For this reason, we do not have access to this data prior to the date when we pull data for the first time.

[26]Unfortunately, due to reasons outside of our control, we do not have the benchmark data for the second quarter of 2020, and thus we will always have to exclude this period from the analysis. In any case, since that quarter was the worst-hit from the COVID pandemic, we would have excluded that period from the baseline analysis anyways.

[27]We restrict our sample to employees in positions with available benchmark information, regardless of whether the information was looked up by the firm or not.

[28]More precisely, in the baseline specification we assume the firm applied the filters by same State and same industry, but only if after applying those filters the benchmark is calculated with at at least 30 employees.

specification, we winsorize the outcome of absolute dispersion at +/-75 percentage points.[29] To minimize concerns about seasonality in hiring of some positions, the baseline specification re-weights observations to maintain the same composition across Standard Occupational Classification (SOC) groups over time.[30] Last, we complement the administrative data with other sources. For example, we categorize positions by mapping the O*NET codes to well-known crosswalks.[31]

For the heterogeneity analysis, we categorize positions by skill levels. We define low-skill positions as those that usually require no more than a High-School diploma, that employ younger employees and with modest pay. More precisely, we construct the low-skill group in two steps. First, we map O*NET codes to identify positions in job zones 1 and 2, which commonly require no more than a high school diploma.[32] Second, we exclude positions in which the average worker is above 31 years of age and has an average annual salary above $30,000. Roughly 42% of the sample is classified as low-skill, and the remaining 58% as high-skill. Some examples of low-skill positions in the sample are Bank Teller, Hand Packer and Receptionist, while some examples of high-skill positions are Ophthalmic Technician, Production Operations Engineer and Software Developer.[33]

3.5 Sample of New Hires

Our theoretical and empirical analysis focuses on new hires. Our survey of hiring managers indicates that one of the primary uses of the tool is setting salaries of new hires. Indeed, this view is supported by the anecdotal accounts of the partner organization and the analysis of utilization data.[34] Most importantly, the focus on new hires simplifies both the theoretical and empirical analysis. For instance, firms must always set a salary when hiring a new employee – in contrast, firms revise the salaries of their existing employees infrequently and, even when doing so, they may be subject to constrains such as downward wage rigidities.[35]

The theoretical framework from Section 2 provides a stylized version of hiring new employees, where employers make a take-it-or-leave-it offer. In that model, the salary benchmarks inform that first offer. In practice, however, the hiring process is more nuanced and, as such, the information on salary benchmarks may be used at different steps of the process. For example, the firm may

[29]Moreover, we drop outlier observations: employees with annual base salaries over $2,000,000 or below $1,000. Moreover, for the analysis of effects on salary levels, we winsorize the base salary at the 2.5 and 97.5 percentiles within the relevant position.

[30]More precisely, for each position type, we compute the distribution of SOC groups in the month before onboarding and re-weight all the other periods to match that distribution.

[31]For more details about the data, see Appendix D.

[32]For 27% of observations there is no job zone classification available. In those cases, we impute education using data from Zippia.com: positions are classified as low-education or high-education depending on whether the share of employees with at most a high school degree is below or above 10%.

[33]For more details and examples on this categorization, see Appendix D.2.

[34]Results presented in Appendix E.

[35]In Appendix H, we present some additional results for a sample of existing employees.

find that information useful later in the hiring process, when deciding whether to respond to a counter-offer.[36] Or the information may also come in handy earlier in the hiring process, to post wages in job advertisements. For example, using data from Burning Glass, Hazell et al. (2021) reports that 17% of the job ads include a posted wage or wage range.

Our main sample of interest consists of new hires from January 2017 through March 2020. We stop at March 2020 for several reasons, most importantly because we want to avoid our baseline results from being affected by the COVID pandemic. In any case, we show that the results hold when we expand the sample to include new hires after March 2020 – for more details, see Appendix F. Since we are interested in what happens around the date when the firm gains access to the tool, we restrict our sample to a window of 10 quarters around the date of onboarding: i.e., up to 5 quarters before the onboarding date, and up to 5 quarters after the onboarding date.

3.6 Firms in the Sample

The salary benchmarking tool is only available to the payroll clients that subscribe to the cloud services, which launched in late 2015.[37] We observe the exact date each client was granted access to the tool. Anecdotally, which firms are granted access to the business analytic tools, and when they do so, depends on many arbitrary factors. During the roll-out, account managers were instructed to introduce the tool to business clients at any opportunity, such as calls pertaining to payroll and other services. Nearly all firms that gain access to the business analytics service did not search for the service or request it, but rather, their account manager introduced them to business analytic services as part of a broader conversation. Our empirical tests comparing the evolution of firm characteristics as a function of adoption corroborate anecdotes that dissemination was as good as random.

Our main sample comprises 586 firms that gained access to the tool, with onboarding dates between December of 2015 and January of 2020. The vast majority of these firms used the tool at least once.[38] Among the firms with access, we have suggestive evidence that the tool was being used by a small set of employees – most likely members of the Human Resources or compensation teams.[39]

We obtained data on an additional 1,419 firms that never gained access to the tool but were selected to match observable characteristics of firms that did get access to the tool: number of

[36] As suggestive evidence that this channel may play a role, 16.4% of the companies surveyed by PayScale (2021) report that they shared their own benchmarking data with their employees.

[37] The benchmarks themselves are based on payroll records for all clients of the payroll company, not just the ones subscribing to the cloud services.

[38] More precisely, among the 586 firms with access to the tool, 561 (96%) conducted at least one search during the period for which we have usage data.

[39] For a subset of the utilization data, we observe an identifier for the person conducting the search. For 50% of the firms with access to the tool, there is a single user conducting the searches. Even in firms with multiple users, the searches are concentrated: if you take a random pair of searches, there is a 58.2% probability that they were conducted by the same user. These results have to be taken with a grain of salt, however, as it is possible that one account is being shared by multiple employees, or that one employee is looking up the data by request of other employees.

employees, state and 6-digit industry codes. We assigned a "hypothetical" on-boarding date to the firms that never gain access to the tool. For each control firm, we find the firm with access that is most similar in observable characteristics, and assign the date when that firm obtained access as the hypothetical access date for the control firm.[40] For example, if Ford gains access but Fiat does not, we assume Ford would have gained access when Fiat did.[41]

Table 1 presents more descriptive statistics for the firms in our sample. Column (1) shows that the average firm employs 503 employees, 45.3% of which are female, and the average employee is 34 years old and earns a salary of $46,945. Columns (2) and (3) breaks down these average characteristics by whether firms that gain access to the tool (i.e., treatment firms) and firms that do not gain access (i.e., the control firms). Due to the large sample sizes, the pairwise differences are often statistically significant. However, these differences tend to be modest or negligible in magnitude. This finding should not be surprising, given that we asked the partner institution to select control firms that are similar to the treatment firms. Columns (4) and (5) break down the treatment firms in the top half and bottom half based on a measure of higher versus lower utilization of the benchmark tool. Again, firms with high vs. low utilization look very similar to each other in these observable characteristics.

Appendix D.1 provides a comparison between the firms in our sample and a representative sample of U.S. firms. In terms of size, measured in number of employees, our sample is most representative of the top quartile of firms in the United States. In terms of salaries, the employees in our sample are representative of the population of U.S. employees, with the exception that our sample has limited coverage of the bottom quartile of the distribution (earning less than $20,000 per year). Our sample also provides broad coverage of all the U.S. industries.

3.7 Classification of New Hires

Based on the utilization data, we assign each new hire to one of the following three groups:

- <u>Searched Positions:</u> positions in treatment firms that were either searched in the compensation explorer prior to the hire date or that they will be eventually searched in the tool.

- <u>Non-Searched Positions:</u> positions in treatment firms that were not searched. One potential concern with the classification is that some searched positions may be incorrectly attributed

[40] More precisely, for each control firm, we restrict to all treatment firms in the same sector, and then select the firm which is closest according to the Mahalanobis distance for firm size and state.

[41] We use Ford and Fiat purely for illustration purposes, as we work with de-identified data and thus do not know the names of any of the companies in our sample.

as Non-Searched. This may be due to the limited window of the searched data,[42] or due to information spillovers. For example, assume a company hires an "accountant" and an "accounting analyst", and searched for the benchmark of "accountant" (and thus this is a Searched position) but not for the "accounting analyst" (the Non-Searched position). Perhaps the two positions are close enough that the company is using the benchmark for "accountant" to set pay for the "accounting analyst" too. In this case, the comparison between Searched and Non-Searched would yield a null effect of the benchmark only because "accounting analyst" is incorrectly being classified as non-searched. To minimize the scope for information spillovers, we exclude from the Non-Searched positions all the new hires in positions "adjacent" (i.e., in the same SOC group) to those new hires that *were* searched in the same month.

- Non-Searchable Positions: all positions in the control firms (i.e., those that never gain access to the tool).

The utilization data shows that while firms have access to the benchmark tool, that does not mean that all firms use it, or that they use it all the time. Consider the 534 firms who had onboarded prior to the last quarter of 2019. During that quarter, 199 (37.3%) of these firms hired in at least one position. These firms searched the benchmark for 20.8% of the positions in which they hired.[43] For this reason, there are substantially more new hires categorized as Non-Searched than as Searched. Also, since our sample includes more control firms than treatment firms, we have an even larger number of new hires in the Non-Searchable category. Our final sample includes 5,266 new hires in the Searched category, 39,686 new hires in the Non-Searched category and 156,865 new hires in the Non-Searchable category.

In our sample of new hires, we observe 329 unique positions in the Searched category. Table 2 lists the 35 most common positions from this sample. The 3,129 hires in these 35 positions account for a majority (66.7%) of the hires in the Searched category. These common Searched positions include all sorts of occupations such as Bank Teller, Hand Packer and Software Developer. Table 2 also reports the number of employees being hired in each position, and number of hiring firms, broken down by whether the hire falls into the categories Searched (column (1)), Non-Searched (column (2)) and Non-Searchable (column (3)). This table shows that there is quite a bit of overlap in the positions that different firms are searching for. For example, the 468 hires for Customer Service Representative in the Searched category are distributed across 44 different firms. This table also shows that there is a lot of overlap across the Searched, Non-Searched and Non-Searchable categories:

[42] For example, it is possible that some positions are being attributed to the Non-Searched category because they were not searched after the start of the usage data (September 2019), yet perhaps they were searched prior to September 2019.

[43] More precisely, around 62.3% of these firms did not search for any of the positions in which they hired; among the remaining firms, they looked up on average 55.2% of the positions in which they hired.

i.e., for each firm that searches for a given position (column (1)), there are many other firms hiring in that position that did not conduct a search because they choose not to (column (2)) or because they didn't even have the choice (column (3)). For example, while there are 468 new hires Customer Service Representative in the Searched category, there are 4,401 hires for that same position in the Non-Searched category and 4,012 in the Non-Searchable category. In other words, these positions are *searched* the most often, largely because those are the positions in which firms *hire* the most often.

Column (1) of Table 3 shows the average characteristics of the employees in the sample of new hires. The average employee is 35 years old, 50.6% of them are female, 81.1% work for an hourly wage, they have an starting annual base salary of $41,359 and an average median market benchmark that is almost identical ($41,412). The salaries differ from their corresponding median benchmarks (in absolute value) by an average of 20.4%. The last rows show the main occupation groups in the sample: 19.8% of the positions are in Office and Administrative Support, 8.0% in Management, 6.6% in Production, 9.3% in Transportation and Material Moving, 4.8% in Building and Grounds Cleaning, and the rest (51.5%) belong to other groups.

Next, we can compare the characteristics across treatment and control groups. As usual in differences-in-differences designs, the key (testable) assumption is that, prior to the onboarding date, the outcome of interest *evolved* similarly between treatment and control groups. As a result, it should not matter whether the treatment and control groups start at difference baselines, or whether they are different in observable characteristics. However, it is always re-assuring to check that the differences between the treatment and control groups are not large. Columns (2) through (4) of Table 3 break down the average characteristics for each of the three categories: Searched, Non-Searched and Non-Searchable. Perhaps the two most important characteristics are the (pre-treatment) salary and its absolute %-difference with respect to the benchmark, because they constitute the outcome variables in the analysis that follows. The differences are economically small. For example, the average salaries are $39,064, $42,013 and $41,405 in the Searched, Non-Searched and Non-Searchable categories, respectively. Due to the large sample sizes, the difference between the Searched and Non-Searchable groups is statistically significant (p-value = 0.013), despite modest differences in economic terms. The difference between the Searched and Non-Searched group is not significant (p-value = 0.617). For the other characteristics, the pairwise differences are again almost always statistically significant, but they tend to be economically small. Some exceptions are that, relative to Non-Searched and Non-Searchable positions, Searched positions have a higher share of female employees and a higher share of office and administrative support positions.

