Online Appendix (For Online Publication Only)

What's My Employee Worth? The Effects of Salary Benchmarking

Cullen, Li and Perez-Truglia (August 26, 2024)

A Proofs of Theorems

A.1 Proof of Theorem 5.1

Lemma A.1. *In any no-benchmark equilibrium η, if a firm with value v hires with positive probability when $Q = q$, then $\eta(v, q) \leq v$.*

Proof. Otherwise the firm with value v can profitably deviate to offer 0. □

Let ω be an arbitrary constant. Suppose that $\eta(v, q) = \omega$ for all $v > \tau_q$; we now argue that η is not a no-benchmark equilibrium. There are two cases to check: $\omega > \tau_q$ and $\omega \leq \tau_q$.

If $\omega > \tau_q$, then some firm makes an offer strictly higher than its value and hires with positive probability, so η is not an equilibrium by Lemma A.1.

Suppose instead that $\omega \leq \tau_q$. Since we assumed that C has a density conditional on $Q = q$ and $V_i = v$, it follows that $\tau_q < \overline{v}$ and, moreover, that under η the firm with value $\overline{v}$ hires with probability strictly less than 1. If offering ω results in a positive probability of hiring, then the firm with value $\overline{v}$ is sometimes rationed under the tie-breaking rule, so it can profitably deviate to offer $\omega + \epsilon$ for small $\epsilon > 0$. If offering ω results in zero probability of hiring, then the firm with value $\overline{v}$ can profitably deviate to offer $\tau_q + \epsilon$ for small $\epsilon > 0$, which by Lemma A.1 hires with probability 1. This completes the proof.

A.2 Proof of Theorem 5.2

Lemma A.2. *$H(c \mid v, q)/h(c \mid v, q)$ is decreasing in v.*

Proof. This follows by affiliation, as in the proof of Lemma 1 in Milgrom and Weber (1982). □

By the usual argument, $L(\cdot \mid v, q)$ defines a probability distribution on $[\tau_q, v]$ that increases stochastically in v, that is $L(\alpha \mid v, q)$ is decreasing in v (Krishna, 2009, p. 95). Thus $w^*(v, q)$ is increasing in v and we have $w^*(v, q) \leq v$.

Lemma A.3. *$w^*(v, q)$ is continuous in v.*

Proof. By construction, $w^*(\cdot, q)$ is continuous at v for all $v \leq \tau_q$. In particular,

$$\lim_{v \downarrow \tau_q} \int_{\tau_q}^{v} \alpha dL(\alpha \mid v, q) = \tau_q \tag{A.1}$$

Suppose that $w^*(\cdot, q)$ is discontinuous at v for some $v > \tau_q$. Then there exists $v > \tau_q$ such that for all $\epsilon > 0$ the first expression of (A.2) is infinite.

$$\begin{aligned} \int_v^{v+\epsilon} \frac{h(\beta \mid \beta, q)}{H(\beta \mid \beta, q)} d\beta &\leq \int_v^{v+\epsilon} \frac{h(\beta \mid v+\epsilon, q)}{H(\beta \mid v+\epsilon, q)} d\beta \\ &= \ln H(v+\epsilon \mid v+\epsilon, q) - \ln H(v \mid v+\epsilon, q). \end{aligned} \tag{A.2}$$

The inequality is by Lemma A.2. The last expression of (A.2) is infinite only if $H(v \mid v+\epsilon, q) = 0$. But we assumed that for all q and all $v > \tau_q$, there exists $\epsilon > 0$ such that $H(v \mid v+\epsilon, q) > 0$, a contradiction. □

A firm facing w^* can hire with positive probability only if it offers a wage of at least τ_q, so if its value is $v \leq \tau_q$, then it is optimal to offer a wage equal to v.

It remains to verify that no firm with value $v > \tau_q$ can profitably deviate. Suppose for the moment that $w^*(v, q)$ is differentiable in v. For $v > \tau_q$, the wage function w^* defined in (5) solves the differential equation

$$w_1^*(v, q) = (v - w^*(v, q)) \frac{h(v \mid v, q)}{H(v \mid v, q)} \tag{A.3}$$

with boundary condition $w^*(\tau_q, q) = q$. Thus we have

$$(\hat{v} - w^*(\hat{v}, q)) \frac{h(\hat{v} \mid \hat{v}, q)}{H(\hat{v} \mid \hat{v}, q)} - w_1^*(\hat{v}, q) = 0 \text{ for all } \hat{v} > \tau_q \tag{A.4}$$

If the firm with $v > \tau_q$ has any profitable deviation, then it has a profitable deviation to an offer in the set $\{w^*(\hat{v}, q) : \hat{v} \geq \tau_q\}$. Thus it remains to verify that

$$v \in \operatorname*{argmax}_{\hat{v} \geq \tau_q} (v - w^*(\hat{v}, q)) H(\hat{v} \mid v, q). \tag{A.5}$$

Taking the derivative of the objective with respect to $\hat{v}$ yields

$$\begin{aligned} &-w_1^*(\hat{v}, q) H(\hat{v} \mid v, q) + (v - w^*(\hat{v}, q)) h(\hat{v} \mid v, q) \\ &= H(\hat{v} \mid v, q) \left[(v - w^*(\hat{v}, q)) \frac{h(\hat{v} \mid v, q)}{H(\hat{v} \mid v, q)} - w_1^*(\hat{v}, q) \right] \end{aligned} \tag{A.6}$$

The right-hand side of (A.6) has the same sign as $\hat{v} - v$, by (A.4) and Lemma A.2. Thus, the firm's objective is maximized at $\hat{v} = v$.

Suppose that $w^*(v, q)$ is not differentiable in v. Since $w^*(v, q)$ is continuous in v and of bounded variation in v, and $[\underline{v}, \overline{v}]$ is bounded, there exists a homeomorphism $\lambda : [\underline{v}, \overline{v}] \to [\underline{v}, \overline{v}]$ such that $\overline{w}(\theta, q) \equiv w^*(\lambda(\theta), q)$ is differentiable in θ (Bruckner and Goffman, 1976). Moreover, we can pick λ to be strictly increasing. Let us define the random variables $\Theta_i \equiv \lambda^{-1}(V_i)$ and

$\overline{C} \equiv \lambda^{-1}(C)$. Affiliation is preserved under increasing transformations (Milgrom and Weber, 1982, Theorem 5), so $(\Theta_i, \overline{C})$ are affiliated conditional on Q. Let $\overline{H}(\bar{c} \mid \theta, q) \equiv P(\overline{C} \leq \bar{c} \mid \Theta_i = \theta, Q = q)$. Now we reformulate (A.5) into the equivalent statement

$$\theta \in \operatorname*{argmax}_{\hat{\theta} \geq \lambda^{-1}(\tau_q)} (\lambda(\theta) - \overline{w}(\hat{\theta}, q))\overline{H}(\hat{\theta} \mid \theta, q). \tag{A.7}$$

The argument we used to prove (A.5) for differentiable w^* applies *mutatis mutandis* to establish (A.7). It follows that w^* is a no-benchmark equilibrium.

A.3 Proof of Theorem 5.5

This proof closely parallels the proof of Theorem 15 of Milgrom and Weber (1982). Let us consider an incentive-compatible mechanism M in which the firm observes V_i and Q, and then reports V_i to the mechanism. Depending on the firm's report and the cutoff C, this results in some probability of hiring and some wage conditional on hiring. Let A be the mechanism that corresponds to the no-benchmark equilibrium w^*, and let B be the mechanism that corresponds to the benchmark equilibrium $\tilde{w}$.

Let $R(\hat{v}, v, q) \equiv vH(\hat{v} \mid v, q)$ denote the expected gross revenue to the firm when it reports value $\hat{v}$, its value is v, and the supply of workers is q.

Let $K^M(\hat{v}, v, q)$ denote the conditional expected wage paid by the firm in mechanism M when it reports value $\hat{v}$, its value is v, the supply of workers is q, and the firm hires a worker. Observe that $K^A(\hat{v}, v, q) = w^*(\hat{v}, q)$ and $K^B(\hat{v}, v, q) = E[C \mid C \leq \hat{v}, V_i = v, Q = q]$.

Facing mechanism $M \in \{A, B\}$, the firm chooses $\hat{v}$ to maximize

$$R(\hat{v}, v, q) - K^M(\hat{v}, z, q)H(\hat{v} \mid v, q). \tag{A.8}$$

By incentive compatibility, the first-order condition holds at $\hat{v} = v$, so for $v > \tau_q$ we have

$$0 = R_1(v, v, q) - K_1^M(v, v, q)H(v \mid v, q) - K^M(v, v, q)h(v \mid v, q). \tag{A.9}$$

The boundary condition is $K^M(\tau_q, \tau_q, q) = \tau_q$.

By inspection, the partial derivative $K_2^A(\hat{v}, v, q)$ is equal to zero. By affiliation and Theorem 5 of Milgrom and Weber (1982), we have $K_2^B(\hat{v}, v, q) \geq 0$. If $K^B(v, v, q) < K^A(v, v, q)$ for some v and q, then by (A.9) we have

$$\frac{d}{dv}K^B(v, v, q) = K_1^B + K_2^B \geq K_1^A + K_2^A = \frac{d}{dv}K^A(v, v, q). \tag{A.10}$$

By Lemma 2 of Milgrom and Weber (1982), it follows that $K^B(v, v, q) \geq K^A(v, v, q)$ for all $v > \tau_q$. Thus, for all $v > \tau_q$ we have

$$w^*(v, q) = K^A(v, v, q) \leq K^B(v, v, q) = E_C\left[\tilde{w}(v, q, C) \mid C \leq v, V_i = v, Q = q\right]. \tag{A.11}$$

The corollary follows by the law of iterated expectations.

B Model Extensions

Next, we discuss some additional extensions of the model that capture more realistic features of the institutional context. In practice, we observe that there are some sources of benchmarking data that are less precise but freely available to all firms (Section 2). This can be incorporated to the model as follows. Let us represent those benchmarks as a random variable B that is partially informative about the state S. All the firms have common knowledge of B. If our technical conditions hold conditional on B, then all the results extend straightforwardly. That is, (V_i, C) are affiliated conditional on Q and B, and there exists a distribution $H(c \mid v, q, b) \equiv P(C \leq c \mid V_i = v, Q = q, B = b)$ with corresponding density, and so on.

Our baseline model assumes that firms offer identical amenities. However, in real labor markets, there can be large differences in amenities between firms. Suppose that each firm has a firm-specific amenity A_i that is exogenous and known to that firm and to all workers. The firms that hire are those that offer the highest total remuneration (i.e., wages plus amenities). Let us define a new random variable $\overline{V}_i \equiv V_i + A_i$, equal to the revenue from filling the position plus the amenity. Moreover, assume that $\overline{V}_i$ is a sufficient statistic for the firm's private information about aggregate conditions, and satisfies the same technical assumptions as V_i in the baseline model, with $\overline{C}$ defined analogously. Then the function w^* can be reinterpreted as the total remuneration offered in equilibrium, as a function of $\overline{V}_i$ and Q.[66] On that interpretation, Theorem 5.3 predicts that learning the state causes compression in total remuneration.

[66] Observe that the payoff to a firm with value V_i and amenity A_i of hiring at total remuneration $w + A_i$ is $V_i - w = V_i + A_i - w - A_i = \overline{V}_i - (w + A_i)$. In particular, firms with the same $\overline{V}_i$ but different A_i will offer different wages, but the same total remuneration.

Supplemental Material

What's My Employee Worth? The Effects of Salary Benchmarking

Cullen, Li and Perez-Truglia (August 26, 2024)

C SHRM Survey

C.1 Pool of Respondents

Table C.1 presents the average characteristics of the sample of respondents. Most of the participants work in the private sector (80.81%), set salaries for both new hires and current employees (91.04%), and have 6 years or more of experience setting salaries (63.26%). The sample includes firms of all sizes and industries.

C.2 Use of Salary Benchmarking Tools

Table C.2 provides statistics on the use of salary benchmark tools. The most common uses of benchmark tools are to set the salary ranges for specific job titles and to change the salaries of current employees – 89.78% and 76.81% of participants, respectively, indicate that they use them for these purposes. Other popular uses are to set precise salaries for new employees (54.07%), in salary negotiations (53.11%), to determine salary in job advertisements (40.89%), and to plan ahead for headcount (25.33%). When asked about when they typically use benchmarks in relation to new hires, most (67.19%) answer that they use them before they publicize the position to include the expected salary in a job advertisement. Some managers also indicate that they use them right before making an offer to the candidate (34.96%), during negotiations after the candidate received the initial offer (22.30%), and when the candidate presents an outside offer (12.44%). When asked the same question but in relation to existing employees, the majority (74.30%) indicate that they use them when adjusting the salary ranges for positions. Many also respond that they use them when the employee goes through an annual review (48.06%), when the employee is up for promotion (47.48%) and when the employee presents an outside offer (33.11%).

Table C.2 shows that the majority (61.67%) of HR managers in our sample report using benchmarks to set salaries for the majority or all of their current employees, 25.89% only for some, 10.01% for a minority and 1.82% for none of their current employees. Similarly, the majority (64.36%) indicate that they use benchmarks to set salaries for a majority or all their new hires, 26.06% for some of their new hires, 6.99% for a minority of hires, and 2.58% for none of their new hires. When asked about why they use benchmarks for some but not all new hires, 68.91% answer they search for the benchmarks only for some positions

and then apply them to all employees, 17.85% that they only search for specific employees within a given position, and 18.10% provide an open-ended response. Among open-ended responses, one common reason is that they search only for new positions in the firm. For instance, one respondent wrote: "For new hires, we primarily use the salaries of current employees rather than salary benchmarks. If it's a position we don't have that for, then we use the salary benchmarks.". Another reason is that they only search for positions that are more competitive or challenging to fill. For example, one respondent wrote: "Certain positions that are becoming more competitive in our industry or market." Some responses also mention a range of other reasons, such as the role of unions or budget constraints.

In relation to the sources used to obtain salary benchmarks, Table C.2 shows that the most popular options are industry surveys and free online data sources (68% and 58.07% of participants, respectively, indicate that they use these sources). Other popular options are government data (37.11%), paid online data sources (34.37%), compensation consultants (26.30%), and payroll data services (23.19%). Among HR managers in our sample, 48.59% have used Glassdoor as their salary benchmark source and 9.48% our partner firm's benchmarking tool. Table C.2 also shows that when looking up benchmark information, a strong majority apply filters for state and industry (84.15% and 87.33%, respectively). Other popular filters are firm size (48%), revenue size (38.96%), and hourly vs. salaried (37.11%). Figure C.1 presents the piece of information they usually care most about when looking for the benchmark for a position. Most (56. 73%) ranked the median salary first. The second most popular piece of information is the average salary (32.59% ranked it 1st), and only the remaining 10.68% choose the 10th percentile, 25th percentile, 75th percentile, or the 90th percentile as the one they care about the most.

C.3 Other Information Frictions

For reference, we included a survey question to assess whether there may be frictions in accessing internal information. More precisely, we asked respondents if they could access information on the median salary that their own company pays to employees in a specific position. Most (79.78%) of the participants indicated that they could access the information easily, 19.26% indicated that it would require some work, and less than 1% indicated that they do not have access. In addition, we asked HR managers how frequently their employees share with them the offers they receive from other firms. Table C.1 shows that 1.7% of the participants responded *Always,* 17.33% *Often*, 61.85% *Sometimes*, 15.56% *Rarely* and 3.56% answered *Never.* This evidence suggests that employers have limited information on competing offers.

C.4 Survey Experiment on the Effects on Salary Dispersion

We embedded an experiment in the SHRM survey. With these data, we can provide complementary experimental evidence on the effects on salary dispersion documented in Section 3. For each participants, we face them with two sequential scenarios. In each scenario, respondents pick a positions for which they plan to hire in the future. Second, we ask them for the annual base salary they are willing to pay for a new hire in that position. Third, we provide them with information on a hypothetical benchmark for that position. After they receive the feedback, we re-elicit the salary they are willing to pay for that position. Participants receive one of two types of feedback: half of the scenarios receive a benchmark that is 15% *below* what they were originally planning to offer, while the other half receive a benchmark that is 15% *above*.

Figure C.2 presents the distribution of the percentage change in the salary offer after receiving feedback. Panel A shows the salary update where the respondent was shown a benchmark 15% above the initial offer, and Panel B corresponds to scenarios where the respondent was shown a benchmark 15% below the initial offer. Figure C.2 shows that, consistent with the results from the payroll data, salaries get compressed toward the benchmark, both from above (Panel A) and below (Panel B). More precisely, when individuals learn that the benchmark is above their planned salary, they react by revising their offer upward; when they find that the benchmark is below their planned salary, they react by revising their offer downward.

An interesting similarity exists between the results of the survey experiment and the results from the natural experiment presented in Section 3. In the natural experiment, it looks like some individuals react to the benchmark information by fully updating (i.e., by bunching at exactly the median benchmark), while other individuals react by updating partially. We observe a similar pattern in the survey experiment. For example, panel A of Figure C.2 indicates that 27.9% of subjects fully updated to the feedback (i.e., they revised their offers upward by exactly 15%) while 44.2% updated only partially (i.e., they revised their offers upward by a number between 0% and 15%). On the other hand, there is one difference between the results of the survey experiment and the results from the natural experiment. In the natural experiment, the compression from above is similarly strong as the compression from above. In the survey experiment, the compression from above (Panel B) is stronger than the compression from below (Panel A). However, there are two natural explanations for this discrepancy. The simplest explanation involves downward wage rigidities: respondents might implicitly assume that the initial salary they selected was already conveyed to the hypothetical candidate, leading to a reluctance to lower the salary amount after receiving the feedback. Alternatively, social desirability bias provides another explanation. Due to

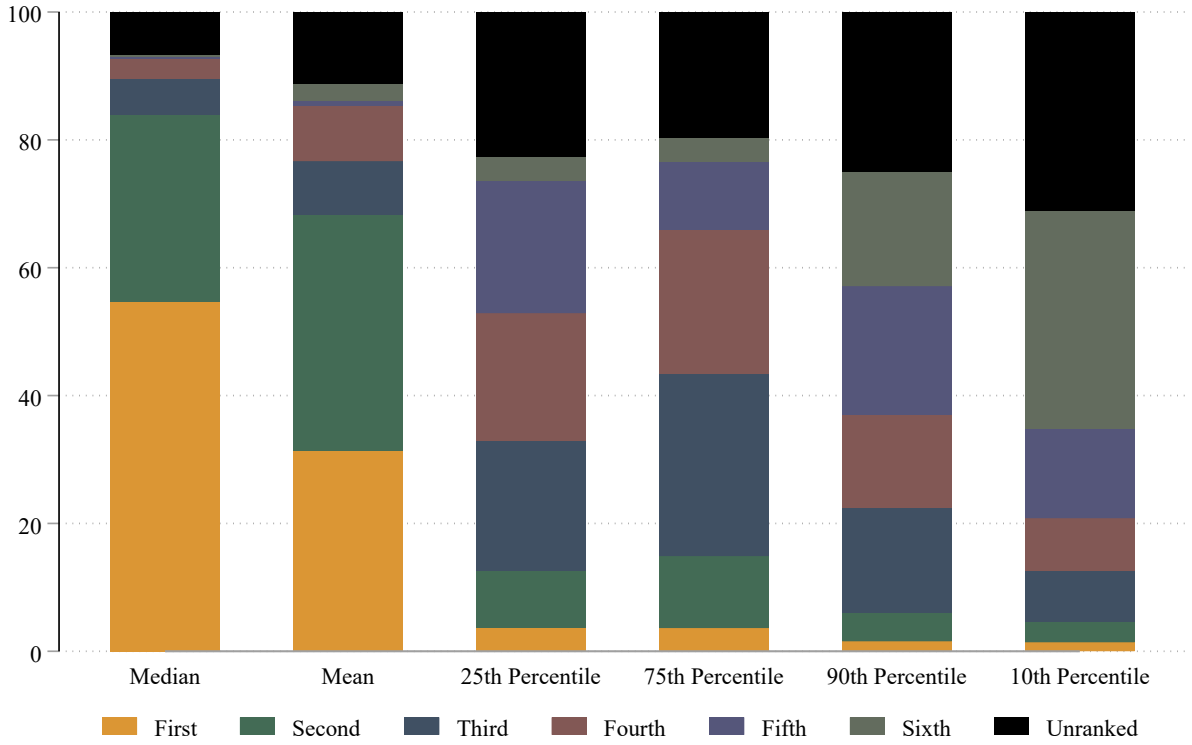
the survey experiment's non-incentivized nature, respondents may wish to appear generous in the eyes of the researchers by being open to increasing salaries but reluctant to reducing them.

C.5 Affiliated Values

The model presented in Section 5 relies on a key assumption of affiliated values. We conducted another survey experiment to provide a test of this assumption. We begin by asking participants to choose the position for which the benchmark is most useful to them and to think of their two closest competitors (firm A and firm B) that also hire in that position. Second, we elicit the maximum salary they think firm A would be willing to pay a new hire in that position. Third, we give them information on a hypothetical salary that firm A is actually willing to pay. The participants are assigned to receive information that the salary is 15% below or 15% above their initial guess. After receiving the information on the salary firm A is willing to offer, we elicit the salary they think firm B would be willing to pay to hire in that position (posterior salary).

The results of this survey experiment are reported in Figure C.3, which shows the distribution of the %-change in their guess after receiving feedback. Panel A presents this outcome when the feedback is 15% above the initial guess, and Panel B when the feedback is 15% below. Consistent with the assumption on affiliated values, Figure C.3 shows compression towards the competitor's salary, both from above (Panel A) and from below (Panel B). Intuitively, when individuals learn that firm A values the worker more (less) than they thought initially, they update upward (downward) what they think firm B would be willing to pay.

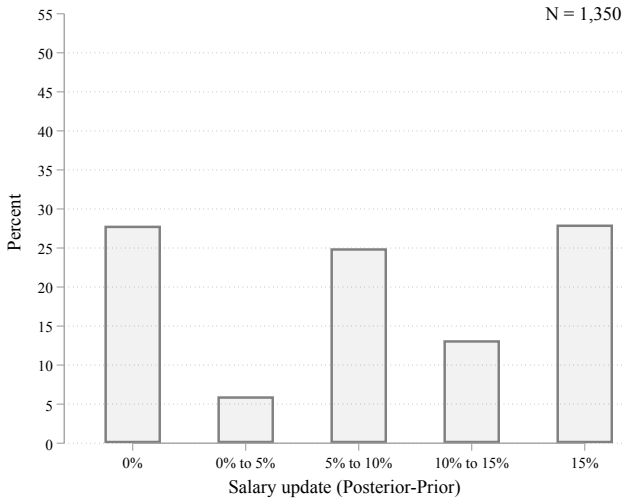
Figure C.1: Ranking of Salary Benchmark Statistics



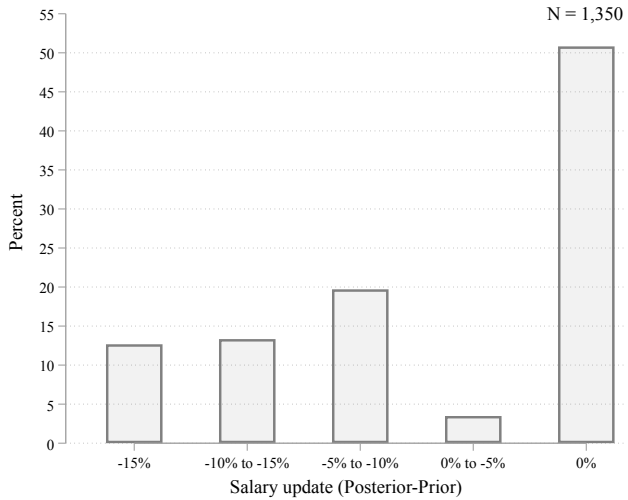
Notes: Responses from a sample of 1,301 HR managers that participated in the SHRM survey. The survey question asks participants to rank these pieces of information from most important to least important.

Figure C.2: Salary Updating with Benchmark Information

PANEL A: When benchmark is 15% above prior



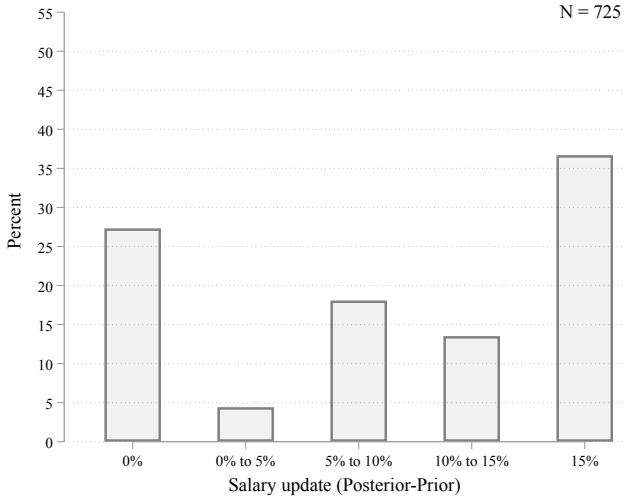
PANEL B: When benchmark is 15% below prior



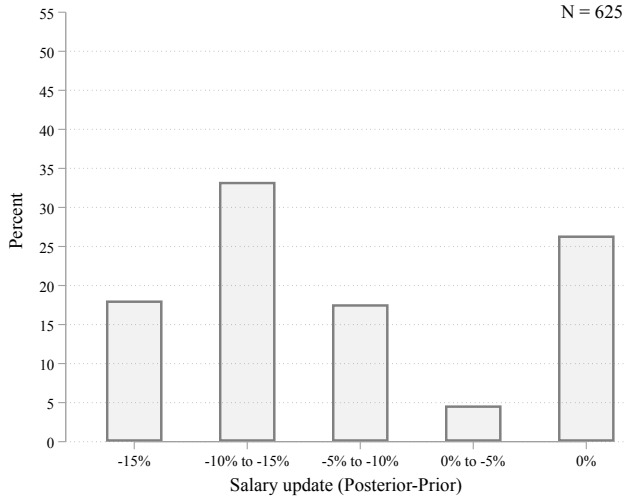
Notes: Salary updating with benchmark information in experimental setting. Participants are asked to choose two positions they expect to hire in the future and to indicate the base salary for the new hires, before and after receiving information on an hypothetical benchmark for each position. The hypothetical benchmark is 15% above their prior salary for one position, and 15% below for the other. Participants are randomly assigned to receive first a higher or lower benchmark. Panel A shows the base salary update for positions where the salary benchmark was 15% above the prior salary, and Panel for positions where the benchmark was 15% below.

Figure C.3: A Test of Affiliated Values

PANEL A: When competitor's salary is 15% above prior



PANEL B: When competitor's salary is 15% below prior



Notes: A test of affiliated values in an experimental setting. Participants are asked to choose the most relevant position for them and to think of their two closest competitors (A and B) that hire in that position. We elicit their beliefs on the maximum amount competitor A will be willing to pay to hire a worker in the position (prior salary), and then we provide them information on the actual (hypothetical) salary that competitor A would be willing to pay. We then elicit the maximum amount they think the competitor B will be willing to pay for to hire a worker in that position (posterior salary). Participants were randomly assigned to receive as information a competitor's salary 15% above their prior or 15% below. Panel A shows the salary update when the competitor's salary was 15% above the prior salary, and Panel B when it was 15% below.

