Worker Beliefs About Outside Offers, Wage Setting, Wage Dispersion, and Sorting

Junjie Guo*

University of Wisconsin-Madison

January 28, 2025

Abstract

This paper shows that employed workers in the U.S. have accurate beliefs about the average wage of outside offers they could potentially receive, and actual wage offers received by an employed worker are not strongly correlated with the worker's current wage conditional on productivity, which suggests that workers have little bargaining power and wages are largely set by firms without considering the outside option of each worker individually. Using each worker's belief and actual wage offers as two noisy measures of the worker's productivity, we find that about 73.7% of the variation in reported wages and 84.7% of the variation in true wages across employed workers are due to heterogeneity in worker productivity, which are larger than most estimates in the literature. Moreover, we find that sorting between worker productivity and the job-specific wage rate is positive but not significantly different from zero, either across all employed workers or for most sub-samples by gender, education and experience.

Keywords: Outside options, worker beliefs, wage setting, wage dispersion, sorting
JEL codes: J01, J31, J41, J64, D83

*Web: `https://sites.google.com/site/econjunjieguo/home`. Department of Economics, University of Wisconsin-Madison. 1180 Observatory Drive, Madison, WI 53706. Email: jguo27@wisc.edu. I thank Andreas I. Mueller for helpful comments.

1 Introduction

Do employed workers have accurate beliefs about their outside options? Are outside wage offers for an employed worker customized by firms making those offers to reflect the worker's wage with the current employer? How much of wage dispersion among employed workers is due to heterogeneity in worker productivity (as opposed to firm and job heterogeneity and labor market frictions)? Are high-productivity workers more or less likely to be employed by high-wage jobs? These questions are critical for our understanding of the labor market and relevant policies. For example, Jäger et al. (2024) show that employed workers in Germany wrongly anchor their beliefs about outside options on their current wages, which keeps overly pessimistic workers stuck in low-wage jobs and gives rise to monopsony power and labor market segmentation. Regarding policy, they find that providing information about outside options could help employed workers correct their beliefs and change their intentions for job search and wage negotiation.

This paper provides evidence suggesting that (1) different from German workers surveyed by Jäger et al. (2024), employed workers in the U.S. have accurate beliefs about the average wage of outside offers they could potentially receive, (2) actual wage offers received by an employed worker are not strongly correlated with the worker's current wage conditional on productivity, which suggests that workers have little bargaining power and wages are largely set by firms without considering the outside option of each worker individually, (3) heterogeneity in worker productivity accounts for about 73.7% of the variation in reported wages and about 84.7% of the variation in true wages across employed workers, which are larger than most estimates in the literature, and (4) sorting between worker productivity and the job-specific wage rate (per unit of worker productivity) is positive but not significantly different from zero, either across all employed workers or for most sub-samples by gender, education and experience.

We use novel data from the Survey of Consumer Expectations (SCE). Every four months (in March, July and November),[1] the survey asks a representative sample of employed workers in the U.S. to report their wages, whether they have looked for new work in the last four weeks, the expected chance of receiving at least one job offer from another firm in the next four months, beliefs about the average wage of outside offers they could potentially receive in the next four months, and the number of job offers received in the last four months, including accepted offers, rejected offers, and offers without a decision by the time of the survey. Workers with a positive number of job offers in the last four months are also asked to report the wages of the three best offers.

[1]SCE is a monthly survey. But the key information used in this paper is available every four months through a rotating module called the Labor Market Survey. Both SCE and its Labor Market Survey include employed as well as non-employed workers. We focus on employed workers in the Labor Market Survey.

We combine two consecutive surveys of the same worker that are four months apart from each other into one observation. This allows us to relate the job offers received in the four months between the two surveys, which are reported in the second survey, to the beliefs and other characteristics reported in the first survey, e.g., the wage and whether the worker has looked for new work in the four weeks before the first survey.

We find that outside offers are prevalent among employed workers, and majority of the offers are rejected and thus not observed in traditional data. 16% of workers employed in the first survey receive at least one job offer in the four months between the two surveys, and about 80% of the offers are rejected. Not surprisingly, rejected offers tend to pay lower wages than the job reported in the first survey, while accepted offers tend to pay higher wages. Moreover, we find that a significant amount of job offers received in the four months between the two surveys are not expected at the time of the first survey. Over 20% of workers employed in the first survey report a zero expected chance of receiving an outside offer in the four months between the two surveys. Among them, 6.3% receive at least one job offer in the four months. Over three quarters of workers employed in the first survey did not look for new work in the four weeks before the survey. Among them, 11.7% receive at least one job offer in the four months between the two surveys. Finally, about two thirds of workers employed in the first survey did not look for new work either in the four weeks before the first survey or in the four weeks before the second survey. Among them, 10% receive at least one job offer in the four months between the two surveys. These evidence is consistent with Faberman et al. (2022), who find that a fair amount of job offers for employed (and non-employed) workers are unsolicited, defined as job offers from potential employers that come about without a direct application by the job seeker.

To evaluate the accuracy of an employed worker's belief about the average wage of outside offers he/she could potentially receive, we focus on workers employed in the first survey who receive 1-3 job offers in the four months between the two surveys. Different from workers with 4 or more offers, who are asked to report the wages of only the three best offers, workers with 1-3 offers are asked to report the wages of all offers received in the four months, so that there is no selection in the reported wage offers. For each worker, we calculate the difference between (1) the average wage of job offers received in the four months between the two surveys and (2) the belief reported in the first survey about the average wage of outside offers that the worker could potentially receive in the same four months between the two surveys. We find that the difference is not significantly different from zero, either among all employed workers with 1-3 offers in the four months or for different sub-samples by gender, education, experience, and the number of offers received in the four months. Moreover, the difference is not significantly different from zero even among employed workers who did not look for new work in the four weeks before the first survey. This suggests that our results are

not driven by workers who applied for a job right before the first survey, learned about the wage of that job during the application process, reported the wage as his/her belief about the average wage of outside offers in the first survey, and received an offer from that job in the four months between the two surveys. Finally, we show that the difference (between actual wage offers and the belief) is not significantly correlated with either the worker's wage or expected chance of receiving an outside offer in the four months, both of which are obtained from the first survey and thus before any offer is received.

Together, the evidence suggests that employed workers in the U.S. have accurate beliefs about the average wage of outside offers they could potentially receive. Different from the evidence of belief anchoring documented by Jäger et al. (2024), where over 40% of employed workers in Germany expect to find a job with exactly the same pay as their current job if they were forced to leave their current job and had three months to find a job at another employer, only about 11% of employed workers in our data expect the average wage of outside offers to be the same as their current wage. Moreover, we show that the value of about 11% in our data could be explained by measurement error. In comparison, supplemental evidence from Jäger et al. (2024) suggests that the value of over 40% in their data cannot be explained fully by the framing effect, where workers are induced to choose the "same pay" option when it is available. Assuming employed workers in the U.S. have more accurate beliefs about their outside options than employed workers in Germany, this difference in worker beliefs could be both a cause and a consequence of the fact that worker mobility across jobs is higher in the U.S. than Germany (Jolivet et al., 2006). On the one hand, as shown by Jäger et al. (2024), if employed workers wrongly believe that their outside options are the same as their current job, they would have little incentive to search for another job, which would result in a low level of worker mobility. On the other hand, a low level of worker mobility means that employed workers have little chance to learn about their outside options, so that their beliefs about outside options are more likely to be inaccurate persistently. Exploring the interaction between worker beliefs and worker mobility and its variation across countries could be a fruitful direction for future work.

Next, we provide some evidence about wage setting by estimating the effect of an employed worker's current wage on the wages of his/her outside offers. A zero effect would suggest that wages are set by firms without considering the outside option of each individual worker, as in the seminal model by Burdett and Mortensen (1998). On the other hand, a significant effect would suggest that firms customize their wage offers for each worker based on the worker's outside options, as in the sequential-auction framework pioneered by Postel-Vinay and Robin (2002). Distinguishing the two cases is important because they have significantly different implications for wage dispersion, wage growth, and labor market transitions. If wages are posted by firms as in Burdett and Mortensen (1998), comparable

workers in the same firm must earn the same wage irrespective of where they are hired from and how many outside offers they have received and then rejected on the job, and a worker may move voluntarily from a high-productivity job to a low-productivity job if the latter offers a higher wage. In comparison, with sequential auctions, comparable workers in the same firm could earn different wages based on where they are hired from and experience different wage growth on the job depending on the outside offers they have received and rejected, and no worker would move voluntarily from a high-productivity job to a low-productivity job because the former could always match the wage offer from the latter.

We start by combining three consecutive surveys of the same worker into one observation. Focusing on employed workers who receive 1-3 job offers both in the four months between the first two surveys and in the four months between the last two surveys, we find that the difference between the wage offers received by the same worker in the two periods is not significantly correlated with the difference in the worker's current wage between the first two surveys. This suggests that the change in an employed worker's wage between the first two surveys has no effect on the wages of his/her outside offers.

Because employed workers in our data have accurate beliefs about the average wage of their outside offers, we could also relate the change in each worker's belief between two consecutive surveys to the change in his/her current wage during the same period, and use the correlation between the two changes to learn about the effect of the current wage on the expected wage of outside offers. Empirically, we find the two changes are significantly positively correlated with each other. However, in addition to the effect of the current wage on the wages of outside offers, the correlation could also arise from two other factors: (1) changes in worker productivity between the two surveys that affect both the current wage and the wages of outside offers, and (2) updating of worker beliefs about the average wage of outside offers in response to changes in the wage rate (per unit of worker productivity) of the current job. After excluding these two factors, the estimates suggest that a 1% increase in an employed worker's current wage rate raises the wages of his/her outside offers by less than 0.2%.

Finally, we use each employed worker's wage, belief about the average wage of outside offers, and the wage of an actual outside offer to estimate the contributions of worker heterogeneity and sorting to wage dispersion among employed workers. Intuitively, if the wage of an outside offer for an employed worker depends on the worker's productivity but not the wage rate (per unit of worker productivity) of the current job, and the worker's belief about the average wage of his/her outside offers is accurate, we could use the outside wage offer and the belief as two noisy but unbiased measures of the worker's productivity, so that the covariance between the two measures of outside offers across employed workers is equal to the variance of worker productivity, which identifies the contribution of worker heterogeneity

to wage dispersion. Given the variance of worker productivity, the covariance between each of the two measures of outside offers and the worker's current wage, which by construction is the sum (in logs) of the worker's productivity and the wage rate of his/her current job, identifies the covariance/sorting between worker productivity and the job-specific wage rate across employed workers and its contribution to wage dispersion. More generally, we show that the contributions of worker heterogeneity and sorting to wage dispersion could be identified from the variance-covariance matrix of the three wages (the current wage and the two measures of outside wage offers) even if the outside wage offer for an employed worker depends on the worker's current wage rate, as long as we know the strength of this effect and the contribution of measurement error to the variation in reported wages.

Empirically, after accounting for the fact that not all workers employed in the first survey receive a job offer in the four months between the two surveys, we find that about 73.7% of the variation in reported wages and about 84.7% of the variation in true wages are due to heterogeneity in worker productivity, which are larger than most estimates in the literature discussed below. The exact estimates depend on but are relatively robust to the effect of the current wage rate on outside wage offers and the contribution of measurement error to the variation in reported wages. Moreover, we find the contribution of worker heterogeneity is larger for women, workers without a bachelor's degree, and workers with more 15 years of experience than men, workers with a bachelor's degree or more, and workers with no more than 15 years of experience, respectively, although many of the differences are small and not significantly different from zero. Finally, we find that sorting between worker productivity and the job-specific wage rate is positive but not significantly different from zero, either across all employed workers or for most sub-samples by gender, education and experience.

1.1 Contributions and Implications

By addressing the four questions, this paper contributes to four strands of literature.

Worker beliefs. The standard assumption that workers have accurate beliefs about their labor market opportunities are challenged by a growing number of empirical studies, which tend to focus on unemployed workers, including their beliefs about the job finding probability (Spinnewijn, 2015; Mueller et al., 2021; Mueller and Spinnewijn, 2023) and the reservation wage (Krueger and Mueller, 2016; Le Barbanchon et al., 2019), which is a reflection of their beliefs about both the probability of receiving a job offer and the wage distribution of job offers they could potentially receive. Conlon et al. (2021) study the beliefs of both non-employed and employed workers about their job offers, how the beliefs vary with employment status, and how the beliefs are updated over time in response to actual job offers.

Following Jäger et al. (2024), we focus on employed workers' beliefs about their outside

options. Different from Jäger et al. (2024), who find employed workers in Germany anchor their beliefs about outside options on their current wages, we find that employed workers in the U.S. have accurate beliefs about the average wage of their outside offers, and the beliefs are updated in response to changes in their current wages. This suggests that the benefit from providing employed workers with information about their outside options is smaller in the U.S. than Germany.

Wage setting. Given the importance of wage-setting protocols mentioned above, many studies have provided relevant evidence from different angles. While some studies find evidence of bargaining (Hall and Krueger, 2012; Card et al., 2016; Biasi and Sarsons, 2022; Roussille, 2024; Caldwell et al., 2025) and the significance of outside options for individual wages (Caldwell and Danieli, 2024), other studies find that wages are posted for the majority of workers (Lachowska et al., 2022) and do not respond to changes in their outside options (Jäger et al., 2020; Guo, 2024b). Presumably, some offers are customized to workers' outside options while others are not (Flinn and Mullins, 2021; Doniger, 2023), and the composition of the two types of offers may vary with labor market conditions (Rabinovich et al., 2023).

Our focus on the effect of an employed worker's current wage on the wages of his/her outside offers is new to the literature. In addition to Jäger et al. (2020), Lachowska et al. (2022) and Guo (2024b) mentioned above, our finding that outside wage offers for an employed worker are not strongly tailored to the worker's current wage is also consistent with Di Addario et al. (2023), who show that where a worker is hired from is relatively inconsequential for his/her wages in comparison to where he/she is employed, and Roussille and Scuderi (2024), who show that, even in a high-wage labor market, a model where firms offer the same wage to all workers with the same level of predicted productivity fits the data better than alternative models where firms customize the wage offer for each worker based on the worker's predicted level of productivity and idiosyncratic preferences for each job. Moreover, our finding is consistent with the evidence of national wage setting documented by Hazell et al. (2022), where firms set the same nominal wage for a job across all their establishments, despite varying local labor market conditions.

Wage dispersion. Existing studies tend to use matched employer-employee data and one of two approaches. The first is the econometric model with two-way fixed effects pioneered by Abowd et al. (1999), where individual wages are assumed to be the sum of a worker fixed effect, a firm fixed effect, and the effects of some time-varying characteristics such as the worker's experience. Relevant studies are reviewed by Card et al. (2018) and Bonhomme et al. (2023), where firm effects are found to account for about 20% of overall wage variation. Excluding this direct effect of firms and the positive correlation between firm effects and worker effects found in most studies, the direct contribution of worker heterogeneity itself is

less than 70%, e.g., 63% in the Portuguese data studied by Card et al. (2018).[2]

The second approach to wage dispersion using matched employer-employee data is to model the observed wages and labor market transitions directly in an equilibrium framework, e.g., Postel-Vinay and Robin (2002), Cahuc et al. (2006), Bagger et al. (2014), Bagger and Lentz (2019), and Taber and Vejlin (2020). These studies tend to find an even smaller contribution of worker heterogeneity, e.g., less than 40% in both Postel-Vinay and Robin (2002) and Bagger and Lentz (2019). One exception is Taber and Vejlin (2020), who find that heterogeneity in pre-market skills account for 59-82% of wage inequality.

Our estimates suggest that the contribution of worker heterogeneity is larger than previously thought. With both accepted and rejected offers, our identification strategy does not rely on worker mobility across firms or jobs. As a result, we do not have to model worker mobility directly or assume it is exogenous conditional on worker and firm fixed effects and observable characteristics, and our estimates are not affected by limited worker mobility. Moreover, we account for the entirety of each worker's productivity at the time when wages are observed using the two noisy measures of outside wage offers. In comparison, existing studies tend to assume that worker productivity is the sum of a worker fixed effect and some parametric function of labor market experience. This may not capture the entirety of a worker's productivity if the parametric function is inadequate in accounting for the change in each worker's productivity over time and the heterogeneity across workers.[3]

Sorting. We focus on sorting between worker productivity and the job-specific wage rate. In comparison, existing studies tend to focus on sorting between workers and firms. A typical measure is the correlation between worker fixed effects and firm fixed effects estimated from the econometric model pioneered by Abowd et al. (1999). Bonhomme et al. (2023) find that the inter-quartile range of the contribution of this measure of sorting to wage dispersion in the literature is -2% to 18%, which become 10% to 20% when they correct the biases from limited worker mobility using one method (correlated random effects) and 5% to 13% using another (heteroskedastic fixed effects). In addition to wages, workers may also sort across firms based on non-wage characteristics (Bagger and Lentz, 2019; Lamadon et al., 2022; Lentz et al., 2023), and they may sort across jobs based on multidimensional skills (Lindenlaub, 2017; Lise and Postel-Vinay, 2020).

[2]Card et al. (2018) and Bonhomme et al. (2023) also discuss the economic and econometric issues related to this approach, e.g., worker mobility across firms is assumed to be exogenous conditional on the two fixed effects and other observable characteristics, the challenge in estimating the effect of time-varying characteristics such as age/experience when time fixed effects are included in addition to worker and firm fixed effects, and the biases due to limited mobility of workers across firms.

[3]The effect of unmeasured worker productivity on measured wage inequality and labor market sorting is demonstrated by Gibbons and Katz (1992) and Gibbons et al. (2005), among others. As discussed later in the paper, our approach to worker heterogeneity is similar to Hall and Mueller (2018) and Mueller et al. (2021), both of which use the covariance between a subjective measure of worker belief and the relevant objective measure of labor market realizations to identify the heterogeneity across unemployed workers.

The rest of this paper proceeds as follows. Section 2 describes the data and documents the prevalence of outside offers among employed workers. Section 3 evaluates the accuracy of the belief that an employed worker has about the average wage of outside offers he/she could potentially receive. Section 4 estimates the effect of an employed worker's current wage on his/her outside wage offers. Section 5 estimates the contributions of worker heterogeneity and sorting to wage dispersion across employed workers. Section 6 concludes. Supplemental results are in an Online Appendix.