4 Effects on Salary Compression

To begin, we will examine the impact of salary benchmarking on the distribution of salaries around the benchmark. Based on our theoretical framework, we expect to see compression from above, meaning that companies that would have otherwise paid higher than the market benchmark will decrease salaries and move closer to the benchmark. On the other hand, the model makes an ambiguous prediction as to whether there would be compression from below as well.

4.1 Non-Parametric Estimates: Histograms

To start, we will conduct a non-parametric analysis of the data using simple histograms. More precisely, we look at the distribution of the difference between the salaries chosen by the employers and the benchmarks they saw (or could have seen) in the benchmark tool. The results of this analysis are presented in Figure 2. Each panel corresponds to a different position type. In each panel, the x-axis denotes the difference between the starting salary and the corresponding benchmark (i.e., the median market pay). For example, the middle bin corresponds to salaries that are close (±2.5%) to the median benchmark, the bins on the left half of the figure correspond to salaries below the benchmark and the bins on the right half correspond to salaries above the benchmark.

Panel A of Figure 2 corresponds to the Searched positions, with gray bars corresponding to employees who were hired before the firm gained access to the benchmark tool (i.e., when the benchmark information *was not* visible to the firm) and the red bars correspond to employees hired after the onboarding date (i.e., when the benchmark information *was* visible to the firm). The comparison between the two histograms from panel A suggest that, after onboarding, salaries are more compressed around the benchmark. More precisely, we observe compression from above as well as compression from below: there is a decline in the probability of observing salaries above the benchmark, as well as a decline in the probability of observing salaries below the benchmark. This evidence is consistent with our theoretical model, which predicted that there should be compression from above and that there could be compression from below as well. We can also compare this result to the responses in our prediction survey: only a minority of experts were able to predict this finding (for details, see Appendix I).[44]

One simple way of summarizing the compression to the benchmark is by noticing that firms are more likely to "hit" the benchmark: the probability that the firm chooses a salary close (±2.5%) to the median benchmark increases from 11.6% before onboarding to 22.1% after onboarding. Another way of summarizing the dispersion around the benchmark is by means of the absolute mean difference. This metric suggests that, among Searched positions, salaries were on average 19.4 pp from the benchmark before the firms gained access to the tool. After gaining access to the tool,

[44]The survey itself is included in Appendix M.

the average distance to the benchmark dropped from 19.8 pp to 14.9 pp, a change that is highly statistically significant (p-value<0.001) and also large in magnitude (equivalent to a 24.7% drop).

For the purposes of placebo exercises, we use the Non-Searched and Non-Searchable positions as two alternative control groups. The results for Non-Searched positions are presented in Panel B of Figure 2. Because the firms never see the benchmark, we should not expect compression towards the benchmark for Non-Searched positions. The dispersion around the benchmark is similar in magnitude in the pre-onboarding period (20.8 pp) to the post-onboarding period (22.0 pp). Due to the large sample sizes this difference is precisely estimated and thus statistically significant (p-value<0.001). However, the difference is small in magnitude, and much smaller than the corresponding difference for the Searched category reported in Panel A. In turn, Panel C of Figure 2 presents the results for the Non-Searchable positions. Because firms cannot see the benchmarks for the Non-Searchable positions, we should not expect compression towards the benchmark for this category. We find that, again, dispersion around the benchmark is similar in magnitude in the pre-onboarding period (21.1 pp) as in the post-onboarding period (21.9 pp). Due to the large sample sizes, the difference is again statistically significant (p-value<0.001). However, most importantly, the difference is negligible in magnitude. We employ a second identification strategy corroborating the result by exploiting quasi-random changes to the benchmark, and tracking the time it takes for treatment and control firms to converge to the new benchmark. For details, see Appendix K.

As an additional robustness check, we embedded an experiment in our SHRM survey. In a nutshell, we ask the participants to pick two positions for which they are planning to hire in the future and we elicit the annual base salary they are willing to offer for these new hires. Next, we provide them (hypothetical) information on the median salary benchmark for that position. Participants are randomly assigned to receive either a benchmark that is 15% above their initial salary offer, or 15% below. After the respondent received the benchmark information, we re-elicit the salary they are willing to offer for that position. The results from this survey experiment are largely consistent with the results presented above. More precisely, the experiment shows that salary offers get compressed towards the benchmark, both from above and from below (see Appendix C for more details).

We find that firms want to "aim" for the median market pay. Ex-ante, one could have expected that, instead, firms would have preferred to be stingy, for example, by "aiming" for the 25th percentile of market pay instead of the median. For a more direct comparison, Appendix F reproduces the analysis but, instead of using the median benchmark, it uses each of the alternative benchmarks: 10th, 25th, 75th and 90th percentile of market pay, and the average too. The results confirm that firms are, for the most part, aiming for the median market pay. This evidence is consistent with the SHRM survey where 56.73% of the HR managers ranked the median salary as the piece of information they care about the most when searching for a position benchmark. [45]

[45] For more details, see Appendix Table C.2.

Last, it is worth noting that our model makes a prediction about the distribution of salaries among those bids that get accepted, and this is precisely what we test with our data. Additionally, it would be interesting to estimate the effects on the distribution of all bids. For instance, it is possible that some firms who were planning to make an offer below the benchmark, after looking up the benchmark information, end up deciding not to hire at all. Unfortunately, we do not have sufficient data to test these additional hypotheses.

4.2 Econometric Model

Next, we extend the above analysis to a more traditional differences-in-differences design. Let subscript t denote time, subscript i index employees, and subscript j index firms. Let $\omega_{i,j,t}$ be the starting base salary of employee i when hired by firm j at time t. And let $\bar{\omega}_{i,t}$ denote the corresponding benchmark: i.e., the median base salary that the search tool indicates for employee i's position at time t. Let $Y_{i,j,t}$ denote the outcome variable. For example, in this section the outcome of interest is the absolute difference between the salary of the employee and the benchmark: $100 \cdot |\frac{\omega_{i,j,t} - \bar{\omega}_{i,t}}{\bar{\omega}_{i,t}}|$. This outcome is multiplied by 100 so that the effects can be interpreted readily in percentage points.

We have two distinct differences-in-differences designs: one based on the comparison between Searched and Non-Searched positions, and the second one based on the comparison between Searched and Non-Searchable positions. For the sake of brevity, we will use Θ_1 to refer to observations categorized as either Searched or Non-Searched, and Θ_2 to the set of observations categorized as either Searched or Non-Searchable. Let $T_{i,j}$ be a dummy variable that takes the value 1 if the employee i's position at firm j was categorized as a Searched position, and 0 if it was categorized as Non-Searched or Non-Searchable. Let $A_{j,t}$ be a dummy variable that takes the value 1 if firm j has access to the benchmark tool in period t and 0 otherwise. This variable will take the value 0 in every month until the month of on-boarding, after which it will always take the value 1.[46] Finally, let δ_t denote year dummies, ψ^k denote position dummies and $X_{i,j,t}$ denote a vector of additional controls consisting of the employee's age, a dummy for gender, and a dummy for hourly pay. And let $\epsilon^k_{i,j,t}$ be the standard error term – unless stated otherwise, all of the analysis in this paper uses standard errors that are clustered at the firm-position-month level. Consider the following regression specification:

$$Y_{i,j,t} = \alpha^k_1 \cdot A_{j,t} \cdot T_{i,j} + \alpha^k_2 \cdot A_{j,t} + \alpha^k_3 \cdot T_{i,j} + X_{i,j,t}\alpha^k_4 + \delta^k_t + \psi^k + \epsilon^k_{i,j,t}, \forall \{i,j,t\} \in \Theta_k \tag{14}$$

When $k=1$, equation (14) boils down to the first identification strategy, based on the comparison between Searched and Non-Searched groups. When $k=2$, equation (14) boils down to the second identification strategy, based on the comparison between Searched and Non-Searched. In both cases, the differences-in-differences coefficient of interest is α^k_1, which measures the effect

[46]In the case of control firms, this would correspond to the "hypothetical" onboarding date.

of the benchmark tool. When $k = 1$, α_1^1 measures the difference in outcomes between Searched (treatment) and Non-Searched (control) in post-onboarding relative to pre-onboarding. When $k = 2$, α_1^2 measures the difference in outcomes between Searched (treatment) and Non-Searchable (control) groups in post-onboarding relative to pre-onboarding.

These two alternative differences-in-differences designs are based on different control groups, and as such they may different advantages and disadvantages. For example, one advantage of using Non-Searchable positions as control group is that it is not subject to the potential concern about misattributing Searched positions as Non-Searched positions (as discussed in Section 3.7 above).[47] While we do not have a strong preference for one strategy versus the other, we believe that being able to compare the results across the two strategies provides a meaninful validation check for the research design.

As a formal test of pre-trends, we follow the standard practice in differences-in-differences designs of introducing a "fake" post-treatment dummy ($A_{j,t}^{\text{fake}}$) that is identical to the true post-treatment dummy ($A_{j,t}^{\text{fake}}$) except that it takes value 1 in the two quarters before the onboarding date. We can expand equation (14) as follows:

$$Y_{i,j,t} = \alpha_1^k \cdot A_{j,t} \cdot T_{i,j} + \alpha_2^k \cdot A_{j,t} + \alpha_3^k \cdot A_{j,t}^{\text{fake}} \cdot T_{i,j} + \alpha_2^k \cdot A_{j,t}^{\text{fake}} + \alpha_4^k \cdot T_{i,j} + X_{i,j,t}\alpha_5^k + \delta_t^k + \psi^k + \epsilon_{i,j,t}^k, \forall \{i,j,t\} \in \Theta_k \tag{15}$$

The coefficient of interest is α_3^k, which measures if the difference in outcomes between Searched (treatment) and Non-Searched (control) groups was already changing even before the onboarding date. Under the null hypothesis of no differences in pre-trends between treatment and control groups, we expect this coefficient to be zero. Furthermore, we can extend the econometric framework to an event-study analysis, by expanding $A_{j,t}$ into a set of dummies. Let $A_{j,t}^s$ be a dummy variable that takes the value 1 if the firm onboarded on period $t - s$. For example, $A_{j,t}^{+1}$ would take the value 1 one quarter post-onboarding, while $A_{j,t}^{-4}$ would take the value 1 four quarters prior to onboarding. And let S be the set of non-zero integers between -5 and +5, except for -1 (the reference category).[48] We expand equation (14) as follows:

$$Y_{i,j,t} = \sum_{s \in S} \alpha_{1,s}^k \cdot A_{j,t}^s \cdot T_{i,j} + \sum_{s \in S} \alpha_{2,s}^k \cdot A_{j,t}^s + \alpha_3^k \cdot T_{i,j} + X_{i,j,t}\alpha_4^k + \delta_t^k + \psi^k + \epsilon_{i,j,t}^k, \forall \{i,j,t\} \in \Theta_k \tag{16}$$

The set of coefficients $\alpha_{1,s}^k \, \forall s \in S$ correspond to the event-study coefficients. For example, $\alpha_{1,+1}^k$ would correspond to the effect one quarter post-onboarding (relative to the base category, one quarter

[47]One example of a potential advantage of using Non-Searched positions as the control group is that it would not be subject to the potential concern of picking up effects from other tools besides the compensation explorer.

[48]In all the analysis, we drop observations for employees who were hired in the exact month of onboarding. Due to the coarseness of the timestamps, it would be impossible for us to distinguish between the hires that were post- vs. pre-onboarding.

pre-onboarding), while $\alpha^k_{1,-4}$ would correspond to the "placebo effect" four quarters pre-onboarding.