Table C.1: Characteristics of the Sample

Questions	Share of responses (%)
How many employees does your company have?	
1-49	22.15
50-99	23.48
100-999	36.67
1000-4999	9.70
5000 or more	8.00
Are you working in the private sector or the public sector?	
Private sector	80.81
Public sector	19.19
What main industry do you operate in?	
Agriculture, Forestry, Fishing and Hunting	0.45
Mining, Quarrying, and Oil and Gas Extraction	0.45
Utilities	0.96
Construction	4.08
Manufacturing	15.43
Wholesale Trade	3.04
Retail Trade	4.01
Transportation and Warehousing	1.71
Information	3.49
Finance and Insurance	7.12
Real Estate and Rental and Leasing	2.15
Professional, Scientific, and Technical Services	18.99
Management of Companies and Enterprises	0.07
Administrative and Support and Waste Management and Remediation Services	3.41
Educational Services	5.04
Health Care and Social Assistance	11.80
Arts, Entertainment, and Recreation	2.23
Accommodation and Food Services	1.56
Other Services (except Public Administration)	7.86
Public Administration	6.16
How many years of experience do you have setting salaries?	
Less than 1 year	2.44
1-5 years	34.30
6-10 years	25.70
11+ years	37.56
Do you participate in salary settings for:	
New hires	6.44
Current employees	2.52
Both	91.04
Do you typically set salaries for:	
Higher-education positions	13.93
Lower-education positions	6.07
Both	80.00
Do you have access to the median salary that your company pays employees in a specific position?	
No, I could not access that data even if I wanted to	0.96
Yes, I can access it easily	79.78
Yes, but it would take some work	19.26
If your employees get an offer from another company, do they share the terms of the offer with you?	
Never	3.56
Rarely	15.56
Sometimes	61.85
Often	17.33
Always	1.70

Notes: Characteristics of the sample of 1,350 HR managers that participated in the SHRM survey.

Table C.2: Use of Salary Benchmarking Tools

Questions	Share of responses (%)
What do you use the salary benchmark for? (Select all that apply) [N = 1,350]	
To set precise salaries for new hires	54.07
To change salaries for current employees	76.81
In salary negotiations	53.11
To set salary ranges for specific job titles	89.78
To determine salary in job advertisement	40.89
To plan ahead for headcount	25.33
How frequently do you use salary benchmarks to set salaries for new hires? [N = 1,316]	
Never	2.58
A minority of hires	6.99
Some of the hires	26.06
A majority of hires	27.81
For every hire	36.55
When do you use salary benchmarks in relation to new hires? (Select all that apply) [N = 1,316]	
Before I publicize the position to include the expected salary in a job advertisement	67.19
Right before I make an offer to the candidate	34.96
After the candidate receives the offer, if the candidate wants to negotiate	22.30
When the candidate presents an outside offer	12.44
How frequently do you use salary benchmarking to change salaries for current employees? [N = 1,263]	
Never	1.82
For a minority of employees	10.61
For some of my employees	25.89
For a majority of my employees	21.77
For all my employees	39.90
When do you use salary benchmarks with current employees? (Select all that apply) [N = 1,240]	
When the employee goes through an annual review	48.96
When the employee is up for promotion	47.48
When the employee presents an outside offer	33.11
When adjusting the salary ranges for positions	74.30
Why do you use salary benchmarks for some but not all employees? (Select all that apply) [N = 801]	
I search for some specific employees (within a given position)	17.85
I search only in some specific positions (and apply it to employees in those positions)	68.91
Other	18.10
Which sources do you use to obtain salary benchmarks? (Select all that apply) [N = 1,350]	
Free online data sources	58.07
Paid online data sources	34.37
Industry surveys	68.00
Government data	37.11
Compensation consultants	26.30
Payroll data services	23.19
Have you ever used Glassdoor as your salary benchmark source? [N = 1,350]	
Yes	48.59
No	51.41
Have you ever used [the Firms's] Data Cloud Compensation Explorer as your salary benchmark source? [N = 1,350]	
Yes	9.48
No	90.52
Which filters would you typically apply when searching for a position salary benchmark? [N = 1,350]	
Industry	87.33
State	84.15
Firm size	48.00
Revenue size	38.96
Hourly vs. salaried	37.11
None of the above, the position-level filter is sufficient	0.00
Which piece of information you typically care about the most? (% that ranked first) [N = 1,301]	
Median salary	56.73
Average salary	32.59
10th percentile	1.46
25th percentile	3.84
75th percentile	3.77
90th percentile	1.61
For which positions are salary benchmarks most useful, higher-education or lower-education? [N = 1,350]	
Not useful for either group	0.30
Most useful for higher-education positions	21.78
Most useful for lower-education positions	3.93
Equally useful for both groups	74.00

Notes: Responses to survey questions on the use of salary benchmarking tools. The subject pool consists of 1,350 HR managers that participated in the SHRM survey. The relevant sample size, which may be different from question to question, is reported in brackets. For questions where more than one option could be selected (i.e., "select all that apply"), we report the share of total responses that selected that option, and thus the percentages can add up to more than 100%.

D Additional Details about the Payroll Data

D.1 Comparison with a Representative Sample of U.S. Firms

Table D.1 provides a comparison between the firms in our sample and a representative sample of U.S. firms (Song et al., 2019). We match the sample restrictions of Song et al. (2019) by excluding firms with fewer than 20 employees, and employees outside of 20 to 60 years of age. In terms of size, measured in the number of employees, our sample is most representative of the top quartile of firms in the United States. This is probably because businesses with fewer than 100 employees do not have enough scale to justify the use of data analytics services. 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).

Table D.2 provides some statistics on the distribution of industries, given by the first 2 digits of the firm's main 6-digit industry code. Columns (1) and (2) compare the distribution of sectors in our sample (column (1)) with the U.S. distribution according to Census data (column (2)). Columns (3) and (4) are the same as columns (1) and (2), except that they are based on the number of employees instead of the number of firms. We should not expect our sample to be perfectly representative of the U.S. industries. For example, as discussed above, the firms in our sample are larger than the U.S. average, and as a result they will be more representative of industries with larger firms. While not perfectly representative of the U.S. average, our sample provides broad coverage of the U.S. industries. Some industries, such as manufacturing and finance, are somewhat overrepresented, while some other industries, such as construction and accommodation and food services, are somewhat underrepresented.

D.2 Timing of Adoption and Utilization

Figure D.1 shows the timing of onboarding of the firms in our dataset. The onboarding dates range from December 2015 to January 2020. Table D.3 presents a list of the most popular position titles in the Searched position. Mechanically, this list includes a lot of positions at the lower end of the company's hierarchies, as those are the positions in which the companies hire most often. The important lesson from Table D.3 is that there is a lot of overlap between the different position types. For instance, while bank tellers fall into the Searched category for some firms, there are plenty of bank tellers in the Non-Searched and Non-Searchable categories as well. The same is true for each of the other position titles listed in this table.

D.3 Classification of Positions by Skill Level

In Table D.4, we provide more details on the classification of the positions as high-skill and low-skill. This table reports some characteristics of the 35 most searched position titles. Column (1) indicates if a position is classified as high-skill, as defined in Section 2.4. Column (2) indicates whether a position is high-education or low-education. Column (3) indicates if the mean age among new hires in a position is equal to or above 31. Column (4) indicates if the mean salary among new hires in a position is equal to or above $30,000. A position is categorized as high-skill if: it is categorized as high-education (column (2)); or if it is low-age (column (3)) and low-earnings (column (4)).[67]

D.4 Comparison between the Proprietary Benchmark and a Free Online Benchmark

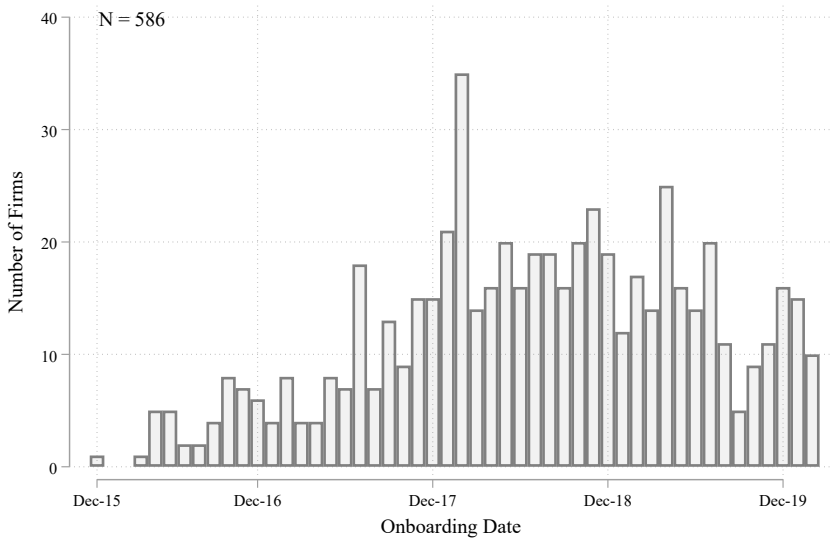
In Figure D.2, we compare our salary benchmark, constructed from proprietary payroll data, with one popular free online salary benchmark. The alternative benchmark is constructed from anonymous users who voluntarily provide their salary information. According to the SHRM survey, 48.59% of HR managers report having used this free online source for their salary benchmarking. To compare these two benchmarks, we use the salient figure provided by the free online tool with the similarly salient median salary reported in our proprietary salary benchmark.[68] We select the 100 most searched positions in our proprietary benchmark tool, and collect the salary benchmark available for these positions in the free online website in 2019 using The Wayback Machine. For each position, we compare the proprietary and the free online benchmark for the same quarter of 2019. Figure D.2 presents a histogram of the percent difference between the proprietary salary benchmark and the free online benchmark. While the propietary benchmark and the online source are similar on average, there is a significant deal of variation. Most of the positions in the public benchmark are more than 15% off from the propietary benchmark.

[67] 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%.

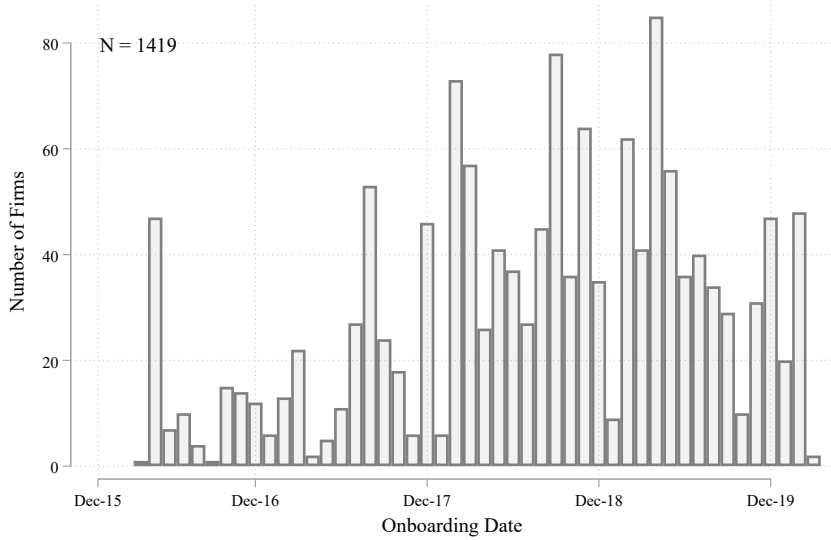
[68] On the free online tool, the salient figure is referred to as the "average salary". However, in the technical notes they report that it is calculated as the median salary.

Figure D.1: Onboarding Dates for the Firms in Our Sample

PANEL A: Treated Firms

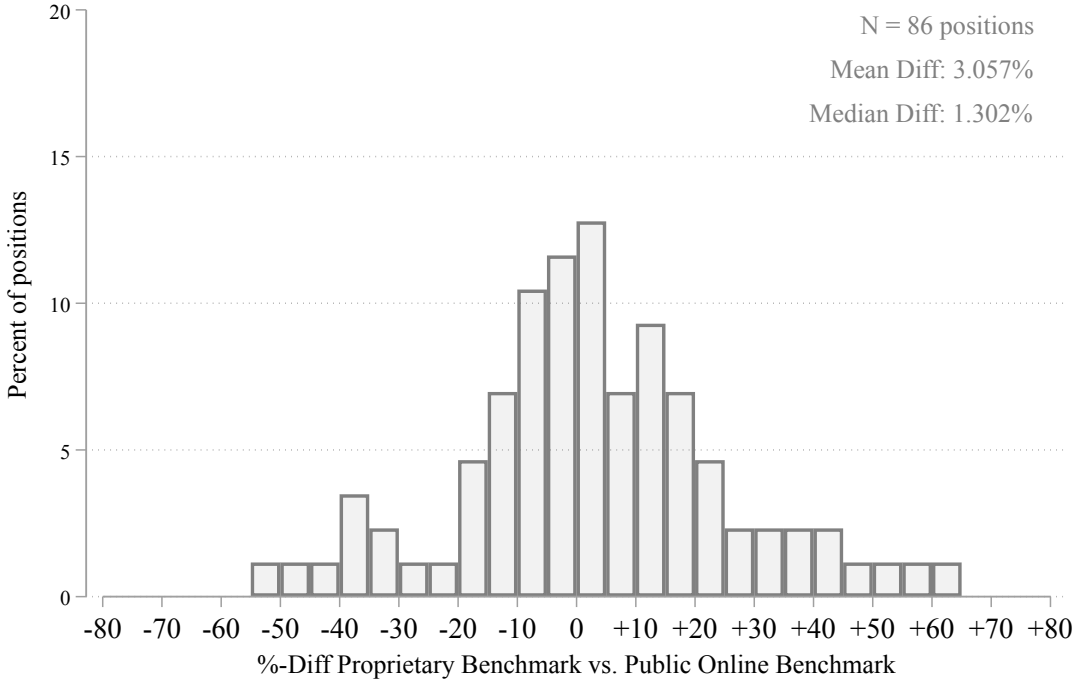


PANEL B: Control Firms



Notes: Panel A presents the onboarding dates of the 586 firms in our main data sample. Panel B presents the "hypothetical" onboarding date assigned to the firm that never gains access to the tool. See section 2.6 for more details.

Figure D.2: Comparison Between Proprietary vs. Public Online Benchmarks



Notes: Histogram of the percent difference between the proprietary median annual base salary benchmark relative to a public online benchmark of median salary. The graph includes salary benchmarks from 2019 for 86 of the 100 most searched positions for which the public benchmark was available. For each position, we calculate the difference between the proprietary and the public online benchmark as (proprietary-public)*100/public. The salaries from the public online benchmark are obtained using The Wayback Machine.

Table D.1: Comparison of Firms in Our Sample vs. Representative Sample of U.S. Firms

	Percentile				
	10th	25th	50th	75th	90th
Number of Employees					
Our Sample	68	109	225	529	1,159
U.S. Representative Sample	22	26	39	79	189
Salary (Annual $)					
Our Sample	20,071	25,468	38,177	64,604	105,689
U.S. Representative Sample	9,820	19,200	36,000	63,200	104,000

Notes: *U.S. Representative Sample* corresponds to the statistics of firms taken from the most recent year (2013) of Song et al. (2019). *Our Sample* denotes to the sample of 2,051 firms in our dataset for the earliest period for which data are available (January 2016). To make the statistics more comparable across the two samples, we match the sample restrictions from Song et al. (2019) by excluding firms with less than 20 employees and employees younger than 20 years old or older than 60 years old. Our *Salary* statistics are based off the distribution of individual annual base salaries across employees in all firms. Song et al. use earnings. To make the two samples more comparable, we converted the salary statistics in our sample to 2013 dollars using the PCE deflator published by the Bureau of Economic Analysis.

Table D.2: Comparison of Sector Representation in Our Sample vs. U.S. Employees & Firms

	Firms (%)		Employees (%)	
	(1)	(2)	(3)	(4)
Sector	Our Sample	U.S.	Our Sample	U.S.
Agriculture, Forestry, Fishing and Hunting	0.25	0.37	0.37	0.13
Mining, Quarrying, and Oil and Gas Extraction	0.44	0.32	0.11	0.45
Utilities	0.44	0.10	0.34	0.50
Construction	2.33	11.58	0.51	5.08
Manufacturing	22.22	4.10	21.94	9.12
Wholesale Trade	8.87	4.92	14.24	4.76
Retail Trade	3.90	10.70	7.82	12.21
Transportation and Warehousing	2.20	3.05	1.25	3.78
Information	2.77	1.32	3.71	2.73
Finance and Insurance	13.91	3.94	11.10	4.98
Real Estate and Rental and Leasing	3.02	5.11	1.58	1.67
Professional, Scientific, and Technical Services	11.83	13.39	8.56	6.93
Management of Companies and Enterprises	1.01	0.45	1.29	2.69
Administrative and Support and Waste Management	4.59	5.74	6.58	9.25
Educational Services	2.64	1.54	2.51	2.87
Health Care and Social Assistance	11.33	10.81	13.42	15.74
Arts, Entertainment, and Recreation	0.57	2.15	0.40	1.84
Accommodation and Food Services	1.95	8.91	1.59	10.96
Other Services (except Public Administration)	5.73	11.50	2.70	4.30

Notes: Percent of firms and employees in each sector in our sample vs. in the U.S. The NAICS code. *Public Administration* excluded from statistics of our sample because the Census does not report data for that code.

Table D.3: 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 D.4: Position Characteristics

Position Title	(1) High Skill	(2) High Educ.	(3) $\bar{Age} \geq 31$	(4) $\bar{Inc.} \geq \$30K$
Customer Service Representative	Y	N	Y	Y
Security Guard	Y	N	Y	Y
Legal Associate Specialist	Y	Y	Y	Y
Patient Care Coordinator	Y	N	Y	Y
Medical Assistant	Y	Y	Y	Y
Welder	Y	N	Y	Y
Registered Nurse	Y	Y	Y	Y
Software Developer/Programmer	Y	Y	Y	Y
Mammographer	Y	Y	Y	Y
Production Operations Engineer	Y	Y	Y	Y
Licensed Practical Nurse	Y	Y	Y	Y
Sales Manager	Y	Y	Y	Y
General Practitioner/Physician	Y	Y	Y	Y
Lawyer	Y	Y	Y	Y
Ophthalmic Technician	Y	Y	Y	Y
Business Development Specialist	Y	Y	Y	Y
Warehouse Manager	Y	N	Y	Y
Other Social Work and Counseling Professional	Y	Y	Y	Y
Operations Officer	Y	Y	Y	Y
Bank Teller	N	N	N	N
Hotel Cleaner	N	N	Y	N
Hand Packer	N	N	Y	N
Receptionist	N	N	N	N
Cook	N	N	Y	N
Waiter/Waitress	N	N	N	N
Delivery Driver	N	N	Y	N
Dish Washer/Plate Collector/Table Top Cleaner	N	N	N	N
Cashier	N	N	N	N
Assembler	N	N	Y	N
Other Housekeeper and Related Worker	N	N	Y	N
Warehouse Laborer	N	N	Y	N
Nursing Assistant	N	N	Y	N
Bartender/Mixologist	N	N	Y	N
Building Caretaker/Watchman	N	N	Y	N
Shipping Clerk	N	N	Y	N

Notes: List of 35 position titles with the most hires in the Searched group. A position is classified as low-skill if: (i) it is classified as low education; (ii) the average age is below 31 years; (iii) the average salary is less than $30,000.

E Effects on Salary Dispersion: Additional Results and Robustness Checks

E.1 Dispersion Around Different Benchmarks

Figure E.1 reproduces Figure 2 using a narrower bin width (+/- 1%). This figure highlights bunching at exactly the median. Panel A of Figure E.2 replicates Panel A of Figure 2, which presents the dispersion of salaries around the median benchmark in Searched positions both before and after onboarding. In panels B through F of Figure E.2, we replicate the analysis but, instead of analyzing the difference between the salaries and the median benchmark, we look at the difference with respect to the other points of the benchmark distribution (e.g., the 10th percentile benchmark). Panel B shows the change in percent difference from the 10th percentile benchmark, both before (denoted by solid gray bins) and after (hollow red bins) the onboarding date. Before onboarding, salaries in the Searched positions were on average 34.3 pp from the 10th percentile benchmark. After the onboarding date, the average distance to the benchmark fell to 29.2 pp (p-value<0.001). While this corresponds to a reduction in salary dispersion, Panel B shows that this decrease is driven by a higher share of salaries 10-20% above the 10th percentile benchmark, rather than a higher share of salaries at the 10th percentile benchmark itself. The same result holds when we analyze compression around the other points of the distribution: the 25th percentile benchmark (Panel C), the average benchmark (Panel D), the 75th percentile benchmark (Panel E), and the 90th percentile benchmark (Panel F).

E.2 Effects on Composition of New Hires

The types of new hires that join a firm after a firm gains access to the salary benchmark may shift. To test this hypothesis, Table E.1 presents difference-in-differences estimates using the characteristics of the employee (instead of their starting salary) as the dependent variable. In column (1), the dependent variable is an indicator for whether the employee is female. In column (2), the outcome is an indicator variable for whether the employee was hired through an hourly contract. In column (3), the dependent variable is the employee's age in years. The coefficients on the post-treatment coefficients are close to zero and statistically insignificant, suggesting that access to benchmarking did not have a significant effect on the composition of new hires. However, this result must be taken with a grain of salt for at least two reasons. First, we do not have sufficient precision to rule out modest effect sizes. Second, these results cover just three basic employee characteristics (gender, contract type and age).

E.3 Other Robustness Checks

Table E.2 and Table E.3 present the same difference-in-differences regressions as in Table 3, but with different clustering for the standard errors. Table E.2 clusters at the firm level and Table E.3 clusters at the firm-position level. The results are largely robust across all these different specifications.

Table E.4 presents additional robustness checks for our difference-in-differences estimates for the effect on absolute dispersion from the benchmark. Column (1) replicates the baseline specification from column (1) of Table 3. Columns (2) through (7) each change a different aspect of the baseline specification. In columns (2) and (3), we use alternative benchmark filters to compute dispersion from the benchmark. In our main specification we use benchmarks for a given position title filtered by state and sector when that benchmark is based on more than 30 employees, and no filters otherwise. In column (2), we use only the unfiltered benchmark and in column (3) we use the filtered benchmark only when it based on more than 100 employees and no filters otherwise. In column (4), we include job titles with low "match scores" (the quality of the mapping between a firm-specific job title and a position title) that are filtered out of the baseline specification. Column (5) uses only Non-Searched and Non-Searchable new hires after September 2019, the start of our search data. Column (6) includes data from August 2020 through July 2021.[69] Finally, column (7) excludes all HR positions (2.28% of the sample). Intuitively, HR employees may be looking their own salaries up due to curiosity, or for their own salary planning, rather than to negotiate with the new hires in HR positions. The results from the baseline specification in column (1) of Table E.4 are qualitatively and quantitatively similar to the results in each of the alternative specifications in columns (2) through (7).

E.4 Additional Heterogeneity Analysis

Figure E.3 presents the event-study results for low-skill and high-skill positions. The panels in the top row of Figure E.3 (i.e., panels A and B) correspond to the low-skill positions, while the panels below (i.e., panels C and D) correspond to high-skill positions. Before firms had access to the tool (i.e., the left side of each panel), for the Searched positions, there was more dispersion among high-skill positions (22.6%) than in low-skill positions (14.2%). The difference-in-differences comparison for Searched vs. Non-Searched low-skill positions (Panel A) suggest the benchmark tool reduced the salary dispersion from 14.2 pp to 7.6 pp (p-value<0.001), equivalent to a 46.5% reduction. For Searched vs. Non-Searchable low-

[69] Starting September 2020, the payroll 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. However, the results from this specification include both the old and the new taxonomies.

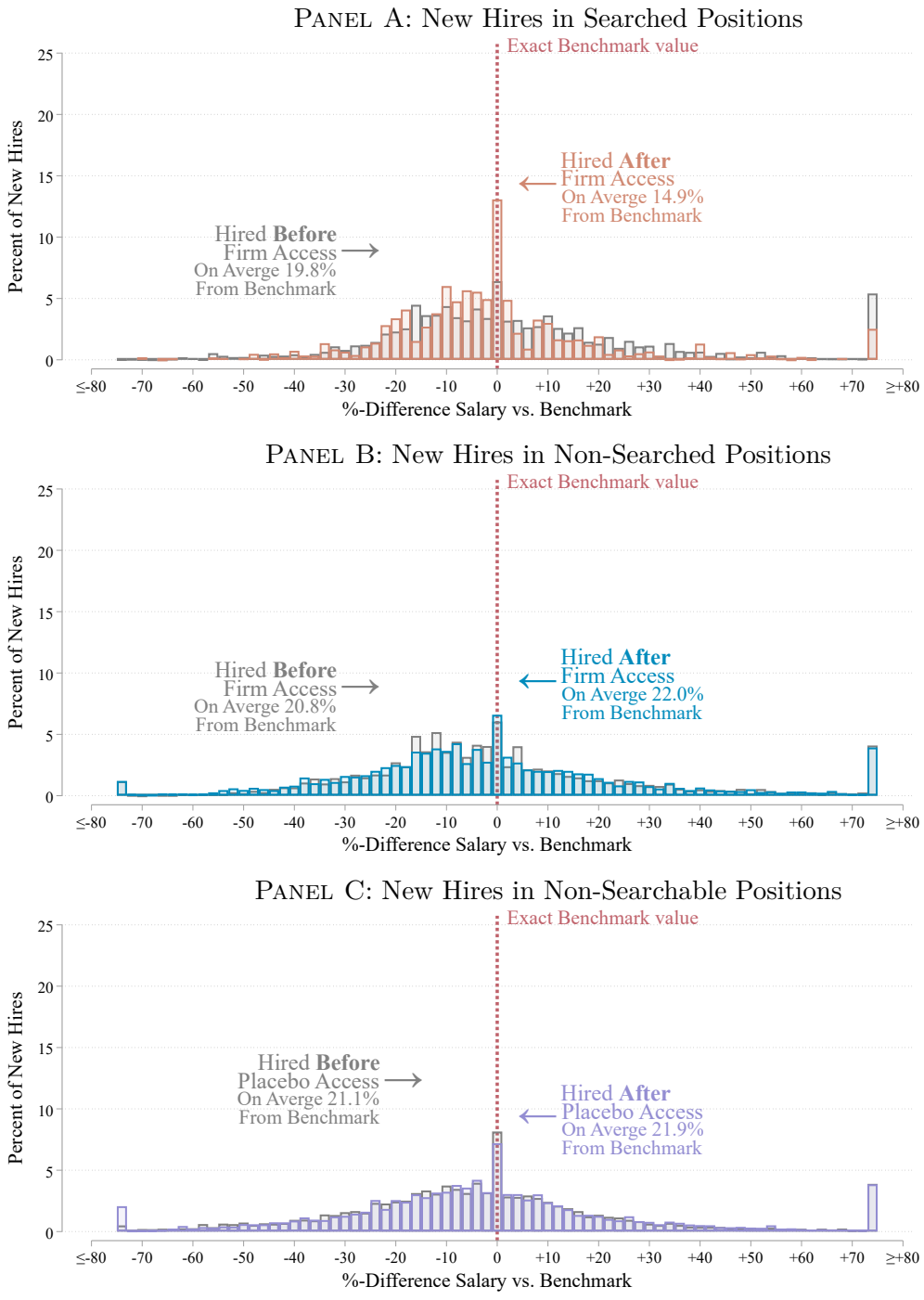
skill positions (Panel B), salary dispersion drops from 14.2 pp to 7.6 pp (p-value<0.001), or a 46.5% reduction. For high-skill positions, the difference-in-differences comparison for Searched vs. Non-Searched (Panel C) suggests that the benchmark tool reduced the salary dispersion from 22.6 pp to 18.9 pp (p-value = 0.019), equivalent to a 16.4% reduction. Finally, comparing Searched vs. Non-Searchable high-skill positions (Panel D) suggests a drop in dispersion from 22.6 pp to 17.2 pp (p-value < 0.001), equivalent to a 23.9% reduction.