2 Data

We use data from the Survey of Consumer Expectations (SCE). Fielded by the Federal Reserve Bank of New York, SCE is an internet-based monthly survey of a rotating panel of approximately 1,300 household heads from across the U.S. Each respondent participates in the panel for up to 12 months, with a roughly equal number of respondents rotating in and out of the panel each month.

A respondent is asked about his/her employment status each month, including whether he/she is working full-time or part-time and whether he/she is self-employed or employed by someone else. Every four months (in March, July and November), active members who had participated in a SCE monthly survey in the prior three months are asked some additional questions through a rotating module called the Labor Market Survey. Because respondents are in SCE for up to 12 months, they may end up taking up to 3 Labor Market Surveys. Our analysis combines two consecutive Labor Market Surveys of the same worker into one observation, which means a worker may have up to two observations. We treat different observations for the same worker as if they belong to different workers. The results are similar if we use one observation for each worker.

Following are some questions from the Labor Market Survey that we use in this paper, with the original question number (and relevant instructions, if any) given at the beginning of each question:

- L3: "*How much do you make before taxes and other deductions at your [main/current] job, on an annual basis? Please include any bonuses, overtime pay, tips or commissions.*"
- JS6: "*Have you done anything in the LAST 4 WEEKS to look for new work?*"
- NL1: "*How many job offers did you receive in the last 4 months? Remember a job offer is not necessarily a job that you accepted.*"

- NL2 (shows up to 3 job offers depending on NL1 response): "*Thinking about the 3 best job offers that you received in the last 4 months, What was their annual salary? And were they for a full-time or a part-time job? Note the best offer is the offer you would be most likely to accept.*"

- NL3 (loop through for up to 3 job offers): "*Did you accept this [job offer/job offer X]?*" Besides yes and no, a respondent could also choose "*I am still deciding*".

- OO2e: "*What do you think is the percent chance that within the coming four months, you will receive at least one job offer from another employer? Remember that a job offer is not necessarily a job you will accept.*"

- OO2a: "*Think about the job offers that you may receive within the coming four months. Roughly speaking, what do you think the average annual salary for these offers will be for the first year?*"

All employed workers are asked about questions L3, JS6, NL1, and OO2e.[4] If a respondent reports a strictly positive number of offers in question NL1, he/she is then asked about questions NL2 and NL3 for each of the three best offers. In other words, if $n_{i,t}$ is the answer to question NL1 by worker i in month t, then $\min\{n_{i,t}, 3\}$ is the number of responses to each of the two questions NL2 and NL3. Similarly, a respondent is asked about question OO2a only if his/her answer to question OO2e is strictly positive (i,e., > 0).

The Labor Market Survey started in March 2014. The last five questions listed above were first asked in November 2014. We use data available by November 2022. Because the survey asks about annual salaries, and there is no information about work hours, we focus on workers employed full-time by someone else (as opposed to part-time workers or the self-employed) throughout the paper, and assume that, in the absence of measurement error, the different salaries reported by the same worker reflect differences in wages rather than hours. In fact, we will refer to them as wages rather than salaries.

Specifically, let $y_{i,t}$ be the wage of worker i in month t (obtained from the worker's response to question L3 in month t), $b_{i,t,t+4}$ be the belief of worker i about the average wage of outside offers that he/she could potentially receive in the four months between t and $t+4$ (obtained from the worker's response to question OO2a in month t), and $q_{i,j,t,t+4}$ be the wage of the jth full-time offer received by worker i in the four months between t and $t+4$

[4]The questions are ordered in the same way as they appear in the survey's questionnaire, with questions in the survey but not used in this paper omitted. For example, each employed worker is asked about L3 before JS6, and there are some questions in between in the survey that are not listed here. While not the focus of this paper, non-employed workers are also asked about some of the questions, e.g., NL1-NL3. See Conlon et al. (2021) for a more detailed discussion of the Labor Market Survey, and Armantier et al. (2017) for an overview of SCE.

(obtained from the worker's response to question NL2 in month $t+4$). The survey does not specify the order in which a respondent with multiple offers should report the wage of each offer. We use random numbers to re-order the offers reported by each worker in each survey, so that the index j in $q_{i,j,t,t+4}$ is random. Although actual outside wage offers $q_{i,j,t,t+4}$ are observed four months later, we use the same value of the Consumer Price Index (CPI) from month t to adjust each of the three wages ($y_{i,t}$, $b_{i,t,t+4}$ and $q_{i,j,t,t+4}$) to 2019 dollars. Adjusting $q_{i,j,t,t+4}$ using a different value of CPI from month $t+4$ would introduce an artificial difference between $q_{i,j,t,t+4}$ and the other two measures. We limit the impact of outliers by focusing on values between \$10,000 and \$500,000 for each of the three wages ($y_{i,t}$, $b_{i,t,t+4}$ and $q_{i,j,t,t+4}$).

2.1 Outside Offers for Employed Workers

While our focus is on the three wages ($y_{i,t}$, $b_{i,t,t+4}$ and $q_{i,j,t,t+4}$), it is worth emphasizing that a unique feature of the SCE Labor Market Survey is that respondents are asked to report all job offers, including accepted offers, rejected offers and offers without a decision by the time of the survey. As can be seen from the questions listed above, respondents are reminded throughout the survey that a job offer is not necessarily a job that they accepted (question NL1) or will accept (question OO2e). This is different from standard labor market surveys, where we could only learn about accepted offers through employment transitions.

Table 1 reports the distribution of the number of offers (from question NL1) received in the 4 months between two consecutive Labor Market Surveys in months t and $t+4$ among workers who were employed full-time in the first survey in month t. The first row shows that, among the 7861 workers who were employed full-time in month t and reported the number of offers received between t and $t+4$, 84.0% received zero offer in the 4 months, 8.9% received one offer, 4.2% received two offers, 1.1% received three offers, and 1.7% received four or more offers. A deviance test for the hypothesis that the number of offers follows a Poisson distribution results in a p-value of essentially zero. This suggests that, inconsistent with the assumption adopted by most job search models, the arrival of job offers does not follow a Poisson process. A similar point is made by Bradley and Gottfries (2021), among others. We make no parametric assumption about the offer arrival process in our analysis.

Women, workers without a bachelor's degree, and workers with at most 15 years of experience (age minus years of schooling minus 6) are more likely to receive an offer than men, workers with a bachelor's degree or more, and workers with more than 15 years of experience, respectively. The differences are not very large, especially between the two groups by education. It is worth noting that 65% (5110/7861) of workers in our sample have a bachelor's degree or more, much larger than the value of around one third among people aged 25 or older in the U.S. There are two reasons. First, it is common for college-educated

Table 1: Number of job offers received in four months

		Number of job offers				
	Observations	0	1	2	3	4+
All	7861	84.0	8.9	4.2	1.1	1.7
Female	3609	82.3	10.6	4.2	1.4	1.5
Male	4249	85.4	7.5	4.3	0.9	1.9
Bachelor's degree: No	2739	83.5	8.6	4.7	1.2	2.0
Bachelor's degree: Yes	5110	84.8	9.2	3.6	1.0	1.4
Experience: ≤ 15 years	2764	81.5	11.3	4.6	1.1	1.5
Experience: > 15 years	5074	85.2	7.8	4.1	1.1	1.8
Looked for new work in last 4 weeks						
No	6085	88.3	6.7	3.0	0.6	1.4
Yes	1776	68.3	16.9	8.8	2.9	3.1
Expected chance of receiving at least one outside offer in 4 months						
0%	1716	93.7	3.8	1.6	0.3	0.5
1-5%	1532	92.9	5.2	1.3	0.2	0.3
6-20%	1930	86.0	8.9	3.2	0.6	1.3
21-50%	1620	78.5	12.2	5.5	1.8	2.0
51-100%	1056	59.5	17.7	12.6	3.9	6.2
Searched before neither t nor $t+4$	5212	90.0	5.8	2.7	0.5	1.0

Notes: This table reports the distribution of the number of job offers received in the four months between t and $t+4$ among workers employed full-time in month t. The values in the last five columns are measured in percentages. For example, the first row shows that 84.0% received 0 offer in the four months, 8.9% received one offer, 4.2% received two offers, etc. Characteristics in the first column of all but the last row are measured in month t before any offer was received. The last row focuses on workers who looked for new work neither in the 4 weeks before t nor in the 4 weeks before $t+4$. Due to missing values, groupings by some characteristics may not add up to the total number of observations. Experience equals age minus years of schooling minus 6.

individuals to be over-represented in internet-based surveys like SCE, probably because they have better access to and are more familiar with the internet. This over-representation could be addressed using the survey weight for each respondent provided by SCE. In particular, Armantier et al. (2017) show that the weighted distribution of SCE respondents by education is almost identical to the corresponding distribution from the American Community Survey. Except for the number of observations, all statistics reported in this paper account for the survey weight. In particular, the weighted fraction of workers with a bachelor's degree or more in our sample is 44.8%, much smaller than the unweighted value of 65% mentioned above but still larger than the value of around one third in the population. The remaining difference arises because we focus on full-time employed workers, and college-educated individuals are more likely to be employed than those without a bachelor's degree (National Center for Education Statistics, 2020).

In our sample of workers who were employed full-time in month t, 22.6% (1776/7861) had looked for new work in the 4 weeks before the survey in month t (question JS6). Not surprisingly, those who had looked for new work in the 4 weeks before t received more offers on average in the four months between t and $t+4$. However, the fraction of workers with one or more offers is as large as 11.7% even among those who did not look for new work in the 4 weeks before t. Moreover, the last row of Table 1 shows that 66.3% (5212/7861) of full-time employed workers in our sample did not look for new work in either the four weeks before t or the four weeks before $t+4$, and 10% of them received at least one offer in the four months between t and $t+4$.

These findings are consistent with Faberman et al. (2022). Using a different SCE supplement with detailed measures of job search behavior and outcomes for both the employed and non-employed, they find that 22.4% of employed workers in their sample actively looked for work in the last four weeks, and the fraction of employed workers who received at least one offer in four weeks (as opposed to four months in our data) is 26.1% among those who looked for work in the four weeks and 5.3% among those who did not look for work in the four weeks. Moreover, they show that the existence of *unsolicited* offers, defined as job offers from potential employers that came about without a direct application by the job seeker, is an important reason why workers who did not look for work also received a fair amount of offers. In particular, 3.0% of all employed workers in their sample received at least one unsolicited offer in four weeks, including 4.4% for those who looked for work in the four weeks and 2.5% for those who did not look for work in the four weeks, and unsolicited offers account for 17% of all offers received by the employed looking for work and 28% of all offers received by all employed. The existence of unsolicited offers could also explain the large fraction of offers received by full-time employed workers in our sample who did not look for new work in the four weeks before t (and $t+4$). This is useful when we evaluate the belief

$b_{i,t,t+4}$ using realized wage offers $q_{i,j,t,t+4}$ in the next section.[5]

In response to question OO2e, 21.8% (1761/7861) of the full-time employed workers in our sample reported in month t that they had a zero chance of receiving a job offer from another employer in the next four months. Among them, 6.3% received one or more offers in the four months. This suggests that many offers received between t and $t+4$ were not expected when workers reported their beliefs about the average wage of outside offers $b_{i,t,t+4}$ in month t. This is also useful when we evaluate the belief $b_{i,t,t+4}$ using realized wage offers $q_{i,j,t,t+4}$ in the next section.

An employed worker's belief about the chance of receiving an outside offer (from question OO2e) is not determined solely by his/her past search behavior (from question JS6). Among the 1776 full-time employed workers in our sample who looked for new work in the four weeks before the survey in month t, 4.3% reported a zero chance of receiving an outside offer in the four months between t and $t+4$. Among them, 21.2% received at least one offer in the four months. On the other hand, among the 6085 full-time employed workers in our sample who did not look for new work in the four weeks before the survey in month t, 71.8% reported a strictly positive chance of receiving an outside offer in the four months between t and $t+4$. Among them, 14.0% received at least one offer in the four months. Moreover, 20.8% (1632 out of 7861) of the full-time employed workers in our sample did not look for new work in the four weeks before t and reported a zero chance of receiving an outside offer in the four months between t and $t+4$. Among them, 5.7% received at least one offer in the four months. This further suggests that a significant fraction of outside offers between t and $t+4$ were not expected when employed workers reported their beliefs $b_{i,t,t+4}$ in month t.

Table 1 shows that workers with higher expected chances of receiving an outside offer (reported in month t) received more offers on average in the four months between t and $t+4$. This suggests that worker beliefs about the probability of receiving an outside offer are informative, consistent with the finding by Conlon et al. (2021) using data from the same survey. Quantitatively, Conlon et al. (2021) show that workers tend to over-estimate their chances of receiving an offer. Instead of worker beliefs about the probability of receiving an outside offer (question OO2e), we focus on worker beliefs about the average wage of outside offers they could potentially receive (question OO2a).

2.2 Accepted and Rejected Offers

Let p be the probability of receiving at least one outside offer in each month, so that $(1-p)^4$ is the probability of receiving zero offer in four months. For our sample of full-time employed

[5]The Job Search Supplement used by Faberman et al. (2022) was administered annually in each October between 2013 and 2021. Because the survey is cross-sectional, it cannot be used for our analysis, which requires panel data as described below.

workers, the first row of Table 1 indicates that $(1-p)^4 = 0.84$, so that $p = 0.043$. In comparison, the monthly employer-to-employer transition probability constructed by Fujita et al. (2024) is almost always below 0.025 during our sample period (July 2014 to November 2022). This suggests that there is indeed a large fraction of rejected offers in the data that are not associated with employer-to-employer transitions.

Figure 1 provides more direct evidence on this. For each worker who was employed full-time in month t and received one or more offers between t and $t+4$, we use the worker's responses to question NL3 to the calculate the number of offers with a decision (either accepted or rejected, excluding those choosing the option "*I am still deciding*"). As mentioned above, a respondent has to answer question NL3 for each of the three best offers, so that each worker could have up to three offers with a decision.

Among workers who were employed full-time in month t and received one or more offers between t and $t+4$, 78.2% had at least one offer with a decision. The blue bars show that workers with one, two and three decisions rejected 64.6%, 85.7% and 89.7% of their offers, respectively, and 79.1% of all offers with a decision were rejected (as opposed to accepted). The red bars focus on workers with a decision for each of the (up to) three best offers, which accounts for 72.8% of all workers who were employed full-time in month t and received at least one offer between t and $t+4$. In particular, workers who received two or more offers in the four months but reported a decision for only one of them in month $t+4$ are included in the first blue bar (from the left) but not the first red bar, which includes only workers who received exactly one offer in the four months and reported a decision for that offer in month $t+4$. The fractions of rejected offers indicated by the red bars are very close to the blue bars, suggesting that the large fraction of rejected offers is robust to missing decisions. Assuming the remaining offers for workers with four or more offers are more likely to be rejected than the three best offers, the true fraction of rejected offers should be even higher than the value of around 79% indicated by the last two bars.

Looking across workers instead of offers, we find that 65.4% of employed workers who reported the decision of each of the (up to) three best offers (whose offers are used for the last red bar in Figure 1) rejected all those offers. This is consistent with Faberman et al. (2022), who find that 32.8% of employed workers in their sample who received at least one offer accepted the best offer. They note that the large fraction of rejections suggests substantial randomness in the job search process, which is inconsistent with models of directed search where all best offers are accepted. This randomness suggests that we could use outside wage offers to measure each worker's productivity and estimate the contribution of worker heterogeneity to wage dispersion across employed workers, as discussed in detail later.

For each full-time wage offer $q_{i,j,t,t+4}$ received by worker i between t and $t+4$, we define the wage premium as $\log q_{i,j,t,t+4} - \log y_{i,t}$. Figure 2 plots the distribution of the wage premium



Figure 1: Fraction of Rejected Offers

Notes: This graph uses workers who were employed full-time in month t, received at least one offer in the four months between t and $t+4$, and reported the decision (either accepted or rejected) of at least one of the (up to) three best offers. The horizontal axis groups workers by the number of offers with a decision, and the vertical axis reports the fraction of rejected offers for each group. The blue bars use all workers with the relevant numbers of decided offers, which accounts for 78.2% of all workers with at least one offer. The red bars use only workers with a decision for each of the (up to) three best offers, which accounts for 72.8% of all workers with at least one offer. In particular, workers who received two or more offers but reported the decision of only one offer are included in the first blue bar (from the left) but not in the first red bar, which includes only workers who received exactly one offer and reported the decision of that offer. Because a worker could report the decisions of up to three best offers, the third blue bar is the same as the third red bar by construction, and both bars use workers who received three or more offers. Assuming the remaining offers for workers with four or more offers are more likely to be rejected than the three best offers, the true fraction of rejected offers is larger than the value of around 79% indicated by the last two bars.

for both accepted and rejected offers. The average wage premium is 0.097 (standard error: 0.017) for accepted offers and -0.036 (standard error: 0.015) for rejected offers. This suggests that the data on wage offers are reasonable and informative. Moreover, the figure shows that the wage premium is negative for a significant fraction of accepted offers and positive for a significant fraction of rejected offers. This suggests that wages are not the only reason why some offers are accepted and others are rejected, consistent with studies finding a significant role for non-wage benefits (Rosen, 1986; Wiswall and Zafar, 2018; Hall and Mueller, 2018; Sorkin, 2018; Taber and Vejlin, 2020). With both accepted and rejected offers, we do not need to model the offer acceptance decision that is necessary for most studies using only accepted wages.