4.3 Differences-in-Differences Estimates

Figure 3 presents the event-study analysis. The left panels (A and C) correspond to the comparison between Searched and Non-Searched categories, while the right panels (B and D) correspond to the comparison between Searched and Non-Searchable categories. In each of the panels, the y-axis corresponds to the salary dispersion around the benchmark. The y-axis starts at 0, which is the minimum value that the outcome can take, corresponding to the extreme case in which all salaries are exactly equal to their corresponding market benchmarks. The higher the value of the y-axis, the more different salaries are from the benchmark. For example, a value of 20 would mean that salaries differ from the benchmark, on average, by 20%. The x-axis corresponds to the time since the date of onboarding, from -5 (i.e., 5 quarters prior to the month of onboarding) to +5 (i.e., 5 quarters after to the month of onboarding). To make the interpretation of the effect sizes more straightforward and intuitive, we follow Hastings and Shapiro (2018) by normalizing the y-axis. In this and all other event-study graphs, all coefficients are shifted by the same constant, as to match the average of the baseline outcome in the pre-treatment period. That's the reason why the coefficient for quarter -1 is the omitted category yet its value is different from 0.

The results from Figure 3 indicate that the effects on salary compression coincide precisely with the timing of access to the benchmark: the dispersion with respect to the benchmark was stable in the quarters before the firm gained access to the tool, dropped sharply in the quarter after the firm gained access, and remained stable at the lower level afterwards. More precisely, Panel A of Figure 3 shows the evolution of the outcome separately for Searched (denoted in red dots) and Non-Searched (blue squares) positions. For the Searched positions, the dispersion with respect to the benchmark was stable at around 19.8 pp prior to the onboarding, but then dropped sharply to around 14.9 pp in the quarter after onboarding and remained stable at that lower level afterwards. In contrast, the compression in Non-Searched positions was stable around 20.8 pp prior to onboarding, and remained stable at a similar level (22.1 pp) after the onboarding date. Panel C of Figure 3 corresponds to the difference between the two series from Panel A. This differences-in-differences comparison suggests that the benchmark tool reduced the salary dispersion from 19.8 pp to 14.8 pp (p-value<0.001), equivalent to a 25.3% reduction.

Regarding the second identification strategy, Panel B of Figure 3 provides to the comparison between Searched (denoted in red dots) and Non-Searchable (purple squares) positions. While the outcome dropped sharply after onboarding for Searched positions, it remained stable around the date of onboarding for Non-Searchable positions. Panel D of Figure 3 correspond to the difference between the two series in panel B. The differences-in-differences estimate suggests that the benchmark tool reduced the salary dispersion from 19.8 pp to 13.6 pp (p-value<0.001). The drop in dispersion from panel D (6.2 pp) is close in magnitude to the corresponding drop from panel C (5 pp) – furthermore, these two effects are statistically indistinguishable from each other. The fact that the results are similar across the two identification strategies is re-assuring about their validity of the research design.

4.4 Robustness Checks

Table 4 presents the differences-in-differences estimates in table form, which summarizes the differences-in-differences results in fewer coefficients. This simpler approach maximizes the statistical power and also is more practical for the purpose of comparing the results across alternative specifications. Panel A of Table 4 presents the post-treatment coefficients (α_1^k from equation (14)). Column (1) of Table 4 corresponds to the baseline specification. The post-treatment coefficients are negative and statistically significant: -4.775 (p-value<0.001) when using Non-Searched positions as control group, and -6.149 (p-value<0.001) when using Non-Searchable positions as control. In turn, Panel B presents the corresponding "pre-treatment" coefficients (α_3^k from equation (14). Consistent with the assumption of no differences in pre-trends, the pre-treatment coefficients in column (1) are close to zero (-0.346 for the comparison with Non-Searched positions and -0.310 for the comparison with Non-Searchable positions), statistically insignificant (p-values of 0.749 and 0.604, respectively) and precisely estimated.

Each of the columns (2) through (12) of Table 4 are identical to column (1), except that they change a different feature of the baseline specification. In columns (2) and (3), we use alternative versions of the dependent variable. In column (2), we measure dispersion using the log difference: $100 \cdot |log(\omega_{i,j,t}) - log(\bar{\omega}_{i,j,t})|$. This outcome is multiplied by 100, just like the outcome from column (1), so that it can be interpreted (approximately) in percentage points. The results from column (2) are qualitatively and quantitatively consistent with the results from column (1). In column (3), we measure dispersion with a dummy variable that takes the value 100 if the salary is over 10% away from the benchmark, and 0 otherwise. Again, the results are both qualitatively and quantitatively similar between columns (1) and (3). For example, the first post-treatment coefficient from column (1) suggests that, relative to the baseline, dispersion dropped by 24.1% ($= \frac{4.775}{19.812}$), while the corresponding coefficient from column (3) suggests a decline of 25.5% ($= \frac{16.270}{63.732}$).

The specification from column (4) of Table 4 is different from column (1) in that it is winsorized at $\pm$ 100% instead of $\pm$ 75%. Column (5) uses heteroskedasticity-robust standard errors instead of clustered standard errors. Column (6) does not include any of the additional control variables. Column (7) adds position fixed effects. Column (8) adds firm fixed effects. Column (9) excludes positions for which the base salary is not a major component of compensation: Waiter/Waitress, Chauffeur, and Bartender/Mixologist. Column (10) restricts the sample to include only the positions that appear in the list of 329 Searched positions. Column (11) does not re-weight observations by SOC groups. Column (12) restricts the sample to individuals aged 21 through 60. In all these alternative specifications, the results are both qualitatively and quantitatively similar to those from column (1).

In Appendix F, we present some additional results and robustness checks. For instance, Appendix F.3 show that the results are robust for a range of additional specifications such as extending the sample after March-2020, and Appendix F.2 shows that there are no significant effects on the composition of new hires.

4.5 Heterogeneity by Skill Level

The above analysis estimates the average effects of salary benchmarking across all sorts of positions, which may mask substantial heterogeneity. In particular, the distinction between low-skill and high-skill positions is quite prominent in interviews with HR professionals. Low-skill positions often involve standardized task, minimal training, and can be easily monitored. For these reasons, as one HR practitioner put it, workers in those jobs are "viewed as interchangeable" (Adler, 2020b). This leads to what Adler (2020b) calls *standardization*: once a candidate is deemed qualified for the job, their pay is a function of the job, not their individual characteristics. As a result, low-skill candidates are offered the market rate as a take-it-or-leave it offer, with an unwillingness to negotiate – efforts by the candidate to ask for more not only rejected, but even viewed as inappropriate (Adler, 2020b).

On the other hand, the value that high-skill employees create may be very different for different employees and for different companies. For example, a Software Engineer may be an excellent fit for some firms, and thus create a lot of value, but a poor fit for others, and thus create less value. In interviews, HR practitioners emphasize the importance of tailoring offers to specific candidates in high-skill jobs (Adler, 2020b). The employer may still look up and use the salary benchmark as a starting point, but there are other factors that can come into play, such as the line manager's opinion of the candidate, the candidate's own salary history, outside offers and salary expectations. Consistent with this view, survey data suggests that high-skill candidates are substantially more likely to engage in salary negotiations relative to low-skill candidates (Hall and Krueger, 2012).

As shown in Section 2 above, the theoretical model can naturally accommodate these differences between low-skill and high-skill jobs. Low-skill jobs can be modeled as auctions for workers whose productively is more common across firms hiring in these positions. As a result, our model predicts that salary benchmarking would have stronger effects for low-skill jobs than for high-skill jobs.

To explore this source of heterogeneity, Figure 4 breaks down the baseline results from Figure 2 by low-skill and high-skill positions. The panels on the left hand side of Figure 4 (A, C and E) correspond to the low-skill positions, while panels on the left hand side (B, D and F) correspond to high-skill positions. The top panels (A and B) correspond to the Searched positions. A comparison between these two panels indicates stark differences by skill level. Even before the firms had access to the tool (gray bins), there was a lot more compression among the low-skill positions (Panel A) than among the high-skill positions (Panel B). This evidence is consistent with the idea of standardization, according to which employers are trying to pay as closely as possible to the market pay in low-skill positions. Most importantly, the drop in compression is markedly sharper for low-skill positions than for high-skill positions. Among low-skill positions (Panel A), dispersion drops from 14.5 pp to 8.7 pp (p-value<0.001), corresponding to a 40% drop. For high-skill positions (Panel B), dispersion falls from 24.0 pp to 20.5 pp (p-value=0.021), corresponding to a drop of just 14.6%.[49]

For the placebo tests, panels C through F of Figure 4 reproduce the analysis for Non-Searched and Non-Searchable positions. As expected, the differences in compression between post-onboarding and pre-onboarding salaries are always small in magnitude – however, due to the large sample sizes, they are sometimes statistically significant.

An alternative explanation for the above heterogeneity is that when firms look up low-skill positions, they are interested in learning about base salary, but when they look up high-skill positions firms may be more interested in other forms of compensation (such as bonuses and commissions). However, this is unlikely to explain our results, considering that base salary comprises the vast majority of compensation in both lower and higher skill positions.[50]

We can compare our heterogeneity findings to the responses to the prediction survey. The experts predicted the opposite of what we find: a majority predicted benchmarking would more strongly influence high-education positions (see Appendix I for more details).

4.6 Interpretation of the Magnitude of the Effects

The effect of benchmarking on salary compression documented above is not only highly statistically significant, but also large in magnitude. Next, we discuss some reasons why those results may under-estimate or over-estimate the true magnitude of the effects.

On the one hand, our results may lead to under-estimation of the effects due to multiple sources of attenuation bias. The first source is that the tool offers many metrics (e.g., median salary, with filters or without filters) but we do not know exactly which metric each employer is paying closest attention to. Another source of attenuation bias is that in some cases, even though an employer hired in a given position, it may have looked up the benchmark for that position not to negotiate with the new hire, but to negotiate the salary of an existing employee in that same position. Likewise, when multiple people get hired in a particular firm-position, our specification

[49] In Appendix F.4 we report a more formal test of the difference by skill levels. Using the differences-in-differences framework, we show that the difference in effects between low-skill and high-skill groups is large both when using Non-Searched and Non-Searchable as control groups. However, while the difference is statistically significant when using the Non-Searched group as control (p-value=0.070), it is statistically insignificant (p-value=0.403) when using the Non-Searchable group as control.

[50] According to the benchmark data, among low-skill positions, 95.2% of the total cash compensation comes in the form of base salary. For high-skill positions, the corresponding figure is 92.9%. One caveat, however, is that our measure of total compensation does not include equity compensation, which may be important at the highest levels of the organizations (e.g., executives) and also in some some specific jobs (e.g., in the technology sector).

is implicitly assuming that the firm will use that information for everyone who gets hired in that position going forward. However, perhaps the manager was looking that information up for one specific new hire (e.g., someone with an outside offer). A last source of attenuation bias is that the tool we study is not the only source of data on market values, so firms in the treatment and control groups may be using other sources of data on market salaries. Therefore, our estimates should be interpreted as intention-to-treat effects from adding one source of benchmark information.

On the other hand, it is possible that our results over-estimate the importance of salary benchmarks. To the extent that the effects can be heterogeneous across positions, we are estimating a treatment effect on the treated. In other words, we estimate the effects of salary benchmarking for positions that end up being searched. Had they been searched, the effects could have been different for positions that were not searched. For example, following the logic of rational inattention, it could be argued that firms are looking up the positions for which they value the information the most. If they value the information the most, they are arguably more likely to use. In that case, our estimates for the positions that are looked up may overestimate the strength of information frictions for the average position. Nevertheless, the fact that we estimate treatment effects on the treated is not necessarily a limitation. On the contrary, for some purposes, the treatment effects on the treated may be the most relevant object of interest. For example, from the perspective of policy implications, the counterfactual of interest is not what would happen if all firms were "forced" to look up every position, but what would happen if all firms had access to look up the positions they want. In that sense, the treatment effects on the treated are the right object of interest.

5 Average Effects of Salary Benchmarking

The above evidence suggests that the use of salary benchmarks has a significant effect on the salary *compression* around the benchmark. Next, we explore the potential effects on the *average* salary level. From the perspective of the theoretical model, the predictions are ambiguous: when one firm gains access to the tool, the effect on the average salary could be negative, positive or zero. Additionally, to shed light on the causal mechanisms, we explore the effects on the retention of new hires.