Figure E.4 displays the event-study analysis by low-dispersion and high-dispersion of market-level salaries. For each combination of position, state and industry, we calculate a metric known as "market dispersion," which is the difference between the 90th and 10th percentiles of market benchmarks, as indicated in the benchmarking tool. Intuitively, significant variations in salaries within a cell imply a substantial variation in skills. The results from Figure E.4 are qualitatively similar to our high- vs. low-skill heterogeneity analysis, though less stark.

Figure E.6 analyzes heterogeneity by monopsonistic power, using the HHI measure computed by Azar et al. (2022). The original measures are based on job ads data from LightCast (previously BurningGlass) and are at the SOC-commuting zone level. We aggregate to the 2-digit SOC major groups and state level and merge with our data. We split observations by whether the HHI index is above or below the median value in our entire sample. The results from Figure E.6 constitute suggestive evidence indicating that the impact of salary benchmarking is more pronounced in more competitive labor markets.

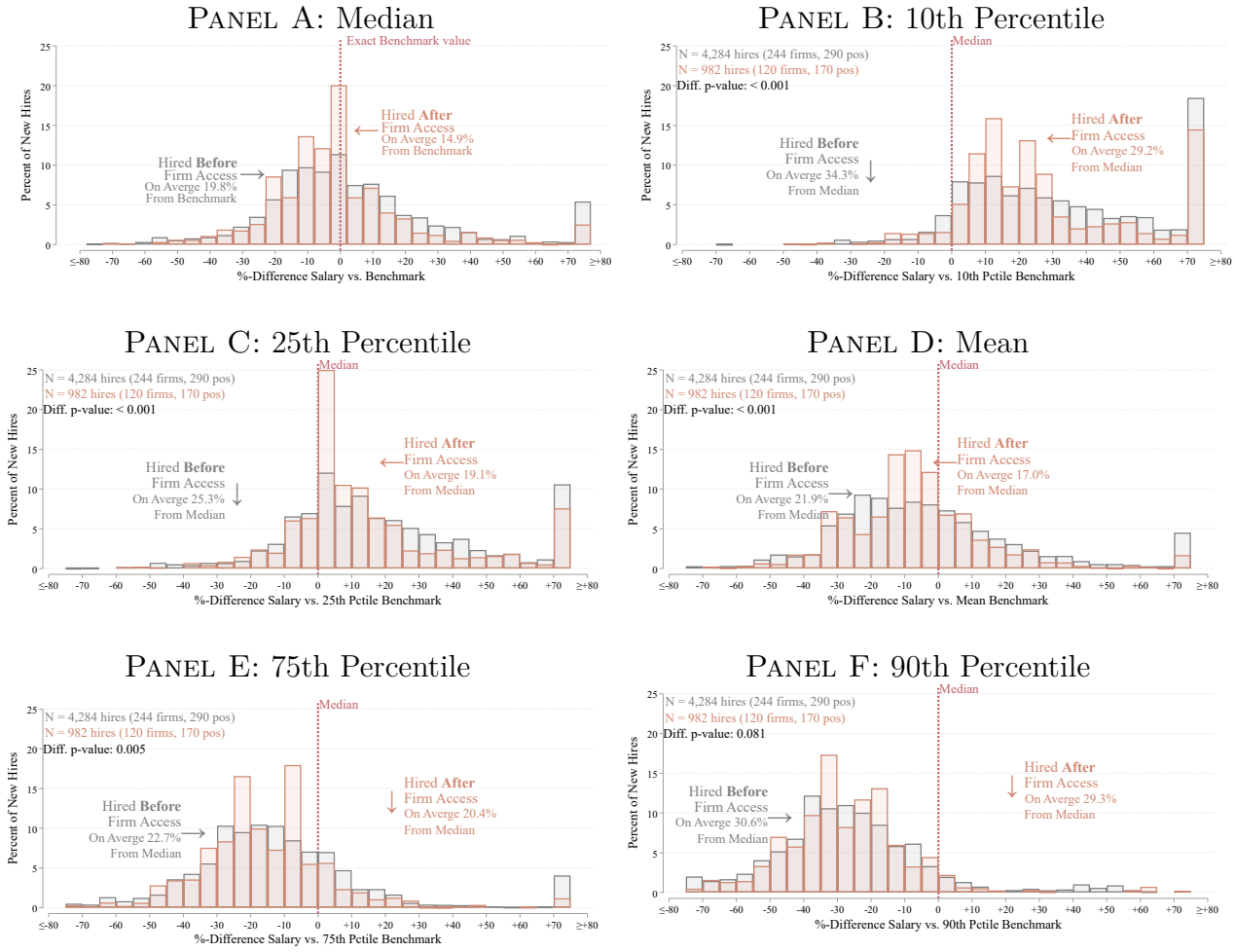
Figure E.5 breaks down the effect on salary dispersion by gender. Panel A and Panel B reproduce Figure 3 with the male subsample, and Panel C and Panel D reproduce Figure 3 with the female subsample. we find differences by gender that are economically small and statistically insignificant.

Figure E.1: The Effects of Benchmarking on Dispersion around the Benchmark: Non-Parametric Analysis



Notes: It reproduces Figure 2 using a narrower bin width (+/- 1%).

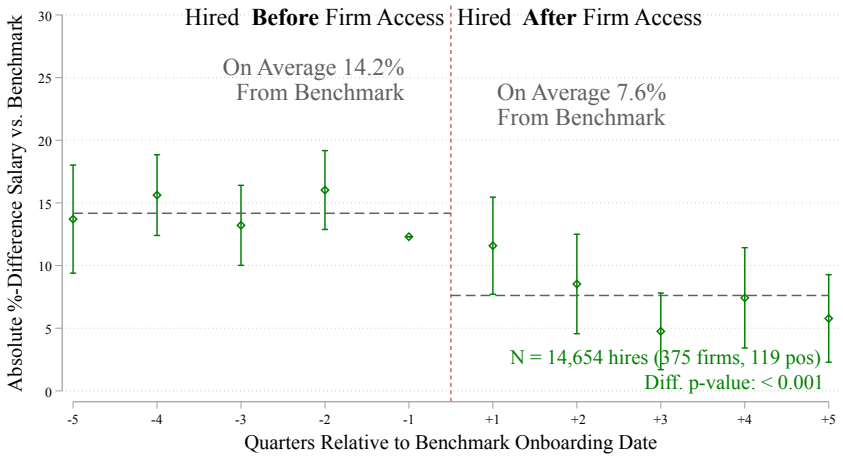
Figure E.2: Effects of the Compensation Benchmark on New Hire Salaries Relative to Different Percentiles of the Benchmark: Non-Parametric Analysis



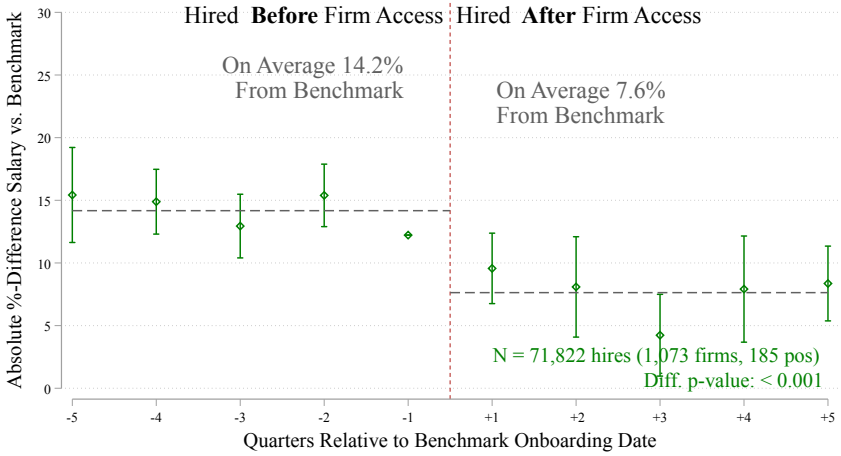
Notes: Panel A is a reproduction of Panel A of Figure 2. All other panels are identical, but using absolute dispersion from the specified percentile of the benchmark distribution rather than the median. For more details, see notes to Figure 2.

Figure E.3: Heterogeneity by Skill: Event-Study Analysis

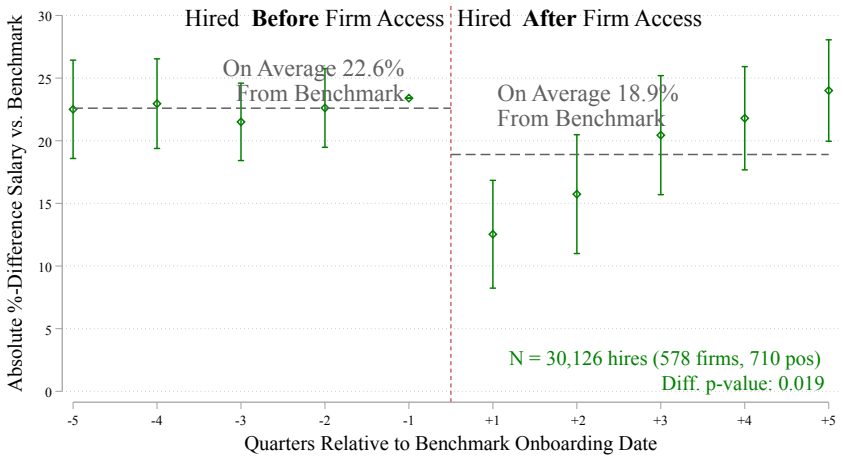
PANEL A: Low-Skill: Searched vs. Non-Searched



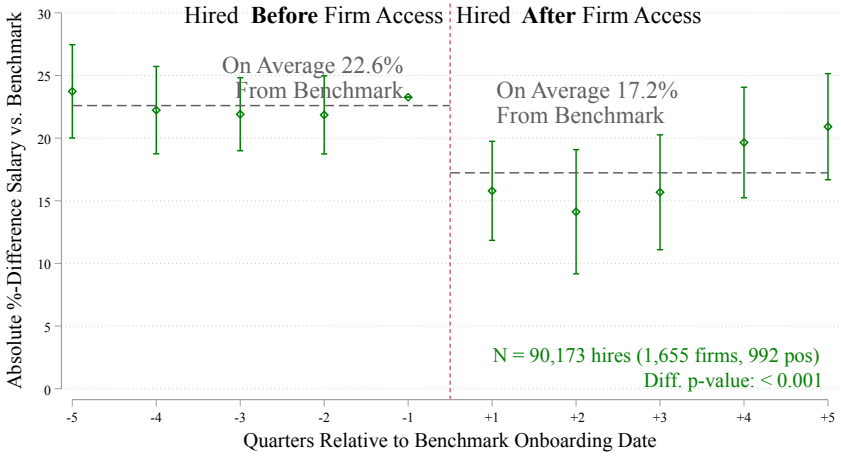
PANEL B: Low-Skill: Searched vs. Non-Searchable



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



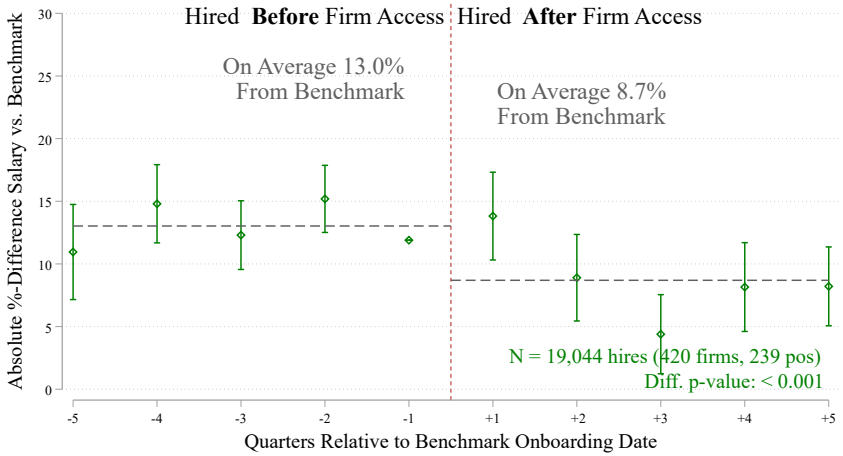
PANEL D: High-Skill: Searched vs. Non-Searchable



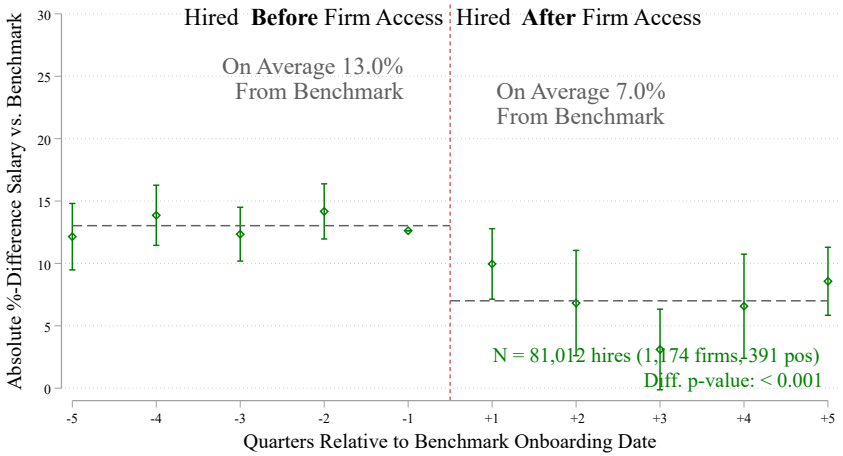
Notes: Panels A and C are a reproduction of Panel C from Figure 3, and Panels B and D are a reproduction of Panel D, but for the specified sub-samples. *Skill* is defined in Section 2.1. See the notes of Figure 3 for more details.

Figure E.4: Heterogeneity by Market Dispersion: The Effects on Pay Dispersion

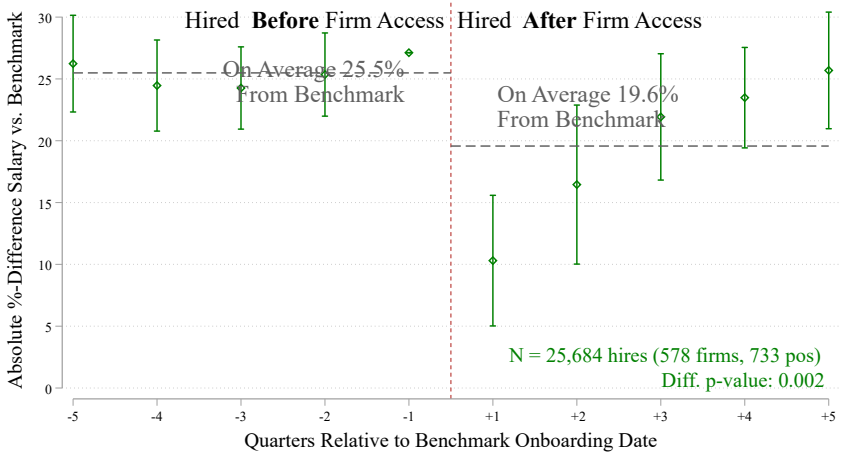
PANEL A: Low Market Dispersion: Searched vs. Non-Searched



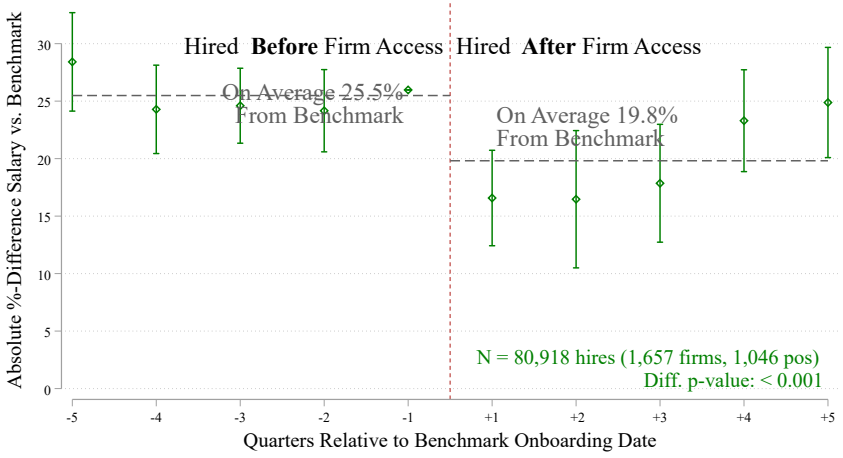
PANEL B: Low Market Dispersion: Searched vs. Non-Searchable



PANEL C: High Market Dispersion: Searched vs. Non-Searched



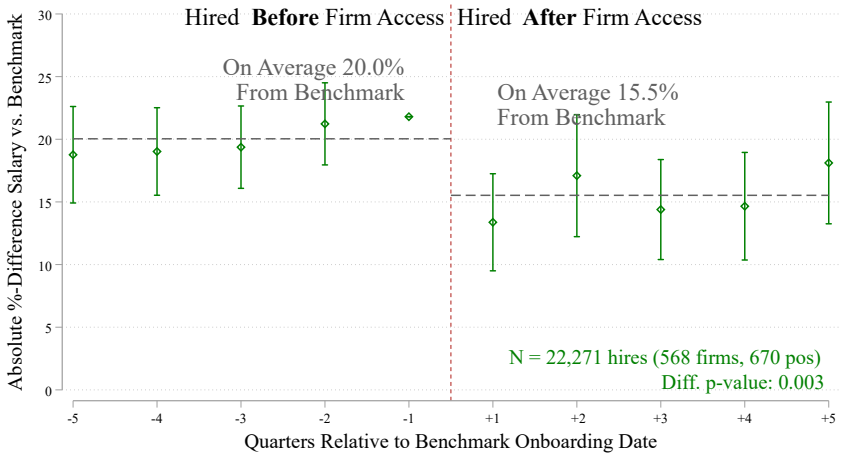
PANEL D: High Market Dispersion: Searched vs. Non-Searchable



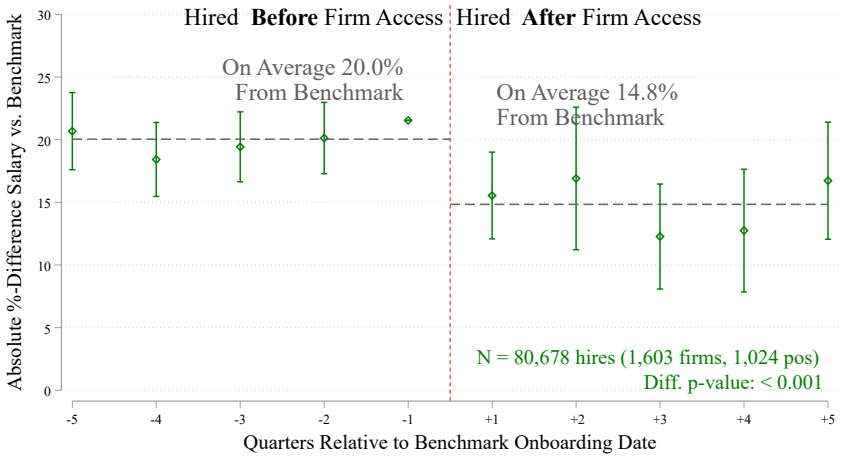
Notes: Panels A and C are a reproduction of Panel C from Figure 3, and Panels B and D are a reproduction of Panel D, but for the specified sub-samples split by the market dispersion. See the notes of Figure 3 for more details.

Figure E.5: Heterogeneity by Gender: Event-Study Analysis

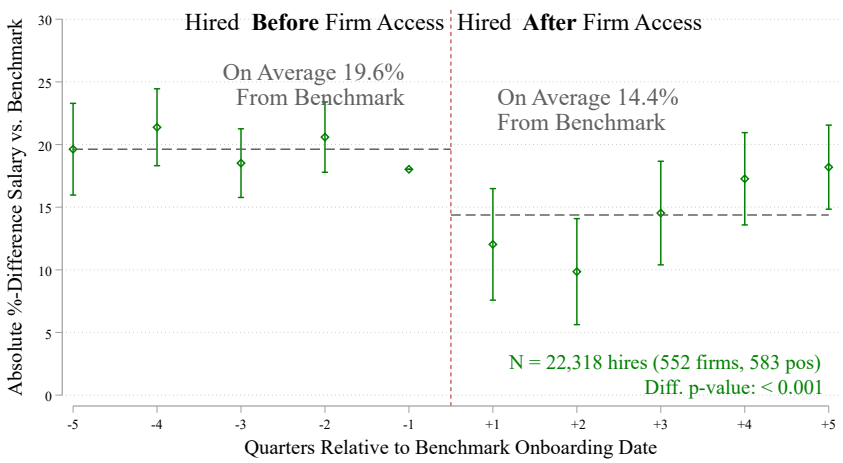
PANEL A: Male: Searched vs. Non-Searched



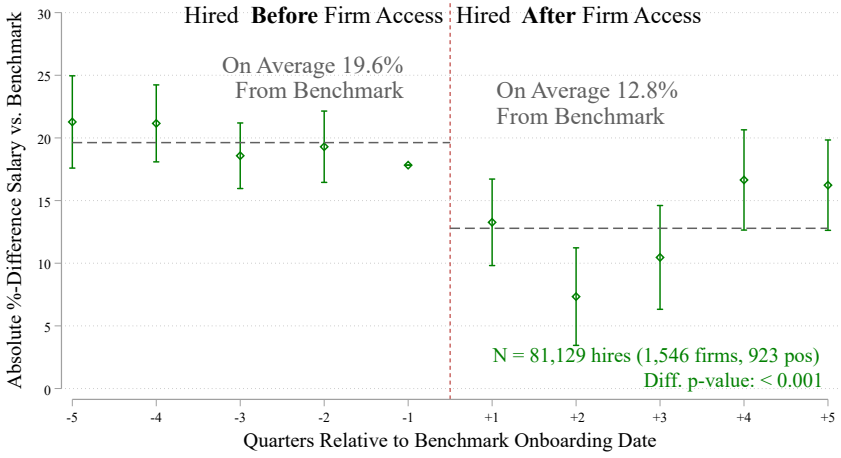
PANEL B: Male: Searched vs. Non-Searchable



PANEL C: Female: Searched vs. Non-Searched



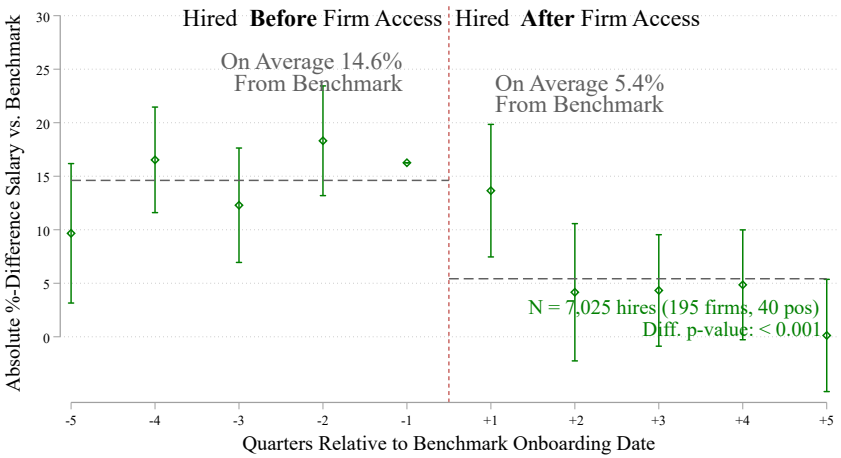
PANEL D: Female: Searched vs. Non-Searchable



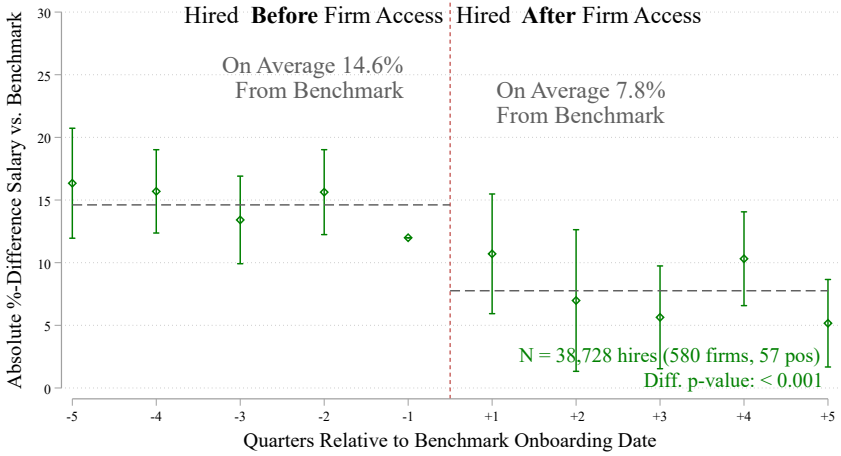
Notes: Panels A and C are a reproduction of Panel C from Figure 3, and Panels B and D are a reproduction of Panel D, but for the specified sub-samples split by gender. See the notes of Figure 3 for more details.

Figure E.6: Heterogeneity by imputed HHI: The Effects on Pay Dispersion (Low Skill)

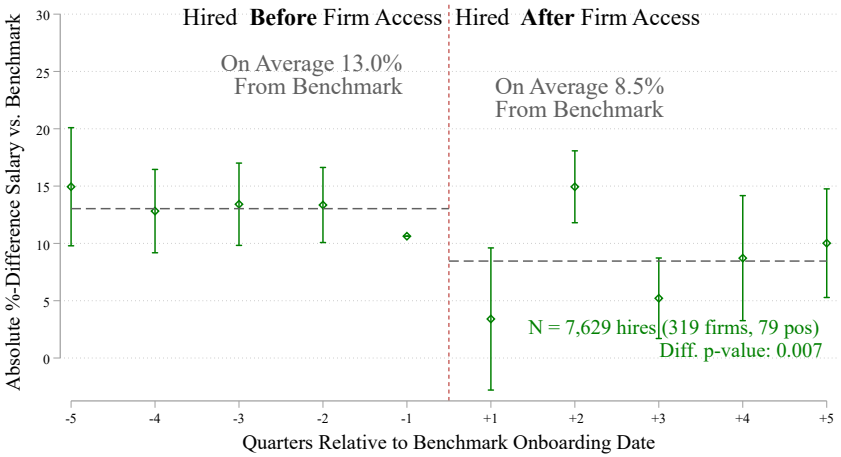
PANEL A: Low HHI: Searched vs. Non-Searched



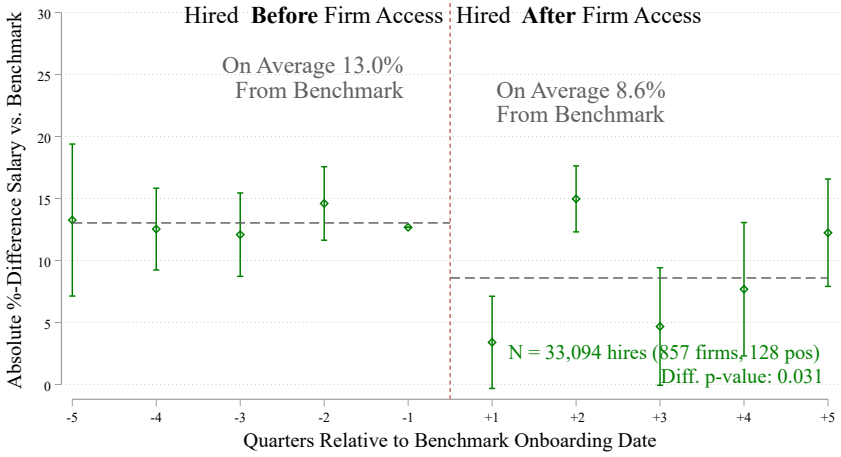
PANEL B: Low HHI: Searched vs. Non-Searchable



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



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



Notes: Panels A and C are a reproduction of Panel C from Figure 3, and Panels B and D are a reproduction of Panel D, but for the specified sub-samples split by the imputed HHI. See the notes of Figure 3 for more details.