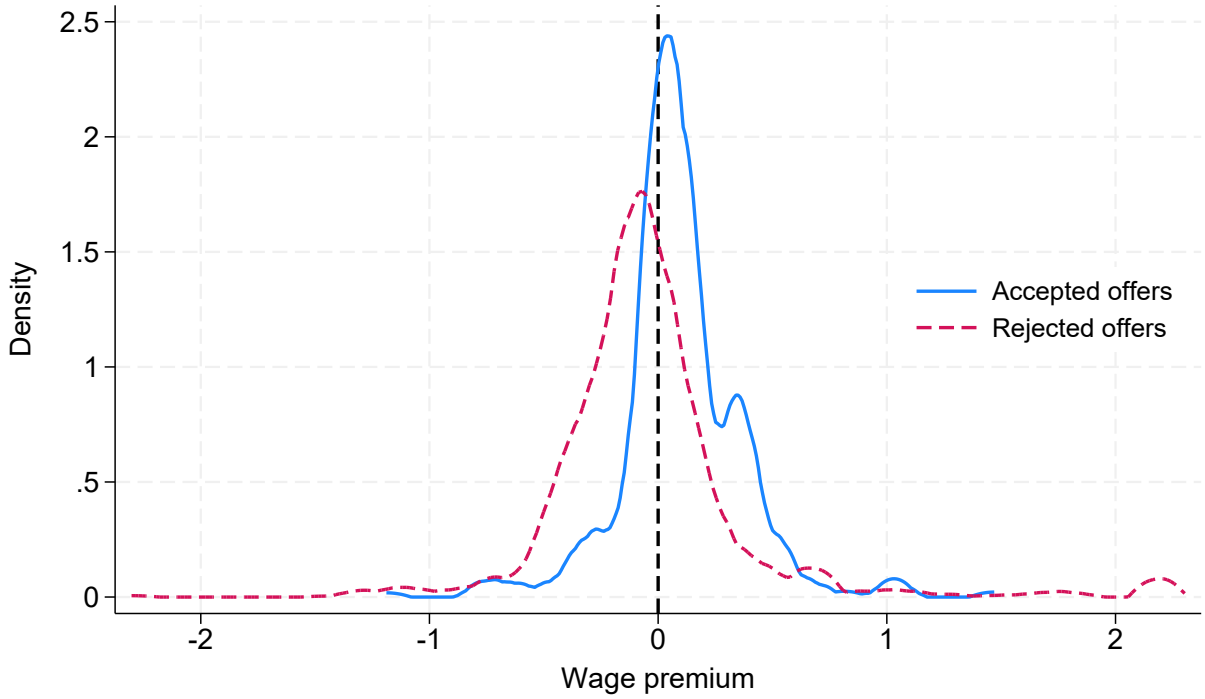


Figure 2: Wage Premium for Accepted and Rejected Offers

Notes: This graph uses workers who were employed full-time in month t, received at least one offer in the four months between t and $t+4$, and reported the annual salary, full-time (vs part-time) status and decision (either accepted or rejected) of at least one of the (up to) three best offers. Focusing on full-time offers with a valid salary and decision, we define the wage premium for each offer as the difference between the log annual salary of the offer and the log annual salary of the worker reported in month t before any offer was received. The average wage premium is 0.097 (standard error: 0.018) for accepted offers and -0.036 (standard error: 0.015) for rejected offers.

3 Worker Beliefs About Outside Offers

This section evaluates worker beliefs $b_{i,t,t+4}$ using realized wage offers $q_{i,j,t,t+4}$. Let $n_{i,t,t+4}$ be the number of offers received by worker i in the four months between t and $t+4$ (from question NL1). We focus on workers with $n_{i,t,t+4} \in \{1,2,3\}$ offers. Because respondents are asked to report wages of only the three best offers (see question NL2), wage offers reported by a worker who received $n_{i,t,t+4} \geq 4$ offers are not representative of all wage offers received by him/her, and thus cannot be used to evaluate his/her belief about the average wage of outside offers. Moreover, we focus only on full-time offers with valid wages $q_{i,j,t,t+4}$ (from question NL2), the number of which is denoted by $n^f_{i,t,t+4}$.

For each worker with $n_{i,t,t+4} \in \{1,2,3\}$ and $n^f_{i,t,t+4} \in \{1,2,3\}$, we define

$$LA_{i,t,t+4} = \log\left(\frac{1}{n^f_{i,t,t+4}} \sum_{j=1}^{n^f_{i,t,t+4}} q_{i,j,t,t+4}\right)$$

$$AL_{i,t,t+4} = \frac{1}{n^f_{i,t,t+4}} \sum_{j=1}^{n^f_{i,t,t+4}} \log q_{i,j,t,t+4}$$

as the log average and average log of outside wage offers received by the worker, respectively.

We consider three measures of the bias in each worker's belief $b_{i,t,t+4}$

$$B^{LA}_{i,t,t+4} = LA_{i,t,t+4} - \log b_{i,t,t+4}$$
$$B^{AL}_{i,t,t+4} = AL_{i,t,t+4} - \log b_{i,t,t+4}$$
$$B^{q_1}_{i,t,t+4} = \log q_{i,1,t,t+4} - \log b_{i,t,t+4}$$

As mentioned above, we use random numbers to re-order the offers reported by each worker, so that the first wage offer $q_{i,1,t,t+4}$ is representative of all wage offers received by worker i in the four months between t and $t+4$.

Table 2 shows that, across the 796 (full-time) employed workers in our sample who received 1-3 (full-time) offers in a four-month period, none of the three measures of biases is significantly different from zero. The same is true when we look at sub-samples by gender, education, experience, and the number of offers. By construction, the three measures are equal to each other for workers with one offer. On the other hand, because the log function is concave, the log average $LA_{i,t,t+4}$ is larger than the average log $AL_{i,t,t+4}$ for workers with two or three offers, so that $B^{LA}_{i,t,t+4}$ is larger than $B^{AL}_{i,t,t+4}$. Empirically, the difference between $B^{LA}_{i,t,t+4}$ and $B^{AL}_{i,t,t+4}$ is small for both workers with two offers and workers with three offers, and neither of the two measures is statistically different from zero for either of the two groups

of workers.[6]

As mentioned in our discussion of Table 1, 22.6% of employed workers in our sample looked for new work in the four weeks before reporting their beliefs $b_{i,t,t+4}$ in month t. It is possible that some of these job seekers may have learned about the wages of some jobs they had just applied to or even interviewed for, and used the information in reporting their beliefs $b_{i,t,t+4}$. In this case, it would not be a surprise to find that the actual wage offers received later are not significantly different from their beliefs.

Table 2 shows that, indeed, none of the three measures of biases is significantly different from zero among employed workers who looked for new work in the four weeks before reporting their beliefs $b_{i,t,t+4}$ in month t. More importantly, there is also no significant bias among the larger number of employed workers who did not look for new work in the four weeks before reporting their beliefs $b_{i,t,t+4}$ in month t, neither is there any significant bias among those who did not look for new work in either the four weeks before t or the four weeks before $t+4$. If anything, all three measures of biases are smaller for workers who did not look for new work than those who looked, probably because the beliefs of those who looked for new work were affected by the specific jobs they had looked at before t, and those jobs were not very representative of the offers they received between t and $t+4$.

The insignificant bias among employed workers who did not look for new work suggests that the additional information received by those looked for new work right before the survey is not the main reason why employed workers have unbiased beliefs about the average wage of outside offers they could potentially receive. As discussed in the previous section, a large fraction of outside offers for employed workers are unsolicited and unexpected, irrespective of whether a worker had looked for new work or not. With unsolicited and unexpected offers, the evidence in Table 2 suggests that employed workers indeed have accurate beliefs about the average wage of outside offers they could potentially receive.

Figure 3 plots the bias of each worker measured by $B^{LA}_{i,t,t+4}$ against the worker's expected chance of receiving an outside offer (from question OO2e) reported in month t. The 95% confidence interval for the kernel-weighted local polynomial fit (red solid line) covers zero (black dashed line) in almost all cases. Regressing the bias $B^{LA}_{i,t,t+4}$ on the vertical axis against the expected chance on the horizontal axis results in an intercept of -0.0003 (standard error: 0.0004) and a slope of 0.034 (standard error: 0.024), neither of which is significantly different from zero. Figure A1 in the Online Appendix shows that the results are almost identical

[6]Across the 796 observations used in Table 2, the minimum of the three pairwise correlations among $LA_{i,t,t+4}$, $AL_{i,t,t+4}$ and $\log q_{i,1,t,t+4}$ is 0.992, and the minimum of the three pairwise correlations among $B^{LA}_{i,t,t+4}$, $B^{AL}_{i,t,t+4}$ and $B^{q_1}_{i,t,t+4}$ is 0.979. The small difference between $B^{LA}_{i,t,t+4}$ and $B^{AL}_{i,t,t+4}$ for workers with two or three offers is a result of the small dispersion of different wage offers for the same worker, which suggests that frictional wage offer dispersion (i.e., the dispersion of wage offers for the same worker due to labor market frictions) is smaller than most estimates in the literature. See Guo (2024a) for details.

Table 2: Biases in Worker Beliefs About Outside Offers

	Observations	$B^{LA}_{i,t,t+4}$	$B^{AL}_{i,t,t+4}$	$B^{q_1}_{i,t,t+4}$
All	796	0.017	0.015	0.012
		(0.014)	(0.014)	(0.014)
Female	362	0.034	0.033	0.028
		(0.021)	(0.021)	(0.021)
Male	434	0.004	0.000	-0.001
		(0.018)	(0.018)	(0.019)
Bachelor's degree: No	260	0.025	0.022	0.021
		(0.025)	(0.025)	(0.026)
Bachelor's degree: Yes	533	0.008	0.005	0.001
		(0.016)	(0.016)	(0.017)
Experience: ≤ 15 years	346	0.001	-0.000	-0.001
		(0.018)	(0.018)	(0.018)
Experience: > 15 years	446	0.026	0.023	0.020
		(0.020)	(0.020)	(0.021)
One wage offer	550	0.013	0.013	0.013
		(0.016)	(0.016)	(0.016)
Two wage offers	196	0.040	0.032	0.024
		(0.032)	(0.032)	(0.035)
Three wage offers	50	-0.034	-0.041	-0.046
		(0.032)	(0.032)	(0.037)
Searched before t: Yes	366	0.026	0.024	0.023
		(0.022)	(0.022)	(0.022)
Searched before t: No	430	0.010	0.007	0.003
		(0.017)	(0.017)	(0.018)
Searched before neither t nor $t+4$	324	0.019	0.015	0.010
		(0.017)	(0.017)	(0.019)

Notes: Using workers who were employed full-time in month t and received 1-3 outside offers in the four months between t and $t+4$, this table reports the means (and standard errors of the means in the parentheses) of three measures of biases in their beliefs reported in month t about the average wage of outside offers they could potentially receive between t and $t+4$. Each measure for worker i is defined as the difference between (1) a measure of the wages of outside offers received by the worker in the four months between t and $t+4$ and (2) the log of the worker's belief reported in month t about the average wage of outside offers that he/she could potentially receive in the four months between t and $t+4$. To measure (1), $B^{LA}_{i,t,t+4}$ uses the log of the average wage of all outside offers received by the worker, $B^{AL}_{i,t,t+4}$ uses the average of the log wage of each outside offer received by the worker, and $B^{q_1}_{i,t,t+4}$ uses the log wage of an offer chosen randomly from all outside offers received by the worker. Due to missing values, groupings by some characteristics do not added up to the total number of observations. Experience is equal to age minus years of schooling minus 6. A worker is said to have searched before t if he/she reported in month t that he/she had looked for new work in the last 4 weeks.

when the other two measures of biases ($B^{AL}_{i,t,t+4}$ and $B^{q1}_{i,t,t+4}$) are used.

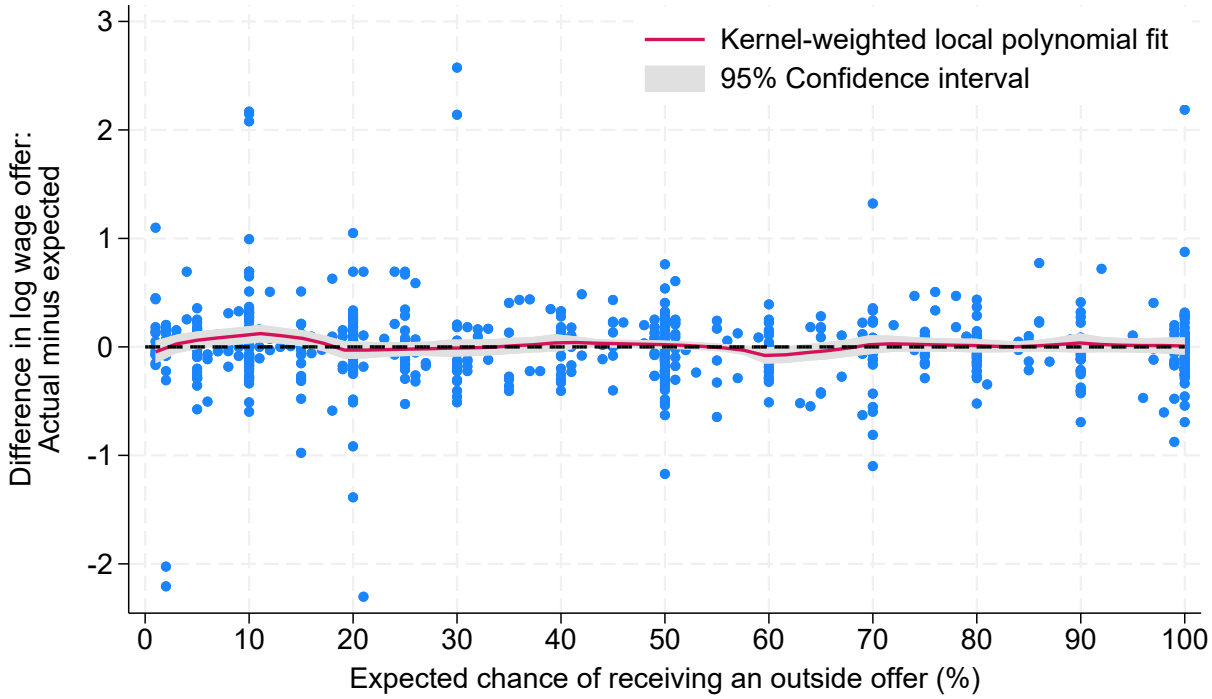


Figure 3: Biases in Worker Beliefs About Outside Offers vs Expected Chance of Receiving an Outside Offer

Notes: This graph uses workers who were employed full-time in month t and received 1-3 full-time offers with valid wages in the next four months. The horizontal axis plots the expected chance of receiving an outside offer in the next four months reported by each worker in month t, and the vertical axis plots the difference between (1) the log of the average wage of all outside offers received by the worker in the four months following t and (2) the log of the worker's belief reported in month t about the average wage of outside offers that he/she could potentially receive in the next four months.

Figure 4 plots the bias of each worker measured by $B^{LA}_{i,t,t+4}$ against $\log y_{i,t}$, the worker's log wage from question L3 in month t. The 95% confidence interval for the kernel-weighted local polynomial fit (red solid line) covers zero (black dashed line) in almost all cases. Regressing the bias $B^{LA}_{i,t,t+4}$ on the vertical axis against the wage $\log y_{i,t}$ on the horizontal axis results in an intercept of 0.017 (standard error: 0.022) and a slope of -0.181 (standard error: 0.244), neither of which is significantly different from zero. Figure A2 in the Online Appendix shows that the results are almost identical when the other two measures of biases ($B^{AL}_{i,t,t+4}$ and $B^{q1}_{i,t,t+4}$) are used.

Together, the evidence suggests that employed workers in our data have accurate beliefs about the average wage of outside offers they could potentially receive. This is consistent with the findings by Conlon et al. (2021). Using data from the same survey, they focus on the

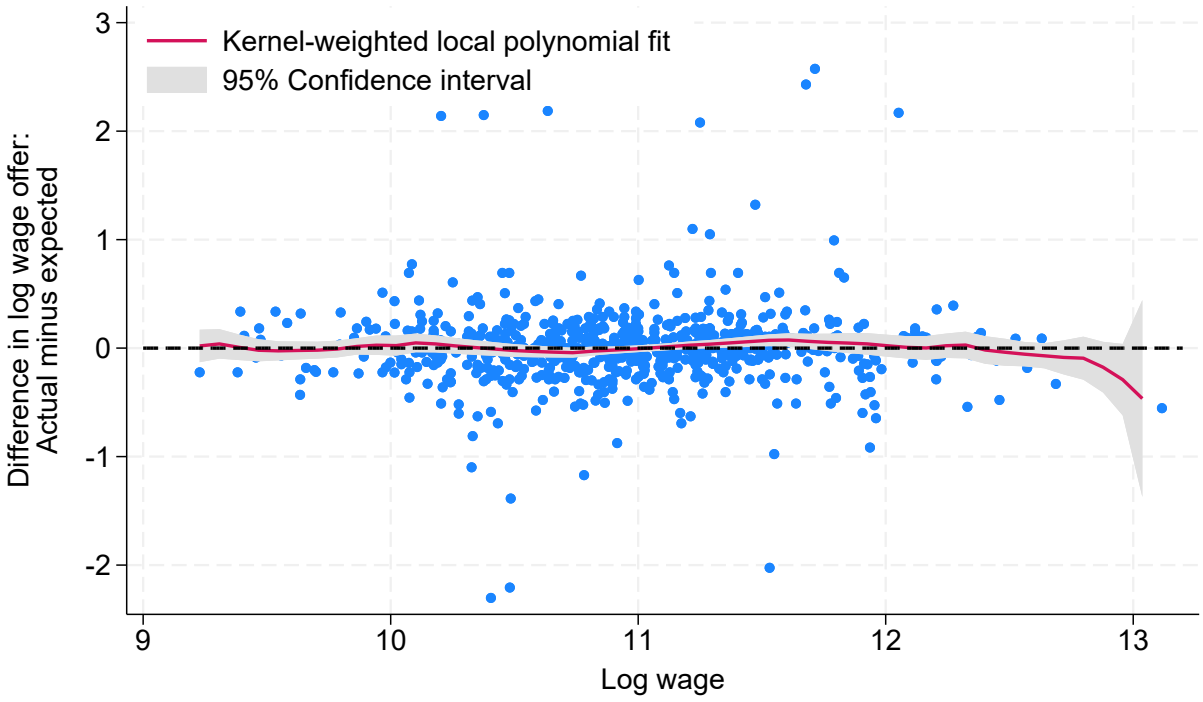


Figure 4: Biases in Worker Beliefs About Outside Offers vs Current Wages

Notes: This graph uses workers who were employed full-time in month t and received 1-3 full-time offers with valid wages in the next four months. The horizontal axis plots the log wage of each worker in month t, and the vertical axis plots the difference between (1) the log of the average wage of all outside offers received by the worker in the four months following t and (2) the log of the worker's belief reported in month t about the average wage of outside offers that he/she could potentially receive in the next four months.

average wage of all offers (both full-time and part-time) received by each worker (employed, unemployed, and out of the labor force) in each period, i.e., $A_{i,t,t+4} \equiv \frac{\sum_{j=1}^{\min\{n_{i,t,t+4},3\}} q_{i,j,t,t+4}}{\min\{n_{i,t,t+4},3\}}$, and show that (1) $A_{i,t,t+4}$ increases almost one-to-one with the belief $b_{i,t,t+4}$ reported by the worker at the beginning of the period and (2) both the median and average difference between $A_{i,t,t+4}$ and $b_{i,t,t+4}$ are close to zero.[7] Both evidence suggests that workers have accurate beliefs about the average wage of job offers they could potentially receive. They also find that the absolute difference between $A_{i,t,t+4}$ and $b_{i,t,t+4}$ is significantly different between the employed and those out of the labor force.