5.1 Effects on Salary Levels

To estimate the average effects of salary benchmarking, we use the same identification strategy from Section 4 above. The key difference is that, instead of using salary dispersion as the dependent variable, we use other outcomes, most importantly the average salary.[51]

[51] When interpreting the magnitude of the effect on average salary, it is useful to note that these results are not subject to one of the sources of attenuation bias described in Section 4.6 above: this analysis does not require data on the benchmarks that the firm saw in the platform, so it is not subject to that source of measurement error.

The event-study results for the salary levels are presented in Figure 5. This figure is identical to Figure 3, except that the y-axis is the average salary (in logs). The left panels (A and C) use the Non-Searched positions as control group, while the right panels (B and D) use the Non-Searchable positions as controls. The evidence from Figure 5 suggests that, for the average employee, salary benchmarking has, if anything, a small positive effect on the average salary. Panel A of Figure 5 corresponds to the comparison between Searched (denoted in red dots) and Non-Searched (blue squares) positions. During the pre-onboarding period, the average salary was stable in both Searched and Non-Searched positions. In the post-onboarding period, both the Searched and Non-Searched positions continued at their pre-onboarding levels. Panel C of Figure 5 corresponds to the difference between the two series in panel A. This differences-in-differences estimate suggests that there is no significant effect of salary benchmarking on the average salary. More precisely, access to the tool had an effect on the average salary that is virtually zero (-0.002 log points, or equivalent to an effect of just 0.2%),[52] and statistically insignificant (p-value=0.756).

Regarding the second identification strategy, Panel B of Figure 5 corresponds to the comparison between Searched (denoted in red diamonds) and Non-Searchable (purple circles) positions. Again, the average salary evolved similarly before and after the onboarding date, both for Searched and Non-Searchable positions. Panel D of corresponds to the difference between the two series in panel B. This differences-in-differences comparison indicates that access to the tool had a slight positive effect on the average salary (0.017 log points, equivalent to a 1.7% increase), but the effect is imprecisely estimated and thus statistically insignificant (p-value=0.308). The fact that the results are roughly similar across the two identification strategies (panels C and D, respectively) is re-assuring about the validity of the findings. Moreover, in Appendix F we present some additional robustness checks. For example, Appendix F.3 shows that these results are robust to a wide range of alternative specifications.

We can also compare these findings to the responses from the prediction survey. A slight majority of experts predicted a null effect on the average salary, making it their most accurate prediction (for details, see Appendix I).

Given the strong heterogeneity in salary compression between low-skill and high-skill positions reported in Section 4 above, it is natural to explore this same heterogeneity for salary levels. The results are presented in Figure 6, which reproduces the results from Panels C and D of Figure 5 but separately for low-skill positions (panels A and B) and high-skill positions (panels C and D). When considering the low-skill positions, the evidence points to a modest increase in their average salary. Depending on whether the control group is comprised of Non-Searched positions (panel A) or Non-Searchable positions (panel B), the gains in average salary are estimated at 5.0% (p-value=0.014) and 6.7% (p-value=0.001), respectively. By contrast, when considering high-skill positions, there is no evi-

[52] To be more precise, the effect is 0.2002% ($=100\cdot(exp(0.002)-1)$). Since the approximation error is so small, in the remainder of the paper we treat log-point effects and percent-effects as interchangeable.

dence of significant effects on the salary level: the average salary drops by 2.9% (in panel C) and 1.6% (panel D), but those effects are statistically insignificant (p-values of 0.119 and 0.288, respectively).

Note that the effects on average salary are largely consistent with the non-parametric analysis presented in Section 4 above. For instance, Figure the histograms in 2 show that, when considering the whole sample, benchmarking leads to compression both from below and from above the benchmark. Those negative and positive effects on salaries largely cancel each other out, resulting in an average effect that is positive but small. In turn, panel A of Figure 4 shows what happens when we focus on low-skill positions. Pre-onboarding (gray bars), the distribution of salaries is skewed towards the left of the benchmark, meaning that firms were systematically under-paying employees in this group. Thus, when salaries get compressed towards the benchmark after onboarding, that naturally ends up raising the average salary for low-skilled positions.

5.2 Effects on Retention Levels

It may seem puzzling at first that benchmarking leads employers to an increase the average salary in low-skill positions. According to our model, employers raise salaries because, while it increases the labor costs, it also improves other outcomes they care about, like the retention rate.[53] To provide evidence of this underlying mechanism, we estimate the effects of salary benchmarking on the retention of new hires. The results are presented in Figure 7. For the sake of brevity, this figure presents the results broken down by low-skill and high-skill positions – the results for the full sample are presented in Appendix G.

Figure 7 is identical to Figure 6 discussed above, except that, instead of the average salary, the dependent variable is the probability that the employee is still working at the firm 12 months after the hiring date. Panels A and B from Figure 7 suggest that, for low-skill positions, the gains in average salary were followed by an increase in retention rates. While panels C and D suggest that, for high-skill positions, for which we do not observe a significant change in average salary, we do not observe a change in retention rates either. The magnitude of the gains in retention are also worth discussing. Depending on whether the Non-Searched or Non-Searchable positions are used as control (panels A and B, respectively), the gains in retention rates for the low-skill positions are estimated at 6.6 pp (p-value=0.101) and 6.8 pp (p-value=0.029), respectively. These effects correspond to 16.1% and 16.6% of the baseline retention rate, respectively. By comparison, the corresponding gains in average salary are estimated at 5.0% (p-value=0.014) and 6.7% (p-value=0.001), respectively. These effects on average salary and retention imply labor supply elasticities of 3.22 ($=\frac{16.1}{5.0}$) and 2.48 ($=\frac{16.6}{6.7}$), respectively. Our estimates are consistent with the range of estimates found in the literature. For instance, the meta-analysis of Sokolova and

[53]Since we do not observe the job offers that were not accepted, we cannot measure the effect on acceptance rates directly. However, in Appendix G.1 we provide related evidence that uses an indirect approach.

Sorensen (2021) reports a (weighted) mean of separations-based labor supply elasticities of 3.05.

6 Conclusions

Most medium and large firms use salary benchmarking in their compensation strategies. Despite their pervasiveness, there is no evidence on the effects of these tools. To fill this gap, we provide theoretical and empirical evidence. Our model makes predictions about the effects of salary benchmarking. We then test those predictions using administrative data from one of the leading providers of payroll services and salary benchmarks in the United States. The evidence suggests that salary benchmarking has a significant effect on pay setting, and in a manner consistent with the predictions of the model. For instance, we find that access to the tool compresses salaries towards the market benchmark quite significantly, and especially in low-skill positions.

We are the first to document how firms use their salary benchmarking tools and, additionally, the effects of these tools on pay-setting. This evidence has two important implications for the understanding of labor markets. First, it shows that salary benchmarking plays an significant role in pay-setting and as such it deserves further study. Second, this constitutes direct evidence that information frictions around salaries are significant, even among medium and large firms with hundreds or thousands of employees. Furthermore, our evidence shows that firms can use big data to ameliorate their information frictions.

Our findings have implications for a current policy debate. While U.S. legislation currently allows employers to use aggregated data on market wages, that practice is regulated and has been under recent scrutiny by policy makers. For example, in 2018 two senators wrote a letter to the Department of Justice and the Federal Trade Commission requesting those agencies to revisit their guidance on salary benchmarking.[54] A 2021 statement from the White House raises similar concenrs, mentioning for example that "workers may also be harmed by existing guidance (...) that allows third parties to make wage data available to employers and not to workers" (White House, 2021).

Despite this renewed interest from policy-makers, there is no evidence about the effects of salary benchmarking and whether it leads to wage suppression. While this is the first attempt to study salary benchmarking and thus more research is needed, we make at least two contributions to this debate. First, our model provides a theoretical foundation for what the policy-makers call procompetitive effects. Second, we provide empirical evidence on the effects of salary benchmarking. While our empirical evidence cannot speak to the equilibrium effects, it shows that when one additional firm gains access to the benchmark information, that leads, if anything, to modest salary gains concentrated in low-skill positions. And in addition to employees, employers may benefit too, as we find that the salary gains are accompanied by gains in employee retention.

[54]See for example: https://www.jdsupra.com/legalnews/in-effort-to-increase-employees-26757/.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High Wage Workers and High Wage Firms. *Econometrica 67*(2), 251–333.

Adler, L. (2020a). From the Job's Worth to the Person's Price: Changes in Pay-Setting Practices since 1950. *Working Paper*.

Adler, L. (2020b). What's a Job Candidate Worth? Status and Evaluation in Pay-Setting Process. *Working Paper*.

Baker, M., Y. Halberstam, K. Kroft, A. Mas, and D. Messacar (2023). Pay Transparency and the Gender Gap. *American Economic Journal: Applied Economics, forthcoming*.

Bennedsen, M., E. Simintzi, M. Tsoutsoura, and D. Wolfenzon (2022). Do Firms Respond to Gender Pay Gap Transparency? *Journal of Finance 77*, 2051–2091.

Berger, L. A. and D. Berger (2008). *The Compensation Handbook.* New York: McGraw-Hill.

Blankmeyer, E., J. LeSage, J. Stutzman, K. Knox, and R. Pace (2011). Peer-group dependence in salary benchmarking: a statistical model. *Managerial and Decision Economics 32*(2), 91–104.

Breza, E., S. Kaur, and Y. Shamdasani (2018). The Morale Effects of Pay Inequality. *The Quarterly Journal of Economics*.

Caldwell, S. and O. Danieli (2021). Outside Options in the Labor Market. *Working Paper*.

Caldwell, S. and N. Harmon (2018). Outside Options, Bargaining and Wages: Evidence from Coworker Networks. *Working Paper*.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and Labor Market Inequality: Evidence and Some Theory. *Journal of Labor Economics 36*(S1), S13–S70.

Card, D., A. Mas, E. Moretti, and E. Saez (2012). Inequality at Work: The Effect of Peer Salaries on Job Satisfaction. *American Economic Review 102*(6), 2981–3003.

Collins, L., D. Fineman, and A. Tsuchida (2017). People analytics: Recalculating the route. *Rewriting the rules for the digital age: 2017 Deloitte Global Human Capital Trends*.

Cullen, Z. and B. Pakzad-Hurson (2016). Equilibrium Effects of Pay Transparency in a Simple Labor Market. *Working Paper*.

Cullen, Z. and R. Perez-Truglia (2018). The Salary Taboo: Privacy Norms and the Diffusion of Information. *NBER Working Paper No. 25145*.

Cullen, Z. and R. Perez-Truglia (2022). How Much Does Your Boss Make? The Effects of Salary Comparisons. *Journal of Political Economy 30*(3), 766–822.

Davenport, T. (2019). Is HR the Most Analytics-Driven Function? *Harvard Business Review Digital Article*.

Davenport, T. and J. Shapiro (2010). Competing on talent analytics. *Harvard Business Review 88*(10), 52–58.

DellaVigna, S. and M. Gentzkow (2019). Uniform Pricing in U.S. Retail Chains. *The Quarterly Journal of Economics 134*(4), 2011–2084.

Dube, A., L. Giuliano, and J. Leonard (2019). Fairness and Frictions: The Impact of Unequal Raises on Quit Behavior. *American Economic Review 109*(2), 620–663.

Dube, A., A. Manning, and S. Naidu (2018). Monopsony and Employer Mis-optimization Explain Why Wages Bunch at Round Numbers. *NBER Working Paper No. 24991*.

Duchini, E., S. Simion, and A. Turrell (2022). Pay Transparency and Cracks in the Glass Ceiling. *CAGE Working Paper No. 482.*.

Duffie, D., P. Dworczak, and H. Zhu (2017). Benchmarks in Search Markets. *The Journal of Finance 72*(5), 1983–2044.

Faulkender, M. and J. Yang (2010). Inside the black box: The role and composition of compensation peer groups. *Journal of Financial Economics 96*(2), 257–270.

Grennan, M. and A. Swanson (2020). Transparency and Negotiated Prices: The Value of Information in Hospital-Supplier Bargaining. *Journal of Political Economy 128*(4), 1234–1268.

Grigsby, J., E. Hurst, and A. Yildirmaz (2021). Aggregate Nominal Wage Adjustments: New Evidence from Administrative Payroll Data. *American Economic Review 111*(2), 428–471.