Table E.1: The Effects of Benchmarking on Employee Composition

	(1) Female	(2) Hourly	(3) Age
Panel A: Post-treatment			
Searched vs. Non-Searched	-2.306	0.487	-0.749
	(2.424)	(1.722)	(0.740)
Searched vs. Non-Searchable	-1.987	2.708	-0.117
	(2.338)	(1.654)	(0.761)
Panel B: Pre-treatment			
Searched vs. Non-Searched	-1.051	1.605	-0.409
	(2.133)	(1.640)	(0.606)
Searched vs. Non-Searchable	0.584	3.241**	0.0461
	(2.009)	(1.651)	(0.529)
Mean Dep. Var. (Baseline)	54.870	76.482	34.598
Observations			
Searched	5,253	5,253	5,253
Non-Searched	39,527	39,527	39,527
Non-Searchable	156,734	156,734	156,734

Notes: Significant at *10%, **5%, ***1%. Robust standard errors in parentheses. All columns follow the specification of column (1) from Table 3, with the exception that here we exclude the additional controls. The dependent variables are a dummy equal to 100 if a new hire is *Female* and zero otherwise in column (1), a dummy equal to 100 if a new hire is an *Hourly* worker and zero otherwise in column (2), and a new hire's *Age* in column (3).

Table E.2: The Effects of Benchmarking on Salary Dispersion: Clustering at the Firm Level

	(1) $\lvert\%\Delta\rvert$	(2) $\lvert log\Delta\rvert$	(3) $\lvert\%\Delta\rvert > 10$	(4) $\lvert\%\Delta\rvert$	(5) $\lvert\%\Delta\rvert$	(6) $\lvert\%\Delta\rvert$	(7) $\lvert\%\Delta\rvert$	(8) $\lvert\%\Delta\rvert$	(9) $\lvert\%\Delta\rvert$	(10) $\lvert\%\Delta\rvert$	(11) $\lvert\%\Delta\rvert$	(12) $\lvert\%\Delta\rvert$
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***
	(1.530)	(1.740)	(5.311)	(1.763)	(0.906)	(1.604)	(2.262)	(2.631)	(1.505)	(1.567)	(2.283)	(1.554)
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***
	(1.545)	(1.712)	(5.919)	(1.715)	(0.824)	(1.561)	(2.104)	(3.582)	(1.533)	(1.575)	(2.143)	(1.603)
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
	(1.398)	(1.517)	(4.324)	(1.559)	(0.751)	(1.424)	(2.035)	(1.750)	(1.379)	(1.441)	(1.758)	(1.416)
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
	(1.294)	(1.406)	(3.710)	(1.463)	(0.643)	(1.304)	(1.978)	(1.610)	(1.285)	(1.311)	(1.689)	(1.301)
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 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 (2), while pre-treatment coefficients in Panel B refer to parameters α_3^k from equation (2) (see Section 3.2 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 column (2) is the log of Δ and in column (3) is a dummy that equals 100 if $\lvert\%\Delta\rvert$ 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 at $\pm$ 75 except in column (4) where it is winsorized at $\pm$ 100. All columns except (6) include additional controls (female dummy, high education dummy, hourly dummy, age, position tenure). Column (7) excludes position fixed effects. Column (8) includes firm fixed effects instead of position fixed effects. 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.

Table E.3: The Effects of Benchmarking on Salary Dispersion: Clustering at the Firm-Position Level

	(1) $\lvert\%\Delta\rvert$	(2) $\lvert log\Delta\rvert$	(3) $\lvert\%\Delta\rvert > 10$	(4) $\lvert\%\Delta\rvert$	(5) $\lvert\%\Delta\rvert$	(6) $\lvert\%\Delta\rvert$	(7) $\lvert\%\Delta\rvert$	(8) $\lvert\%\Delta\rvert$	(9) $\lvert\%\Delta\rvert$	(10) $\lvert\%\Delta\rvert$	(11) $\lvert\%\Delta\rvert$	(12) $\lvert\%\Delta\rvert$
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***
	(1.551)	(1.763)	(5.439)	(1.806)	(0.906)	(1.610)	(2.227)	(2.453)	(1.535)	(1.579)	(2.240)	(1.560)
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***
	(1.524)	(1.730)	(5.497)	(1.769)	(0.824)	(1.532)	(2.139)	(3.147)	(1.519)	(1.549)	(2.072)	(1.578)
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
	(1.252)	(1.416)	(3.941)	(1.397)	(0.751)	(1.272)	(1.927)	(1.506)	(1.237)	(1.290)	(1.784)	(1.275)
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
	(1.198)	(1.354)	(3.544)	(1.361)	(0.643)	(1.205)	(1.860)	(1.295)	(1.197)	(1.218)	(1.679)	(1.231)
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 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 (2), while pre-treatment coefficients in Panel B refer to parameters α_3^k from equation (2) (see Section 3.2 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 column (2) is the log of Δ and in column (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 at $\pm$ 75 except in column (4) where it is winsorized at $\pm$ 100. All columns except (6) include additional controls (female dummy, high education dummy, hourly dummy, age, position tenure). Column (7) excludes position fixed effects. Column (8) includes firm fixed effects instead of position fixed effects. 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.

Table E.4: The Effects of Benchmarking on Salary Dispersion: Additional Robustness Checks

	(1) |%Δ|	(2) |%Δ|	(3) |%Δ|	(4) |%Δ|	(5) |%Δ|	(6) |%Δ|	(7) |%Δ|
Panel A: Post-treatment							
Searched vs. Non-Searched	-4.775***	-6.876***	-5.167***	-4.231***	-3.588***	-4.786***	-4.996***
	(1.143)	(1.870)	(1.113)	(1.104)	(1.218)	(1.094)	(1.153)
Searched vs. Non-Searchable	-6.149***	-8.163***	-6.018***	-5.701***	-6.288***	-6.582***	-6.420***
	(1.070)	(1.498)	(1.041)	(1.064)	(1.152)	(1.008)	(1.096)
Panel B: Pre-treatment							
Searched vs. Non-Searched	-0.346	-0.549	-0.669	0.152	-0.301	-0.369	-0.578
	(1.167)	(1.594)	(1.172)	(1.163)	(1.134)	(1.168)	(1.194)
Searched vs. Non-Searchable	-0.310	-0.154	-0.703	0.293	0.207	-0.315	-0.454
	(1.055)	(1.432)	(1.041)	(1.109)	(1.032)	(1.055)	(1.083)
No Filters		✓					
Filtered Benchmark $\geq$ 100			✓				
Include Match Outliers				✓			
Restricted Sample					✓		
After Aug-2020						✓	
Exclude HR Positions							✓
Mean Dep. Var. (Baseline)	19.812	24.290	20.163	20.478	19.812	19.812	19.790
Observations							
Searched	5,253	5,253	5,253	6,150	5,246	5,414	5,080
Non-Searched	39,527	39,527	39,527	50,943	15,958	46,435	38,156
Non-Searchable	156,734	156,734	156,734	196,768	83,348	160,595	151,778

Notes: Column (1) follows the specification of column (1) from Table 3. Column (2) uses the same specification as column (1), except using absolute dispersion from the unfiltered median benchmark, as opposed to using the state and sector filtered benchmark when available, as the outcome. Column (3) uses only filtered benchmarks computed using 100 or more employees, as opposed to the baseline threshold of 30 employees, and unfiltered benchmarks otherwise. Column (4) drops new hires who's organization specific job title has a low match score to the designated position title (scores less than the 20th percentile of scores in that quarter). Column (5) using the *Restricted Sample* uses only control observations after September 2019, the start of our search data. Column (6) adds data from Aug-2020 to July-2021 to the sample. Column (7) excludes new hires in HR positions. See Table 3 for more details.

F Effects on Average Salary and Retention: Additional Results and Robustness Checks

F.1 Main Robustness Checks

Regarding the effects of salary benchmarking on the average salary, the difference-in-differences estimates are presented in Table F.1. The post-treatment coefficients (α_1^k, from equation (1)) are presented in Panel A. Column (1) of Table F.1 corresponds to the baseline specification. The post-treatment coefficients are positive: 0.003 log points (p-value=0.745) when using Non-Searched as a control group and 0.019 log points (p-value=0.918) when using Non-Searchable as control group.

Columns (2) through (11) of Table F.1 are identical to column (1), except that they change a different feature of the baseline specification. In column (2) we use salary as the dependent variable: i.e., in $s, without the log transformation. The results from column (2) are qualitatively consistent with the results from column (1): the post-treatment coefficients are modest (-$724.51 and $271.15 for the comparison to Non-Searched and Non-Searchable, respectively) and statistically insignificant (p-values of 0.829 and 0.871). The results are consistent in magnitude too. For example, the first post-treatment coefficient from column (1) suggests a -1.6% ($= \frac{-\$724.51}{\$44{,}146.85}$) increase in average salary relative to the baseline, while the corresponding coefficient from column (2) suggests an increase of 0.6% ($= \frac{\$271.15}{\$44{,}146.85}$).

The specification from column (3) of Table F.1 is identical to the baseline specification from column (1), except that the dependent variable is winsorized at 10% and 90% of the benchmark instead of the 2.5 and 97.5 percentile by position title. Column (4) is identical to column (1), except that the standard errors are not clustered. Column (5) is identical to column (1), except that it does not include any of the additional control variables. Column (6) excludes position fixed effects. Column (7) includes firm fixed effects instead of position fixed effects. Column (8) is identical to column (1), except that it excludes positions for which the base salary is not a major component of compensation: Waiter/Waitress, Chauffeur, and Bartender/Mixologist. Column (9) is identical to column (1), except that it restricts to the 329 positions that are searched at least once in the sample. Column (10) is identical to column (1), except that it does not re-weight by SOC groups. Last, column (11) is identical to column (1), except that it only includes new hires aged 21 through 60. In all these alternative specifications, the results are both qualitatively and quantitatively similar to those from column (1). Panel B of Table F.1 presents the corresponding "pre-treatment" coefficients. As expected, with few exceptions, these coefficients are close to zero, statistically insignificant and precisely estimated.

F.2 Additional Robustness Checks

In Table E.4, we show that the effects on salary dispersion are robust to a wide range of alternative specifications. In this Appendix, we show that the effects on salary levels are also robust to this same range of alternative specifications. The results are presented in Table F.2. Columns (1)–(5) measure the effect on the log of average base salary, with column (1) identical to column (1) of Table F.1. Columns (2)–(5) each change one feature of the baseline specification. In column (2), we include *Match Outliers* (positions with low match scores, indicating low match quality between the firm-specific job title and the position title) that are excluded from the main analysis. In column (3), we restrict the sample to only include Non-Searched and Non-Searchable positions that are after September 2019, the start of our search data. In column (4), we include new hires from August 2020 through July 2021. In column (5), we exclude HR positions, as HR professionals are the most common users of the benchmarking tool.

One potential concern is that firms may be reacting by changing beyond base salary, through bonuses, commissions or even hours worked. In addition to the base salary, our employee data includes the monthly gross wage: this is how much money the firm effectively pays to the employee each month, which reflects not only the base salary but also a myriad of other factors such as hours worked, tax withholdings, commission, bonuses and reimbursements. The last columns of Table F.2 measures the effects on average gross pay (instead of average base salary). In the first specification of column (6), we define the annual gross wage as the average monthly gross pay during the first three months working at the firm, then multiplied by 12 to transform it to an annual basis (i.e., so that it is comparable to the base salary outcome).[70] The base salary and the gross compensation are highly correlated, but not perfectly so (correlation coefficient of 0.848, p-value<0.001). One minor shortcoming with the gross pay data is that it is missing for 11.2% of the observations, for a variety of reasons.[71] In any case, as shown in column (9) of Table F.2, the results are also similar if we impute these missing values.

The results using the (log) gross pay outcome are presented in columns (6)–(9) of Table F.2. First of all, notice that the coefficients for gross pay are much less precisely estimated than the corresponding coefficients for base salary. For example, the standard errors of the post-treatment coefficients for gross pay (0.042 and 0.043, from column (6)) are 2 times as large as the corresponding coefficients for base salary (0.017 and 0.016, from column (1)).

[70] We compute the average starting on the 1st day of the month following the hire date, to make it more comparable across different employees. For employees who work fewer than three months at the firm, the average will be based on the one or two months they worked at the firm.

[71] For example, payroll data are not available for 7 firms, and for other firms it is missing for some employees for a variety of reasons such as failure of data entry from the manager.

This should be expected: relative to the base salary outcome, the gross pay outcome is more volatile because it includes a myriad of factors such as differences in tax withholdings, commissions and so on. The point coefficients for the gross pay outcome (0.002 and 0.037, from column (6)) are similar in magnitude to the corresponding coefficients for base salary (0.003 and 0.019, from column (1)), and statistically indistinguishable from each other.

F.3 Heterogeneity by Skill

Figure F.1 displays the event-study analysis of salary levels, reproducing Figure 5, but for the high-skill sub-sample. When compared to the Non-Searched control group, the Searched group have an average salary level that is 0.029 lower on average (p-value = 0.119) in the post period. Compared to the Non-Searchable control group, the post-period salary level is 0.016 log points lower (p-value = 0.288). These differences are economically modest and statistically insignificant, suggesting that the salary levels before and after gaining access to the salary benchmarking tool remain steady among the high-skill sub-sample. The difference-in-differences estimates under all the different specifications are presented in Table F.3 and Table F.4, respectively for low-skill and high-skill sub-samples. The results are largely robust across specifications.

F.4 Heterogeneity by Gender

The analysis of heterogeneity by the gender of the employee is presented in Figure F.2. Panel A and Panel B reproduce Panel C and Panel D of Figure 5 for the male subsample, and Panel C and Panel D use female subsample only. We find gender differences that are small in magnitude and statistically insignificant.

F.5 Heterogeneity by Market Dispersion

Figure F.3 displays the event-study analysis of salary levels split by above- and below-median market dispersion. Market dispersion is defined by the dispersion in salaries across all employees within the same position, state and industry as the new hire in our sample. Low dispersion markets are likely markets with more standardized or homogeneous labor, and therefore more compressed pay. Comparing Panels A and B (low-dispersion markets), with Panels C and D (high-dispersion markets) reveals that salary-levels rise modestly after accessing the salary benchmark tool primarily within low-dispersion markets. Among low-dispersion markets, the average salary level is 0.035 higher on average (p-value = 0.254) in the post period when comparing Searched vs. Non-Searched groups, and 0.032 higher (p-value = 0.056) when com-

paring Searched vs. Non-Searchable. Among high-dispersion markets, the respective change in average salary levels are -0.030 (p-value = 0.159) and -0.004 (p-value = 0.922), suggesting that the salary levels do not rise in high-dispersion markets.

F.6 Effects on Retention Levels

In Figure F.4, we present the event-study analysis for the retention outcome in the entire sample. Specifically, the dependent variable is whether an employee hired in a given month is still employed at the same firm one year later. The effects of benchmarking on retention are not statistically significant for the full sample. Panel C of Figure F.4 presents the difference between the Searched and Non-Searched groups. This analysis suggests that the retention rate rose from 52.0 pp to 56.1 pp after onboarding (p-value=0.163). Panel D presents the same difference for Searched and Non-Searchable positions. After onboarding, the retention rate rose from 52.0 pp to 53.4 pp (p-value=0.619). The corresponding difference-in differences estimates are presented in Table F.5, and the results are robust across specifications.

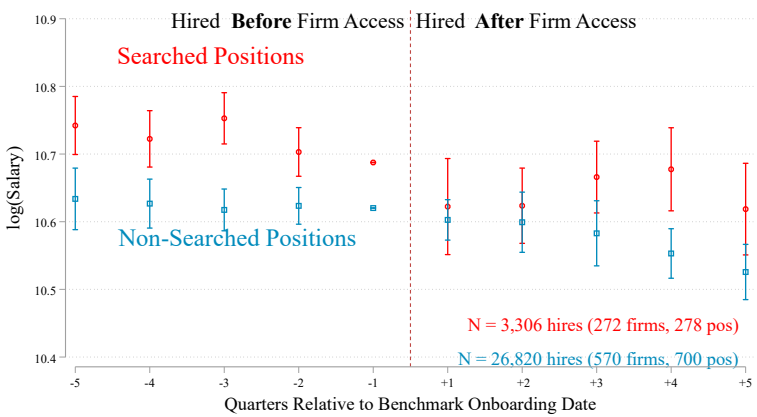
Figure F.5 presents the results for the high-skill sample (for reference, the results for the low-skill sample are presented in Figure 7 above). Consistent with the fact that we do not find significant effects on the average salaries in high-skill positions, we do not find any significant effects on the average retention rate either. More precisely, Panel C of Figure F.5 shows that, when using Non-Searched as control group, the average retention level increases by 3.8 pp (p-value = 0.304) after onboarding. Panel D shows that, when using Non-Searchable as the control group, the average retention drops by 0.2 pp (p-value = 0.976) after onboarding. These differences are economically small and statistically insignificant.[72]

The heterogeneity analysis by market dispersion is presented in Figure F.6. Commensurate with our findings on salary levels, retention rises in low-dispersion markets after onboarding, and remains largely unchanged in high-dispersion markets. Panels A and B present the low-dispersion markets. Compared to the Non-Searched control group, the Searched group experiences a 7.5 pp increase (p-value = 0.042) in the likelihood of being at the same employer a year after being hired, off a base of 41.8%, and a 6.9 pp increase (p-value = 0.019) when compared to the Non-Searchable control group. By contrast, Panels C and D present the high-dispersion markets. The estimated retention effect remains within 2 pp and statistically insignificant when restricting our sample to the high-dispersion markets.

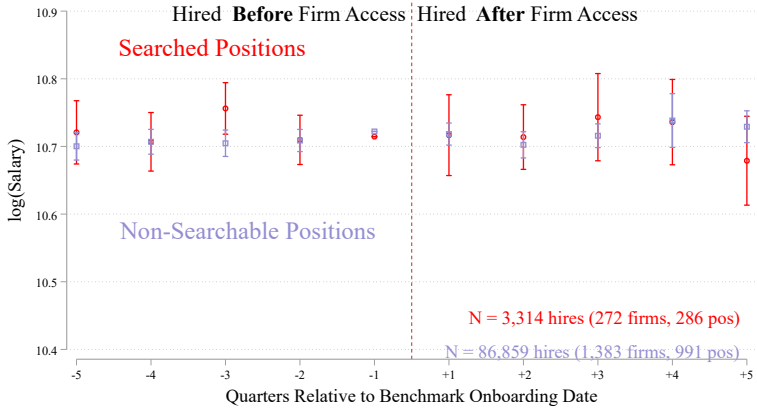
[72]The corresponding difference-in-differences estimates are presented in Table F.6 (for low-skill positions) and Table F.7 (for high-skill positions).

Figure F.1: The Effects of Salary Benchmarking on Salary Levels: High-Skill Subsample

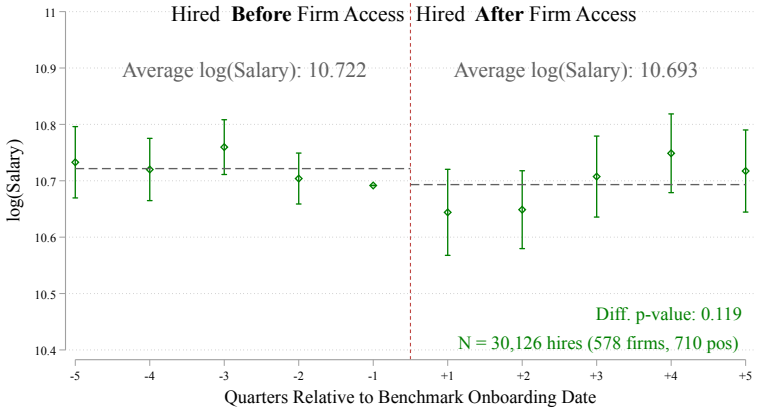
PANEL A: Searched vs. Non-Searched



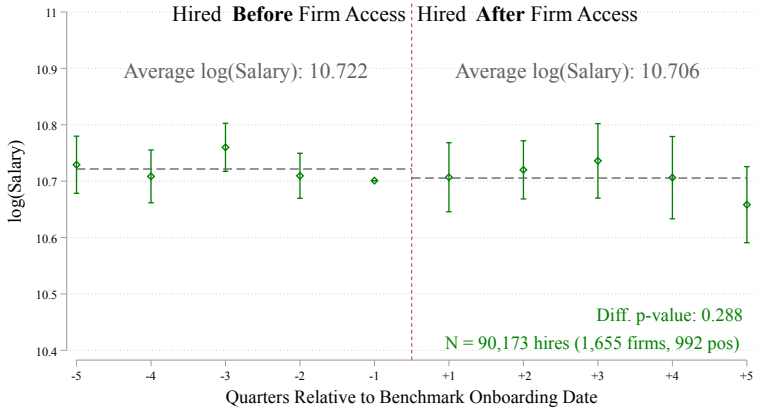
PANEL B: Searched vs. Non-Searchable



PANEL C: Difference Searched *minus* Non-Searched



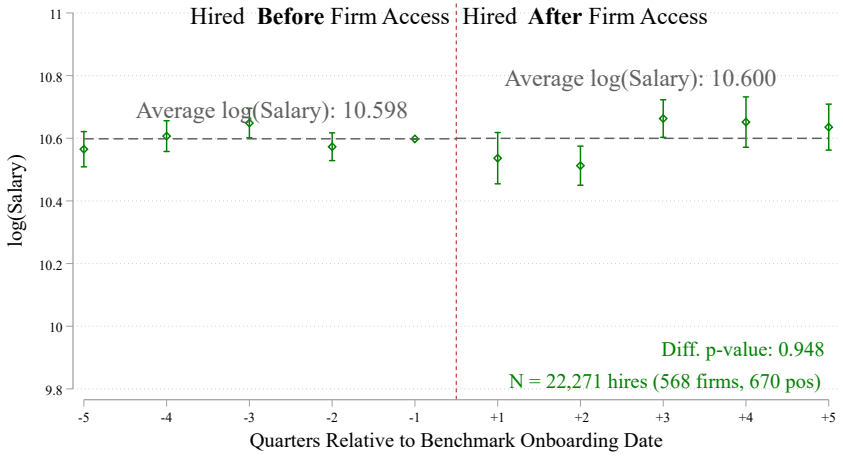
PANEL D: Difference Searched *minus* Non-Searchable



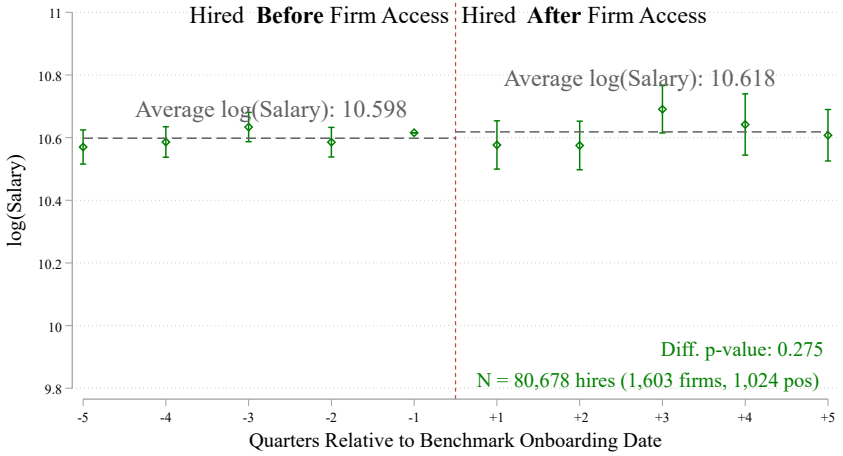
Notes: This is a reproduction of Figure 5, but with the sample being high-skilled positions. For more details, see notes to Figure 5.