Instead of the difference by employment status, we are interested in whether the potential biases in worker beliefs are correlated with worker productivity, which is important not only by itself but also for our analysis of wage setting, wage dispersion, and sorting in the next two sections. To this end, we (1) limit the impact of employment status and work hours by focusing on full-time employed workers and their full-time offers, (2) limit the impact of offer selection by focusing on workers with at most three offers, and (3) estimate the bias by different worker characteristics, e.g., gender, education, experience, number of offers, job search status, expected chance of receiving an outside offer, and wages. The evidence suggests that, among full-time employed workers, the bias in the belief about the average wage of outside offers is neither significantly different from zero on average nor systematically correlated with worker characteristics.

3.1 Belief Anchoring

Jäger et al. (2024) asked a sample of employed workers in Germany the following question: "*Imagine that you were forced to leave your current job and that you had three months to find a job at another employer in the same occupation. Do you think that you would find a job that would offer you a higher overall pay, the same pay, or a lower pay?*" Workers who did not choose the "*same pay*" option were further asked about the size of the expected increase/decrease. Using each worker's wage and responses to these questions about the outside option, Jäger et al. (2024) construct the expected wage change as the difference between the belief about the outside option wage (own wage plus expected wage increase/decrease) in logs and the log of the own wage. To check the accuracy of the expected wage change, they use administrative data to construct an objective measure of the wage change for each worker using coworker moves to another employer that involve an intervening unemployment

[7] Conlon et al. (2021) convert annual salaries reported in the data into hourly wages by assuming (1) full-time jobs and offers entail 52 weeks a year and 40 hours a week, (2) part-time jobs and offers entail 52 weeks a year and 20 hours a week, (3) the belief $b_{i,t,t+4}$ reported by a part-time worker is about part-time offers, and (4) the belief $b_{i,t,t+4}$ reported by other workers (full-time employed, unemployed, and out of the labor force) is about full-time offers.

spell. Assuming the objective measure from coworker moves is a good predictor of the actual wage change faced by each worker, which is shown to be the case using past moves, regressing the expected wage change against the objective measure from coworkers should result in a coefficient of 1 if worker beliefs are accurate. Empirically, they obtain a coefficient of 0.107 (standard error: 0.040). Because the estimate is much closer to 0 than 1, and the expected wage change is zero for 40.6% of their sample, Jäger et al. (2024) conclude that, instead of accurate beliefs, workers anchor their beliefs about outside options on their current wages.

There is much less anchoring in our data. The belief $b_{i,t,t+4}$ is equal to the wage $y_{i,t}$ for 11.8% of the 10,406 worker-month observations where the full-time employed worker reported both measures in month t, regardless of whether they were surveyed again in month $t+4$. The value is even smaller at 10.8% among the 768 workers who also reported the wage of at least one outside offer $q_{i,1,t,t+4}$ in month $t+4$ and thus are used in Figure 4, which is slightly smaller than the 796 observations used in Table 2 because some workers there did not report a valid wage $y_{i,t}$.

Figure 5 plots the expected wage premium of an outside offer measured by $\log b_{i,t,t+4} - \log y_{i,t}$ against the observed wage premium measured by $LA_{i,t,t+4} - \log y_{i,t}$. This is our analogue of the headline figure 3 in Jäger et al. (2024), where the expected wage change is plotted against the actual wage change observed from coworker moves. For clarity, we use only observations where both wage premiums are between -1 and 1, which account for 740 of the 768 observations mentioned above. The slope of the fitted line is 0.550 (standard error: 0.027), much larger than the estimate of 0.107 in Jäger et al. (2024).[8]

The slope should be 1 if worker beliefs are accurate, and both the expected wage premium on the vertical axis and the actual wage premium on the horizontal axis are accurate measures of the true wage premium. While our estimate is still far from 1, the difference could be explained by measurement error and the randomness in the actual wage offers received by each worker. For evidence, instead of the wage premium, Figure 6 plots the expected wage of outside offers measured by $\log b_{i,t,t+4}$ against the observed wage (of outside offers) measured by $LA_{i,t,t+4}$. The figure uses all 796 observations described in Table 2. Without the wage $\log y_{i,t}$ and its measurement error, the slope of the fitted line is now much larger at 0.793 (standard error: 0.020).[9] The remaining gap between this estimate and 1 could be explained by measurement errors in $\log b_{i,t,t+4}$ and $LA_{i,t,t+4}$, and the randomness in each wage offer $q_{i,j,t,t+4}$ and thus the log average $LA_{i,t,t+4}$.

To see this, assume worker beliefs are accurate, and both $\log b_{i,t,t+4}$ and $LA_{i,t,t+4}$ are noisy

[8]The slope is 0.541 (standard error: 0.028) if the observed wage premium is measured by $\log q_{i,1,t,t+4} - \log y_{i,t}$, and it is 0.553 (standard error: 0.027) if the observed wage premium is measured by $AL_{i,t,t+4} - \log y_{i,t}$.

[9]The slope is 0.776 (standard error: 0.021) if the observed wage is measured by $\log q_{i,1,t,t+4}$ and 0.791 (standard error: 0.021) if the observed wage is measured by $AL_{i,t,t+4}$.

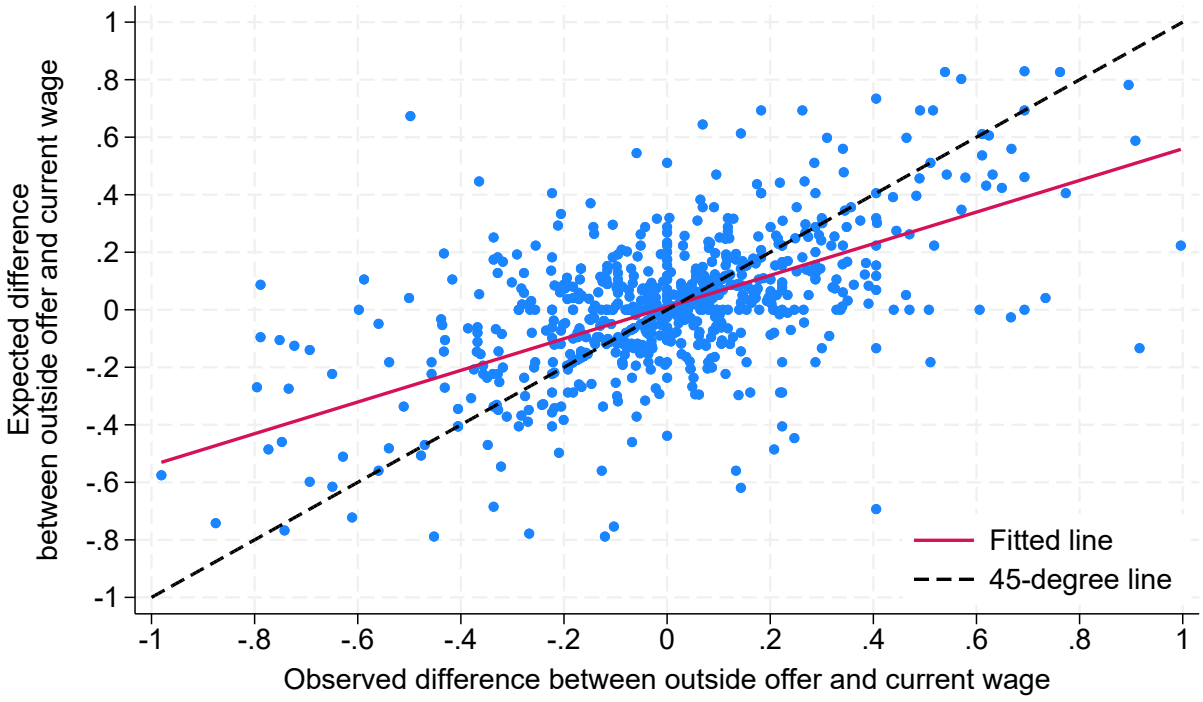


Figure 5: Expected vs Observed Wage Premium of An Outside Offer

Notes: This graph uses workers who were employed full-time in month t and received 1-3 full-time offers with valid wages in the four months between t and $t+4$. The vertical axis plots the difference $\log b_{i,t,t+4} - \log y_{i,t}$, where $b_{i,t,t+4}$ the the worker's belief reported in month t about the average wage of outside offers that he/she could potentially receive in the four months between t and $t+4$, and $y_{i,t}$ is the worker's wage in month t. The horizontal axis plots the difference $LA_{i,t,t+4} - \log y_{i,t}$, where $LA_{i,t,t+4}$ is the log of the average wage of all outside offers received by the worker in the four months between t and $t+4$. The slope of the fitted line is 0.550 (standard error: 0.027).

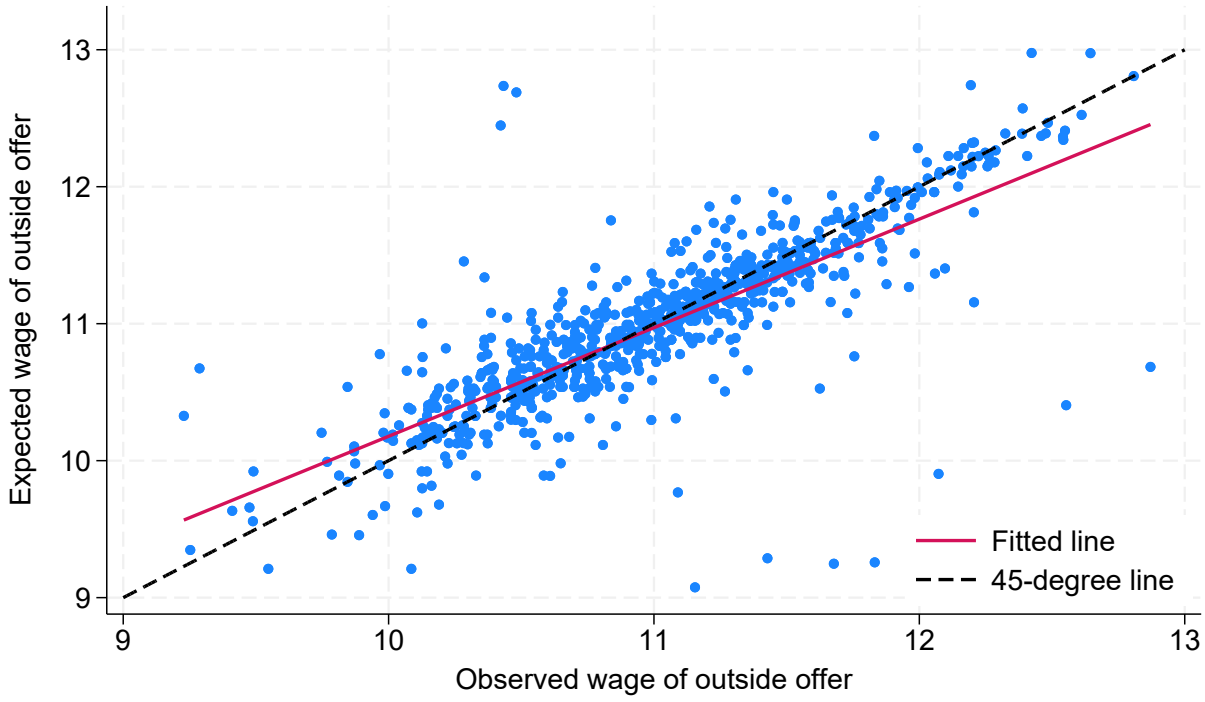


Figure 6: Expected vs Observed Wage of An Outside Offer

Notes: This graph uses workers who were employed full-time in month t and received 1-3 full-time offers with valid wages in the four months between t and $t+4$. The vertical axis plots $\log b_{i,t,t+4}$, where $b_{i,t,t+4}$ the the worker's belief reported in month t about the average wage of outside offers that he/she could potentially receive in the four months between t and $t+4$. The horizontal axis plots $LA_{i,t,t+4}$, the log of the average wage of all outside offers received by the worker in the four months between t and $t+4$. The slope of the fitted line is 0.793 (standard error: 0.020).

measures of the true average wage of outside offers $W_{i,t,t+4}$ as follows

$$\log b_{i,t,t+4} = W_{i,t,t+4} + \iota^1_{i,t,t+4}$$
$$LA_{i,t,t+4} = W_{i,t,t+4} + d_{i,t,t+4} + \iota^2_{i,t,t+4}$$

where $\iota^1_{i,t,t+4}$ and $\iota^2_{i,t,t+4}$ are measurement error, and $d_{i,t,t+4}$ is the deviation between the actual average wage of outside offers received by the worker ($LA_{i,t,t+4} - \iota^2_{i,t,t+4}$) and the true average wage $W_{i,t,t+4}$. The deviation $d_{i,t,t+4}$ arises from the randomness of wage offers $q_{i,j,t,t+4}$ received by the worker, and it reflects the fact that the average of a finite number of draws from a non-degenerate distribution may not equal the mean of the distribution.

Assuming the deviation $d_{i,t,t+4}$ and the two measurement errors $\iota^1_{i,t,t+4}$ and $\iota^2_{i,t,t+4}$ are independent of each other and the true wage $W_{i,t,t+4}$, and let σ^2_x be the variance of $x_{i,t,t+4}$ for any $x \in \{W, d, \iota^2\}$, the slope of the fitted line in Figure 6 is given by

$$\frac{\sigma^2_W}{\sigma^2_W + \sigma^2_d + \sigma^2_{\iota^2}} = \frac{\sigma^2_W}{\sigma^2_W + \sigma^2_d}\left(1 - \frac{\sigma^2_{\iota^2}}{\sigma^2_W + \sigma^2_d + \sigma^2_{\iota^2}}\right)$$

which is smaller than 1 if either $\sigma^2_d > 0$ or $\sigma^2_{\iota^2} > 0$ or both.

Using annual earnings reported by workers in the Current Population Survey (CPS) and the corresponding true earnings from administrative data, Bound and Krueger (1991) find that measurement error accounts for 18% of the earnings reported by men and 8% for women. This suggests that the term in the parentheses is between 0.82 and 0.92. Assume it is 0.87 ($\frac{0.82+0.92}{2}$), the slope of 0.793 for the fitted line in Figure 6 is consistent with accurate beliefs if $\frac{\sigma^2_W}{\sigma^2_W+\sigma^2_d} = \frac{0.793}{0.87} = 0.911$. This is possible given the large range of wage offer dispersion (the dispersion of wage offers for the same worker) estimated in the literature (Hall and Mueller, 2018; Conlon et al., 2021; Guo, 2024a), which determines σ^2_d and thus $\frac{\sigma^2_W}{\sigma^2_W+\sigma^2_d}$.

We focus on the three measures ($B^{LA}_{i,t,t+4}$, $B^{AL}_{i,t,t+4}$, $B^{q_1}_{i,t,t+4}$) discussed above because they are more robust to measurement error and the randomness in the actual wage offers received by each worker. Nevertheless, the slope coefficients in Figures 5 and 6 confirm that, different from Jäger et al. (2024) where worker beliefs are anchored on their current wages, workers in our data have accurate beliefs about the average wage of their outside offers.

One explanation for the difference between the two sets of results is the framing effect, where the answer to a question depends on how the question is asked (Stalans, 2012). In particular, the presence of the "*same pay*" option in the question by Jäger et al. (2024) could induce some workers to choose it even though their beliefs about outside options are not exactly the same as their current wages. Jäger et al. (2024) provide some evidence consistent with this effect. Using an online survey with different wordings, they find that the fraction of workers whose beliefs about outside options are the same as their current wages

is significantly smaller (around 30% as opposed to over 40%) when the "*same pay*" option is removed and workers are asked to report the levels of their outside options directly.[10] Unfortunately, this online survey cannot be matched with their main data, so that Jäger et al. (2024) cannot use it to evaluate the accuracy of worker beliefs elicited from alternative wordings. Instead, they emphasize that the overall distributions of worker beliefs "are similar across many alternative question wordings", which suggests that their results are robust to the framing effect, at least qualitatively.

Because the fraction of German workers surveyed by Jäger et al. (2024) whose beliefs about outside options are the same as their current wages is still around 30% in the absence of the "*same pay*" option, compared to about 11% of U.S. workers in SCE whose beliefs about the average wage of outside offers are the same as their current wages, it is likely that the framing effect accounts for some but not all of the differences between the finding by Jäger et al. (2024) that employed workers in Germany anchor their beliefs about outside options on their current wages and our finding that employed workers in the U.S. have accurate beliefs about the average wage of their outside offers. Accounting for the remaining differences is beyond the scope of this paper but an important direction for future work. Another direction, as mentioned in the Introduction, is to explore the interaction between worker beliefs and worker mobility and its variation across countries.