Hall, R. and A. Krueger (2012, 10). Evidence on the Incidence of Wage Posting, Wage Bargaining, and On-the-Job Search. *American Economic Journal: Macroeconomics 4*(4), 56–67.

Hastings, J. and J. M. Shapiro (2018). How Are SNAP Benefits Spent? Evidence from a Retail Panel. *American Economic Review 108*(12), 3493–3540.

Hazell, J., C. Patterson, H. Sarsons, and B. Taska (2021). National Wage Setting. *Working Paper.*.

Hjort, J., X. Li, and H. Sarsons (2020). Across-Country Wage Compression in Multinationals. *NBER Working Paper No. 26788*.

Jäger, S., C. Roth, N. Roussille, and B. Schoefer (2021). Worker Beliefs About Outside Options and Rents. *Working paper*.

Jensen, R. (2007). The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector. *The Quarterly Journal of Economics 122*(3), 879–924.

Kaur, S. (2019). Nominal Wage Rigidity in Village Labor Markets. *American Economic Review 109*(10), 3585–3616.

Larcker, D., C. McClure, and C. Zhu (2019). Peer Group Choice and Chief Executive Officer Compensation. *Stanford University, Graduate School of Business Working Paper No. 3767*.

Luco, F. (2019). Who Benefits from Information Disclosure? The Case of Retail Gasoline. *American Economic Journal: Microeconomics 11*(2), 277–305.

Mas, A. (2016). Does Disclosure affect CEO Pay Setting? Evidence from the Passage of the 1934 Securities and Exchange Act. *Working Paper*.

Mas, A. (2017). Does Transparency Lead to Pay Compression? *Journal of Political Economy 125*(5), 1683–1721.

Milgrom, P. and R. J. Weber (1982a). The value of information in a sealed-bid auction. *Journal of Mathematical Economics 10*(1), 105–114.

Milgrom, P. R. and R. J. Weber (1982b). A Theory of Auctions and Competitive Bidding. *Econometrica 50*(5), 1089–1122.

Mortensen, D. T. (2005). *Wage Dispersion: Why Are Similar Workers Paid Differently?* Cambridge: MIT Press.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research 6*(1), 58–73.

PayScale (2021). 2021 Compensation Best Practices Report. Technical report.

Perez-Truglia, R. (2020). The Effects of Income Transparency on Well-Being: Evidence from a Natural Experiment. *American Economic Review 110*, 1019–54.

Postel-Vinay, F. and J.-M. Robin (2002). The Distribution of Earnings in an Equilibrium Search Model with State-Dependent Offers and Counteroffers. *International Economic Review 43*(4), 989–1016.

Roussille, N. (2021). The Central Role of the Ask Gap in Gender Pay Inequality. *Working Paper*.

Schiemann, W. A., J. H. Seibert, and M. H. Blankenship (2018). Putting human capital analytics to work: Predicting and driving business success. *Human Resource Management 57*(3), 795–807.

Securities and Exchange Commission (2006). SEC final rules 33-8732a, Item 402(b)(2)(xiv).

Sokolova, A. and T. Sorensen (2021). Monopsony in Labor Markets: A Meta-Analysis. *Industrial & labor relations review 74*(1), 27–55.

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. Von Wachter (2019). Firming up inequality. *Quarterly Journal of Economics 134*(1), 1–50.

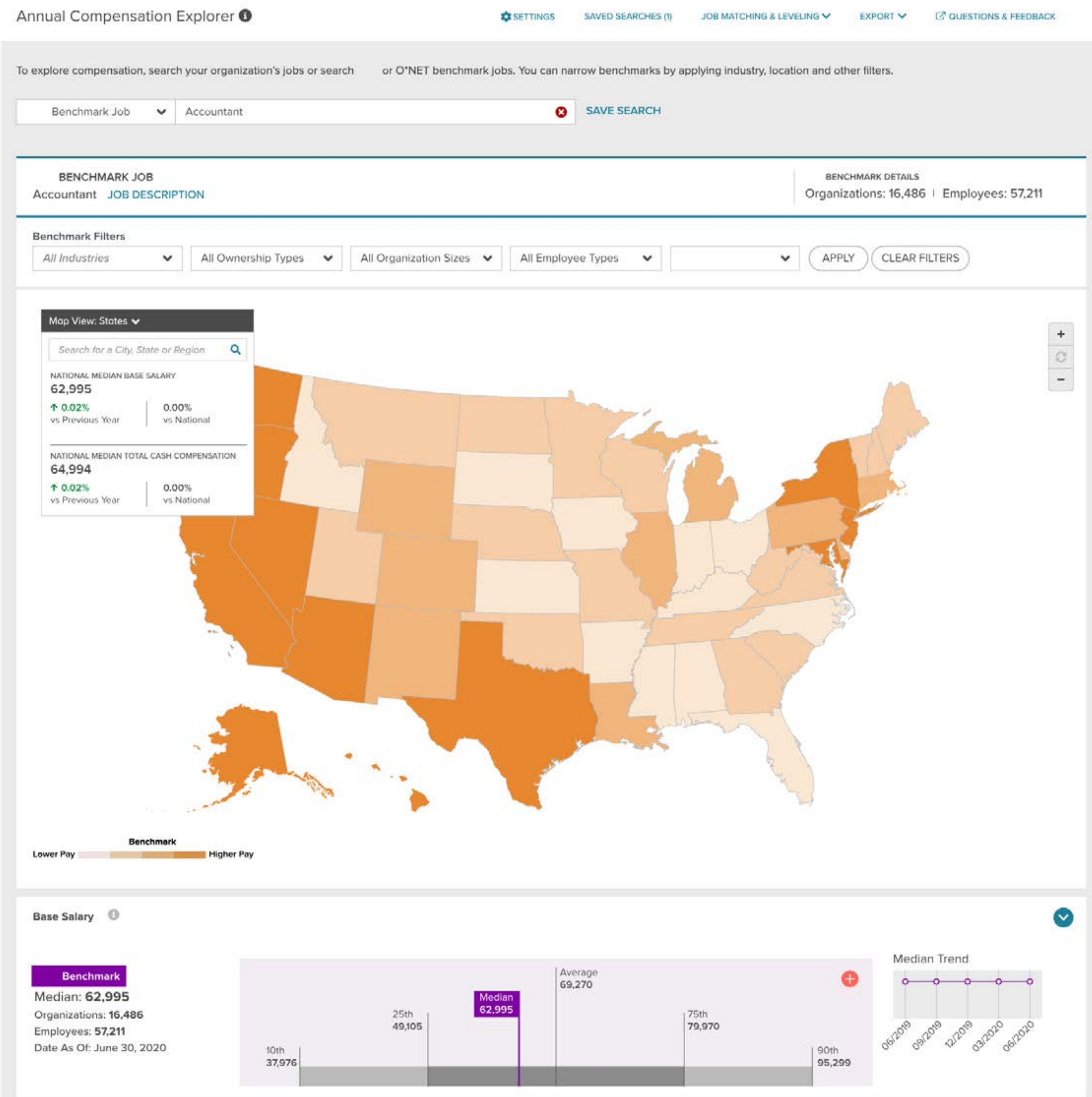
Tadelis, S. and F. Zettelmeyer (2015). Information Disclosure as a Matching Mechanism: Theory and Evidence from a Field Experiment. *American Economic Review 105*(2), 886–905.

Thom, M. and T. Reilly (2015). Compensation Benchmarking Practices in Large U.S. Local Governments. *Public Personnel Management 44*(3), 340–355.

White House (2021). Fact Sheet: Executive Order on Promoting Competition in the American Economy. *Statements and Releases from the White House, July 9, 2021*.

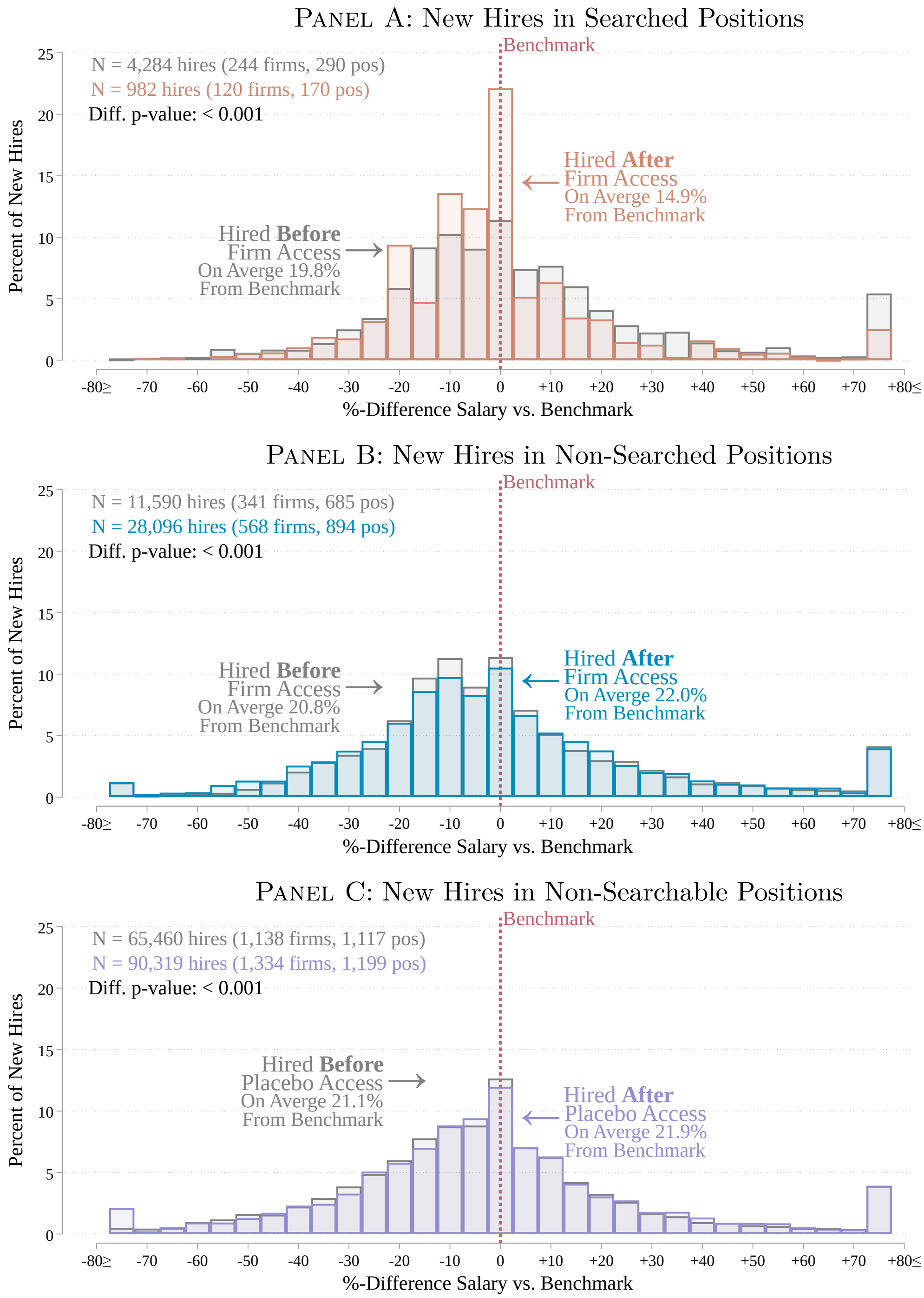
Zeuch, M. (2016). *Handbook of Human Resources Management.* Berlin: Springer.