Figure F.2: Heterogeneity by Gender: Event-Study Analysis

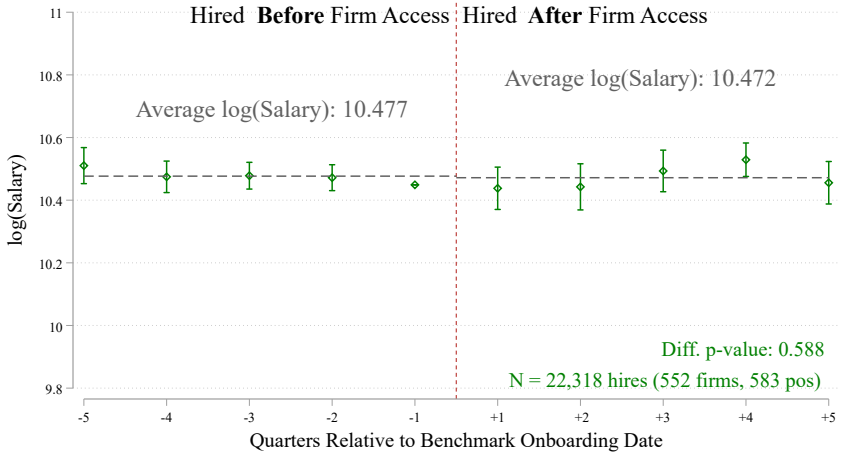
PANEL A: Male: Searched vs. Non-Searched



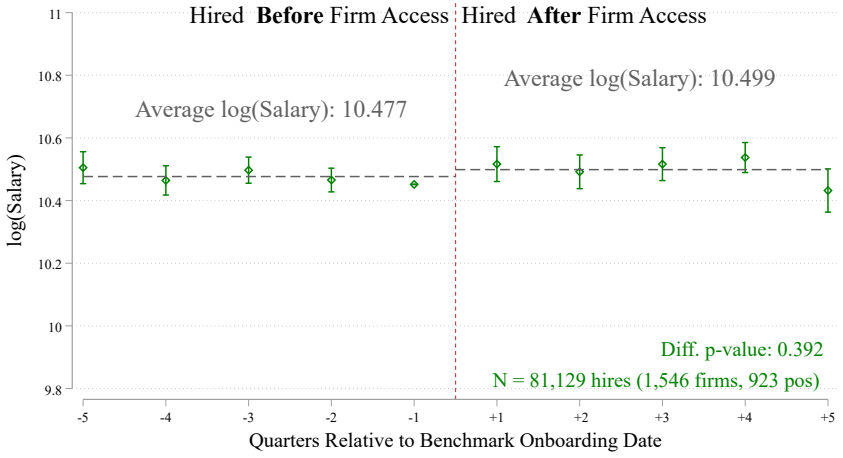
PANEL B: Male: Searched vs. Non-Searchable



PANEL C: Female: Searched vs. Non-Searched



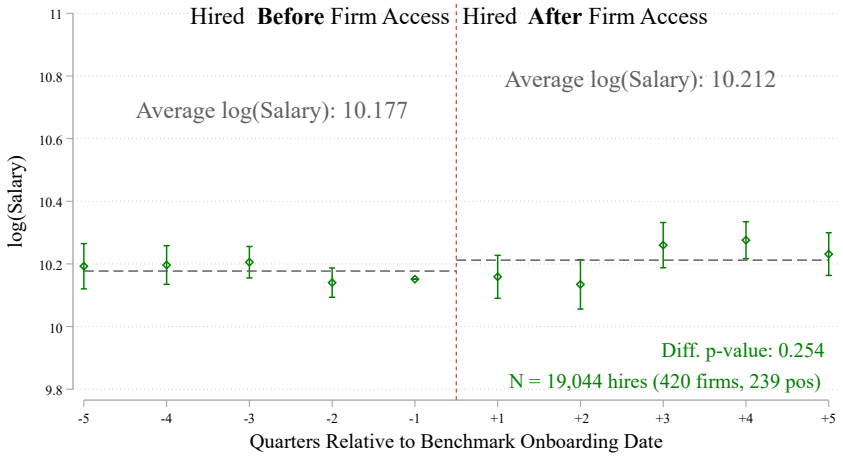
PANEL D: Female: Searched vs. Non-Searchable



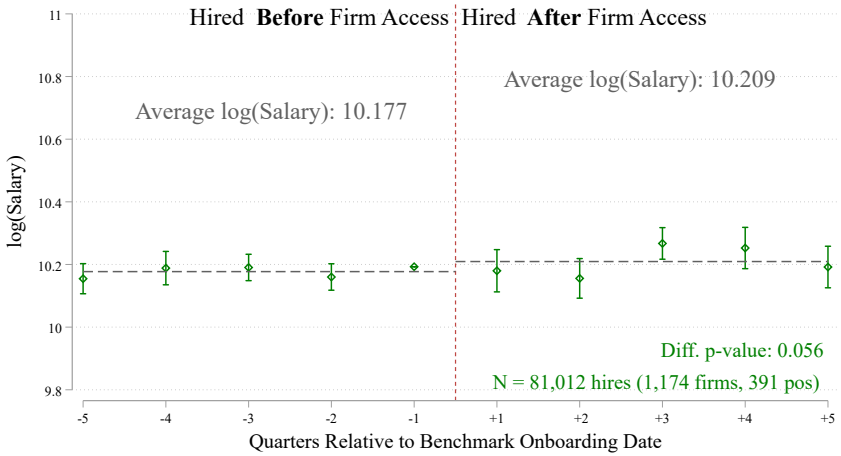
Notes: This is a reproduction of Figure 5 Panel C and D, split by the gender of employees. See the notes of Figure 5 for more details.

Figure F.3: Heterogeneity by Market Dispersion: The Effects on Salary Levels

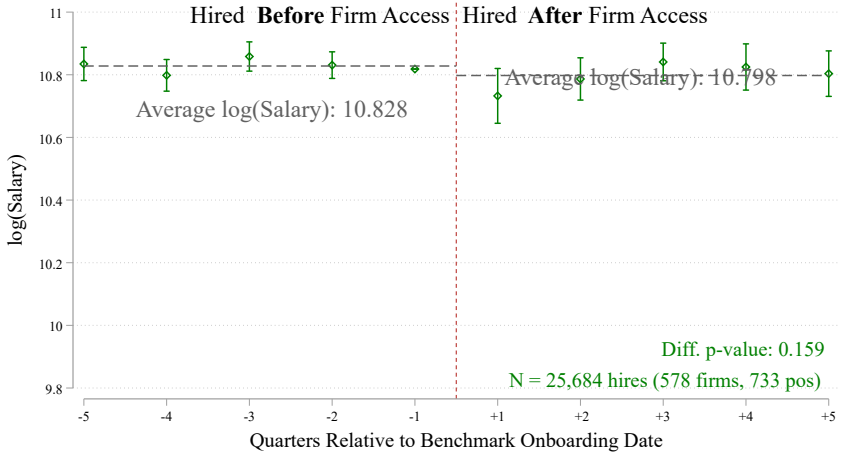
PANEL A: Low Market Dispersion: Searched vs. Non-Searched



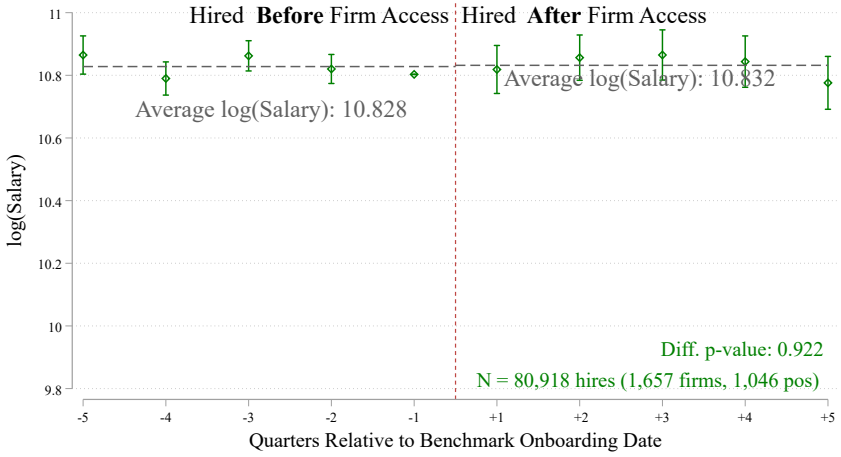
PANEL B: Low Market Dispersion: Searched vs. Non-Searchable



PANEL C: High Market Dispersion: Searched vs. Non-Searched

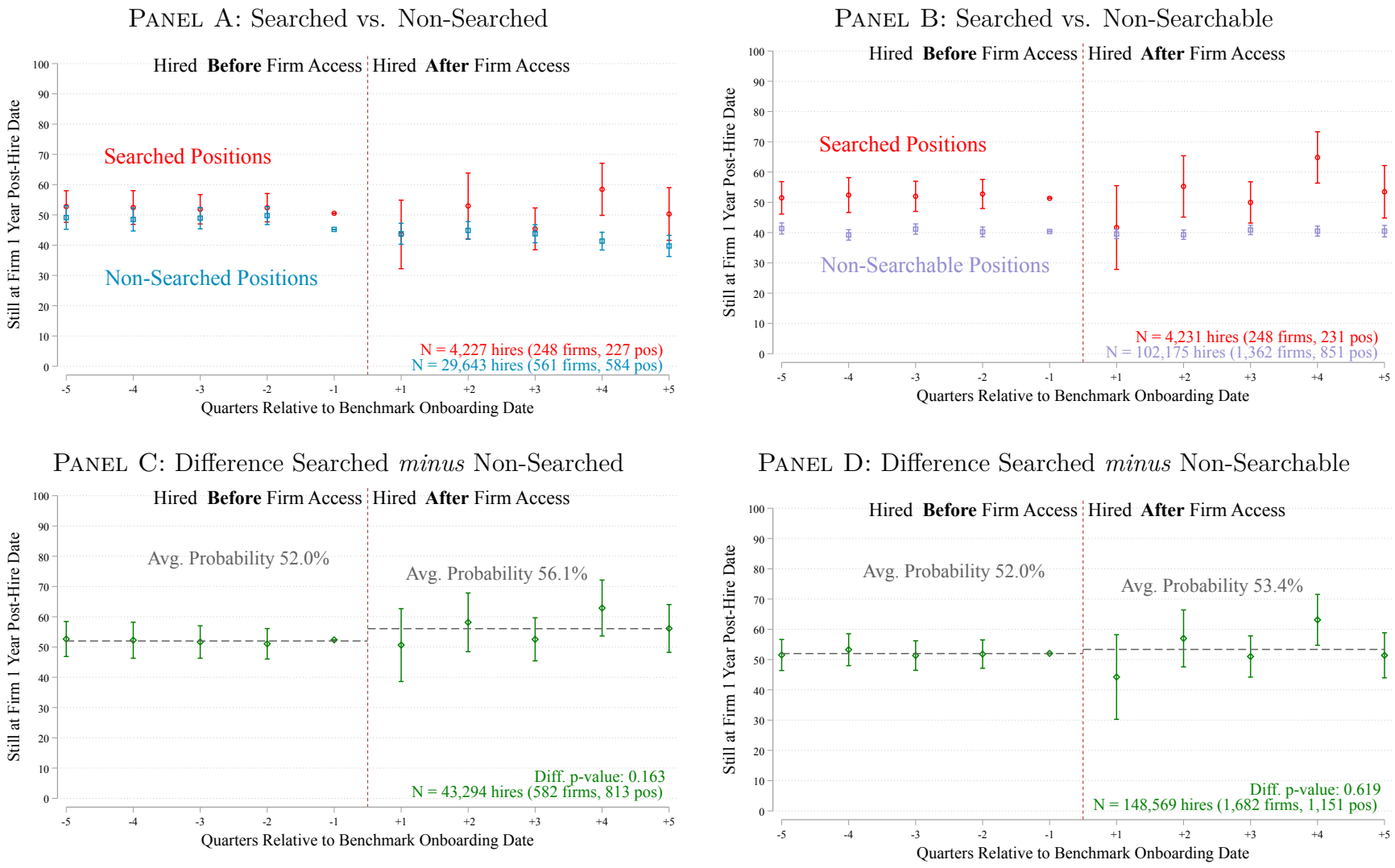


PANEL D: High Market Dispersion: Searched vs. Non-Searchable



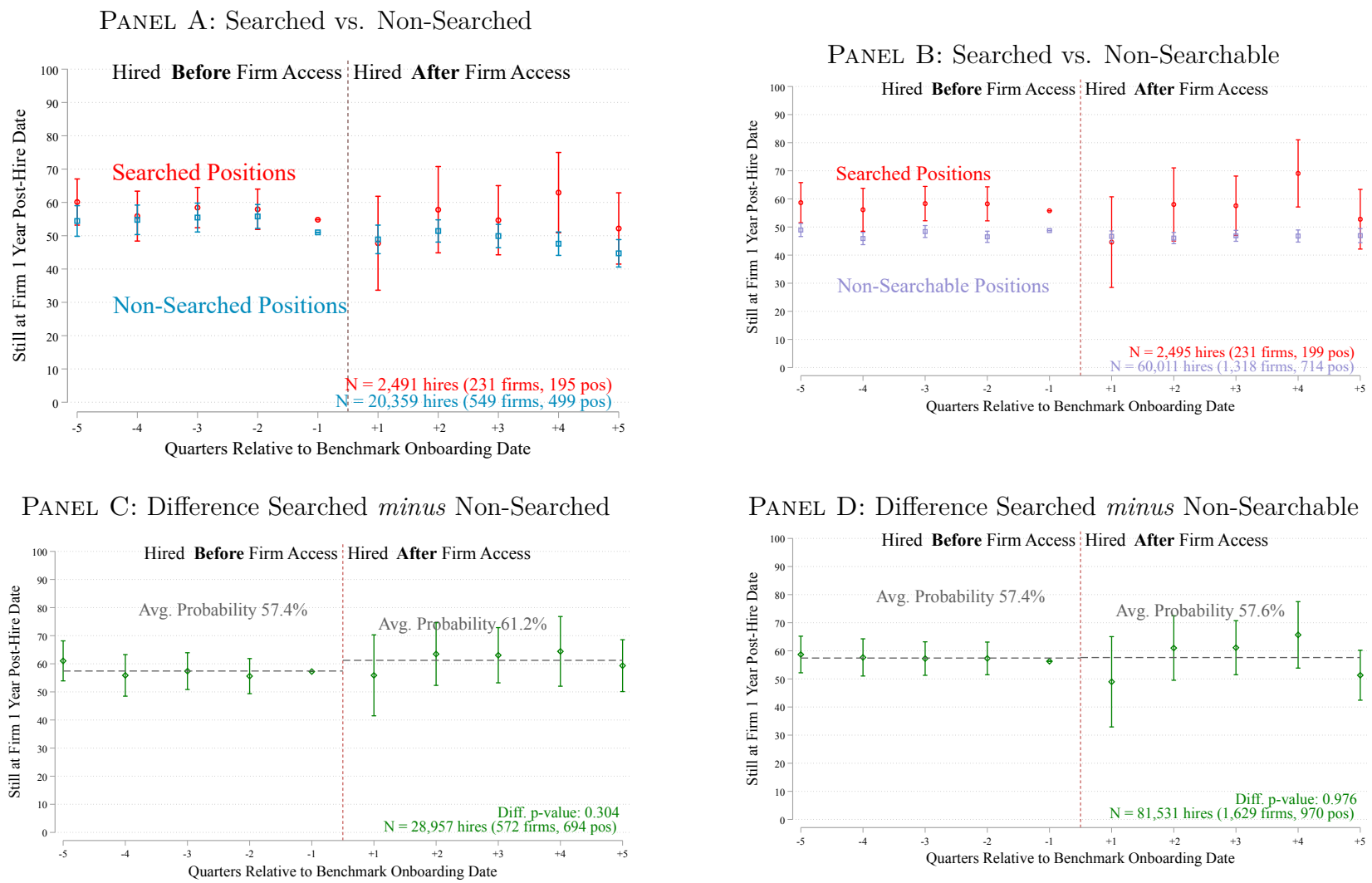
Notes: This is a reproduction of Figure 5 Panel C and D, split by the market dispersion. See the notes of Figure 5 for more details.

Figure F.4: Retention: Event-Study Analysis



Notes: This is a reproduction of Figure 5, 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 5.

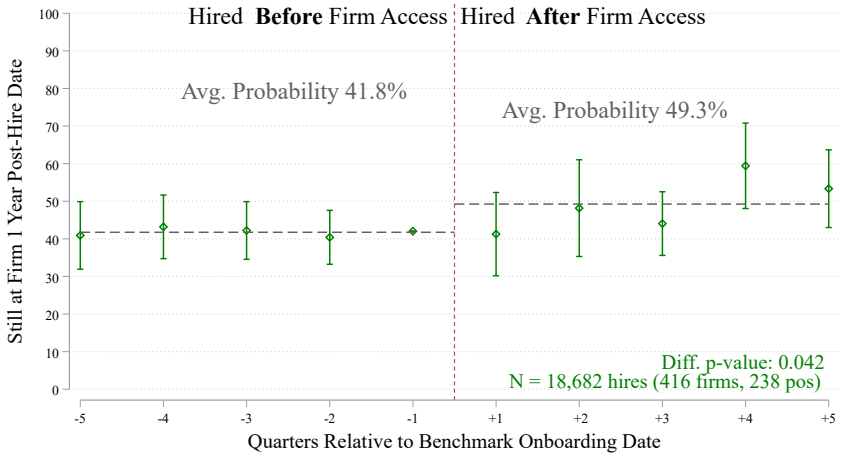
Figure F.5: The Effects of Salary Benchmarking on Retention Rates: High-Skill Subsample



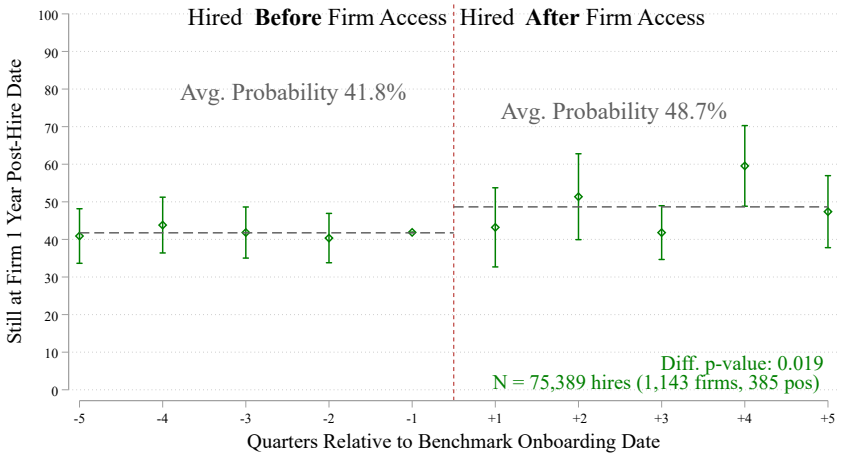
Notes: This is a reproduction of Figure F.4, but with the sample being high-skilled positions. For more details, see notes to Figure F.4.

Figure F.6: Heterogeneity by Market Dispersion: The Effects on Retention Rates

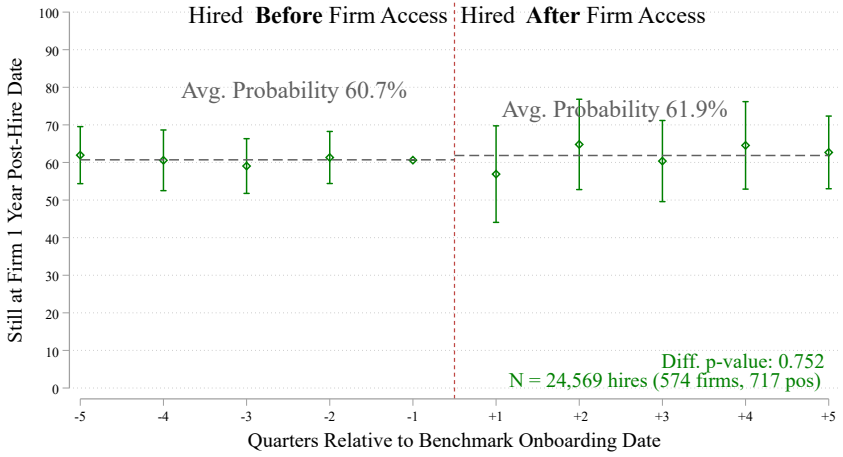
PANEL A: Low Market Dispersion: Searched vs. Non-Searched



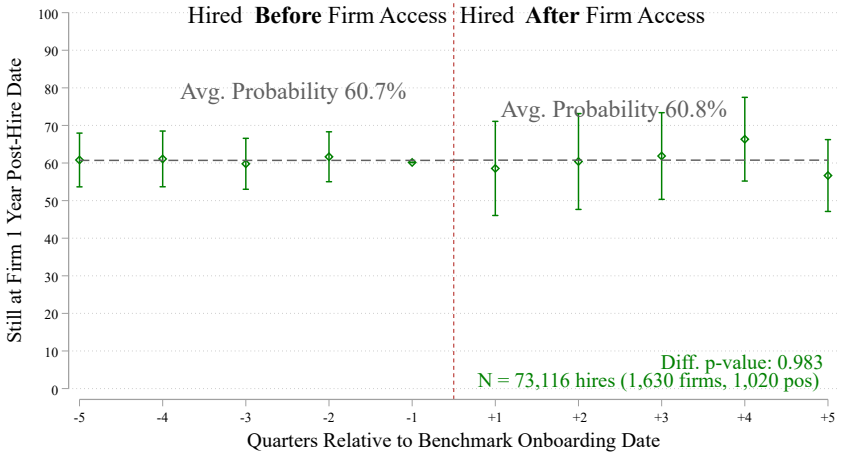
PANEL B: Low Market Dispersion: Searched vs. Non-Searchable



PANEL C: High Market Dispersion: Searched vs. Non-Searched



PANEL D: High Market Dispersion: Searched vs. Non-Searchable



Notes: This is a reproduction of Figure F.4 Panel C and D, split by the market dispersion. See the notes of Figure F.4 for more details.

Table F.1: The Effects of Benchmarking on Salary Levels: Full Sample

	(1) log(*Salary*)	(2) Salary	(3) log(*Salary*)	(4) log(*Salary*)	(5) log(*Salary*)	(6) log(*Salary*)	(7) log(*Salary*)	(8) log(*Salary*)	(9) log(*Salary*)	(10) log(*Salary*)	(11) log(*Salary*)
Panel A: Post-treatment											
Searched vs. Non-Searched	0.003	-724.514	0.004	0.003	0.001	-0.001	-0.000	-0.000	0.005	0.017	-0.001
	(0.017)	(861.430)	(0.017)	(0.012)	(0.019)	(0.032)	(0.019)	(0.018)	(0.018)	(0.017)	(0.018)
Searched vs. Non-Searchable	0.019	271.150	0.021	0.019*	0.009	0.015	0.021	0.015	0.017	0.030*	0.012
	(0.016)	(926.594)	(0.016)	(0.011)	(0.018)	(0.025)	(0.023)	(0.016)	(0.016)	(0.018)	(0.017)
Panel B: Pre-treatment											
Searched vs. Non-Searched	-0.022	-1166.949	-0.021	-0.022**	-0.028	-0.030	0.006	-0.019	-0.019	-0.013	-0.022
	(0.018)	(870.587)	(0.018)	(0.011)	(0.019)	(0.038)	(0.016)	(0.018)	(0.018)	(0.019)	(0.017)
Searched vs. Non-Searchable	-0.005	-686.358	-0.005	-0.005	-0.017	-0.014	0.000	-0.011	-0.004	-0.004	-0.002
	(0.017)	(806.078)	(0.016)	(0.010)	(0.019)	(0.027)	(0.013)	(0.016)	(0.017)	(0.018)	(0.017)
Alternate Winsorization			✓								
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)	10.532	44146.850	10.523	10.532	10.532	10.532	10.532	10.547	10.532	10.506	10.565
Observations											
Searched	5,253	5,253	5,253	5,253	5,253	5,266	5,262	5,105	5,253	5,316	4,611
Non-Searched	39,527	39,527	39,527	39,527	39,527	39,686	39,673	37,841	34,954	39,645	34,338
Non-Searchable	156,734	156,734	156,734	156,734	156,734	156,865	156,817	148,521	127,145	156,883	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 (2), while pre-treatment coefficients in Panel B refer to parameters α_3^k from equation (2) (see Section 3.2 for details). All columns include year fixed effects. In columns (1) and (3)–(11) the dependent variable is the log of annual base salary. The dependent variable in column (2) is the annual base salary (in $s). Log salary and salary are winsorized at the 2.5 and 97.5 percentiles of all salaries for their position. The exception is column (3) where wages are winsorized at ± 90% of the median benchmark. All columns except (5) include additional controls (female dummy, high education dummy, hourly dummy, age, position tenure). Column (6) excludes position fixed effects. Column (7) includes firm fixed effects instead of position fixed effects. Column (8) excludes the three positions where gross pay most exceeds base pay: Waiter/Waitress, Chauffeur, and Bartender/Mixologist. Column (9) restricts the sample to only titles of Non-Searched or Non-Searchable new hires in positions that are searched and hired by firms in the data.

Table F.2: The Effects of Benchmarking on Salary Levels: Additional Robustness Checks

	(1) log(*Salary*)	(2) log(*Salary*)	(3) log(*Salary*)	(4) log(*Salary*)	(5) log(*Salary*)	(6) log(*Gross*)	(7) log(*Gross*)	(8) log(*Gross*)	(9) log(*Gross*)
Panel A: Post-treatment									
Searched vs. Non-Searched	0.003	0.003	-0.007	0.004	0.002	0.002	-0.001	-0.005	-0.004
	(0.017)	(0.016)	(0.019)	(0.017)	(0.018)	(0.042)	(0.043)	(0.041)	(0.040)
Searched vs. Non-Searchable	0.019	0.016	0.000	0.018	0.018	0.037	0.031	0.026	0.033
	(0.016)	(0.016)	(0.017)	(0.015)	(0.017)	(0.043)	(0.043)	(0.042)	(0.040)
Panel B: Pre-treatment									
Searched vs. Non-Searched	-0.022	-0.037**	-0.016	-0.022	-0.018	0.017	0.026	0.024	0.020
	(0.018)	(0.017)	(0.017)	(0.018)	(0.018)	(0.035)	(0.035)	(0.035)	(0.033)
Searched vs. Non-Searchable	-0.005	-0.009	-0.011	-0.005	-0.001	0.011	0.008	0.019	0.020
	(0.017)	(0.017)	(0.016)	(0.017)	(0.017)	(0.031)	(0.031)	(0.031)	(0.029)
Include Match Outliers		✓							
Restricted Sample			✓						
After Aug-2020				✓					
Exclude HR Positions					✓				
3 Month Window						✓			
2 Month Window							✓		
6 Month Window								✓	
Imputed									✓
Mean Dep. Var. (Baseline)	10.532	10.527	10.532	10.532	10.507	10.382	10.381	10.394	10.378
Observations									
Searched	5,253	6,150	5,246	5,414	5,080	4,869	4,864	4,875	5,253
Non-Searched	39,527	50,943	15,958	46,435	38,156	35,884	35,844	35,915	39,527
Non-Searchable	156,734	196,768	83,348	160,595	151,778	138,178	138,002	138,409	156,734

Notes: Columns (1)–(5) look at effects on the log of annual base salary. Column (1) is exactly column (1) from Table F.1. Column (2), *Include Match Outliers*, reproduces column (1), but including new hires whose organization specific job title has a low match score to the designated position title (scores less than the 20th percentile of the scores in that quarter). Column (3) using the *Restricted Sample* uses only control observations after September 2019, the start of our search data. Column (4) includes data from August 2020 through July 2021. Column (5) excludes new hires in HR positions. Columns (6)–(9) look at effects of the log of annual gross wages (as described in Section F.2). Column (6) uses a 3 month window to compute gross wages, while columns (7) and (8) use a 2 and 6 month window, respectively. Column (9) is equivalent to column (6) with missing values imputed. See Table F.1 for more details.