4 Current Wages and Outside Offers

Assume the wage $y_{i,t}$ of worker i in month t is given by

$$\log y_{i,t} = h_{i,t} + w_{i,t} + \varepsilon_{i,t} \tag{1}$$

where $h_{i,t}$ is the worker's (log) productivity or human capital, $w_{i,t}$ is the (log) wage rate per unit of human capital paid by the worker's employer, and $\varepsilon_{i,t}$ is measurement error. This section studies the impact of a worker's current wage rate $w_{i,t}$ on the wages of outside offers that the worker could potentially receive.

Specifically, assume that the wage of the jth (randomly ordered) outside offer received by worker i in the four months between t and $t+4$ is given by

$$\log q_{i,j,t,t+4} = h_{i,t} + \kappa w_{i,t} + v_{i,j,t,t+4} + \eta_{i,j,t,t+4} \tag{2}$$

[10] See the first panel of Figure E.2 in their Online Appendix E for details. The framing effect due to the "*same pay*" option is more salient in the second panel regarding worker beliefs about coworker wage changes, i.e., wage changes for "the typical employee with work experience that switches from your current employer to another employer". With the "*same pay*" option, the fraction of workers who expect a zero wage change for their coworkers is about 60%, compared to essentially 0% when the "*same pay*" option is removed and workers are asked to report their beliefs directly.

where κ measures the degree to which the wage rate of the outside offer depends on the worker's current wage rate $w_{i,t}$, $v_{i,j,t,t+4}$ is the idiosyncratic component of the wage rate of the jth outside offer, and $\eta_{i,j,t,t+4}$ is measurement error. We normalize the mean of $v_{i,j,t,t+4}$ to zero. We could do this without loss of generality, because neither human capital $h_{i,t}$ nor the wage rates ($w_{i,t}$ and $v_{i,j,t,t+4}$) have a natural unit. If the mean of $v_{i,j,t,t+4}$ is $m_{i,t} \neq 0$, we could redefine $v_{i,j,t,t+4}$, $h_{i,t}$ and $w_{i,t}$ as $v_{i,j,t,t+4} - m_{i,t}$, $h_{i,t} + \frac{m_{i,t}}{1-\kappa}$ and $w_{i,t} - \frac{m_{i,t}}{1-\kappa}$, respectively, and everything would go through.

Equation (2) assumes there is no change in the worker's human capital or wage rate between the survey in month t and the time when the jth offer is received, which is likely to be the case given the short period of time in consideration.[11] If there are significant changes in the worker's human capital during the short period of time, and this change is reflected on the outside wage offer $q_{i,j,t,t+4}$ but not anticipated by the worker in month t and thus not reflected on the belief $b_{i,t,t+4}$, the worker's belief would then be biased, which is inconsistent with the evidence documented in the previous section. The same argument suggests that there is no significant change in the worker's wage rate as long as κ, the parameter of interest in this section, is not zero.

Conceptually, κ could be viewed as a measure of the worker's bargaining power. We would have $\kappa > 0$ if firms take the outside option of each worker into consideration, as in sequential-auction models of the labor market pioneered by Postel-Vinay and Robin (2002). Otherwise, we would have $\kappa = 0$ if wages are posted by firms without considering the outside option of each worker, e.g., Burdett and Mortensen (1998). Distinguishing the two cases is important because they have different implications for wage dynamics and labor market transitions, as discussed in more detail in the Introduction.

Ideally, if we could observe two outside offers for the same worker (or two workers with the same productivity h) but different wage rates w, we would be able to identify κ by relating the difference between the two outside offers to the difference between the two wage rates w. In the absence of this ideal case, we approximate it using the data we have.

Suppose worker i received 1-3 outside offers (in terms of both the total number of offers $n_{i,t,t+4}$ and the number of full-time offers $n^f_{i,t,t+4}$ defined in the previous section) both in the four months before t and in the four months after t. We take the difference of equation (2) between the two periods using the first (randomly ordered) offer received in each period

$$\Delta \log q_{i,1,t,t+4} = \Delta h_{i,t} + \kappa \Delta w_{i,t} + \Delta v_{i,1,t,t+4} + \Delta \eta_{i,1,t,t+4} \tag{3}$$

where Δ is a difference operator such that $\Delta \log q_{i,1,t,t+4} = \log q_{i,1,t,t+4} - \log q_{i,1,t-4,t}$, $\Delta h_{i,t} =$

[11] The survey does not ask about the exact time each offer was received, but we know it must be less than 4 months after the survey in month t.

$h_{i,t} - h_{i,t-4}$, and so on. Similarly, we take the difference of equation (1) and obtain

$$\Delta \log y_{i,t} = \Delta h_{i,t} + \Delta w_{i,t} + \Delta \varepsilon_{i,t} \tag{4}$$

We estimate the following equation

$$\Delta \log q_{i,1,t,t+4} = \alpha_q + \beta_q \Delta \log y_{i,t} + \xi^q_{i,t} \tag{5}$$

where $\xi^q_{i,t}$ is the error term. The equation is estimated using workers who (1) were employed full-time in both month $t-4$ and month t and (2) received 1-3 outside offers both in the four months before t and in the four months after t.

Table 3 shows that there are 102 such workers. Among them, β_q is estimated to be 0.084 (standard error: 0.167). The second column replaces the dependent variable $\Delta \log q_{i,1,t,t+4}$ with $\Delta AL_{i,t,t+4} = AL_{i,t,t+4} - AL_{i,t-4,t}$, the difference in the average log wage offer between the two periods. β_q is now estimated to be 0.050 (standard error: 0.121). The last column replaces the dependent variable with $\Delta LA_{i,t,t+4} = LA_{i,t,t+4} - LA_{i,t-4,t}$, the difference in the log average wage offer between the two periods. β_q is now estimated to be 0.038 (standard error: 0.106). None of the three estimates of β_q is significantly different from zero.

Table 3: Current Wages and Outside Offers

	$\Delta \log q_{i,1,t,t+4}$	$\Delta AL_{i,t,t+4}$	$\Delta LA_{i,t,t+4}$
$\Delta \log y_{i,t}$	0.084	0.050	0.038
	(0.167)	(0.121)	(0.106)
Constant	-0.050	-0.025	-0.020
	(0.030)	(0.022)	(0.019)
Observations	102	102	102

Notes: This table uses workers who (1) were employed full-time in both month $t-4$ and month t and (2) received 1-3 full-time outside offers both in the four months between $t-4$ and t and in the four months between t and $t+4$. $\Delta \log y_{i,t}$ is the difference between $\log y_{i,t}$ and $\log y_{i,t-4}$, the worker's log wage in month t and $t-4$, respectively. $\Delta \log q_{i,1,t,t+4}$ is the difference between $\log q_{i,1,t,t+4}$ and $\log q_{i,1,t-4,t}$, the log wage of the first (randomly ordered) offer received in the four months after and before t, respectively. $\Delta AL_{i,t,t+4}$ is the difference between $AL_{i,t,t+4}$ and $AL_{i,t-4,t}$, the average of the log wage offers received in the four months after and before t, respectively. $\Delta LA_{i,t,t+4}$ is the difference between $LA_{i,t,t+4}$ and $LA_{i,t-4,t}$, the log of the average wage of offers received in the four months after and before t, respectively.

Figure 7 plots each of the three dependent variables ($\Delta \log q_{i,1,t,t+4}$, $\Delta AL_{i,t,t+4}$, and $\Delta LA_{i,t,t+4}$) against the independent variable $\Delta \log y_{i,t}$. It confirms the estimates in Table 3 that changes in outside wage offers for an employed worker are not significantly correlated with changes in the worker's current wage.

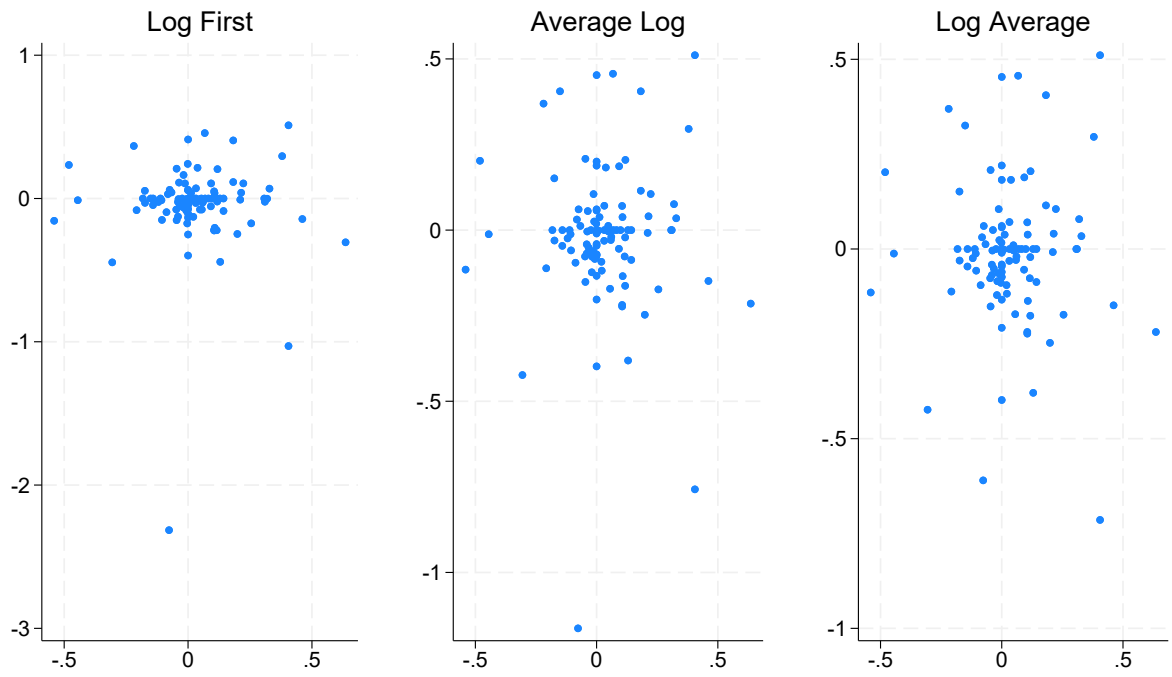


Figure 7: Changes in Outside Offers vs Changes in Current Wages

Notes: This graph uses workers who (1) were employed full-time in both month $t-4$ and month t and (2) received 1-3 full-time outside offers both in the four months between $t-4$ and t and in the four months between t and $t+4$. The horizon axis in each panel is the difference between $\log y_{i,t}$ and $\log y_{i,t-4}$, the worker's log wage in month t and $t-4$, respectively. The vertical axis for the left panel is the difference between $\log q_{i,1,t,t+4}$ and $\log q_{i,1,t-4,t}$, the log wage of the first (randomly ordered) offer received in the four months after and before t, respectively. The vertical axis for the middle panel is the difference between $AL_{i,t,t+4}$ and $AL_{i,t-4,t}$, the average of the log wage offers received in the four months after and before t, respectively. The vertical axis for the right panel is the difference between $LA_{i,t,t+4}$ and $LA_{i,t-4,t}$, the log of the average wage of offers received in the four months after and before t, respectively.

To learn about κ from the estimates of β_q, we assume that the terms on the right-hand sides of equations (3) and (4) are independent of each other. Let $\sigma_{\Delta x}$ be the standard deviation of $\Delta x_{i,t}$ for each $x \in \{h, w, \varepsilon\}$. We have

$$\beta_q = \frac{\sigma^2_{\Delta h} + \kappa\sigma^2_{\Delta w}}{\sigma^2_{\Delta h} + \sigma^2_{\Delta w} + \sigma^2_{\Delta \varepsilon}} = \frac{\sigma^2_{\Delta h} + \kappa\sigma^2_{\Delta w}}{\sigma^2_{\Delta h} + \sigma^2_{\Delta w}} \lambda_{\Delta y} \tag{6}$$

where $\lambda_{\Delta y} = \frac{\sigma^2_{\Delta h} + \sigma^2_{\Delta w}}{\sigma^2_{\Delta h} + \sigma^2_{\Delta w} + \sigma^2_{\Delta \varepsilon}}$ is the contribution of true wage growth $(\Delta h_{i,t} + \Delta w_{i,t})$ to the variation in reported wage growth $\Delta \log y_{i,t}$.

Using the difference between two annual earnings reported by the same worker in CPS and the difference between corresponding true earnings from administrative data, Bound and Krueger (1991) find that $\lambda_{\Delta y}$ is 0.648 for men and 0.814 for women. Assuming $\lambda_{\Delta y}$ is significantly positive in our data, equation (6) implies

$$\kappa \leq \frac{\beta_q}{\lambda_{\Delta y}} \tag{7}$$

where the equality holds if and only if $\sigma_{\Delta h} = 0$, which is the ideal case mentioned above where worker productivity h is constant between two outside offers. In general, the gap between κ and the upper bound is increasing in $\sigma_{\Delta h}$.

Given the upper bound, and assuming $\kappa \geq 0$, the insignificant estimates of β_q reported in Table 3 suggest that κ is not significantly different from zero, i.e., outside wage offers for an employed worker does not dependent significantly on the worker's current wage conditional on productivity.

4.1 Belief Updating

One concern for the estimates in Table 3 is that the number of observations is relatively small. For more evidence, we look at the response of worker beliefs $b_{i,t,t+4}$ to wage growth $\Delta \log y_{i,t}$. Following the specification of actual wage offers in equation (2), we assume

$$\log b_{i,t,t+4} = h_{i,t} + \kappa w_{i,t} + \mu_{i,t,t+4} + \zeta_{i,t,t+4} \tag{8}$$

where $\mu_{i,t,t+4}$ is the worker's belief in month t about $v_{i,j,t,t+4}$, the idiosyncratic component of the wage rate for each outside offer that the worker could potentially receive between t and $t+4$, and $\zeta_{i,t,t+4}$ is measurement error. Evidence from the previous section suggests that the difference between $v_{i,j,t,t+4}$ and $\mu_{i,t,t+4}$ is not significantly different from zero or systematically correlated with worker characteristics. Applying the difference operator Δ to equation (8),

we have

$$\Delta \log b_{i,t,t+4} = \Delta h_{i,t} + \kappa \Delta w_{i,t} + \Delta \mu_{i,t,t+4} + \Delta \zeta_{i,t,t+4} \tag{9}$$

We estimate the following analogue of equation (5)

$$\Delta \log b_{i,t,t+4} = \alpha_b + \beta_b \Delta \log y_{i,t} + \xi^b_{i,t} \tag{10}$$

where $\xi^b_{i,t}$ is the error term. If the terms on the right-hand sides of equations (4) and (9) are independent of each other, β_b will be equal to the β_q given by equation (6), and κ will be bounded from above by $\frac{\beta_b}{\lambda_{\Delta y}}$.

Figure 8 plots the 4-month changes in worker beliefs $\Delta \log b_{i,t,t+4}$ against the changes in their wages $\Delta \log y_{i,t}$ in the same 4 months. There are 4413 workers for whom both measures are available. The two measures are significantly positively correlated with each other, although the relationship is far from perfect. The slope of the fitted line is given by β_b, which is estimated to be 0.250 (standard error: 0.022) in the first column of Table 4. Assuming $\lambda_{\Delta y} = 0.731$, which is the average of the two values from Bound and Krueger (1991) mentioned above, equation (7) implies $\kappa \leq \frac{0.250}{0.731} = 0.342$.

Table 4: Current Wages and Beliefs about Outside Offers

				Bachelor's degree		Years of experience	
	All	Women	Men	No	Yes	≤ 15	> 15
$\Delta \log y_{i,t}$	0.250	0.207	0.295	0.274	0.223	0.351	0.220
	(0.022)	(0.033)	(0.029)	(0.037)	(0.027)	(0.040)	(0.026)
Constant	0.004	0.003	0.004	0.006	0.002	0.000	0.007
	(0.004)	(0.008)	(0.005)	(0.008)	(0.005)	(0.007)	(0.006)
Observations	4413	1825	2585	1311	3092	1757	2641

Notes: This table uses workers who were employed full-time and reported their beliefs about the average wage of outside offers in both month $t-4$ and month t. The dependent variable is the change in a worker's log belief about the average wage of outside offers in the four months. The independent variable $\Delta \log y_{i,t}$ is the change in the worker's log wage in the four months. Groupings by some characteristics may not add up to the total number of observations due to missing values.

The remaining columns of Table 4 report the estimates of β_b by gender, education and experience. All estimates are significantly different from zero, and they range from 0.207 for women to 0.351 for workers with at most 15 years of experience.[12] Using $\kappa \leq \frac{\beta_b}{\lambda_{\Delta y}}$ and

[12]Table A1 in the Online Appendix shows that estimates of β are robust when we control for each worker's gender, education, experience, number of offers received in the four months between $t-4$ and t, and fixed effect for the survey month $t-4$.



Figure 8: Changes in Beliefs about Outside Offers vs Changes in Current Wages

Notes: This graph uses workers who were employed full-time and reported their beliefs about the average wage of outside offers in both month $t-4$ and month t. The horizontal axis is the change in a worker's log wage in the four months. The vertical axis is the change in the worker's log belief about the average wage of outside offers in the four months. The slope of the fitted line is 0.250 (standard error: 0.022).

the two values of $\lambda_{\Delta y}$ from Bound and Krueger (1991) mentioned above, the estimates of β_b imply that $\kappa \leq \frac{0.207}{0.814} = 0.254$ for women, $\kappa \leq \frac{0.295}{0.648} = 0.455$ for men, and $\kappa \leq \frac{0.351}{0.713} = 0.492$ for workers with at most 15 years of experience.

In addition to the sample, the differences between β_q estimated in Table 3 and β_b in Table 4 could also arise if $\Delta\mu_{i,t,t+4}$ is positively correlated with $\Delta w_{i,t}$, which could happen if workers update their beliefs about outside offers in the same direction as their wages. While the evidence in the previous section suggests that worker beliefs are accurate on average, it does not mean that the belief is accurate for each individual worker, or that there is no updating over time. Using data from the same survey, Conlon et al. (2021) find that workers tend to update their beliefs $b_{i,t,t+4}$ over time in response to the wages of their outside offers $q_{i,j,t,t+4}$. It is thus likely that workers may also update their beliefs $b_{i,t,t+4}$ over time in response to changes in their current wages $y_{i,t}$.