Figure 1: Screenshot of the Salary Benchmarking Tool



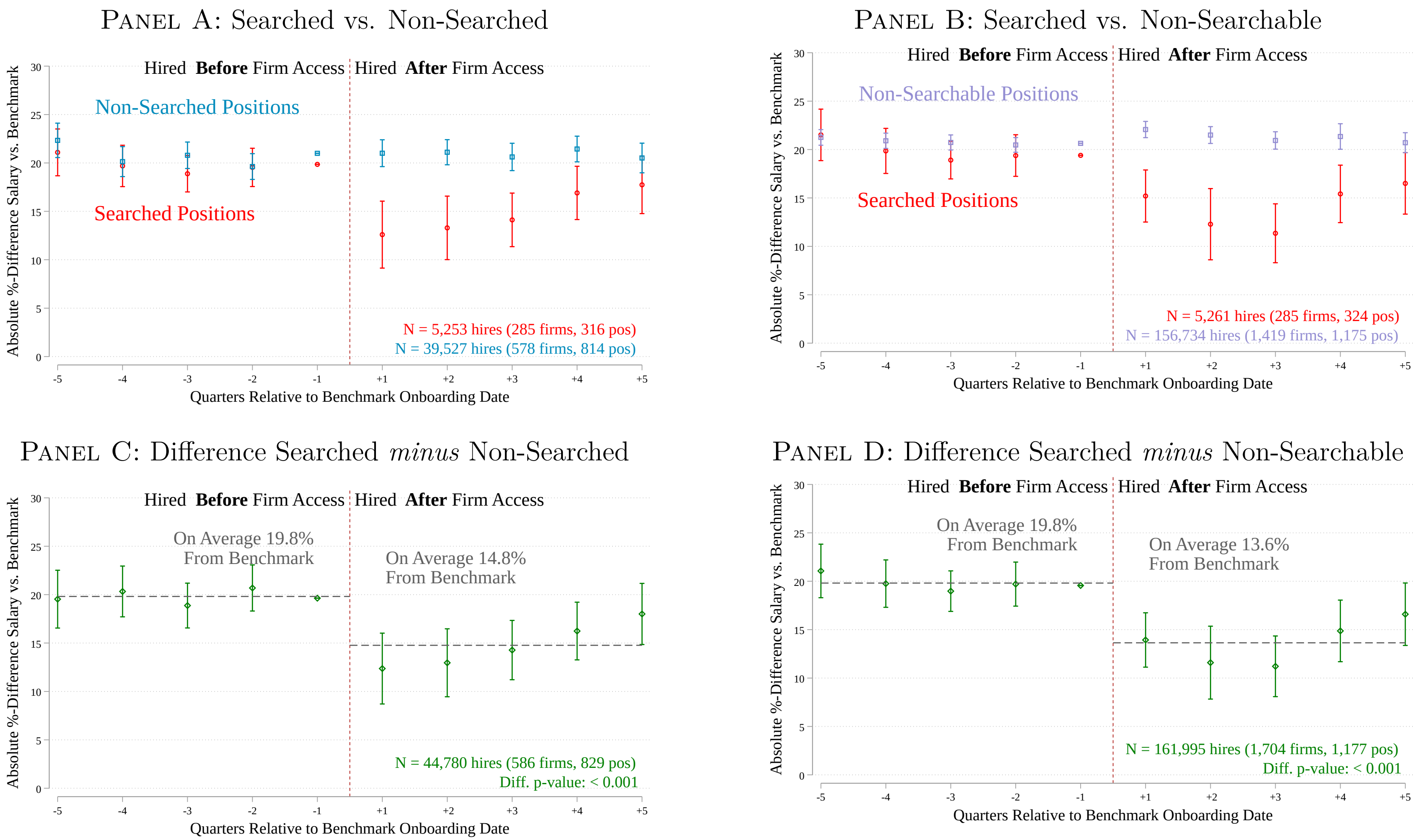
<u>Notes:</u> This is a screenshot of the pay benchmarking tool. It has been slightly altered to conceal the identity of the firm. This is the top of the screen. If you scroll down, you can see panels similar to the bottom panel titled *Base Salary* but for *Bonus*, *Overtime*, and *Total Compensation*.

Figure 2: The Effects of Benchmarking on Compression around the Benchmark: Non-Parametric Analysis



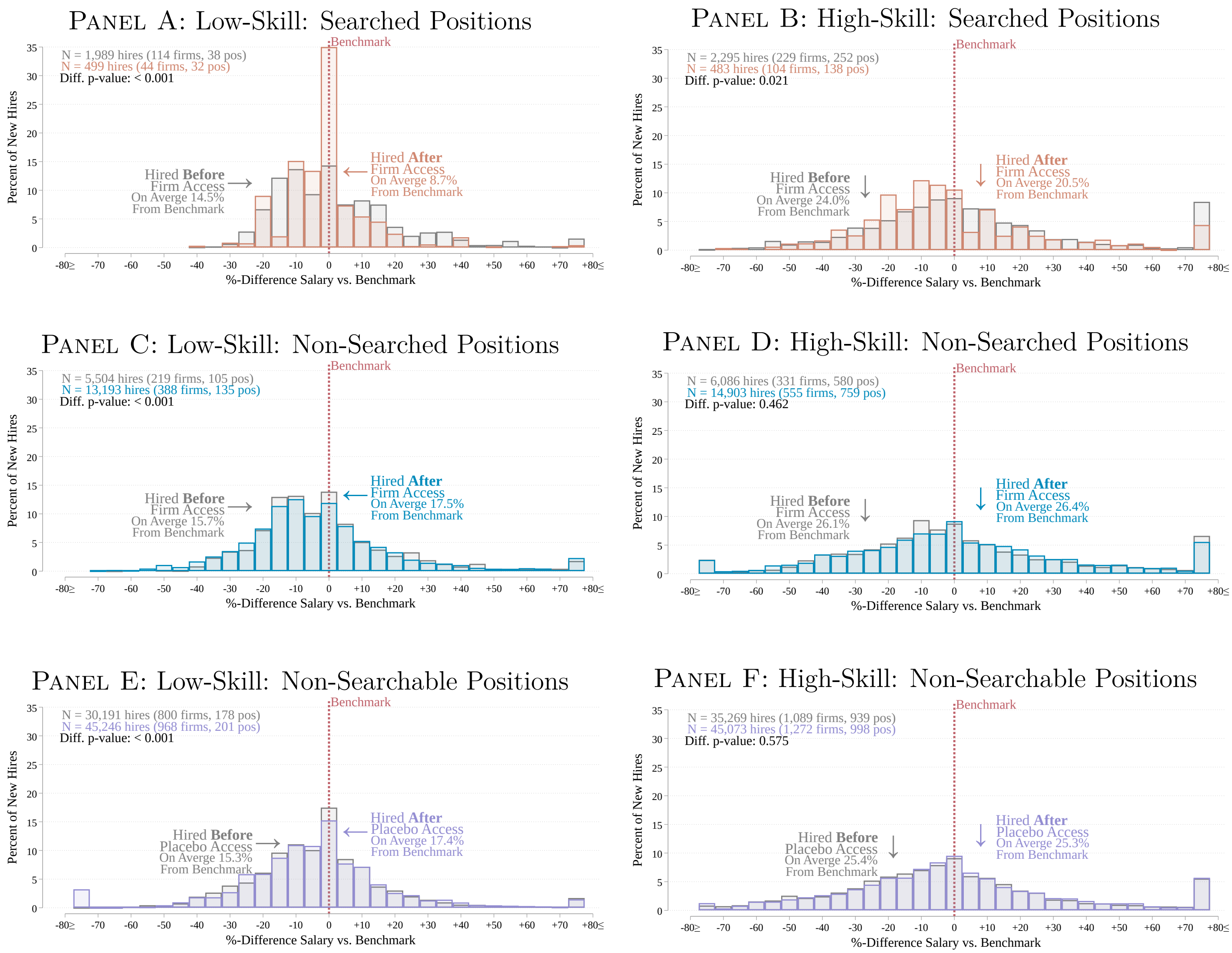
Notes: Histograms of the starting base salary relative to the corresponding external benchmark (winsorized at ± 75%). Each panel corresponds to a different set of positions: panel A for *Searched* positions (i.e., positions in firms with access to the benchmark tool that are eventually searched for by the firm), panel B for *Non-Searched* positions (i.e., positions in firms with access to the benchmark tool that are not eventually searched for by the firm), and panel C for *Non-Searchable* positions (i.e., positions in firms without access to the benchmark tool). In each panel, the solid and hollow bars correspond to the observations before and after the firm gains access to the benchmark tool, respectively (and in panel C, that date corresponds to the "placebo" onboarding date assigned to the firm that never gains access to the tool).

Figure 3: Event-Study Analysis: Effects on Pay Compression



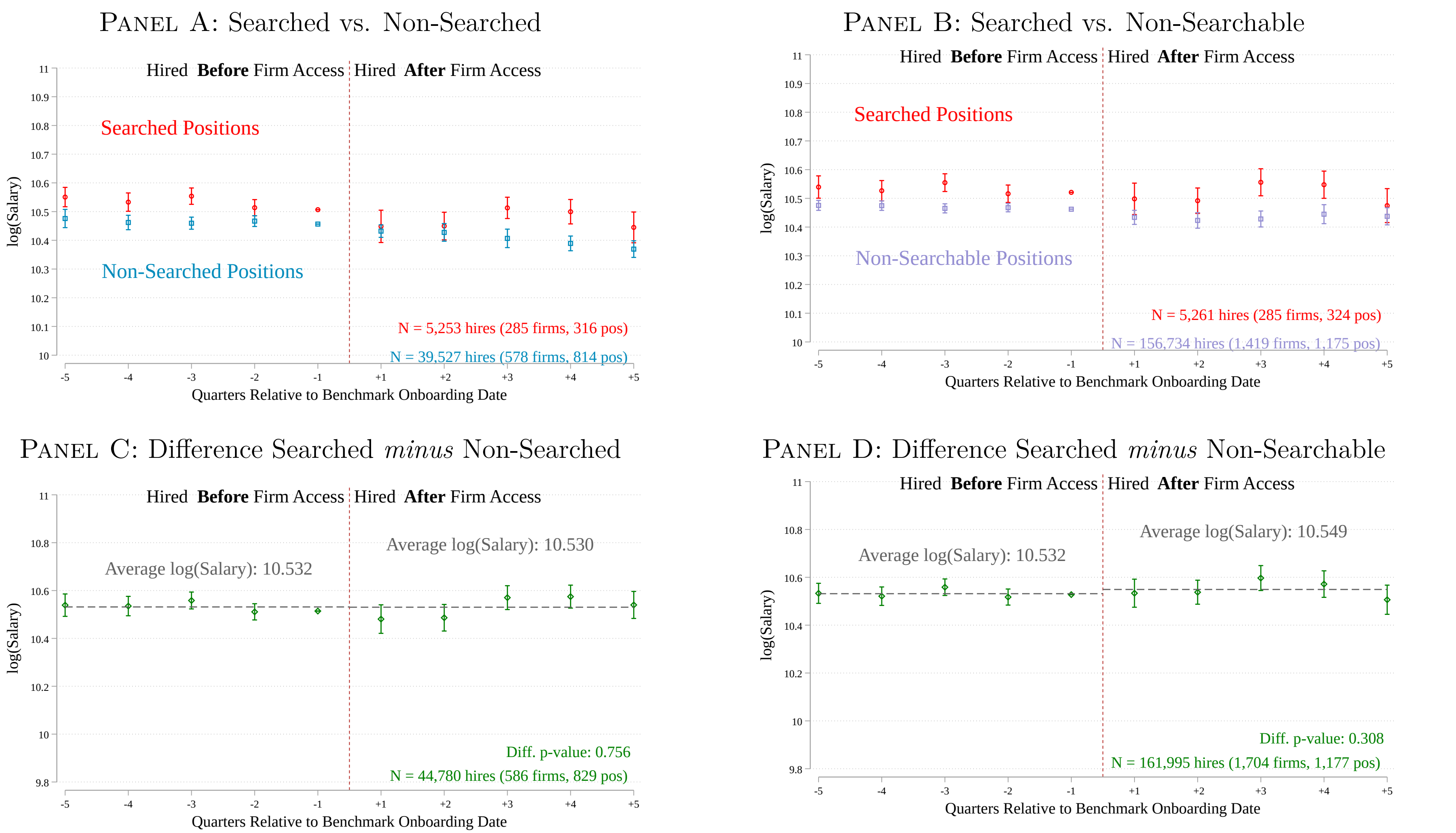
Notes: Point estimates with 90% confidence intervals in brackets, using standard errors clustered at the firm-position-month level. Panels A and C are based off one regression for Searched and Non-Searched positions, while panel A presents the estimates for each position type and panel C presents the difference. Panels B and D are analogous for Searched vs. Non-Searchable positions. All coefficients are shifted such that the pre-treatment coefficients average to the pre-treatment mean of the absolute dispersion outcome. Coefficients in panels C and D refer to parameters $\alpha^k_{1,s}$ $\forall s \in S$ from equation (16) (see Section 4.2 for details).