Table F.3: The Effects of Benchmarking on Salary Levels: Low-Skill Subsample

	(1) log(Salary)	(2) Salary	(3) log(Salary)	(4) log(Salary)	(5) log(Salary)	(6) log(Salary)	(7) log(Salary)	(8) log(Salary)	(9) log(Salary)	(10) log(Salary)	(11) log(Salary)
Panel A: Post-treatment											
Searched vs. Non-Searched	0.073***	1618.196***	0.058***	0.073***	0.070***	0.088***	0.034**	0.071***	0.078***	0.080***	0.066***
	(0.020)	(541.012)	(0.020)	(0.013)	(0.020)	(0.021)	(0.016)	(0.019)	(0.020)	(0.025)	(0.020)
Searched vs. Non-Searchable	0.087***	1862.423***	0.076***	0.087***	0.079***	0.090***	0.038**	0.087***	0.095***	0.089***	0.081***
	(0.021)	(509.221)	(0.021)	(0.013)	(0.021)	(0.021)	(0.015)	(0.019)	(0.022)	(0.025)	(0.022)
Panel B: Pre-treatment											
Searched vs. Non-Searched	0.002	17.476	-0.008	0.002	0.002	0.015	0.016	0.016	0.005	0.005	-0.001
	(0.023)	(593.343)	(0.025)	(0.012)	(0.025)	(0.025)	(0.012)	(0.024)	(0.024)	(0.024)	(0.022)
Searched vs. Non-Searchable	0.035	335.014	0.030	0.035***	0.030	0.037	0.020**	0.018	0.038	0.037	0.046
	(0.030)	(700.088)	(0.029)	(0.014)	(0.031)	(0.028)	(0.008)	(0.028)	(0.031)	(0.031)	(0.030)
Alternate Winsorization			✓								
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)	10.147	25929.204	10.147	10.147	10.147	10.147	10.147	10.150	10.147	10.145	10.168
Observations											
Searched	1,947	1,947	1,947	1,947	1,947	1,947	1,938	1,799	1,947	1,962	1,545
Non-Searched	12,707	12,707	12,707	12,707	12,707	12,724	12,671	11,021	11,015	12,715	10,211
Non-Searchable	69,875	69,875	69,875	69,875	69,875	69,890	69,755	61,662	55,625	69,903	56,380

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 (2), while pre-treatment coefficients in Panel B refer to parameters α_3^k from equation (2) (see Section 3.2 for details). All columns include year fixed effects. In columns (1) and (3)–(11) the dependent variable is the log of annual base salary. The dependent variable in column (2) is the annual base salary (in $s). Log salary and salary are winsorized at the 2.5 and 97.5 percentiles of all salaries for their position. The exception is column (3) where wages are winsorized at ± 90% of the median benchmark. All columns except (5) include additional controls (female dummy, high education dummy, hourly dummy, age, position tenure). Column (6) excludes position fixed effects. Column (7) includes firm fixed effects instead of position fixed effects. Column (8) excludes the three positions where gross pay most exceeds base pay: Waiter/Waitress, Chauffeur, and Bartender/Mixologist. Column (9) restricts the sample to only titles of Non-Searched or Non-Searchable new hires in positions that are searched and hired by firms in the data.

Table F.4: The Effects of Benchmarking on Salary Levels: High-Skill Subsample

	(1) log(Salary)	(2) Salary	(3) log(Salary)	(4) log(Salary)	(5) log(Salary)	(6) log(Salary)	(7) log(Salary)	(8) log(Salary)	(9) log(Salary)	(10) log(Salary)	(11) log(Salary)
Panel A: Post-treatment											
Searched vs. Non-Searched	-0.030	-1739.161	-0.021	-0.030*	-0.033	-0.045	-0.006	-0.030	-0.029	-0.025	-0.027
	(0.023)	(1236.904)	(0.022)	(0.017)	(0.026)	(0.048)	(0.026)	(0.023)	(0.024)	(0.021)	(0.023)
Searched vs. Non-Searchable	-0.022	-706.932	-0.011	-0.022	-0.032	-0.022	-0.019	-0.022	-0.028	-0.014	-0.026
	(0.020)	(1336.488)	(0.019)	(0.014)	(0.023)	(0.037)	(0.033)	(0.020)	(0.020)	(0.023)	(0.020)
Panel B: Pre-treatment											
Searched vs. Non-Searched	-0.039	-1935.302	-0.030	-0.039**	-0.048*	-0.072	-0.017	-0.039	-0.036	-0.025	-0.036
	(0.025)	(1312.840)	(0.024)	(0.016)	(0.027)	(0.058)	(0.021)	(0.025)	(0.025)	(0.026)	(0.024)
Searched vs. Non-Searchable	-0.026	-1380.412	-0.024	-0.026**	-0.042*	-0.042	-0.017	-0.026	-0.025	-0.019	-0.023
	(0.020)	(1174.559)	(0.020)	(0.013)	(0.024)	(0.044)	(0.019)	(0.020)	(0.020)	(0.023)	(0.020)
Alternate Winsorization			✓								
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)	10.722	53147.502	10.709	10.722	10.722	10.722	10.722	10.722	10.722	10.711	10.730
Observations											
Searched	3,306	3,306	3,306	3,306	3,306	3,319	3,316	3,306	3,306	3,354	3,066
Non-Searched	26,820	26,820	26,820	26,820	26,820	26,962	26,947	26,820	23,939	26,930	24,127
Non-Searchable	86,859	86,859	86,859	86,859	86,859	86,975	86,904	86,859	71,520	86,980	78,671

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 (2), while pre-treatment coefficients in Panel B refer to parameters α_3^k from equation (2) (see Section 3.2 for details). All columns include year fixed effects. In columns (1) and (3)–(11) the dependent variable is the log of annual base salary. The dependent variable in column (2) is the annual base salary (in $s). Log salary and salary are winsorized at the 2.5 and 97.5 percentiles of all salaries for their position. The exception is column (3) where wages are winsorized at ± 90% of the median benchmark. All columns except (5) include additional controls (female dummy, high education dummy, hourly dummy, age, position tenure). Column (6) excludes position fixed effects. Column (7) includes firm fixed effects instead of position fixed effects. Column (8) excludes the three positions where gross pay most exceeds base pay: Waiter/Waitress, Chauffeur, and Bartender/Mixologist. Column (9) restricts the sample to only titles of Non-Searched or Non-Searchable new hires in positions that are searched and hired by firms in the data.

Table F.5: The Effects of Benchmarking on Retention: Full Sample

	Dep. Var.: Still at Firm 12 Months Later (=100)								
	(1) % retention	(2) % retention	(3) % retention	(4) % retention	(5) % retention	(6) % retention	(7) % retention	(8) % retention	(9) % retention
Panel A: Post-treatment									
Searched vs. Non-Searched	5.509*	5.509*	5.342*	5.433*	-0.367	4.923*	5.347*	2.655	6.008*
	(2.868)	(2.956)	(2.829)	(2.919)	(2.629)	(2.916)	(2.931)	(2.347)	(3.088)
Searched vs. Non-Searchable	2.466	2.466	1.673	2.573	1.138	1.957	2.321	2.437	3.237
	(2.691)	(2.806)	(2.604)	(2.717)	(2.648)	(2.725)	(2.701)	(2.146)	(2.746)
Panel B: Pre-treatment									
Searched vs. Non-Searched	0.460	0.460	0.391	0.174	1.659	0.749	0.595	1.982	0.336
	(2.509)	(2.282)	(2.537)	(2.511)	(2.367)	(2.550)	(2.559)	(2.567)	(2.614)
Searched vs. Non-Searchable	0.055	0.055	-0.469	0.016	1.784	0.078	0.030	0.494	-0.553
	(2.293)	(2.012)	(2.315)	(2.298)	(2.222)	(2.330)	(2.314)	(2.375)	(2.371)
No Clustering		✓							
No Additional Controls			✓						
Position FE				✓					
Firm FE					✓				
Exclude High-Tip Jobs						✓			
Searched Positions Only							✓		
No Re-weighting								✓	
Ages 21-60									✓
Restricted Sample	51.999	51.999	51.999	51.999	51.999	52.880	51.999	49.835	53.117
Mean Dep. Var. (Baseline)									
Observations	5,111	5,111	5,111	5,111	5,121	4,974	5,111	5,189	4,478
Searched	38,183	38,183	38,183	38,189	38,339	36,569	33,785	38,472	33,091
Non-Searched	143,450	143,450	143,450	143,453	143,528	135,826	116,071	143,708	123,157

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.

Table F.6: The Effects of Benchmarking on Retention: Low-Skill Subsample

	Dep. Var.: Still at Firm 12 Months Later (=100)								
	(1) % retention	(2) % retention	(3) % retention	(4) % retention	(5) % retention	(6) % retention	(7) % retention	(8) % retention	(9) % retention
Panel A: Post-treatment									
Searched vs. Non-Searched	6.126	6.126	7.903*	5.123	-0.417	4.627	4.661	4.415	6.253
	(4.528)	(5.179)	(4.111)	(4.700)	(4.048)	(4.753)	(4.718)	(3.621)	(5.380)
Searched vs. Non-Searchable	4.732	4.732	4.988	4.323	5.271	3.622	4.730	3.668	4.443
	(3.822)	(4.989)	(3.743)	(3.914)	(3.509)	(3.975)	(3.852)	(2.966)	(4.671)
Panel B: Pre-treatment									
Searched vs. Non-Searched	1.982	1.982	3.250	0.493	1.221	3.042	1.643	3.295	1.777
	(4.345)	(3.833)	(4.478)	(4.447)	(4.166)	(4.588)	(4.462)	(4.410)	(4.617)
Searched vs. Non-Searchable	1.883	1.883	2.628	1.394	4.424	2.158	1.525	3.289	1.585
	(4.015)	(3.369)	(4.066)	(4.079)	(3.692)	(4.222)	(4.075)	(4.001)	(4.144)
No Clustering		✓							
No Additional Controls			✓						
Position FE				✓					
Firm FE					✓				
Exclude High-Tip Jobs						✓			
Searched Positions Only							✓		
No Re-weighting								✓	
Ages 21-60									✓
Restricted Sample	41.085	41.085	41.085	41.085	41.085	42.687	41.085	36.986	41.397
Mean Dep. Var. (Baseline)									
Observations	1,882	1,882	1,882	1,883	1,874	1,745	1,882	1,898	1,487
Searched	12,449	12,449	12,449	12,450	12,415	10,834	10,778	12,473	9,992
Non-Searched	65,152	65,152	65,152	65,156	65,033	57,526	51,747	65,202	52,397

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.

Table F.7: The Effects of Benchmarking on Retention: High-Skill Subsample

	Dep. Var.: Still at Firm 12 Months Later (=100)								
	(1) % retention	(2) % retention	(3) % retention	(4) % retention	(5) % retention	(6) % retention	(7) % retention	(8) % retention	(9) % retention
Panel A: Post-treatment									
Searched vs. Non-Searched	5.291	5.291	4.478	5.333	2.050	5.291	5.932*	2.957	5.687
	(3.519)	(3.368)	(3.627)	(3.590)	(3.336)	(3.519)	(3.573)	(2.956)	(3.595)
Searched vs. Non-Searchable	1.379	1.379	0.245	1.486	-0.864	1.379	1.143	1.873	2.522
	(3.535)	(3.323)	(3.420)	(3.583)	(3.672)	(3.535)	(3.531)	(2.906)	(3.336)
Panel B: Pre-treatment									
Searched vs. Non-Searched	-0.242	-0.242	-0.618	-0.299	0.456	-0.242	0.191	-0.504	-0.513
	(3.057)	(2.848)	(3.071)	(3.025)	(2.899)	(3.057)	(3.112)	(3.141)	(3.160)
Searched vs. Non-Searchable	-1.448	-1.448	-2.373	-1.325	-0.723	-1.448	-1.392	-1.936	-2.005
	(2.797)	(2.515)	(2.833)	(2.783)	(2.811)	(2.797)	(2.818)	(3.049)	(2.887)
No Clustering		✓							
No Additional Controls			✓						
Position FE				✓					
Firm FE					✓				
Exclude High-Tip Jobs						✓			
Searched Positions Only							✓		
No Re-weighting								✓	
Ages 21-60									✓
Restricted Sample	57.433	57.433	57.433	57.433	57.433	57.433	57.433	57.095	58.039
Mean Dep. Var. (Baseline)									
Observations	3,228	3,228	3,228	3,228	3,239	3,228	3,228	3,291	2,991
Searched	25,729	25,729	25,729	25,739	25,868	25,729	23,002	25,999	23,099
Non-Searched	78,295	78,295	78,295	78,297	78,331	78,295	64,322	78,506	70,760

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.

G Additional Results: Sample of Existing Employees

Due to its simplicity, our main theoretical and empirical analysis focuses on new hires. For the sake of completeness, we provide some additional results for a sample of existing employees. For this sample, however, we must keep in mind some limitations with the data. The main limitation is that we cannot be certain which employee the information pertains to precisely. As an example, suppose the firm looks up the benchmark for bank tellers, of which there are 100 existing employees. The data challenge is that we only observe whether the firm looked up the benchmark for "bank teller", but not which of their 100 bank tellers the data was relevant for. It may be that the benchmark was looked up to adjust the salary ranges for all 100 tellers. Or maybe the firm only needs the information to respond to the outside offer of one particular teller. As a result, if we assume that the information should affect all 100 existing employees, the estimates could suffer from massive attenuation bias. For these reasons, the results shown below must be taken with a grain of salt.[73]

The outcome of interest when thinking of existing employees is not the salary level, but the changes in salaries. Firms review the salaries of their existing employees infrequently and, even when doing so, they are subject to strong downward wage rigidities (Kaur, 2019; Grigsby et al., 2021). These rigidities must be taken into account when interpreting the results: if hiring managers find it harder to lower the wages of existing employers than to raise them, that could mechanically lead to an increase in average salaries.

We constructed an annual panel of existing employees and calculated the percent change in their salary from January to December of each year. Figure G.1 presents the distribution of annual salary changes for all position types in the pre-onboarding period. Typically, employees do not experience a salary revision. When they do experience a revision, it tends to be positive and small, with a notable right tail of significant revisions that might accompany promotion. And, consistent with downward wage rigidities, negative salary revisions are extremely rare, occurring only in 1% of the employee-year observations.

Figure G.2 extends our event-study framework to study the effect of salary benchmarks on how firms change salaries for existing employees. Panels A and B show the estimates for Searched, Non-Searched, and Non-Searchable existing employees, while panels C and D show the difference-in-differences estimates. The post-treatment coefficients are positive and statistically significant: 1.407 (p-value<0.001) using Non-Searched positions as control group and 0.911 (p-value= 0.005) using Non-Searchable positions as control. The modest gains in

[73] This source of attenuation bias still exists for the analysis of new hires, but is much less severe. By construction, the firm has to actively set a salary for every one of their new employees. As a result, if a firm looks up the information for a position in which they are hiring, it is likely that they will use that information in setting the salary of that new hire. To the extent that typically one or a few employees are hired at the same time, there is much less scope for misattribution.

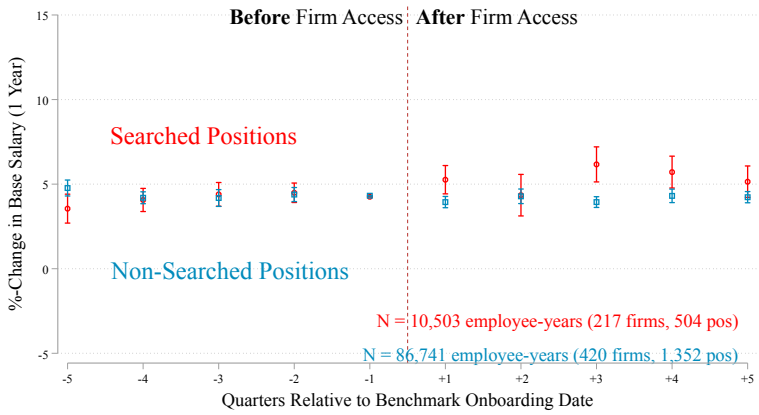
salaries among existing employees are consistent with the gains in average salaries observed in the analysis of new hires. However, as discussed above, this result for existing employees must be interpreted in light of nominal wage rigidities: i.e. after looking up the benchmark some employers may desire to cut salaries of some employees but are unable to do so because of nominal wage rigidities.

Figure G.1: Analysis of Existing Employees: Annual Percent Change in Salary

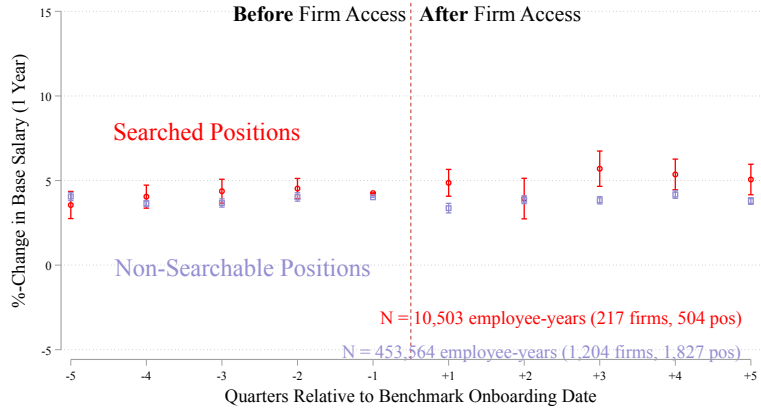
Notes: Distribution of the annual percent change in salary for existing employees of all position types before onboarding. Winsorized at -5 and +20.

Figure G.2: Analysis of Existing Employees: Event-Study Analysis of the Effects on Annual Percent Change in Salary

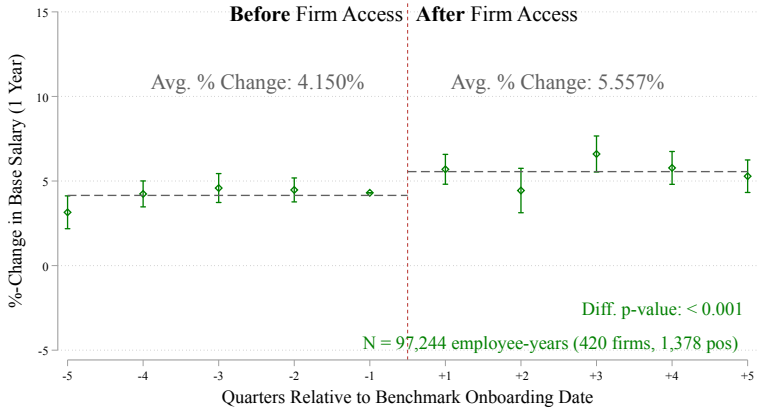
PANEL A: Searched vs. Non-Searched



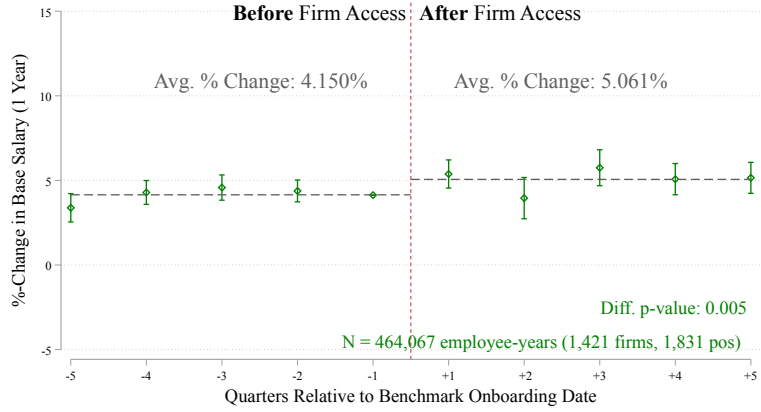
PANEL B: Searched vs. Non-Searchable



PANEL C: Difference Searched *minus* Non-Searched



PANEL D: Difference Searched *minus* Non-Searchable

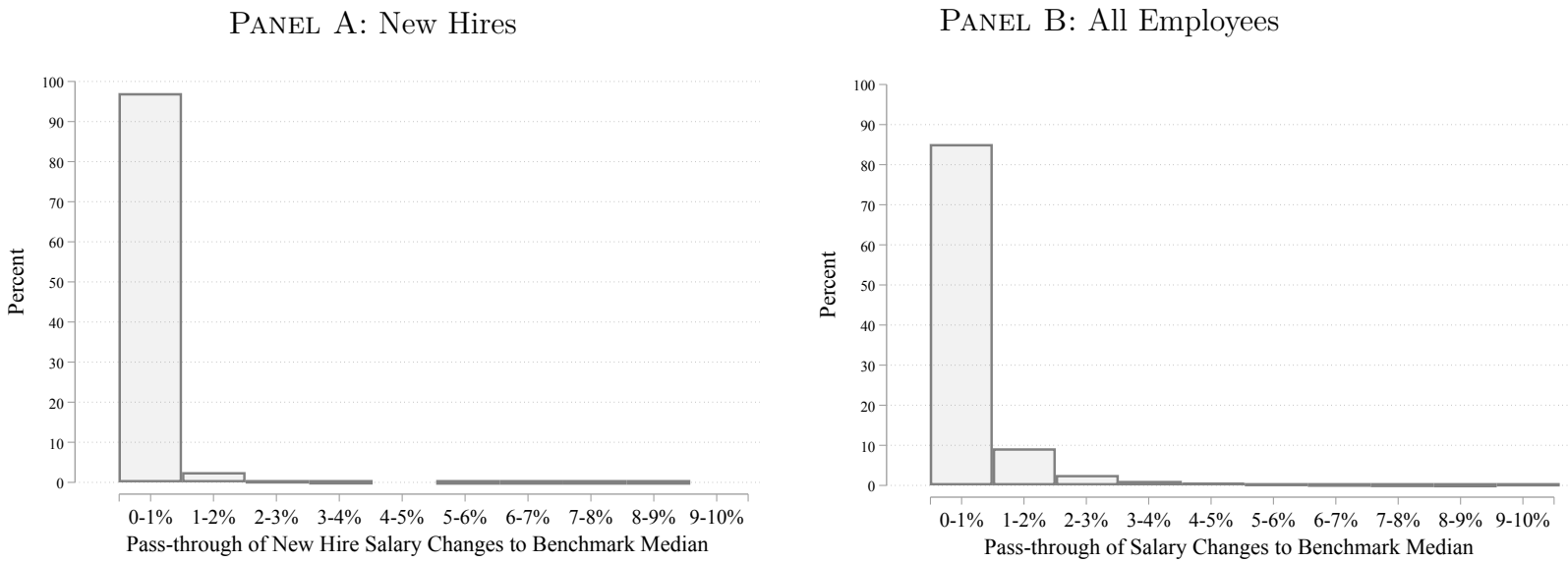


Notes: Each panel is a reproduction of the corresponding panel of Figure 3, except looking at the annual percent change in salary for existing employees and excluding new hires. For more details, see notes to Figure 3.

H Pass-through from Salaries to Benchmark

Recall, in our baseline model, each firm treats the benchmark as exogenous. However, in practice, some employers may be large enough to influence the market benchmarks. Their wages "pass-through" to the benchmark and hence affect the information that other firms use to make offers. Here we explore empirically how common these cases might be. Specifically, we investigate whether some firms hire a sufficiently large share of the employees in a labor market to shift the median of the benchmark when they adjust the salaries of their employees. In Figure H.1, we simulate how much the median of the salary benchmark for a position-industry-state would shift in the following quarter, when the benchmark is recalculated, if a firm decided to raise the salaries of all new hires (Panel A) and all existing employees (Panel B) by 10%. The result is very stark, shifting the salaries of all new hires by 10% would shift the benchmark median by 0.23% on average, and only 3% of firm-positions could shift the median by more than 1%. Even if a firm were to raise the salaries of all its employees by 10%, the median of the benchmark would only shift on average 0.59%. However, there are some firms that are large enough to have a meaningful impact on the market benchmarks.

Figure H.1: Salary Pass-through to Benchmark



Notes: Distribution of the rate of pass-through for position-state-sector-firms with access to the benchmark. Panel A reflects the pass-through of raising all new hire salaries 10%, while Panel B shows the same but for all employees, not just new hires. Mean (median) pass-through for new hires is 0.23 (0.12) and for all employees is 0.59 (0.25).

I Wage Responses to Benchmark Shocks

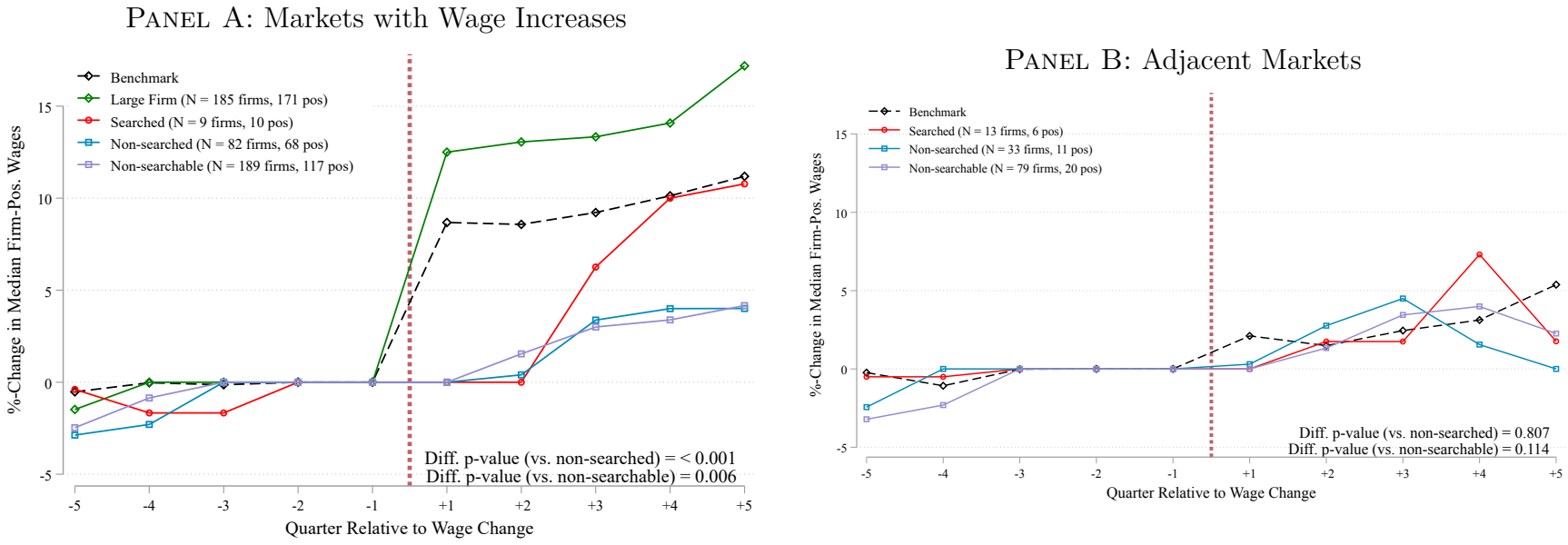
We corroborate our main results on salary dispersion using an alternative identification strategy. Inspired by Derenoncourt et al. (2021), we leverage rare cases where large firms change wages in a particular position by 10% or more. Due to the granular mappings of position titles, this constitutes a sudden and localized shock to the benchmark information: the benchmark for one position may change significantly while holding constant the benchmarks for other adjacent positions. We then track the time it takes for other firms hiring in the same labor market to converge to the new benchmark as a function of whether they searched for that particular benchmark versus an adjacent benchmark. For example, if a large firm increases wages for employees in the position title say "Fulfillment Center Workers", the benchmark will be affected; however the benchmark for an adjacent position title, say "Warehouse Laborers", will not be affected even though the two positions are quite similar. We compare firms that then Searched (vs. Non-Searched and Non-Searchable) Fulfillment Center Workers versus firms that Searched (vs. Non-Searched and Non-Searchable) Warehouse Laborers, and compare the wage evolution with their respective benchmarks.

Figure I.1 shows how both the benchmark and salaries respond to the shock. We identify large firms as those with more than 95th percentile pass-through rates, as defined in Appendix H above. Panel A corresponds to the case where a large firm raises the salaries in a specific position-state-sector market by at least 10% in a single quarter. The dashed black line illustrates that the salary benchmarks associated with these wage changes also rises, closely tracking the changes made by the large firms. Panel B reproduces Panel A, but for adjacent markets. Adjacent markets are classified by taking the closest position title in the same SOC group by restricting to the same skill level categorization and selecting the title with the most similar average tenure and 2017Q1 benchmark. Panel B suggests that the large firm wage changes do not mechanically affect the benchmark of adjacent positions. The dashed black line representing the salary benchmark rises gradually over the subsequent 5 quarters rather than discretely around the time of the wage change. Even by the 5^{th} quarter, the adjacent benchmark has only risen modestly by comparison.