Let $\sigma_{\Delta w,\Delta\mu}$ be the covariance between $\Delta w_{i,t}$ and $\Delta\mu_{i,t,t+4}$, which is not zero if workers do update their beliefs $\mu_{i,t,t+4}$ over time in response to changes in their wage rates $w_{i,t}$. We have

$$\beta_b = \frac{\sigma^2_{\Delta h} + \kappa\sigma^2_{\Delta w} + \sigma_{\Delta w,\Delta\mu}}{\sigma^2_{\Delta h} + \sigma^2_{\Delta w} + \sigma^2_{\Delta\varepsilon}} = \frac{\sigma^2_{\Delta h} + \kappa\sigma^2_{\Delta w} + \sigma_{\Delta w,\Delta\mu}}{\sigma^2_{\Delta h} + \sigma^2_{\Delta w}}\lambda_{\Delta y} \tag{11}$$

which is larger than the β_q in equation (6) as long as $\sigma_{\Delta w,\Delta\mu} > 0$. In this case, the gap between κ and the upper bound $\frac{\beta_b}{\lambda_{\Delta y}}$ is larger than the gap between κ and $\frac{\beta_q}{\lambda_{\Delta y}}$, because the latter reflects only human capital accumulation $\sigma^2_{\Delta h}$ while the former reflects both human capital accumulation and belief updating $\sigma_{\Delta w,\Delta\mu}$.

Assuming the three terms in the numerator of equation (11) are equal to each other, κ is equal to one third of the upper bound $\frac{\beta_b}{\lambda_{\Delta y}}$. Given estimates of the upper bound mentioned above, κ is equal to 0.114 for all workers, 0.085 for women, 0.152 for men, and 0.164 for workers with at most 15 years of experience, which is also the largest value among the seven groups estimated in Table 4.

Overall, the estimates in this section suggest that κ is not very large. A small κ implies that, when a worker moves directly from one job to another, the wage of the old job has little effect on the starting wage of the new job. This is consistent with Di Addario et al. (2023), who find that, inconsistent with sequential-auction models of the labor market, where a worker is hired from is relatively inconsequential for their wages in comparison to where they are currently employed. They obtain this result using the starting wage of each worker at each firm, which by definition is the wage of an accepted offer. If the wage of an outside offer that an employed worker ends up accepting does not respond significantly to the worker's current wage, it is likely that wages of rejected offers, which are the majority as documented above, do not respond significantly to the worker's current wage either.

5 Wage Dispersion and Sorting

Assuming the measurement error $\varepsilon_{i,t}$ in equation (1) is independent of the worker's productivity $h_{i,t}$ and wage rate $w_{i,t}$, we have

$$\sigma_y^2 = \sigma_h^2 + 2\sigma_{h,w} + \sigma_w^2 + \sigma_\varepsilon^2 \tag{12}$$

where σ_y, σ_h, and σ_w are the standard deviations of $\log y_{i,t}$, $h_{i,t}$, and $w_{i,t}$ across employed workers, respectively, and $\sigma_{h,w}$ is the covariance between $h_{i,t}$ and $w_{i,t}$, a measure of sorting between worker productivity and the job-specific wage rate. The goal of this section is to decompose wage dispersion among employed workers (measured by σ_y^2) into the four terms on the right-hand side of equation (12).

5.1 Identification

Assume for now that the three wages ($\log y_{i,t}$, $\log q_{i,1,t,t+4}$, $\log b_{i,t,t+4}$) are observed for each employed worker, and they are given by equations (1), (2) and (8), respectively. Assume further that the measurement errors ($\varepsilon_{i,t}$, $\eta_{i,j,t,t+4}$, $\zeta_{i,t,t+4}$) for the three wages are independent of each other and all other terms on the right-hand sides of the three equations, so are the idiosyncratic component $v_{i,j,t,t+4}$ for $\log q_{i,j,t,t+4}$ in equation (2) and the corresponding $\mu_{i,t,t+4}$ for $\log b_{i,t,t+4}$ in equation (8). As mentioned in the previous section, the mean of $v_{i,j,t,t+4}$ is normalized to zero. An unbiased belief is not necessary for the analysis in this section. As long as $\mu_{i,t,t+4}$ is independent of all other terms on the right-hand sides of the three equations, everything would go through even if its mean is not zero.

Let $\sigma_{b,q}$ be the covariance between $\log b_{i,t,t+4}$ and $\log q_{i,1,t,t+4}$, $\sigma_{b,y}$ be the covariance between $\log b_{i,t,t+4}$ and $\log y_{i,t}$, and $\sigma_{q,y}$ be the covariance between $\log q_{i,1,t,t+4}$ and $\log y_{i,t}$. The three equations (1), (2) and (8) imply

$$\sigma_{b,q} = \sigma_h^2 + 2\kappa\sigma_{h,w} + \kappa^2\sigma_w^2 \tag{13}$$

$$\sigma_{x,y} = \sigma_h^2 + (1+\kappa)\sigma_{h,w} + \kappa\sigma_w^2 \text{ for } x \in \{b, q\} \tag{14}$$

The left-hand sides of equations (12)-(14) can be estimated directly from data on ($\log y_{i,t}$, $\log q_{i,1,t,t+4}$, $\log b_{i,t,t+4}$). There are five parameters on the right-hand sides of the three equations: (σ_h, $\sigma_{h,w}$, σ_w, σ_ε, κ). The five parameters cannot be identified jointly from the three equations without further assumptions.[13] To proceed, we focus on the first three parameters

[13]Technically, there are four equations, because equation (14) holds for both $x \in \{b, q\}$. However, because the two cases share the same right-hand side, they are only useful as one restriction. Theoretically, we could use either of the two cases. In practice, we use both of them. While we use $\log q_{i,1,t,t+4}$ as a measure of the actual wage offer in this section, because it is easy conceptually for some discussions, the results reported

(σ_h, $\sigma_{h,w}$, σ_w) and estimate them using different values of the other two parameters.

Informed by estimates in the previous section, we consider three values for κ: 0, 0.1, and 0.2. Moreover, we assume $\sigma_\varepsilon^2 = (1 - \lambda_y)\sigma_y^2$, where λ_y is the contribution of true wages $h_{i,t} + w_{i,t}$ to the variation in reported wages $\log y_{i,t}$. As mentioned previously in section 3, Bound and Krueger (1991) find that measurement error accounts for 18% of the variation in wages reported by men and 8% for women, which implies λ_y is 0.82 for men and 0.92 for women. Based on these values, Hall and Mueller (2018) assume that measurement error accounts for 13% ($\frac{18\%+8\%}{2}$) of the variation of wages in their data, which implies $\lambda_y = 0.87$. We consider all three values for λ_y: 0.82, 0.87, and 0.92.

Given a set of values for (κ, λ_y),

- Worker heterogeneity σ_h^2 is mainly identified from $\sigma_{b,q}$, the covariance between the two measures of outside wage offers ($\log b_{i,t,t+4}$ and $\log q_{i,1,t,t+4}$). In particular, equation (13) implies that $\sigma_h^2 = \sigma_{b,q}$ when $\kappa = 0$, so that worker heterogeneity σ_h^2 is independent of the contribution of measurement error (1-λ_y). Intuitively, $\kappa = 0$ implies that outside wage offers for an employed worker are determined by and thus informative of only the worker's productivity $h_{i,t}$, so that worker heterogeneity σ_h^2 could be identified solely from the covariance between the two noisy measures of outside wage offers.

- Sorting $\sigma_{h,w}$ is mainly identified from $\sigma_{x,y} - \sigma_{b,q}$ for $x \in \{b, q\}$. In particular, equations (13) and (14) imply that $\sigma_{h,w} = \sigma_{x,y} - \sigma_{b,q}$ when $\kappa = 0$, so that sorting $\sigma_{h,w}$ is also independent of the contribution of measurement error (1-λ_y). Intuitively, when outside wage offers for an employed worker are determined only by the worker's productivity $h_{i,t}$, while the current wage is determined by both productivity $h_{i,t}$ and the wage rate $w_{i,t}$ of the current job, the covariance $\sigma_{x,y}$ between a measure of outside wage offers and the current wage is equal to the sum of worker heterogeneity σ_h^2 and sorting $\sigma_{h,w}$. With worker heterogeneity σ_h^2 identified from $\sigma_{b,q}$, sorting $\sigma_{h,w}$ could be identified directly from $\sigma_{x,y} - \sigma_{b,q}$.

- Given worker heterogeneity σ_h^2 and sorting $\sigma_{h,w}$, the dispersion in wage rates σ_w^2 is mainly identified from the dispersion in reported wages σ_y^2. With $\sigma_\varepsilon^2 = (1 - \lambda_y)\sigma_y^2$, equation (12) implies $\sigma_w^2 = \lambda_y\sigma_y^2 - \sigma_h^2 - \sigma_{h,w}$, so that σ_w^2 depends on λ_y even when σ_h^2 and $\sigma_{h,w}$ do not. Essentially, σ_w^2 is treated as a residual of σ_y^2.

Assume that worker productivity $h_{i,t}$ is given by

$$h_{i,t} = \theta_i + x_{i,t}\pi + \omega_{i,t} \tag{15}$$

below are similar when the other two measures ($AL_{i,t,t+4}$ and $LA_{i,t,t+4}$) are used. This is not surprising given the high correlations among the three measures mentioned above.

where θ_i is a fixed effect, $x_{i,t}$ is a vector of time-varying characteristics such as experience, and $\omega_{i,t}$ is the error term. Existing studies of wage dispersion and sorting tend to focus only on the first two terms using panel data on accepted wages, where sorting is measured by the correlation between the worker fixed effect θ_i and some fixed effects for firms. We account for the entirety of $h_{i,t}$ non-parametrically using outside wage offers, majority of which are rejected and thus not observed in traditional data, and measure sorting using the correlation between worker productivity $h_{i,t}$ and the job-specific wage rate $w_{i,t}$ across employed workers. With both accepted and rejected offers, we do not need to model either the offer arrival process or the offer acceptance decision and the resulting mobility patterns, which are critical for existing studies using only accepted wages.

On the other hand, our identification strategy for worker heterogeneity σ_h^2 is similar to Hall and Mueller (2018), who use the covariance between each unemployed worker's reservation wage and wage offer to identify the heterogeneity in productivity across unemployed workers, and Mueller et al. (2021), who use the covariance between each job seeker's expected chance of finding a job and actual job finding to identify heterogeneity in true job-finding rates across job seekers. As mentioned above, we also follow Hall and Mueller (2018) by setting the contribution of measurement error $(1 - \lambda_y)$ to external estimates from Bound and Krueger (1991). The difference is that we focus on employed workers while both Hall and Mueller (2018) and Mueller et al. (2021) focus on unemployed workers.

5.2 Samples and Weights

In practice, the three wages ($\log y_{i,t}$, $\log q_{i,1,t,t+4}$, $\log b_{i,t,t+4}$) are not available for all employed workers. In particular, $\log q_{i,1,t,t+4}$ is only available for workers who received 1-3 outside offers in the four months between t and $t+4$ and reported the wage of at least one full-time offer, and $\log b_{i,t,t+4}$ is only available for workers who reported a strictly positive chance of receiving at least one outside offer in the four months between t and $t+4$. Among the 7861 full-time employed workers described in Table 1, 7508 reported their wages $y_{i,t}$, and 768 reported all three wages ($y_{i,t}$, $b_{i,t,t+4}$, $q_{i,1,t,t+4}$). From now on, we will refer to the 7508 observations with wages $y_{i,t}$ as the full sample, and the 768 observations with all three wages as the estimation sample.

Because Table 1 combines two consecutive Labor Market Surveys of the same worker into one observation, our full sample excludes full-time employed workers who participated in only one Labor Market Survey. This could be a problem if workers who participated in two consecutive surveys are systematically different from those who participated in only one survey. Empirically, this does not seem to be the case, at least in terms of wages $\log y_{i,t}$, which is the variable of interest in this section. Among the 14,128 worker-month observations where

the full-time employed worker reported a valid wage $y_{i,t}$, the mean and standard deviation of $\log y_{i,t}$ are 10.953 and 0.613, respectively. The corresponding values in our full sample are 10.950 and 0.615. A Kolmogorov-Smirnov test of the equality of the distribution of $\log y_{i,t}$ between the two samples results in a $p-$value of 0.601, which suggests that our full sample is representative of the larger sample. We focus on the full sample for consistency with the previous sections, and because it is easy to interpret some of the statistics reported below, such as the fraction of observations with an outside offer $q_{i,1,t,t+4}$.

We could use the estimation sample directly if it is representative of the full sample. This would be the case if the probability that an observation in the full sample belongs to the estimation sample is independent of the worker's human capital $h_{i,t}$ and wage rate $w_{i,t}$. To check this, we divide workers in the full sample into ten deciles based on their wages $y_{i,t}$, and calculate for each decile the fraction of workers who reported both their beliefs $b_{i,t,t+4}$ and actual wage offers $q_{i,1,t,t+4}$ and thus are in the estimation sample. The blue solid line in Figure 9 shows that the fraction declines from 12.5% among workers in the second wage decile to 8.7% in the fifth decile and then rises to 10.6% in the eighth decile. The pattern is mainly driven by the fraction of workers with an actual wage offer $q_{i,1,t,t+4}$ (red dashed line), as opposed to the fraction of workers who reported their beliefs $b_{i,t,t+4}$ (black dashed line). Overall, the fraction of workers in the estimation sample (blue solid line) is in a relatively narrow range of 9-11% in most cases. Regressing an dummy indicating whether a worker is in the estimation sample against the worker's wage $\log y_{i,t}$ results in an estimate of -0.006 (standard error: 0.006).[14] This suggests that, while not perfectly representative of the full sample, the estimation sample is not very selective either.

Figure 10 plots the distribution of wages $\log y_{i,t}$ in both the full sample (blue solid line) and the estimation sample (red dashed line). Despite some differences in the middle, the two distributions are very close to each other. The first column of Table 5 shows that the mean and variance of $\log y_{i,t}$ in the full sample are 10.950 and 0.378, respectively, and they are 10.929 and 0.382 in the estimation sample. The remaining columns of Table 5 show that the two statistics are also not very different between the full sample and the estimation sample when we look at sub-samples by gender, education and experience. This again suggests that the estimation sample is not very different from the full sample.

To make the estimation sample even more representative of the full sample, we adjust the weight of each observation in the former using Kennan (2006). Let F be the cumulative distribution function of a continuous distribution, and $\hat{F}$ be a cumulative distribution function across a set of fixed points $\{x_n\}_{n=1}^N$. Assuming the value of each point x_n is in-

[14]The estimate is 0.031 (standard error: 0.008) when the dependent variable is an dummy indicating whether the belief $b_{i,t,t+4}$ is observed, and it is -0.012 (standard error: 0.006) when the dependent variable is an dummy indicating whether the actual wage offer $q_{i,1,t,t+4}$ is observed. The latter suggests that high-wage workers are less likely to receive an offer, consistent with the evidence by Faberman et al. (2022).

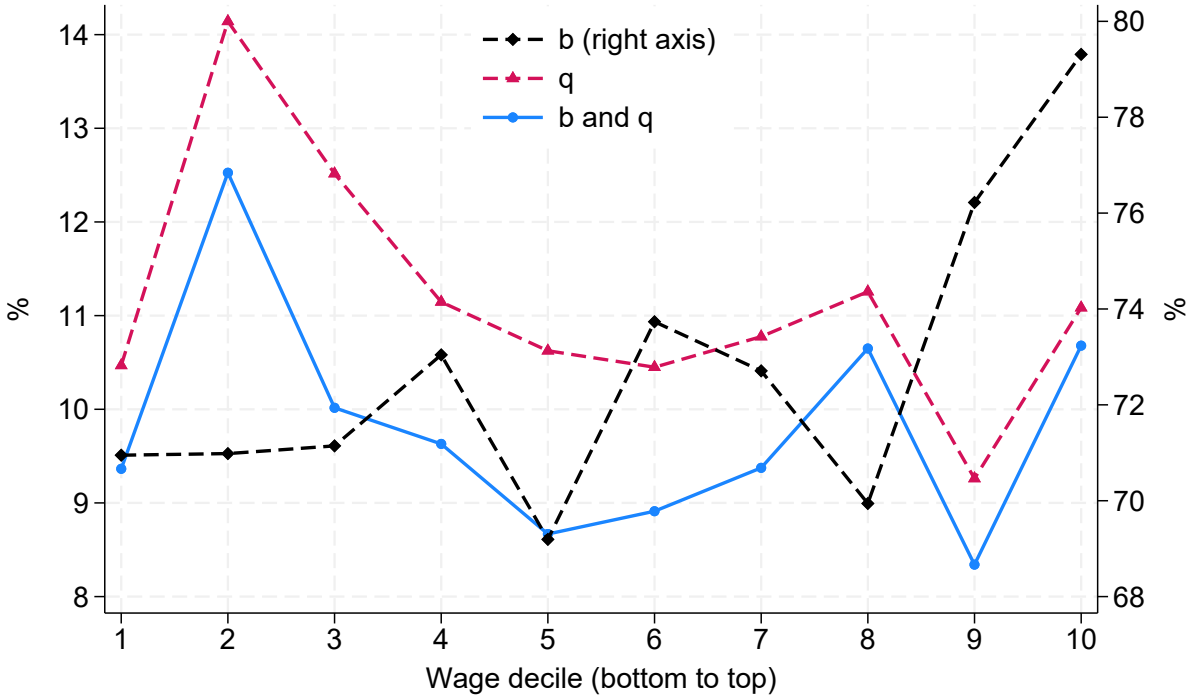


Figure 9: Percentages of Workers with Beliefs and Actual Wage Offers

Notes: This graph uses workers who were employed full-time in month t and reported the number of job offers received in the four months between t and $t+4$. The horizontal axis divides them into ten groups based on their wages reported in month t. For each group, the vertical axis plots the percentages of workers with non-missing $b_{i,t,t+4}$, $q_{i,1,t,t+4}$, and both, where $b_{i,t,t+4}$ is the belief of worker i in month t about the average wage of outside offers that the worker could potentially receive in the four months between t and $t+4$, and $q_{i,1,t,t+4}$ is the wage of the first (randomly ordered) full-time offer received by the worker in the four months between t and $t+4$. $b_{i,t,t+4}$ is only available if the worker reported in month t that he/she had a strictly positive chance of receiving at least one outside offer in the four months between t and $t+4$. $q_{i,1,t,t+4}$ is only available if the worker received 1-3 job offers in the four months between t and $t+4$ and reported the wage of at least one full-time offer.