Figure 4: Heterogeneity: Non-Parametric Analysis



Notes: All figures are a reproduction of the corresponding panel of Figure 2 for low-skill positions (left) and high-skill positions (right).

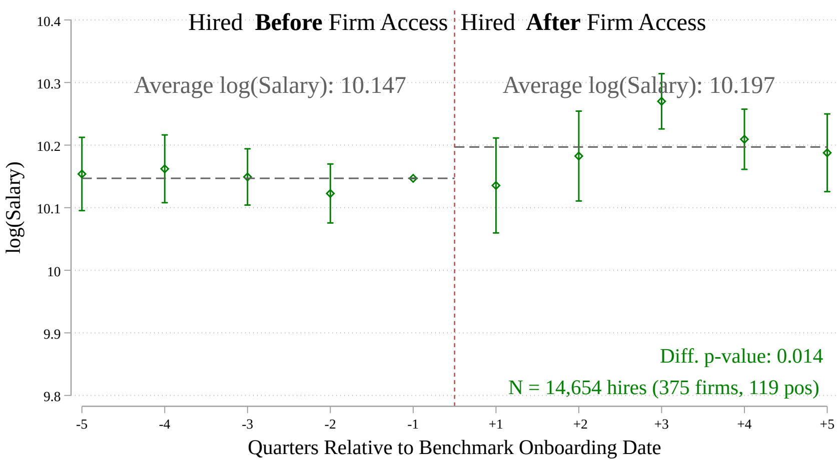
Figure 5: Event-Study Analysis: The Effects on Salary Levels



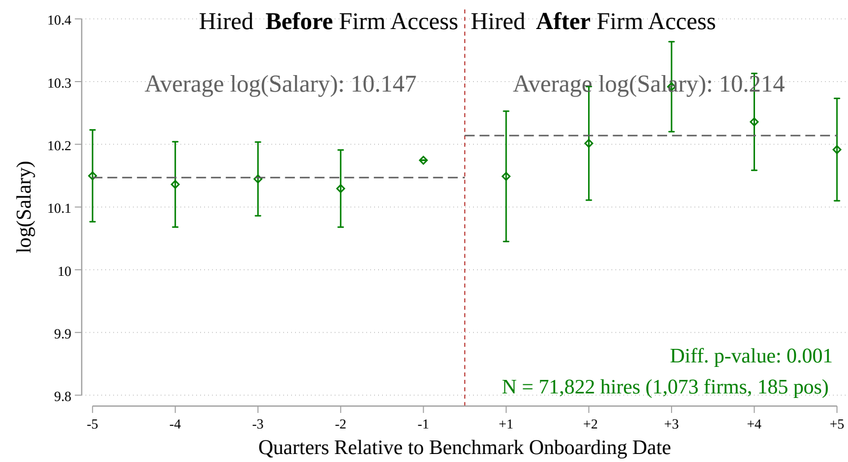
Notes: Point estimates with 90% confidence intervals in brackets, using robust standard errors. Panels A and C are based off one regression for Searched and Non-Searched positions, while panel A presents the estimates for each position type and panel C presents the difference. Panels B and D are analogous for Searched vs. Non-Searchable positions. All coefficients are shifted such that the pre-treatment coefficients average to the pre-treatment mean of log salary. Coefficients in panels C and D refer to parameters $\alpha^k_{1,s}$ $\forall s \in S$ from equation (16) (see Section 4.2 for details).

Figure 6: Heterogeneity by Skill: The Effects on Salary Levels

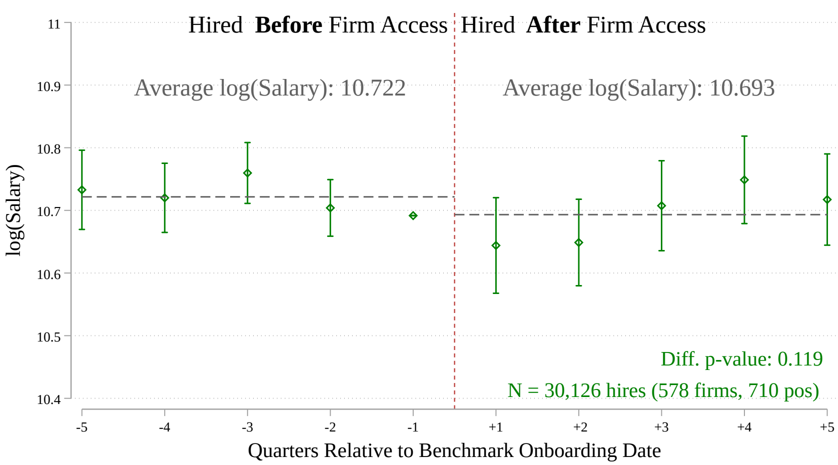
Panel A: Low-Skill: Searched vs. Non-Searched



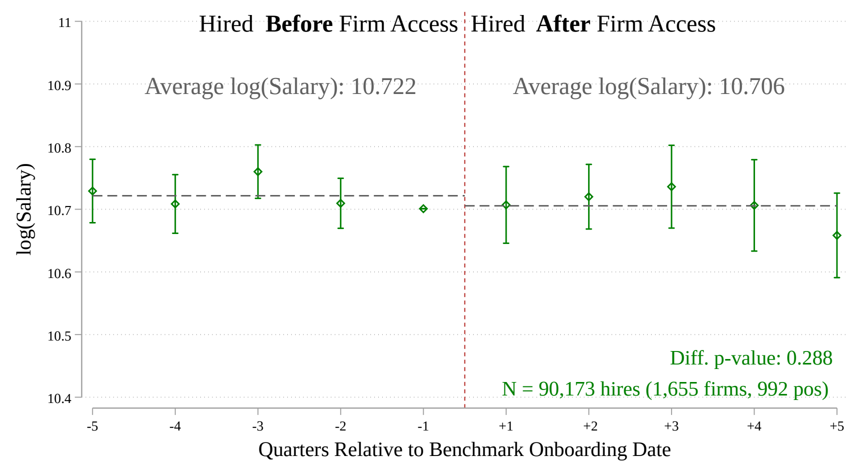
Panel B: Low-Skill: Searched vs. Non-Searchable



Panel C: High-Skill: Searched vs. Non-Searched



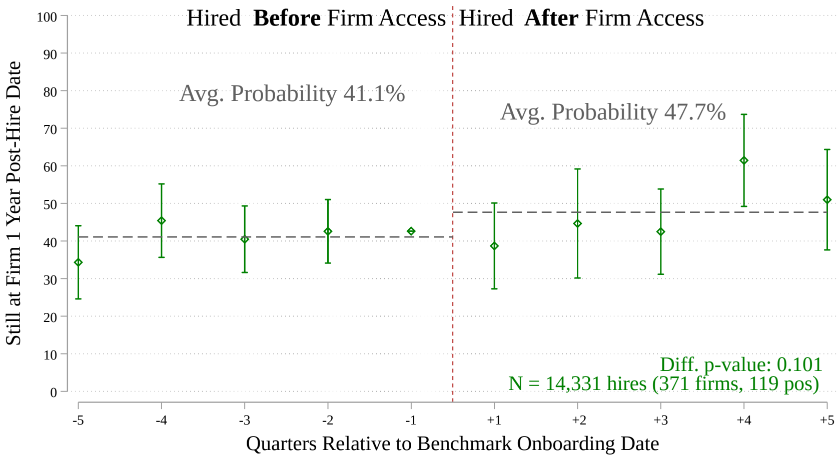
Panel D: High-Skill: Searched vs. Non-Searchable



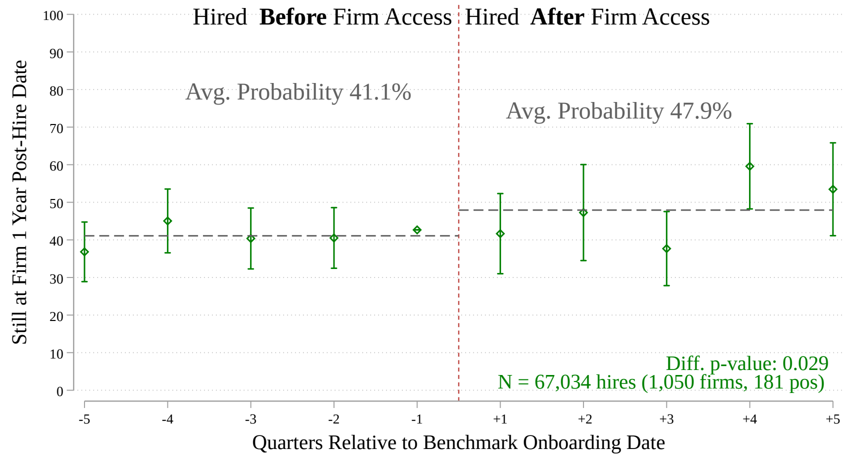
Notes: Panels A and C are a reproduction of panel C from Figure 5, and panels B and D are a reproduction of panel D, but for the specified sub-samples. *Skill* is defined in Section 3.4. See the notes of Figure 5 for more details.

Figure 7: Heterogeneity by Skill: The Effects on Retention Rates

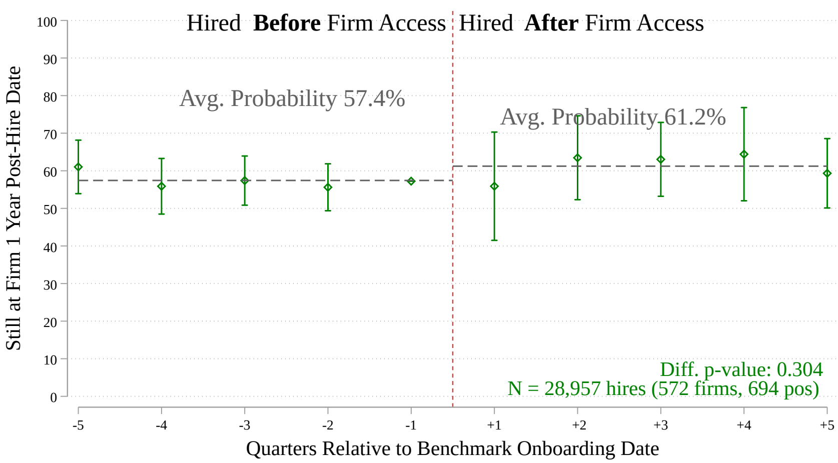
PANEL A: Low-Skill: Searched vs. Non-Searched



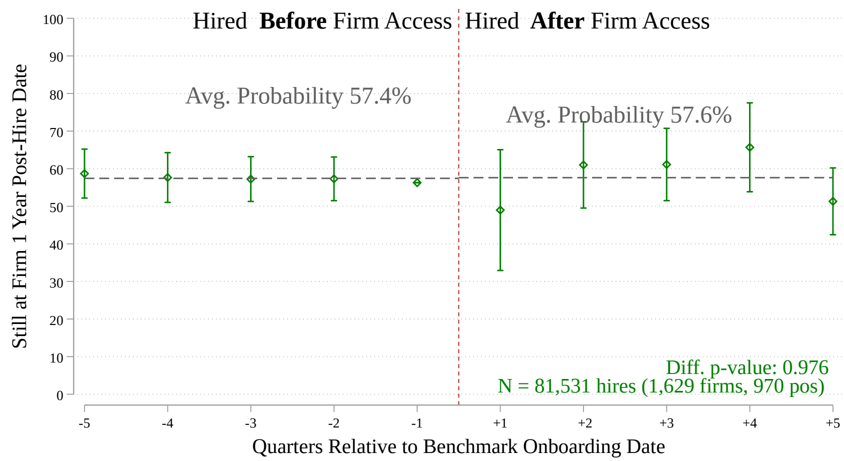
PANEL B: Low-Skill: Searched vs. Non-Searchable



PANEL C: High-Skill: Searched vs. Non-Searched



PANEL D: High-Skill: Searched vs. Non-Searchable



Notes: This is a reproduction of Figure 6, but with the outcome being a dummy equal to 100 if a new hire in a given month is still at the same firm 1 year later. Because our main sample ends in March 2020 and our data ends in July 2021, we observe this outcome for all new hires in our main sample. For more details, see notes to Figure 6.