We next compare firms Searched (vs. Non-Searched and Non-Searchable) positions for benchmarks that were affected, with Searched (vs. Non-Searched and Non-Searchable) positions for the adjacent positions. In Panel A, we see that five quarters after the wage change, the median wages of employees in Searched positions have converged toward the affected benchmark, reaching an average 10% wage increase (equivalent to the benchmark) 5 quarters after the wage change. By contrast, Panel B displays wages in adjacent markets. In these adjacent markets, wages of Search positions rise less than 5% on average 5 quarters out,

converging to the lower benchmark. The fact that we see divergent salary paths between Searched (vs. Non-Searched and Non-Searchable) in Panel A, but not in Panel B, suggesting that it is unlikely that the differential convergence is due to factors unrelated to the benchmark access.

Figure I.1: Wage Responses



Notes: Panel A plots cases where we observed a firm (*Large Firm*) with greater than 95th percentile (24.5%) pass-through in a specific position-state-sector market raising their median wage in that market by at least 10%. We plot the median percent change in wages from the reference period and compare the quarterly benchmark and other position types in the same position-state-sector market, centered around the month of the wage change. P-values test the difference in medians in the percent change from the reference period in quarter 5 between Searched and control positions. Panel B shows those progressions in "Adjacent Markets", or similar position titles in the same state and sector. We identify similar position titles based on SOC group, average tenure, skill classification.

J Expert Prediction Survey

J.1 Survey Design

To assess whether the experimental results are surprising, we conduct a forecast survey with a sample of experts. A sample of the full survey instrument is attached as Appendix L. In this survey, which follows best practices (DellaVigna et al., 2020), we start by describing the benchmarking tool and the context. Then, we outline a hypothetical experiment where some firms are randomly given access to salary benchmarks and other firms are not – we opt for this simpler version because the full quasi-experimental design would have added too much complexity to the survey. We then elicit beliefs about the effects of access to the benchmarking tool on the distribution of salaries around the benchmark and the average salary level. We also included two questions about heterogeneous effects: one by education and one by gender. For each forecast, we elicit how confident the respondent feels about his or her answer, and we also include an open-ended question so that they can explain their choices.

J.2 Implementation

We collected responses from experts in two ways. First, we posted the survey on the Social Science Prediction Platform from May 6, 2022, to June 24, 2022. Second, on May 9, 2022, we emailed an invitation to the prediction survey directly to a list of 500 professors with publications related to the topic, and gave them 7 days to complete the survey. We excluded respondents who are not academics (8 respondents) or who had already seen our study (4 subjects). The final sample includes 68 experts. Of these, 11.8% responded to the survey through the Social Science Prediction Platform, and the remaining 88.2% responded through our direct invitation. This final sample is comprised of 90.7% professors, 2.9% PhD students, and 7.4% researchers. Around 91.2% of the respondents in the sample are economists, and 85.3% report having done research on labor economics.

J.3 Survey Results

The first result worth noting is that a majority of respondents did not feel confident about their own predictions. This is consistent with the fact that, prior to our study, there was little economics research on salary benchmarking and thus the experts do not have prior literature to base their predictions upon. Figure J.1 shows the distribution of certainty for each of the questions in the survey. If we pool all four predictions, the majority (64.3%) of responses were *Not Confident at All* or *Slightly Confident*, few respondents (4.0%) felt *Very*

Confident and no one ever responded *Extremely Confident*. There was more confidence in the responses to some questions than others; for example, 45.6% of respondents were *Very* or *Somewhat* confident in their response to the question about education heterogeneity.

To elicit beliefs about the effects on the distribution of salaries around the benchmark, we show six histograms (reproduced in Figure J.2) and ask the respondent to pick the histogram they think is most likely to represent the real data. One of those histograms looks like the results we see in the analysis (with compression from above and below the benchmark). The other histograms present alternative situations, such as no effect, compression only from above, compression only from below, and so on. Figure J.3 displays the frequency with which the experts chose each of the six histograms. A minority (30.9%) of respondents selected the histogram corresponding to our results, which showed compression from above and below.

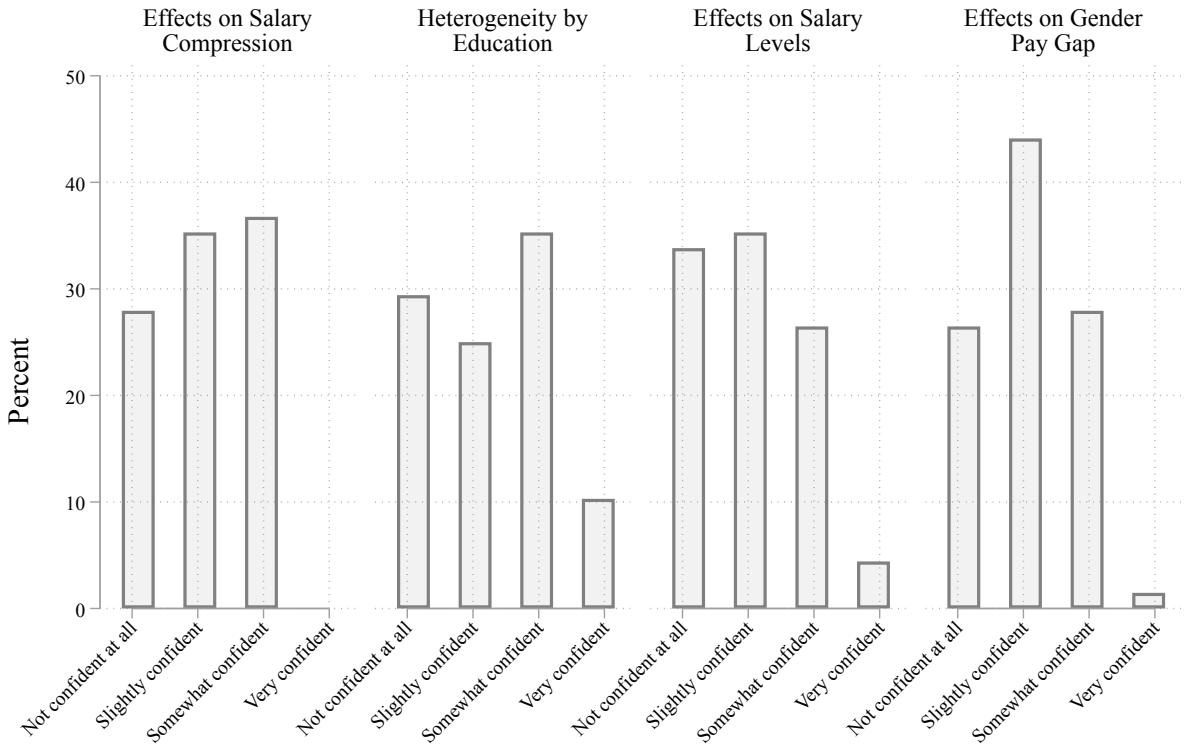
To elicit beliefs about the effect of benchmarking on the average salary, we first asked respondents whether they expect positive, negative or no effects on this outcome. If they responded that they expect positive or negative effects, we then elicit the effect size in percent terms. The results for this prediction are presented in Figure J.4. A slight majority of respondents predicted a null effect or close to a null effect. For this question, the predictions are the most accurate: 58.8% of predictions fall within the 90% confidence intervals of our estimates from Table F.1, overlaid on Figure J.4. Many of the open-ended responses to this question echo the sentiment of one respondent who reasoned that it would be "equally likely that [employers] would revise their salary up or down given the information from the benchmarking."

We also elicited beliefs about heterogeneous effects. We asked respondents whether high or low education positions will be more strongly affected by benchmarking. Panel A of Figure J.5 shows that a majority (61.8%) of experts predicted that high-education positions would be most strongly affected by a salary benchmark. This goes against our findings, according to whichlow-skill positions are more strongly impacted. In the open-ended responses, respondents often noted that there should be less compression at baseline among high-education positions, which also goes against our findings. One common reason why experts believed that high-education positions would be more strongly affected was that for that type of positions the "information about the true distribution should be more valuable."

Last, we asked respondents whether salary benchmarking would increase the gender pay gap, decrease it or leave it unchanged. Panel B of Figure J.5 shows that a majority (66.2%) of experts responded that the wage gap would be reduced. Open-ended responses among those who predicted a reduction in the gender pay gap often mentioned that "bargaining becomes less important if an external source anchors the salary" or that "the employer may rely less on individual negotiations and biases, which often work against women." This prediction also

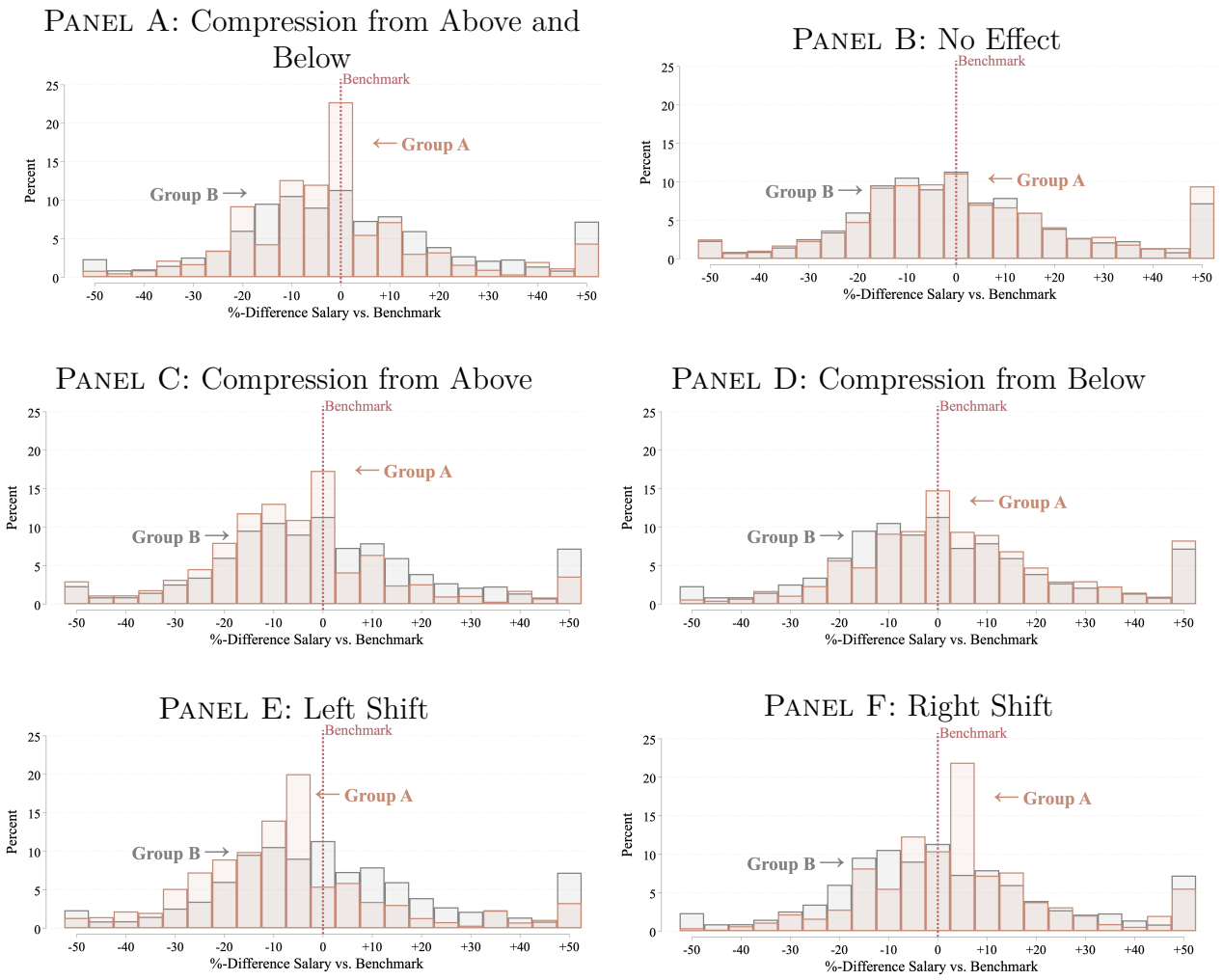
goes against our finding, in that we do not find any significant effects of salary benchmarking on the gender pay gap. However, this comparison must be taken with a grain of salt, as we do not have enough statistical power to rule out small negative or positive effects.

Figure J.1: Experts' Confidence In Their Own Predictions



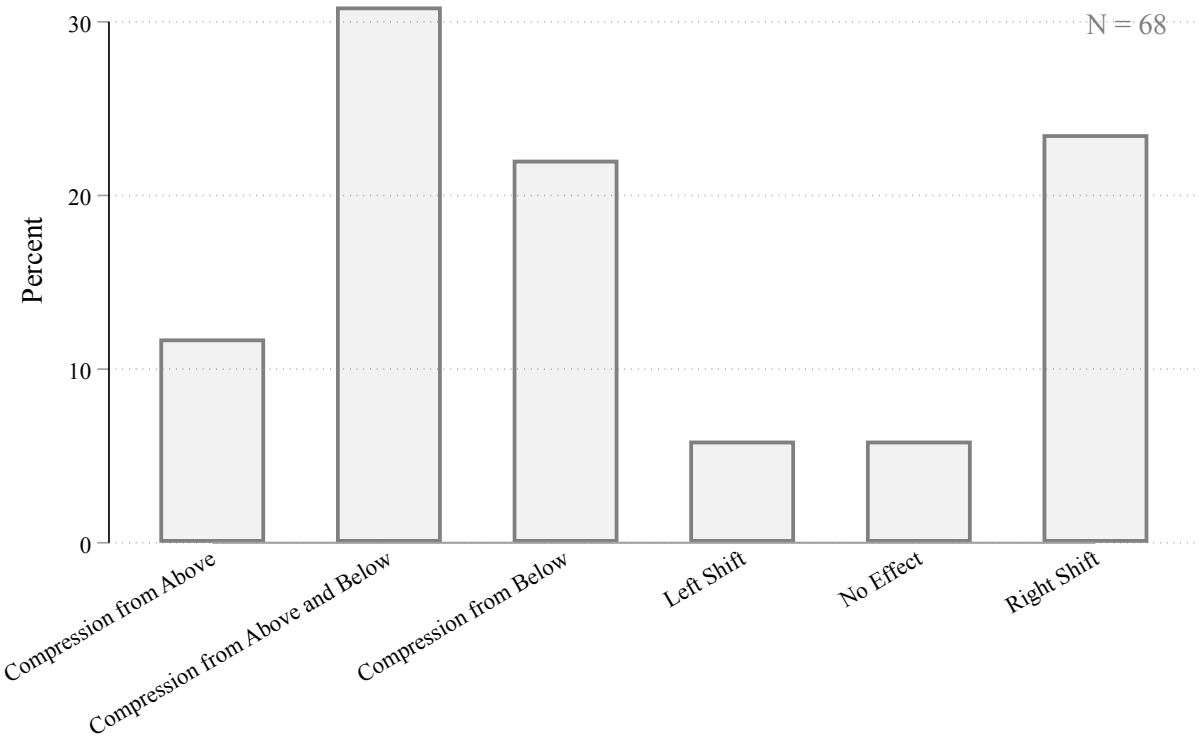
Notes: N = 68. Histogram of the response certainty for each question in the survey. Possible answers are *Not Confident at All*, *Slightly Confident*, *Somewhat Confident*, *Very Confident* and *Extremely Confident*.

Figure J.2: Expert Prediction Choice Set Regarding Salary Dispersion



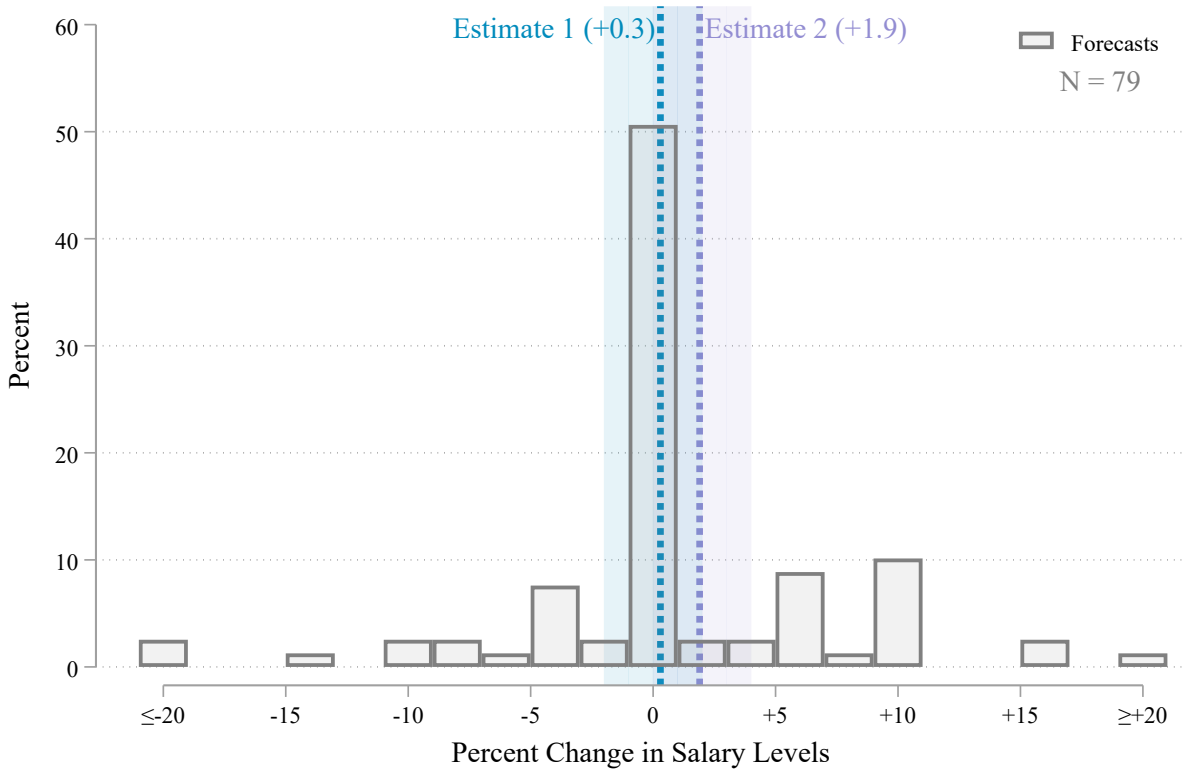
Notes: These are the images respondents could choose from when considering the compression effects of the benchmarking tool. Each figure is intended to show the effect described in the panel title. Panel A is an altered reproduction of Panel A of Figure 2.

Figure J.3: Expert Predictions Regarding Salary Dispersion



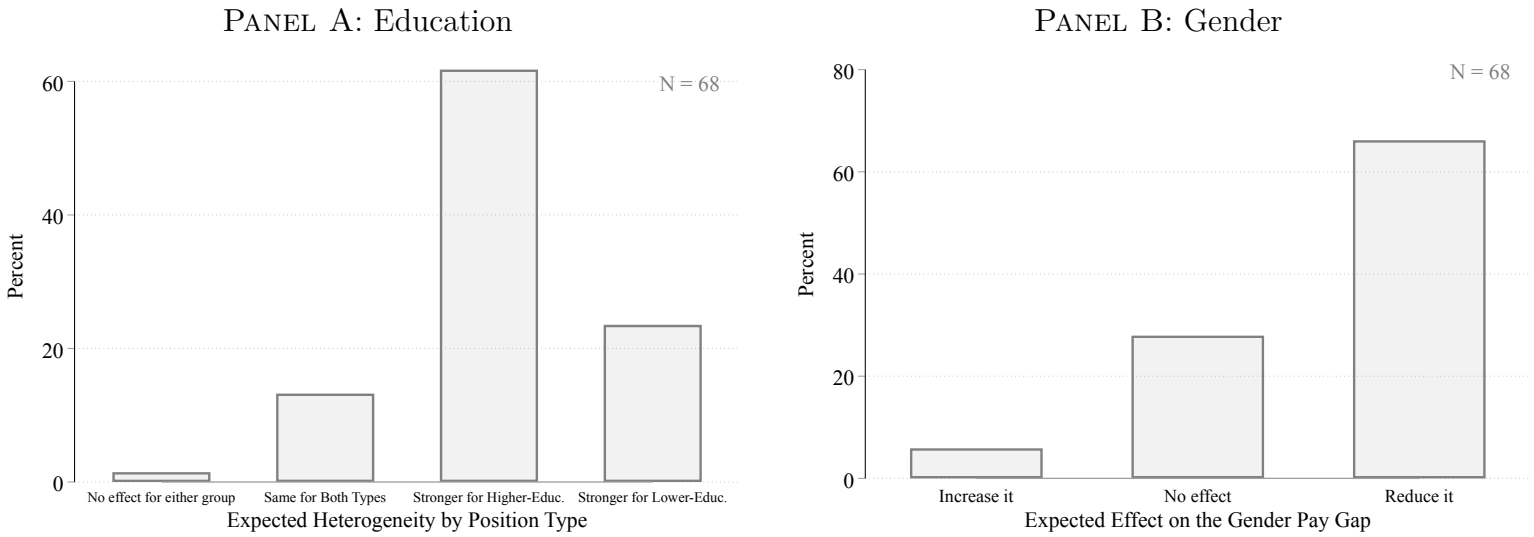
Notes: Histogram of the responses to survey question about the effects of benchmarking on compression. The six possible choices are displayed in Figure J.2.

Figure J.4: Expert Predictions Regarding the Average Salary



Notes: Histogram of the responses to survey questions about percent change in salary levels. Respondents were presented with a text box if they predicted salaries would go up or down. If they responded salary levels would stay the same, we include that here as 0% change. *Estimate 1* is the Searched vs. Non-Searched estimate of salary level effects from Table F.1 and *Estimate 2* is the Searched vs. Non-Searchable estimate. Displayed are the 90% confidence intervals.

Figure J.5: Expert Predictions Regarding Sources of Heterogeneity



Notes: Histogram of the responses to survey questions about heterogeneity by education (Panel A) and gender (Panel B) in the effects of benchmarking.

K SHRM Survey

The following is a short summary of this study to help you decide whether or not to be a part of this research.

Here is some Key Information about the study:

- We are asking you to take part in a research study because you have hiring expertise.
- If you agree to be in this study you will be asked to complete a 10-minute online survey. In the survey, you will answer questions about compensation. At the start of the survey, you will be asked a screening question to determine your eligibility for this study.
- Your participation is completely voluntary. You can choose not to participate, or you can agree to participate and change your mind later and your decision will not be held against you. Your refusal to participate will not result in any consequences or any loss of benefits that you are otherwise entitled to receive. You can ask all the questions you want before you decide.
- If you have questions, concerns, or complaints, or think the research has hurt you, talk to Professor Cullen. She can be reached at 617-495-1876, or zcullen@hbs.edu.

○ Yes, I agree to take the survey

→

Do you participate in setting the salaries for employees?

- ○ Yes
- ○ No

→

How many years of experience do you have setting salaries?

→

How many employees does your company have (please consider all locations)?

- ○ 1-49
- ○ 50-99
- ○ 100-999
- ○ 1,000-4,999
- ○ 5,000 or more

→

What main industry do you operate in? (start typing, then select a category that best describes your business.)

Please select a category from the list below to continue.

Are you working in the private sector or the public sector?

- ○ Private sector
- ○ Public sector

→

How would you describe your current role?

- ○ Human Resources Professional
- ○ Chief Human Resources Officer
- ○ Executive (outside HR division)
- ○ Manager (outside HR division)
- ○ Recruiter (outside HR division)
- ○ Other

→

Do you participate in salary settings for:

- ○ New hires
- ○ Current employees
- ○ Both

→

Suppose you wanted to know the median salary that your company pays employees in a specific position. Would you be able to access that data?

- ○ Yes, I can access it easily
- ○ Yes, but it would take quite a bit of work
- ○ No, I could not access that data even if I wanted to

→

When setting the compensation of their employees, some organizations use aggregate data on the market salaries for specific positions. This type of data is typically referred to as ***salary benchmarks***.

In your organization, do you use ***salary benchmarks***?

- ○ Yes
- ○ No

→

Which sources do you use to obtain ***salary benchmarks***? (Select all that apply.)

- ☐ Payroll data services
- ☐ Industry surveys
- ☐ Free online data sources
- ☐ Compensation consultants
- ☐ Paid online data sources
- ☐ Government data

→

What do you use the ***salary benchmark*** for? (Select all that apply.)

- [ ] To set salary ranges for specific job titles
- [ ] To plan ahead for headcount
- [ ] To determine salary in job advertisement
- [ ] To change salaries for current employees
- [ ] To set precise salaries for new hires
- [ ] In salary negotiations
- [ ] Other

→

How frequently do you use ***salary benchmarks*** to set salaries for new hires?

- ◯ For every hire
- ◯ A majority of hires
- ◯ Some of the hires
- ◯ A minority of hires
- ◯ Never

→

When do you use **salary benchmarks** *in relation* to new hires? (Select all that apply.)

- [ ] Before I publicize the position to include the expected salary in a job advertisement
- [ ] Right before I make an offer to the candidate
- [ ] After the candidate receives the offer, if the candidate wants to negotiate
- [ ] When the candidate presents an outside offer
- [ ] Other

→

Please explain briefly how you typically use ***salary benchmarks*** to set the salaries of new hires?

→

How frequently do you use ***salary benchmarking*** to change salaries for current employees?

- ◯ For all my employees
- ◯ For a majority of my employees
- ◯ For some of my employees
- ◯ For a minority of employees
- ◯ Never

→

When do you use ***salary benchmarks*** with current employees? (Select all the apply)

- [ ] When the employee goes through an annual review
- [ ] When the employee is up for promotion
- [ ] When the employee presents an outside offer
- [ ] When adjusting the salary ranges for positions
- [ ] Other

→

Can you please explain briefly how you typically use ***salary benchmarks*** to set the salaries of current employees?

→

Next, we'd like you to pick two different positions for which you are expecting to be hiring soon and tell us how you would set the salary for the new hire in each position.

Pick a position for which you are expecting to hire soon (start typing, then select a category - kindly allow a few seconds for the bold arrow to re-appear to continue.)

Please select a category from the list below to continue.

Pick a **second *different* position** for which you are expecting to hire soon (start typing, then select a category - kindly allow a few seconds for the bold arrow to re-appear to continue.)

Please select a category from the list below to continue.

→

Think about a future new hire in the role of Sales Engineers. What would be the **annual base salary** that you set for this person? (Please provide your best guess, and do not use commas.)

$

→

Suppose you look up the salary benchmark for this position using a highly accurate, up-to-date, compensation database and discover the median annual base salary is $93500. Upon reviewing that information, what salary would you pick?

$

→

Why did you use (or not use) the salary benchmark information in this compensation decision?

→

Think about a future new hire in the role of Sales Managers. What would be the **annual base salary** that you set for this person for a full-time position? (Please provide your best guess, and do not use commas.)

$

→

Suppose you look up the salary benchmark for this position using a highly accurate, up-to-date, compensation database and discover the median annual base salary is $92000. Upon reviewing that information, what salary would you pick?

$

→

Why did you use (or not use) the salary benchmark information in this compensation decision?

→

For which positions are ***salary benchmarks*** most useful?