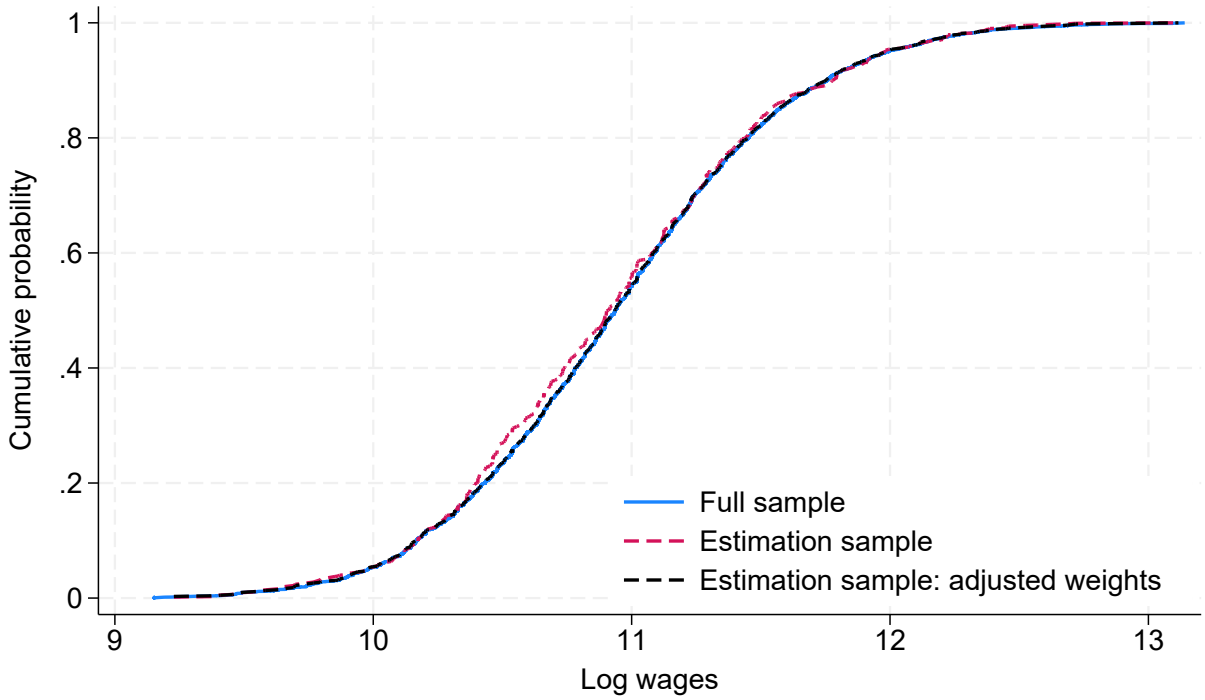


Figure 10: The Distribution of Wages in the Full and Estimation Samples

Notes: This graph plots the cumulative distribution of the log wage $\log y_{i,t}$ of each worker i in month t. Full sample: workers who were employed full-time in month t with annual salary $y_{i,t}$ and reported the number of job offers received in the four months between t and $t+4$. Estimation sample: workers in the full sample who reported both $b_{i,t,t+4}$ and $q_{i,1,t,t+4}$, where $b_{i,t,t+4}$ is the belief of worker i in month t about the average wage of outside offers that the worker could potentially receive in the four months between t and $t+4$, and $q_{i,1,t,t+4}$ is the wage of the first (randomly ordered) full-time offer received by the worker in the four months between t and $t+4$. $b_{i,t,t+4}$ is only available if the worker reported in month t that he/she had a strictly positive chance of receiving at least one outside offer in the four months between t and $t+4$. $q_{i,1,t,t+4}$ is only available if the worker received 1-3 job offers in the four months between t and $t+4$ and reported the wage of at least one full-time offer. The black dashed line adjusts the weight of each observation in the estimation sample to make it representative of the full sample.

Table 5: Statistics of the Full and Estimation Samples

				Bachelor's degree		Experience	
	All	Female	Male	No	Yes	≤ 15	> 15
				Full sample			
Observations	7508	3412	4093	2563	4933	2661	4824
Mean $\log y_{i,t}$	10.950	10.767	11.093	10.714	11.231	10.951	10.951
σ_y^2	0.378	0.322	0.374	0.296	0.328	0.358	0.387
				Estimation sample			
Observations	768	343	425	245	520	336	428
Mean $\log y_{i,t}$	10.929	10.756	11.057	10.689	11.177	10.924	10.943
σ_y^2	0.382	0.367	0.355	0.298	0.339	0.348	0.399
				Estimation sample: adjusted weights			
Mean $\log y_{i,t}$	10.949	10.765	11.091	10.710	11.230	10.949	10.950
σ_y^2	0.378	0.324	0.369	0.291	0.330	0.357	0.387
$\sigma_{b,q}$	0.281	0.256	0.269	0.235	0.230	0.239	0.321
$\sigma_{b,y}$	0.291	0.266	0.285	0.229	0.243	0.263	0.328
$\sigma_{q,y}$	0.297	0.266	0.285	0.244	0.239	0.266	0.318

Notes: Full sample: workers who were employed full-time in month t with annual salary $y_{i,t}$ and reported the number of job offers received in the four months between t and $t+4$. Estimation sample: workers in the full sample who reported both $b_{i,t,t+4}$ and $q_{i,1,t,t+4}$, where $b_{i,t,t+4}$ is the belief of worker i in month t about the average wage of outside offers that the worker could potentially receive in the four months between t and $t+4$, and $q_{i,1,t,t+4}$ is the wage of the first (randomly ordered) full-time offer received by the worker in the four months between t and $t+4$. $b_{i,t,t+4}$ is only available if the worker reported in month t that he/she had a strictly positive chance of receiving at least one outside offer in the four months between t and $t+4$. $q_{i,1,t,t+4}$ is only available if the worker received 1-3 job offers in the four months between t and $t+4$ and reported the wage of at least one full-time offer. The bottom panel adjusts the weight of each observation in the estimation sample to make it representative of the full sample. σ_y^2 is the variance of $\log y_{i,t}$. $\sigma_{b,q}$ is the covariance between $\log b_{i,t,t+4}$ and $\log q_{i,1,t,t+4}$. $\sigma_{b,y}$ is the covariance between $\log b_{i,t,t+4}$ and $\log y_{i,t}$. $\sigma_{q,y}$ is the covariance between $\log q_{i,1,t,t+4}$ and $\log y_{i,t}$.

creasing in its location n, Kennan (2006) shows that the best discrete approximation $\hat{F}$ to the continuous distribution F using the fixed support points $\{x_n\}_{n=1}^N$ is given by $\hat{F}(x_N) = 1$ and $\hat{F}(x_n) = \frac{1}{2}[F(x_n) + F(x_{n+1})]$ for $1 \leq n \leq N-1$. In our case, F is the distribution of $\log y_{i,t}$ in the full sample, and x_n is the nth lowest value of $\log y_{i,t}$ in the estimation sample. Given the distribution F and the set of points $\{x_n\}_{n=1}^N$, it is easy to obtain the best discrete approximation $\hat{F}$ and, in turn, the adjusted weight for each observation in the estimation sample.

The black dashed line in Figure 10 plots the distribution of $\log y_{i,t}$ in the estimation sample using the adjusted weights. It is essentially identical to the distribution of $\log y_{i,t}$ in the full sample given by the blue solid line. Moreover, the bottom panel in Table 5 reports the mean and variance of $\log y_{i,t}$ in the estimation sample using the adjusted weights. Relative to the estimates using the original survey weights reported in the middle panel, the new estimates are much closer to the corresponding estimates from the full sample reported in the top panel in all but one case (the variance of $\log y_{i,t}$ among workers without a bachelor's degree, where the number of observations is relatively small). In most cases, estimates from the adjusted weights are almost identical to the corresponding estimates from the full sample.

The last three rows of Table 5 report the estimates of $(\sigma_{b,q},\, \sigma_{b,y},\, \sigma_{q,y})$ from the estimation sample using the adjusted weights. As mentioned above, these three moments are used to identify the three parameters $(\sigma_h,\, \sigma_{h,w},\, \sigma_w)$. Consistent with equation (14), estimates of $\sigma_{b,y}$ and $\sigma_{q,y}$ are not significantly different from each other in any of the seven columns. Moreover, neither $\sigma_{b,y}$ nor $\sigma_{q,y}$ is significantly different from $\sigma_{b,q}$ in most cases. This suggests that sorting $\sigma_{h,w}$, which is mainly identified from $\sigma_{b,y} - \sigma_{b,q}$ and $\sigma_{q,y} - \sigma_{b,q}$, does not contribute significantly to wage dispersion σ_y^2. On the other hand, $\sigma_{b,q}$ is large relative to σ_y^2 in most cases. This suggests that a large fraction of wage dispersion σ_y^2 is due to worker heterogeneity σ_h^2, as described in more details below.

5.3 Estimation

Instead of the three parameters $(\sigma_h,\, \sigma_{h,w},\, \sigma_w)$, we estimate $s_h = \frac{\sigma_h^2}{\sigma_y^2}$ and $s_{h,w} = \frac{2\sigma_{h,w}}{\sigma_y^2}$, which are the contributions of worker heterogeneity and sorting to observed wage dispersion, respectively. With the contribution of measurement error given by $\frac{\sigma_\varepsilon^2}{\sigma_y^2} = 1 - \lambda_y$, equation (12) implies that the contribution of σ_w^2 is $\frac{\sigma_w^2}{\sigma_y^2} = \lambda_y - s_h - s_{h,w}$. Relative to $(\sigma_h,\, \sigma_{h,w},\, \sigma_w)$, these normalized parameters are easier to interpret and compare across different samples/studies.

Let m_y, m_b, and m_q be the mean of $\log y_{i,t}$, $\log b_{i,t,t+4}$, and $\log q_{i,1,t,t+4}$, respectively. We

construct seven expressions for each observation in the estimation sample:

$$\log y_{i,t} - m_y;\ \ \log b_{i,t,t+4} - m_b;\ \ \log q_{i,1,t,t+4} - m_q$$
$$\tfrac{N}{N-1}(\log y_{i,t} - m_y)^2 - \sigma_y^2;$$
$$\tfrac{N}{N-1}(\log b_{i,t,t+4} - m_b)(\log q_{i,1,t,t+4} - m_q) - \sigma_y^2[s_h + \kappa s_{h,w} + \kappa^2(\lambda_y - s_h - s_{h,w})];$$
$$\tfrac{N}{N-1}(\log b_{i,t,t+4} - m_b)(\log y_{i,t} - m_y) - \sigma_y^2[s_h + 0.5(1+\kappa)s_{h,w} + \kappa(\lambda_y - s_h - s_{h,w})];$$
$$\tfrac{N}{N-1}(\log q_{i,1,t,t+4} - m_q)(\log y_{i,t} - m_y) - \sigma_y^2[s_h + 0.5(1+\kappa)s_{h,w} + \kappa(\lambda_y - s_h - s_{h,w})]$$

where N is the number of observations in the estimation sample, and the last four expressions are based on equations (12) to (14). Setting the expected values of the seven expressions to zero gives us seven moment conditions. Given a set of values for (κ, λ_y) discussed above, we estimate the six parameters (m_y, m_b, m_q, σ_y^2, s_h, $s_{h,w}$) jointly using the Generalized Method of Moments, with the adjusted weight applied to each observation.

5.4 Results

The top panel in Table 6 reports the estimates of s_h and $s_{h,w}$ using all observations in the estimation sample. When $\kappa = 0$, the first row shows that $s_h = 0.743$ (standard error: 0.036) and $s_{h,w} = 0.066$ (standard error: 0.047), which are independent of λ_y as discussed above. Across the nine sets of values for (κ, λ_y), s_h is estimated to be in a narrow range of 0.727 to 0.743, with a median of 0.737. This suggests that about 73.7% of the variation in reported wages across employed workers is due to differences in their productivity.

The contribution of worker heterogeneity to the variation in true wages is $\frac{s_h}{\lambda_y}$, which ranges from 79.8% when $\kappa = 0.2$ and $\lambda_y = 0.92$ to 90.6% when $\kappa = 0$ and $\lambda_y = 0.82$, with a median of 84.7% when $\kappa = 0.1$ and $\lambda_y = 0.87$. This suggests that about 84.7% of the variation in true wages across employed workers is due to differences in their productivity. As mentioned in the Introduction, this is larger than most estimates in the literature, probably because our approach accounts for the entirety of worker productivity $h_{i,t}$, as opposed to the sum of the fixed effect θ_i and some observable time-varying characteristics $x_{i,t}$.

On the other hand, $s_{h,w}$ is estimated to be between 0.014 and 0.076, with a median of 0.066. The contribution of sorting to the variation in true wages across employed workers is $\frac{s_{h,w}}{\lambda_y}$, which ranges from 1.5% when $\kappa = 0.2$ and $\lambda_y = 0.92$ to 9.3% when $\kappa = 0.2$ and $\lambda_y = 0.82$, with a median of 7.2% when $\kappa = 0$ and $\lambda_y = 0.92$. While always positive, none of the nine estimates of $s_{h,w}$ is significantly different from zero. This suggests that sorting between worker productivity and the job-specific wage rate does not contribute significantly to wage dispersion across employed workers. As mentioned above, this is mainly due to the relatively small values of $\sigma_{b,y} - \sigma_{b,q}$ and $\sigma_{q,y} - \sigma_{b,q}$ shown in Table 5.

The middle and bottom panels of Table 6 show that estimates of s_h and $s_{h,w}$ do not vary

Table 6: Wage Dispersion and Sorting

	$\lambda_y = 0.82$		$\lambda_y = 0.87$		$\lambda_y = 0.92$	
	s_h	$s_{h,w}$	s_h	$s_{h,w}$	s_h	$s_{h,w}$
			All			
$\kappa = 0$	0.743	0.066	0.743	0.066	0.743	0.066
	(0.036)	(0.047)	(0.036)	(0.047)	(0.036)	(0.047)
$\kappa = 0.1$	0.736	0.071	0.737	0.059	0.737	0.046
	(0.039)	(0.059)	(0.039)	(0.059)	(0.039)	(0.059)
$\kappa = 0.2$	0.727	0.076	0.731	0.045	0.734	0.014
	(0.044)	(0.077)	(0.044)	(0.077)	(0.044)	(0.077)
			Women			
$\kappa = 0$	0.790	0.062	0.790	0.062	0.790	0.062
	(0.059)	(0.071)	(0.059)	(0.071)	(0.059)	(0.071)
$\kappa = 0.1$	0.782	0.077	0.783	0.065	0.784	0.052
	(0.066)	(0.084)	(0.066)	(0.084)	(0.066)	(0.084)
$\kappa = 0.2$	0.772	0.098	0.775	0.066	0.778	0.035
	(0.076)	(0.103)	(0.076)	(0.103)	(0.076)	(0.103)
			Men			
$\kappa = 0$	0.728	0.087	0.728	0.087	0.728	0.087
	(0.053)	(0.067)	(0.053)	(0.067)	(0.053)	(0.067)
$\kappa = 0.1$	0.719	0.095	0.719	0.083	0.720	0.071
	(0.058)	(0.084)	(0.058)	(0.084)	(0.058)	(0.084)
$\kappa = 0.2$	0.707	0.106	0.710	0.074	0.713	0.043
	(0.065)	(0.108)	(0.065)	(0.108)	(0.065)	(0.108)

Notes: κ is the impact of an employed worker's current wage rate $w_{i,t}$ on the wage rates of outside offers. s_h, $s_{h,w}$, and $1 - \lambda_y$ are the contributions of worker productivity $h_{i,t}$, sorting between worker productivity $h_{i,t}$ and the wage rate $w_{i,t}$, and measurement error to the variation in reported wages $\log y_{i,t}$ across workers who were employed full-time in month t. The remaining part, $\lambda_y - s_h - s_{h,w}$, is due to the variation in the wage rate $w_{i,t}$. Robust standard errors are in the parentheses.

significantly by gender. If anything, s_h is slightly larger for women, while $s_{h,w}$ is slightly larger for men. The top two panels of Table 7 report the estimates of s_h and $s_{h,w}$ by education. The estimates suggest that s_h is larger for workers without a bachelor's degree, while $s_{h,w}$ is larger for workers with a bachelor's degree or more, although none of the estimates of $s_{h,w}$ is statistically different from zero. Finally, the bottom two panels of Table 7 report the estimates of s_h and $s_{h,w}$ by experience. The estimates suggest that s_h is larger for workers with more than 15 years of experience, while $s_{h,w}$ is larger for workers with at most 15 years of experience, for whom the estimates of $s_{h,w}$ is significantly positive when $\kappa = 0$.

6 Conclusion

Using novel data on job offers received by a sample of workers in a short period of time, including both accepted and rejected offers, this paper shows that employed workers in the U.S. have accurate beliefs about the average wage of outside offers they could potentially receive, and actual wage offers received by an employed worker are not strongly correlated with the worker's current wage conditional on productivity, which suggests that workers have little bargaining power and wages are largely set by firms without considering the outside option of each worker individually.

Using each worker's belief and actual wage offers as two noisy measures of the worker's productivity, we find that about 73.7% of the variation in reported wages and about 84.7% of the variation in true wages are due to heterogeneity in worker productivity, which are larger than most estimates in the literature. The estimates are larger for women, workers without a bachelor's degree, and workers with more 15 years of experience than men, workers with a bachelor's degree or more, and workers with no more than 15 years of experience, respectively. On the other hand, we find that sorting between worker productivity and the job-specific wage rate is positive but not significantly different from zero in most cases.