Table 1: Summary Statistics for Firms with vs. without Access

		Has Access?		By Usage	
	(1)	(2)	(3)	(4)	(5)
	All	No	Yes	Higher	Lower
Average Firm Characteristics					
Average Employment	503.3	509.8	483.2	525.7	444.5
	(28.1)	(33.2)	(52.1)	(50.3)	(88.4)
Turnover Rate (%)†	2.424	2.438	2.382	2.392	2.374
	(0.061)	(0.070)	(0.126)	(0.159)	(0.192)
Business Services Sector (%)	17.27	16.73	18.94	14.62	22.87
	(0.99)	(1.13)	(2.07)	(2.71)	(3.07)
Hospitality Sector (%)	2.62	2.83	1.95	2.34	1.60
	(0.42)	(0.50)	(0.73)	(1.16)	(0.92)
Retail & Wholesale Trade Sector (%)	12.04	11.97	12.26	16.37	8.51
	(0.85)	(0.98)	(1.73)	(2.84)	(2.04)
Health Care Sector (%)	8.47	7.95	10.03	11.70	8.51
	(0.73)	(0.82)	(1.59)	(2.46)	(2.04)
Banking Sector (%)	7.16	7.13	7.24	7.02	7.45
	(0.68)	(0.78)	(1.37)	(1.96)	(1.92)
Other Sector (%)	52.44	53.38	49.58	47.95	51.06
	(1.31)	(1.51)	(2.64)	(3.83)	(3.66)
Average Employee Characteristics					
Salary (annual \$)†	46,945	46,439	48,488	45,232	51,449
	(794)	(956)	(1,356)	(1,632)	(2,103)
External Benchmark (annual \$)†	47,643	47,008	49,579	46,491	52,389
	(652)	(752)	(1,307)	(1,650)	(1,977)
Abs. %-Diff. Salary vs. Benchmark†	22.16	22.46	21.26	19.41	22.95
	(0.38)	(0.45)	(0.68)	(0.84)	(1.04)
Age	34.40	34.30	34.72	34.36	35.04
	(0.18)	(0.22)	(0.32)	(0.42)	(0.48)
Share Female (%)	45.29	46.39	41.92	44.74	39.36
	(1.29)	(1.48)	(2.57)	(3.78)	(3.51)
Share High Education (%)	56.92	55.30	61.84	57.89	65.43
	(1.28)	(1.49)	(2.53)	(3.74)	(3.42)
Share Hourly (%)	71.89	73.08	68.25	71.35	65.43
	(1.17)	(1.33)	(2.44)	(3.47)	(3.44)
Number of Firms	2,005	1,419	586	183	403

Notes: Average characteristics in the main sample of new hires, with robust standard errors in parentheses. Variables marked with † are computed using only pre-onboarding data. *Higher Usage* are firms that search at least once and *Lower Usage* are firms with access that never search. *Turnover Rate* is defined as number of employee departures in a month over the number of employees employed at the firm during that month. *Business Services Sector* through *Other Sector* correspond to the distribution of industry sectors. *Salary* is the annual base salary at the time of hire. *External Benchmark* is the median annual base salary benchmark in the position of the new hire during the quarter of the hire date.

Table 2: Most Common Searched Position Titles

Position Title	(1) Searched	(2) Non-Searched	(3) Non-Searchable
Bank Teller	539 [12]	287 [24]	1,976 [87]
Customer Service Representative	468 [44]	4,401 [170]	4,012 [385]
Security Guard	286 [6]	139 [44]	6,263 [95]
Hotel Cleaner	208 [2]	379 [5]	1,058 [17]
Legal Associate Specialist	163 [1]	7 [4]	14 [9]
Hand Packer	155 [4]	234 [17]	1,957 [55]
Patient Care Coordinator	117 [3]	103 [14]	133 [29]
Receptionist	93 [15]	310 [86]	2,911 [238]
Cook	86 [6]	334 [21]	1,606 [85]
Waiter/Waitress	84 [7]	1,113 [18]	2,986 [87]
Delivery Driver	79 [5]	34 [9]	744 [26]
Dish Washer/Plate Collector/Table Top Cleaner	69 [5]	187 [18]	1,350 [67]
Medical Assistant	69 [10]	370 [17]	889 [55]
Welder	66 [8]	112 [27]	652 [59]
Cashier	65 [2]	175 [11]	2,706 [48]
Registered Nurse	64 [11]	244 [22]	2,699 [110]
Assembler	60 [9]	606 [26]	3,823 [90]
Other Housekeeper and Related Worker	59 [5]	173 [17]	948 [63]
Software Developer/Programmer	59 [23]	403 [78]	1,285 [173]
Warehouse Laborer	59 [10]	761 [43]	3,025 [116]
Mammographer	55 [1]	9 [1]	3 [2]
Nursing Assistant	51 [4]	662 [13]	7,346 [65]
Bartender/Mixologist	49 [2]	228 [12]	611 [46]
Production Operations Engineer	49 [1]	41 [16]	68 [29]
Licensed Practical Nurse	48 [9]	189 [23]	1,605 [69]
Sales Manager	48 [18]	166 [67]	693 [181]
General Practitioner/Physician	46 [2]	143 [17]	340 [28]
Lawyer	43 [5]	17 [10]	268 [52]
Ophthalmic Technician	42 [2]	4 [1]	34 [4]
Business Development Specialist	41 [2]	124 [27]	447 [41]
Warehouse Manager	40 [7]	133 [23]	430 [72]
Other Social Work and Counseling Professional	39 [1]	1 [1]	32 [9]
Building Caretaker/Watchman	38 [2]	288 [59]	917 [139]
Operations Officer	37 [2]	73 [18]	108 [36]
Shipping Clerk	37 [4]	39 [19]	218 [63]

Notes: New hires in each position [firms hiring in each position]. Tabulations across all new hires for the 35 Searched *Position Titles* with the most new hires.

Table 3: Summary Statistics by Position Type

		by Position Type		
	(1)	(2)	(3)	(4)
	All	Searched	Non-Searched	Non-Searchable
Salary (annual \$)†	41,359	39,064	42,013	41,405
	(146)	(462)	(390)	(166)
External Benchmark (annual \$)†	41,412	38,649	41,092	41,672
	(113)	(409)	(295)	(128)
Abs. %-Diff. Salary vs. Benchmark†	20.36	17.36	21.03	20.45
	(0.08)	(0.28)	(0.21)	(0.09)
Age	34.77	34.53	34.54	34.83
	(0.05)	(0.22)	(0.13)	(0.06)
Share Female (%)	50.63	60.14	51.01	49.87
	(0.20)	(0.83)	(0.53)	(0.23)
Share High Education (%)	42.21	34.49	42.28	42.76
	(0.20)	(0.80)	(0.52)	(0.23)
Share Hourly (%)	81.11	82.94	80.13	81.16
	(0.16)	(0.64)	(0.42)	(0.18)
Occupation Groups				
Office and Administrative Support (%)	19.84	32.44	28.97	17.23
	(0.16)	(0.79)	(0.48)	(0.17)
Building and Grounds Cleaning (%)	4.77	5.22	2.58	5.14
	(0.09)	(0.38)	(0.17)	(0.10)
Management (%)	8.04	8.10	9.21	7.81
	(0.11)	(0.46)	(0.31)	(0.12)
Production (%)	6.59	6.48	6.35	6.64
	(0.10)	(0.42)	(0.26)	(0.11)
Transportation and Material Moving (%)	9.30	6.62	9.72	9.42
	(0.12)	(0.42)	(0.31)	(0.13)
Other (%)	51.47	41.14	43.16	53.75
	(0.20)	(0.83)	(0.52)	(0.23)
Number of Firms	2,005	285	578	1,419
Number of Positions	1,406	329	973	1,306
Observations	201,817	5,266	39,686	156,865

Notes: Average characteristics in the main sample of new hires, with robust standard errors in parentheses. Variables marked with † are computed using only pre-onboarding data. *Salary* is the annual base salary at the time of hire. *External Benchmark* is the median annual base salary benchmark in the position of the new hire during the quarter of the hire date. Variables under *Occupation Groups* correspond to a new hire's SOC group.

Table 4: The Effects of Benchmarking on Absolute %-Distance from the Benchmark

	(1) $\|\%\Delta\|$	(2) $\|log\Delta\|$	(3) $\|\%\Delta\| > 10$	(4) $\|\%\Delta\|$	(5) $\|\%\Delta\|$	(6) $\|\%\Delta\|$	(7) $\|\%\Delta\|$	(8) $\|\%\Delta\|$	(9) $\|\%\Delta\|$	(10) $\|\%\Delta\|$	(11) $\|\%\Delta\|$	(12) $\|\%\Delta\|$
Panel (a): Post-treatment												
Searched vs. Non-Searched	-4.775***	-5.155***	-16.270***	-5.148***	-4.775***	-4.786***	-5.324***	-4.950***	-4.421***	-4.887***	-4.880***	-4.564***
	(1.143)	(1.266)	(3.626)	(1.338)	(0.906)	(1.198)	(1.282)	(1.286)	(1.153)	(1.165)	(1.276)	(1.178)
Searched vs. Non-Searchable	-6.149***	-7.118***	-13.861***	-6.836***	-6.149***	-6.128***	-7.494***	-7.450***	-5.714***	-6.163***	-5.044***	-5.934***
	(1.070)	(1.211)	(3.681)	(1.220)	(0.824)	(1.076)	(1.233)	(1.576)	(1.078)	(1.087)	(1.231)	(1.127)
Panel (b): Pre-treatment												
Searched vs. Non-Searched	-0.346	-0.129	-5.872	-0.233	-0.346	-0.488	-1.646	-2.062*	-0.714	-0.144	-2.205	-0.199
	(1.167)	(1.313)	(3.690)	(1.289)	(0.751)	(1.185)	(1.514)	(1.200)	(1.133)	(1.199)	(1.528)	(1.174)
Searched vs. Non-Searchable	-0.310	0.156	-4.221	-0.513	-0.310	-0.318	0.021	-1.029	0.241	-0.247	-0.754	-0.500
	(1.055)	(1.175)	(3.246)	(1.184)	(0.643)	(1.057)	(1.375)	(1.116)	(1.046)	(1.069)	(1.342)	(1.105)
Winsorizing at +/- 100%				✓								
No Clustering					✓							
No Additional Controls						✓						
No Position FE							✓					
Firm FE								✓				
Exclude High-Tip Jobs									✓			
Searched Positions Only										✓		
No Re-weighting											✓	
Ages 21-60												✓
Mean Dep. Var. (Baseline)	19.812	20.590	63.732	21.004	19.812	19.812	19.812	19.812	19.430	19.812	19.802	19.903
Observations												
Searched	5,253	5,253	5,253	5,253	5,253	5,253	5,266	5,262	5,105	5,253	5,331	4,611
Non-Searched	39,527	39,527	39,527	39,527	39,527	39,527	39,686	39,673	37,841	34,954	39,810	34,338
Non-Searchable	156,734	156,734	156,734	156,734	156,734	156,734	156,865	156,817	148,521	127,145	157,018	135,051

Notes: Significant at *10%, **5%, ***1%. Standard errors clustered at the firm-position-month level in parentheses. Each column corresponds to two regressions: one for Searched vs. Non-Searched new hires and one for Searched vs. Non-Searchable new hires. Post-treatment coefficients in panel (a) refer to parameters α_1^k from equation (14), while pre-treatment coefficients in panel (b) refer to parameters α_3^k from equation (15) (see Section 4.2 for details). All columns include year fixed effects. In columns (1) and (4)–(12) the dependent variable is the absolute percent difference between the annual base salary and median benchmark (Δ). The dependent variable in col (2) is the log of Δ and in col (3) is a dummy that equals 100 if $|\%\Delta|$ is greater than 10% and zero otherwise. We multiply $\%\Delta$ and $log(\Delta)$ by 100 so that the effects can be interpreted as percentage points. Δ is winsorized to $\pm$ 75 except in column (4) where it is winsorized to $\pm$ 100. All columns except (6) include additional controls (female dummy, high education dummy, hourly dummy, age, position tenure). Column (9) excludes the three positions where gross pay most exceeds base pay: Waiter/Waitress, Chauffeur, and Bartender/Mixologist. Column (10) restricts the sample to only positions of Non-Searched or Non-Searchable new hires in positions that are searched and hired by firms in the data.