- ○ Most useful for lower-education positions
- ○ Most useful for higher-education positions
- ○ Equally useful for both groups
- ○ Not useful for either group

→

Can you please explain briefly why salary benchmarks are most useful for higher-education positions?

→

Can you please explain briefly why salary benchmarks are most useful for lower-education positions?

→

Assume you are using a salary benchmark tool. The tool allows you to look at benchmark salaries, and to apply filters. When choosing filters there is a trade-off: applying filters can allow you to focus on a more relevant subgroup, but at the cost of smaller sample sizes and thus statistically imprecise benchmarks. Taking this into account, please select any filters from the set below that you would typically apply *after filtering for a particular position* (you can select more than one if you wish).

- ☐ Hourly vs Salaried
- ☐ State
- ☐ Industry
- ☐ Firm Size
- ☐ Revenue Size
- ☐ None of the above, the position-level filter is sufficient

→

Assume you are using a salary benchmark tool. It tells you the certain pieces of information about the salaries for that position. Please rank which information you typically care about the most (**drag and drop the options you care about, and then order them from most important to least important**). (1) Median salary (2) 10th Percentile (3) 25th Percentile (4) 50th Percentile (5) 75th Percentile (6) 90th Percentile (7) Average salary

Items	Order from most to least important (1=most important)
Median salary	
10th percentile	
25th percentile	
50th percentile	
75th percentile	
90th percentile	
Average salary	

→

For which position is the salary benchmark most useful for you? (start typing, then select a category - kindly allow a few seconds for the bold arrow to re-appear to continue.)
Please select a category from the list below to continue.

→

Think of your two closest competitors who also hire Retail Salespersons. For anonymity reasons, we'll refer to your competitors as firm A and firm B.

What is the maximum annual salary you think firm A would be willing to pay to hire in the full-time role of Retail Salespersons? (Please do not use commas)

$

→

Let’s say you find out that firm A would be willing to pay a maximum salary of $57500? After reviewing that information, what is the maximum annual salary you think firm B would be willing to pay to hire in the role of Retail Salespersons?

$

→

Suppose you lowered the salaries of your new hires by 10% and your competitors learned this information, what do you expect them to do with their new hires?

- ◯ Nothing, salaries of competitors would stay the same
- ◯ Salaries of competitors would fall between 1-5%
- ◯ Salaries of competitors would fall between 5-10%
- ◯ Salaries of competitors would fall 10%
- ◯ Salaries of competitors would fall by more than 10%

→

Suppose you raised the salaries of your new hires by 10% and your competitors learned this information, what do you expect them to do with their new hires?

- ◯ Nothing, salaries of competitors would stay the same
- ◯ Salaries of competitors would rise between 1-5%
- ◯ Salaries of competitors would rise between 5-10%
- ◯ Salaries of competitors would rise 10%
- ◯ Salaries of competitors would rise by more than 10%

→

If you adjusted the salaries of your new hires by 10%, do you think it would affect the salient numbers of the most commonly used salary benchmarks?

- ○ No, the popular salary benchmarks would stay the same or shift by a negligible amount
- ○ Yes, somewhat (eg. the median would shift by 1-2%)
- ○ Yes, a significant amount (eg. the median would shift by >2%)

→

Have you ever used Glassdoor as your salary benchmark source?

- ○ Yes
- ○ No

Have you ever used [redacted] Data Cloud Compensation Explorer as your salary benchmarking source?

- ○ Yes
- ○ No

→

What share of your competitors do you think use [redacted] Data Cloud Compensation Explorer as a salary benchmarking source?

- ○ The vast majority of my competitors
- ○ Some of my competitors
- ○ Very few of my competitors

→

Please share any feedback you have for us on the survey!

→

L Expert Prediction Survey

We are conducting an empirical study on labor markets. Due to your research record, you have been identified as an expert on the topic. We would love to elicit your expectations about the results of our analysis.

Please read the consent form below and click "I Agree" when you are ready to start the survey.

This survey involves no more than minimal risk to participants (i.e., the level of risk encountered in daily life). Participation typically takes between 5 and 10 minutes and is strictly confidential. Many individuals find participation in this survey enjoyable, and no adverse reactions have been reported thus far. Participation is voluntary, and participants may withdraw from the survey at any time.

Yes, I would like to take the survey

→

We would like to begin by providing some background information about this study.

When setting salaries, U.S. legislation prohibits employers from directly sharing compensation information with each other. However, employers are still allowed to use aggregated compensation data (e.g., median salary by position) provided by third parties. The practice of using aggregated market data is known as salary benchmarking.

We study the effects of salary benchmarking on the pay-setting of new hires. More precisely, we collaborated with a company that offers an advanced salary benchmarking tool that allows employers to look up market salaries in specific positions.

Employers in the Sample

Our sample covers a total of 1,982 firms from the United States, 583 of which gain access to the salary benchmarking tool. The average firm in our sample has 517 employees, ranging from 3 to 19,370 employees. Our firms cover all the main sectors in the U.S. economy, with the most common sectors being Manufacturing (21% of firms) and Finance and Insurance (14% of firms).

Employees in the Sample

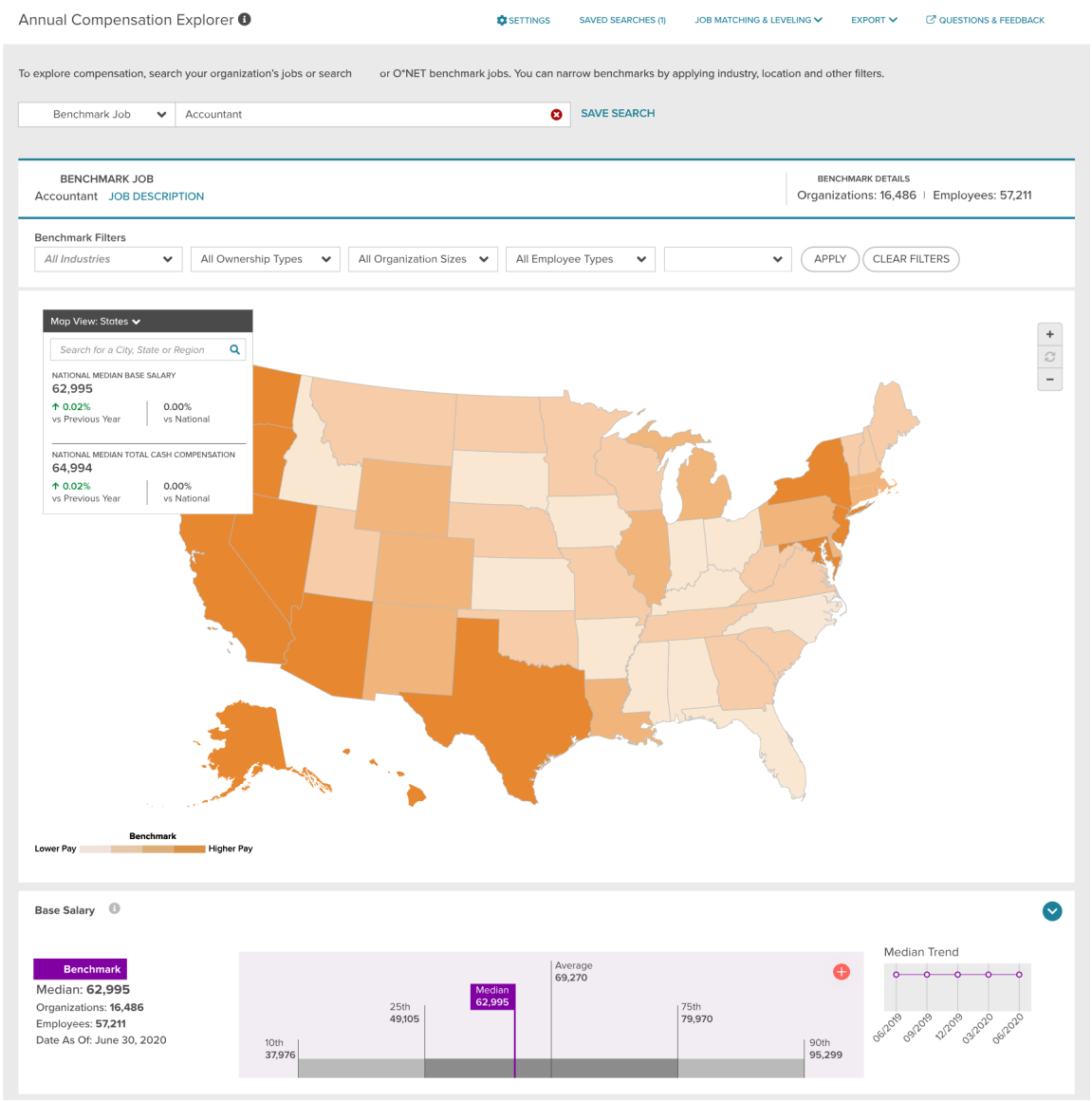
The average employee in our sample earns $41,441 per year in base salary. On average, base salary accounts for the vast majority (92.7%) of the total compensation.

There are over three hundred unique positions that are looked up in the salary benchmarking tool. Some of the most commonly searched positions are Bank Teller, Customer Service Representative, Patient Care Coordinator and Software Developer.

→

Benchmarking Tool

To give you a bit more context, find below a screenshot of the benchmarking tool:



Employers can look up a any position (e.g., in the above screenshot, "Accountant"). Employers can apply filters to see the aggregate statistics within a specific state or industry, among other user-friendly features. The search results display the median annual base salary for the position, along with other key statistics (e.g., the 25th and 75th percentiles).

The benchmarks shown to the employers are of the highest quality. They are calculated using accurate payroll records from hundreds of thousands of firms and tens of millions of employees. As a result, the benchmarks are precisely estimated: e.g., in the above screenshot, the distribution of Accountants' salaries is based on 57,211 unique employees working at 16,486 unique firms. Moreover, the monthly frequency of the payroll records provides the most up-to-date benchmarks.

→

Are you familiar with the results from our research (e.g. have you seen it presented in a seminar)?

Yes

No

→

First, we want to elicit your forecasts about the effects of salary benchmarking on the average salary of new hires.

Consider the following thought experiment. Two employers (A and B) who just hired a new employee (e.g., a bank teller). The two employers are otherwise identical, except that Employer A was randomly chosen to gain access to a salary benchmarking tool, while Employer B was not chosen to receive access to the tool. As a result, employer A looked up the market pay before setting the salary of the new employee, while Employer B did not have access to that information at the time of setting the salary of the new employee.

Relative to Employer B (without salary benchmarking), do you think the average salary set by Employer A (with salary benchmarking) will be higher, lower, or about the same?

About the same

Lower

Higher

→

How confident are you regarding the previous forecast?

- Not confident at all
- Slightly confident
- Somewhat confident
- Very confident
- Extremely confident

Can you please explain briefly why you expect salaries to be lower, on average, for Employer A?

→

How much lower do you expect the average salary of Employer A to be (in percent terms)? Please enter a number between 0% and 100%.

0 %

→

In the previous question, we asked you about the effects of salary benchmarking on the **average** salary.

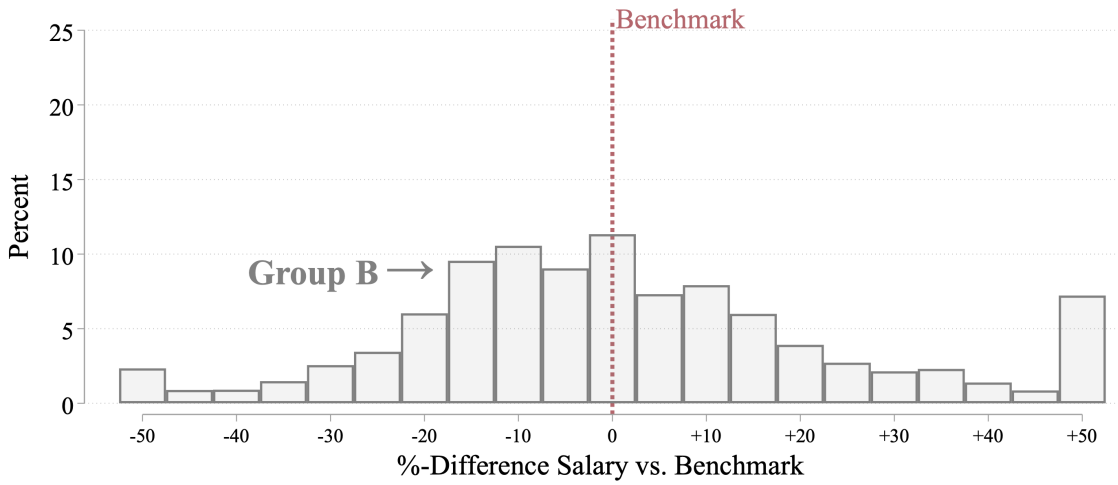
Next, we will ask you to forecast the effects of salary benchmarking on the **distribution** of salaries.

Consider two groups of employers:

Group A corresponds to employers ***with*** salary benchmarking: i.e., those who have access to the benchmark tool and look up the benchmarks before hiring a new employee.

Group B corresponds to employers ***without*** salary benchmarking: i.e., those who do not have access to the benchmark tool and thus cannot look up the benchmarks before hiring a new employee.

Consider the salaries of new hires relative to their corresponding benchmark (the median market salary for the position). For example, this is what the distribution of salaries looks like in Group B (without salary benchmarking):

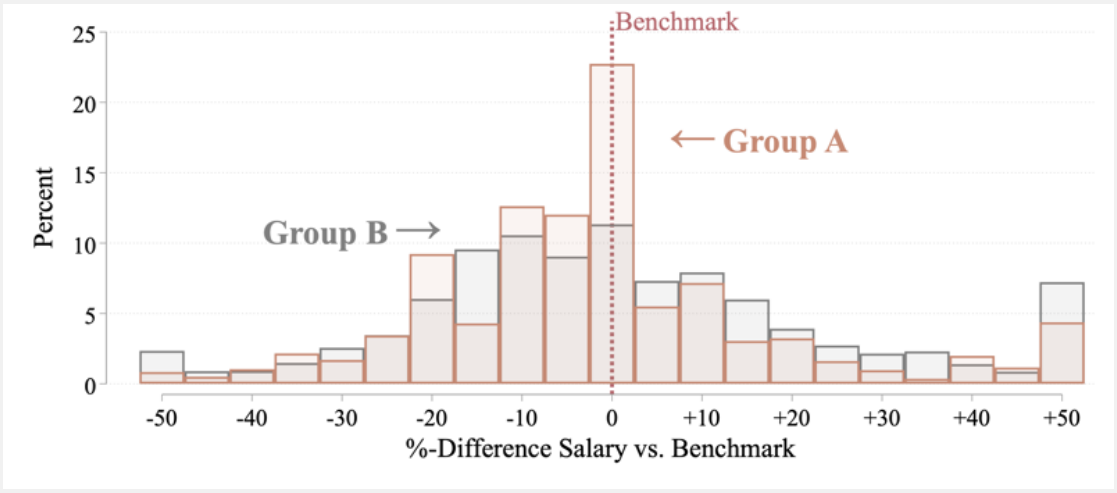


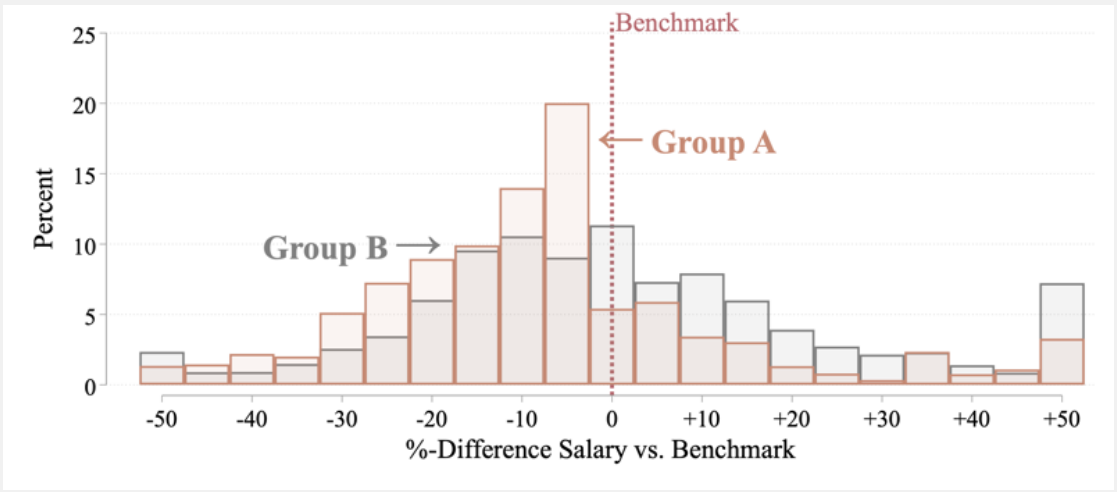
The middle bar corresponds to salaries that are close (i.e., within 2.5%) of the benchmark. The bars to the left of the middle bar correspond to new hires who are paid below the market benchmark, while the bars to the right of the middle bar correspond to new hires who are paid above the market benchmark.

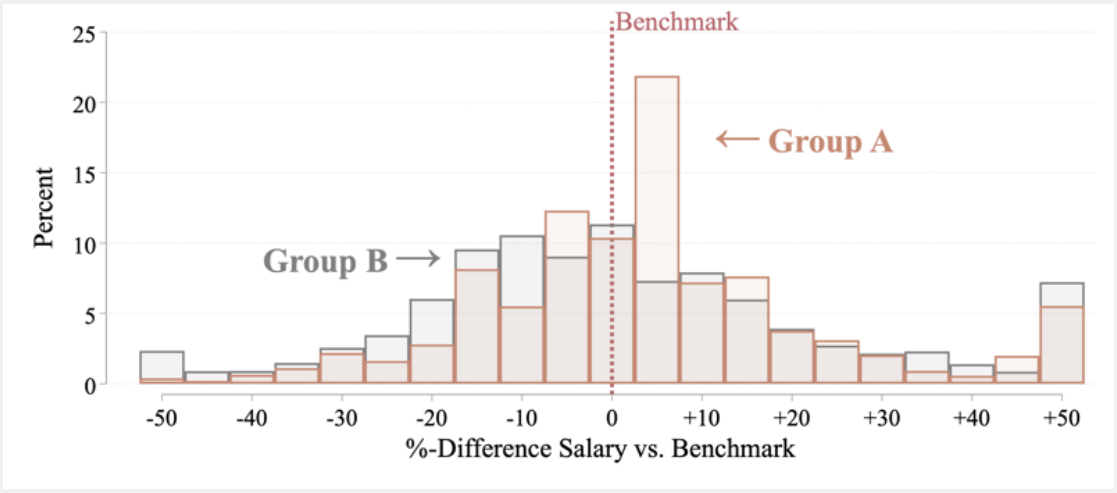
Now that you have seen what the distribution of salaries looks like in Group B (without salary benchmarking), we want you to predict what the distribution would look like for Group A (with salary benchmarking).

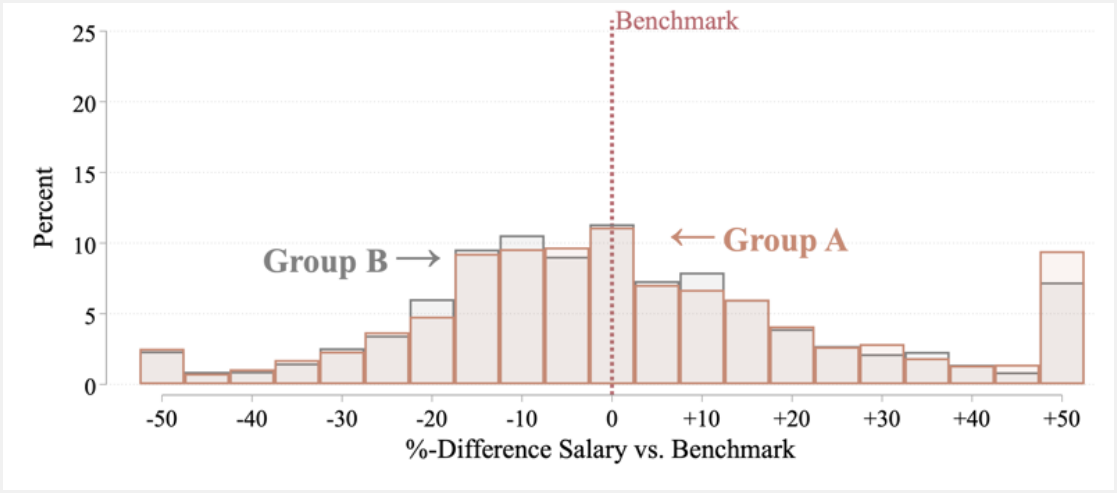
Find below six histograms. In each of them, the gray bars denote the salaries in Group B (without salary benchmarking), while the red bars correspond to salaries of Group A (with salary benchmarking).

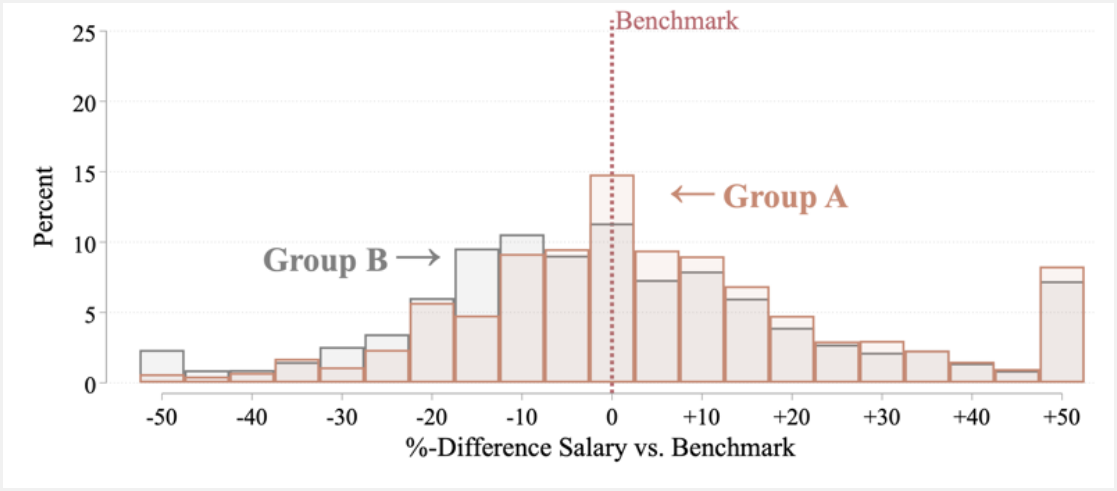
In your opinion, which of the following histograms best describes the effects of salary benchmarking?

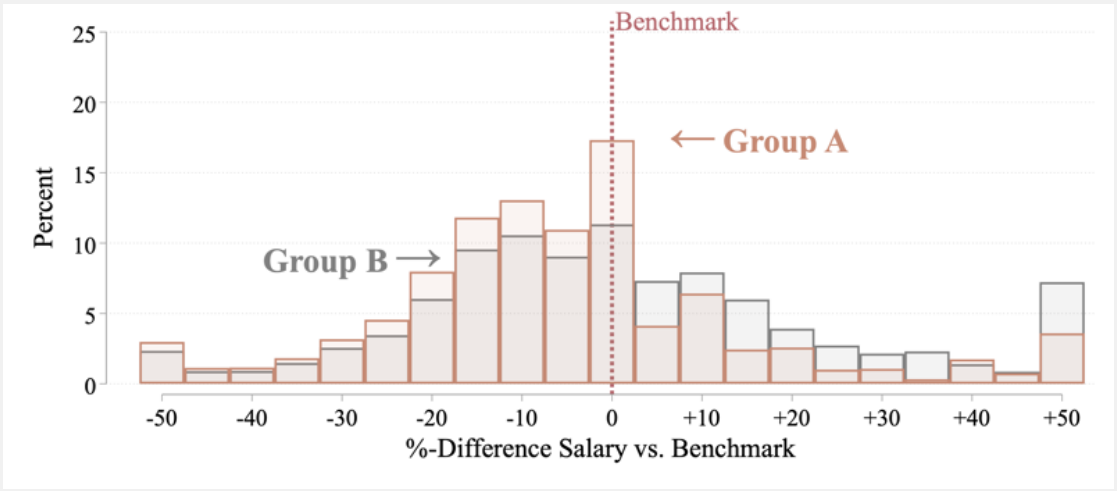





Benchmark
Group A
Group B
Percent
25
20
15
10
5
0
-50
-40
-30
-20
-10
0
+10
+20
+30
+40
+50
%-Difference Salary vs. Benchmark


Benchmark
Group A
Group B
Percent
25
20
15
10
5
0
-50
-40
-30
-20
-10
0
+10
+20
+30
+40
+50
%-Difference Salary vs. Benchmark


Benchmark
Group A
Group B
Percent
25
20
15
10
5
0
-50
-40
-30
-20
-10
0
+10
+20
+30
+40
+50
%-Difference Salary vs. Benchmark


Benchmark
Group A
Group B
Percent
25
20
15
10
5
0
-50
-40
-30
-20
-10
0
+10
+20
+30
+40
+50
%-Difference Salary vs. Benchmark

How confident are you regarding the previous forecast?

- Not confident at all
- Slightly confident
- Somewhat confident
- Very confident
- Extremely confident

Can you please explain briefly why you think the histogram you selected best represents the effects of salary benchmarking?

→

Now, we want you to forecast which type of positions (if any) will be most strongly affected by salary benchmarking.

Consider lower-education vs. higher-education positions. The lower-education positions require little training and no more than a high school degree. The higher-education positions require more training and a College degree or more. Some common examples of lower-education positions are Bank Teller, Receptionist and Delivery Driver. Some common examples of higher-education positions are Legal Associate Specialist, Registered Nurse and Software Developer.

Do you expect the effects of salary benchmarking to differ between lower-education and higher-education positions?

- No effect for either group
- Stronger for lower-education positions
- Stronger for higher-education positions
- Equally strong for both groups

→

How confident are you regarding the previous forecast?

- Not confident at all
- Slightly confident
- Somewhat confident
- Very confident
- Extremely confident

Can you please explain briefly why you think the effects will be stronger for lower-education positions?

→

Do you expect salary benchmarking to affect the gender pay gap?

No, it will not affect the gender pay gap

Yes, it will reduce the gender pay gap

Yes, it will increase the gender pay gap

How confident are you regarding the previous forecast?

Not confident at all

Slightly confident

Somewhat confident

Very confident

Extremely confident

→

Can you please explain briefly why you think salary benchmarking will not affect the gender pay gap?

→

This is the last section of the survey. We would appreciate if you could share some information about yourself.

Are you currently one of the following: graduate student (either Master level or PhD level), faculty, post-doc or non-academic researcher?

Yes

No

→

Which of the following describes your current position?

- Professor (Associate or Full)
- Assistant Professor
- Post-doc
- Researcher
- PhD Student
- Master Student

Please select your discipline

- Economics
- Management
- Political Science
- Psychology
- Sociology
- Other

Do you have research experience in the following fields? Please select all that apply:

- Labor Economics
- Personnel Economics
- Public Economics
- Behavioral Economics
- Organizational Economics
- None of the above

→

This is the end of the survey. We thank you for taking the time to provide your forecasts!

If you have any comments for us, please leave them below:

→