Table 7: Wage Dispersion and Sorting: by Education and Experience

	$\lambda_y = 0.82$		$\lambda_y = 0.87$		$\lambda_y = 0.92$	
	s_h	$s_{h,w}$	s_h	$s_{h,w}$	s_h	$s_{h,w}$
	Without a bachelor's degree					
$\kappa = 0$	0.798	-0.001	0.798	-0.001	0.798	-0.001
	(0.043)	(0.089)	(0.043)	(0.089)	(0.043)	(0.089)
$\kappa = 0.1$	0.798	-0.007	0.799	-0.019	0.800	-0.031
	(0.048)	(0.116)	(0.048)	(0.116)	(0.048)	(0.116)
$\kappa = 0.2$	0.800	-0.015	0.803	-0.047	0.806	-0.078
	(0.057)	(0.157)	(0.057)	(0.157)	(0.057)	(0.157)
	With a bachelor's degree					
$\kappa = 0$	0.694	0.072	0.694	0.072	0.694	0.072
	(0.055)	(0.072)	(0.055)	(0.072)	(0.055)	(0.072)
$\kappa = 0.1$	0.687	0.067	0.687	0.054	0.688	0.042
	(0.060)	(0.091)	(0.060)	(0.091)	(0.060)	(0.091)
$\kappa = 0.2$	0.679	0.056	0.682	0.025	0.686	-0.006
	(0.067)	(0.119)	(0.067)	(0.119)	(0.067)	(0.119)
	Years of experience ≤ 15					
$\kappa = 0$	0.672	0.140	0.672	0.140	0.672	0.140
	(0.049)	(0.062)	(0.049)	(0.062)	(0.049)	(0.062)
$\kappa = 0.1$	0.657	0.154	0.658	0.141	0.658	0.129
	(0.052)	(0.081)	(0.052)	(0.081)	(0.052)	(0.081)
$\kappa = 0.2$	0.638	0.170	0.641	0.139	0.644	0.108
	(0.057)	(0.108)	(0.057)	(0.108)	(0.057)	(0.108)
	Years of experience > 15					
$\kappa = 0$	0.839	-0.004	0.839	-0.004	0.839	-0.004
	(0.044)	(0.059)	(0.044)	(0.059)	(0.044)	(0.059)
$\kappa = 0.1$	0.839	-0.001	0.840	-0.013	0.841	-0.025
	(0.048)	(0.073)	(0.048)	(0.073)	(0.048)	(0.073)
$\kappa = 0.2$	0.839	0.004	0.842	-0.027	0.845	-0.058
	(0.055)	(0.095)	(0.055)	(0.095)	(0.055)	(0.095)

Notes: κ is the impact of an employed worker's current wage rate $w_{i,t}$ on the wage rates of outside offers. s_h, $s_{h,w}$, and $1 - \lambda_y$ are the contributions of worker productivity $h_{i,t}$, sorting between worker productivity $h_{i,t}$ and the wage rate $w_{i,t}$, and measurement error to the variation in reported wages $\log y_{i,t}$ across workers who were employed full-time in month t. The remaining part, $\lambda_y - s_h - s_{h,w}$, is due to the variation in the wage rate $w_{i,t}$. Robust standard errors are in the parentheses.

References

Abowd, John M, Francis Kramarz, and David N Margolis, "High wage workers and high wage firms," *Econometrica*, 1999, *67* (2), 251–333.

Armantier, Olivier, Giorgio Topa, Wilbert Van der Klaauw, and Basit Zafar, "An overview of the Survey of Consumer Expectations," *Economic Policy Review*, 2017, pp. 51–72.

Bagger, Jesper and Rasmus Lentz, "An empirical model of wage dispersion with sorting," *The Review of Economic Studies*, 2019, *86* (1), 153–190.

_ , François Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin, "Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics," *American Economic Review*, 2014, *104* (6), 1551–1596.

Biasi, Barbara and Heather Sarsons, "Flexible wages, bargaining, and the gender gap," *The Quarterly Journal of Economics*, 2022, *137* (1), 215–266.

Bonhomme, Stéphane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler, "How much should we trust estimates of firm effects and worker sorting?," *Journal of Labor Economics*, 2023, *41* (2), 291–322.

Bound, John and Alan B Krueger, "The extent of measurement error in longitudinal earnings data: Do two wrongs make a right?," *Journal of Labor Economics*, 1991, *9* (1), 1–24.

Bradley, Jake and Axel Gottfries, "A job ladder model with stochastic employment opportunities," *Quantitative Economics*, 2021, *12* (4), 1399–1430.

Burdett, Kenneth and Dale T Mortensen, "Wage differentials, employer size, and unemployment," *International Economic Review*, 1998, pp. 257–273.

Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin, "Wage bargaining with on-the-job search: Theory and evidence," *Econometrica*, 2006, *74* (2), 323–364.

Caldwell, Sydnee and Oren Danieli, "Outside Options in the Labour Market," *Review of Economic Studies*, 2024, p. rdae006.

_ , Ingrid Haegele, and Jörg Heining, "Bargaining and inequality in the labor market," Working Paper, National Bureau of Economic Research 2025.

Card, David, Ana Rute Cardoso, and Patrick Kline, "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women," *The Quarterly Journal of Economics*, 2016, *131* (2), 633–686.

_ , _ , Joerg Heining, and Patrick Kline, "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

Conlon, John J, Laura Pilossoph, Matthew Wiswall, and Basit Zafar, "Labor market search with imperfect information and learning," Working paper, National Bureau of Economic Research 2021.

Di Addario, Sabrina, Patrick Kline, Raffaele Saggio, and Mikkel Sølvsten, "It ain't where you're from, it's where you're at: hiring origins, firm heterogeneity, and wages," *Journal of Econometrics*, 2023, *233* (2), 340–374.

Doniger, Cynthia L, "Wage dispersion with heterogeneous wage contracts," *Review of Economic Dynamics*, 2023, *51*, 138–160.

Faberman, R Jason, Andreas I Mueller, Ayşegül Şahin, and Giorgio Topa, "Job search behavior among the employed and non-employed," *Econometrica*, 2022, *90* (4), 1743–1779.

Flinn, Christopher and Joseph Mullins, "Firms' Choices of Wage-Setting Protocols," 2021.

Fujita, Shigeru, Giuseppe Moscarini, and Fabien Postel-Vinay, "Measuring employer-to-employer reallocation," *American Economic Journal: Macroeconomics*, 2024, *16* (3), 1–51.

Gibbons, Robert and Lawrence Katz, "Does unmeasured ability explain inter-industry wage differentials?," *The Review of Economic Studies*, 1992, *59* (3), 515–535.

_ **, Lawrence F Katz, Thomas Lemieux, and Daniel Parent**, "Comparative advantage, learning, and sectoral wage determination," *Journal of Labor Economics*, 2005, *23* (4), 681–724.

Guo, Junjie, "Estimation of the Wage Offer Distribution Using Both Accepted and Rejected Offers," *Available at SSRN 4662267*, 2024.

_ , "The Response of Wages to Rejected Offers," *Available at SSRN 4533973*, 2024.

Hall, Robert E and Alan B Krueger, "Evidence on the incidence of wage posting, wage bargaining, and on-the-job search," *American Economic Journal: Macroeconomics*, 2012, *4* (4), 56–67.

_ **and Andreas I Mueller**, "Wage dispersion and search behavior: The importance of nonwage job values," *Journal of Political Economy*, 2018, *126* (4), 1594–1637.

Hazell, Jonathon, Christina Patterson, Heather Sarsons, and Bledi Taska, "National Wage Setting," Technical Report, National Bureau of Economic Research 2022.

Jäger, Simon, Benjamin Schoefer, Samuel Young, and Josef Zweimüller, "Wages and the Value of Nonemployment," *The Quarterly Journal of Economics*, 2020, *135* (4), 1905–1963.

_ **, Christopher Roth, Nina Roussille, and Benjamin Schoefer**, "Worker beliefs about outside options," *The Quarterly Journal of Economics*, 2024, p. qjae001.

Jolivet, Grégory, Fabien Postel-Vinay, and Jean-Marc Robin, "The empirical content of the job search model: labor mobility and wage distributions in Europe and the US," *European Economic Review*, 2006, *50* (4), 877–907.

Kennan, John, "A Note on Discrete Approximations of Continuous Distributions," Technical Report, University of Wisconsin-Madison and NBER 2006.

Krueger, Alan B and Andreas I Mueller, "A contribution to the empirics of reservation wages," *American Economic Journal: Economic Policy*, 2016, *8* (1), 142–179.

Lachowska, Marta, Alexandre Mas, Raffaele Saggio, and Stephen A Woodbury, "Wage posting or wage bargaining? a test using dual jobholders," *Journal of Labor Economics*, 2022, *40* (S1), S469–S493.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler, "Imperfect competition, compensating differentials, and rent sharing in the US labor market," *American Economic Review*, 2022, *112* (1), 169–212.

Le Barbanchon, Thomas, Roland Rathelot, and Alexandra Roulet, "Unemployment insurance and reservation wages: Evidence from administrative data," *Journal of Public Economics*, 2019, *171*, 1–17.

Lentz, Rasmus, Suphanit Piyapromdee, and Jean-Marc Robin, "The Anatomy of Sorting—Evidence From Danish Data," *Econometrica*, 2023, *91* (6), 2409–2455.

Lindenlaub, Ilse, "Sorting multidimensional types: Theory and application," *The Review of Economic Studies*, 2017, *84* (2), 718–789.

Lise, Jeremy and Fabien Postel-Vinay, "Multidimensional skills, sorting, and human capital accumulation," *American Economic Review*, 2020, *110* (8), 2328–2376.

Mueller, Andreas I and Johannes Spinnewijn, "Expectations data, labor market, and job search," *Handbook of Economic Expectations*, 2023, pp. 677–713.

_ , _ , and Giorgio Topa, "Job seekers' perceptions and employment prospects: Heterogeneity, duration dependence, and bias," *American Economic Review*, 2021, *111* (1), 324–363.

National Center for Education Statistics, "Employment and Unemployment Rates by Educational Attainment," in "The Condition of Education 2020" 2020.

Postel-Vinay, Fabien and Jean-Marc Robin, "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica*, 2002, *70* (6), 2295–2350.

Rabinovich, Stanislav, Brenda Samaniego de la Parra, and Ronald Wolthoff, "Wage Setting Protocols and Labor Market Conditions: Theory and Evidence," 2023.

Rosen, Sherwin, "The theory of equalizing differences," *Handbook of Labor Economics*, 1986, *1*, 641–692.

Roussille, Nina, "The role of the ask gap in gender pay inequality," *The Quarterly Journal of Economics*, 2024, p. qjae004.

_ **and Benjamin Scuderi**, "Bidding for talent: A test of conduct in a high-wage labor market," 2024.

Sorkin, Isaac, "Ranking firms using revealed preference," *The Quarterly Journal of Economics*, 2018, *133* (3), 1331–1393.

Spinnewijn, Johannes, "Unemployed but optimistic: Optimal insurance design with biased beliefs," *Journal of the European Economic Association*, 2015, *13* (1), 130–167.

Stalans, Loretta J, "Frames, framing effects, and survey responses," *Handbook of Survey Methodology for the Social Sciences*, 2012, pp. 75–90.

Taber, Christopher and Rune Vejlin, "Estimation of a roy/search/compensating differential model of the labor market," *Econometrica*, 2020, *88* (3), 1031–1069.

Wiswall, Matthew and Basit Zafar, "Preference for the workplace, investment in human capital, and gender," *The Quarterly Journal of Economics*, 2018, *133* (1), 457–507.

Online Appendix for

"Worker Beliefs About Outside Offers, Wage Setting, Wage Dispersion, and Sorting"

by Junjie Guo, University of Wisconsin - Madison

A Additional Figures and Tables

Table A1: Current Wages and Beliefs about Outside Offers: Robustness

				Bachelor's degree		Experience	
	All	Women	Men	No	Yes	≤ 15	> 15
$\Delta \log y_{i,t}$	0.259	0.238	0.293	0.293	0.229	0.360	0.220
	(0.022)	(0.033)	(0.030)	(0.038)	(0.028)	(0.040)	(0.027)
Female	-0.004			-0.015	0.005	-0.014	-0.001
	(0.009)			(0.016)	(0.011)	(0.014)	(0.012)
Bachelor's degree	-0.007	0.004	-0.012			0.011	-0.021
	(0.010)	(0.016)	(0.012)			(0.016)	(0.012)
Experience/10	0.001	-0.001	0.003	0.009	-0.006	-0.024	0.004
	(0.004)	(0.007)	(0.005)	(0.007)	(0.005)	(0.018)	(0.007)
One offer	-0.004	-0.043	0.023	-0.018	0.017	0.010	-0.018
	(0.015)	(0.023)	(0.019)	(0.026)	(0.018)	(0.021)	(0.020)
Two offers	-0.034	-0.066	-0.020	-0.075	0.019	0.010	-0.058
	(0.020)	(0.035)	(0.024)	(0.032)	(0.027)	(0.032)	(0.025)
Three or more offers	-0.007	-0.033	0.012	-0.025	0.015	0.022	-0.029
	(0.024)	(0.040)	(0.029)	(0.039)	(0.032)	(0.042)	(0.029)
Constant	0.010	0.013	0.004	-0.004	0.008	0.016	0.010
	(0.013)	(0.021)	(0.015)	(0.022)	(0.012)	(0.025)	(0.023)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4396	1818	2578	1311	3085	1755	2641

Notes: This table uses workers who were employed full-time and reported their beliefs about the average wage of outside offers in both month $t-4$ and month t. The dependent variable is the change in a worker's log belief about the average wage of outside offers in the four months. The key independent variable is $\Delta \log y_{i,t}$, the change in the worker's log wage in the four months. One offer, two offers, and three or more offers are three dummies that equal to one if the worker received the corresponding number of offers in the four months between $t-4$ and t. Time fixed effects are dummies for each worker's survey month $t-4$.

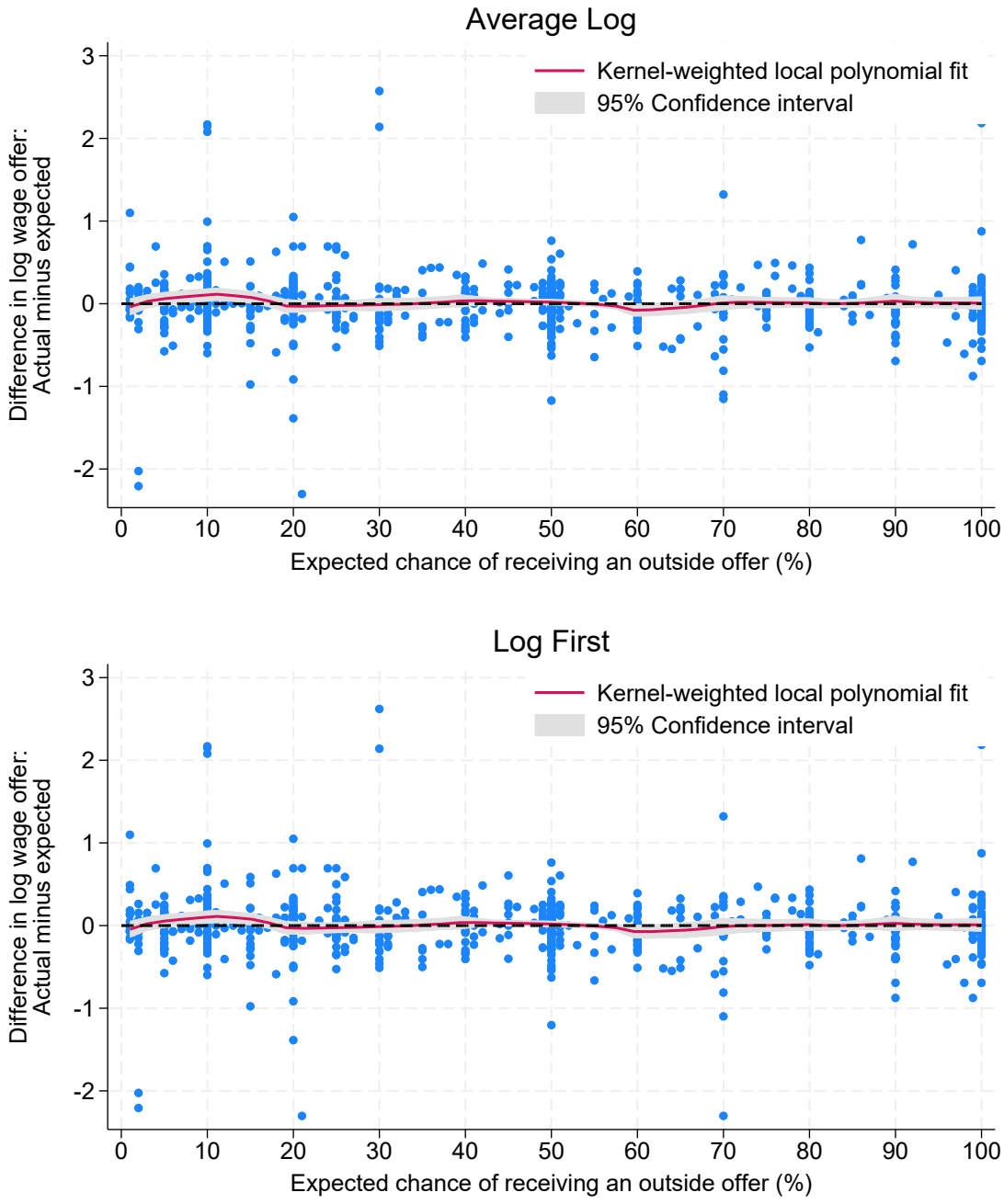


Figure A1: Alternative Measures of Biases in Worker Beliefs About Outside Offers vs Expected Chance of Receiving an Outside Offer

Notes: The two panels are similar to Figure 3 in the paper. The only difference is the variable on the vertical axis, which is $B^{AL}_{i,t,t+4}$ for the top panel, $B^{q_1}_{i,t,t+4}$ for the bottom panel, and $B^{LA}_{i,t,t+4}$ for Figure 3. See the main text and the note of Figure 3 for the definitions of the three variables and other details.

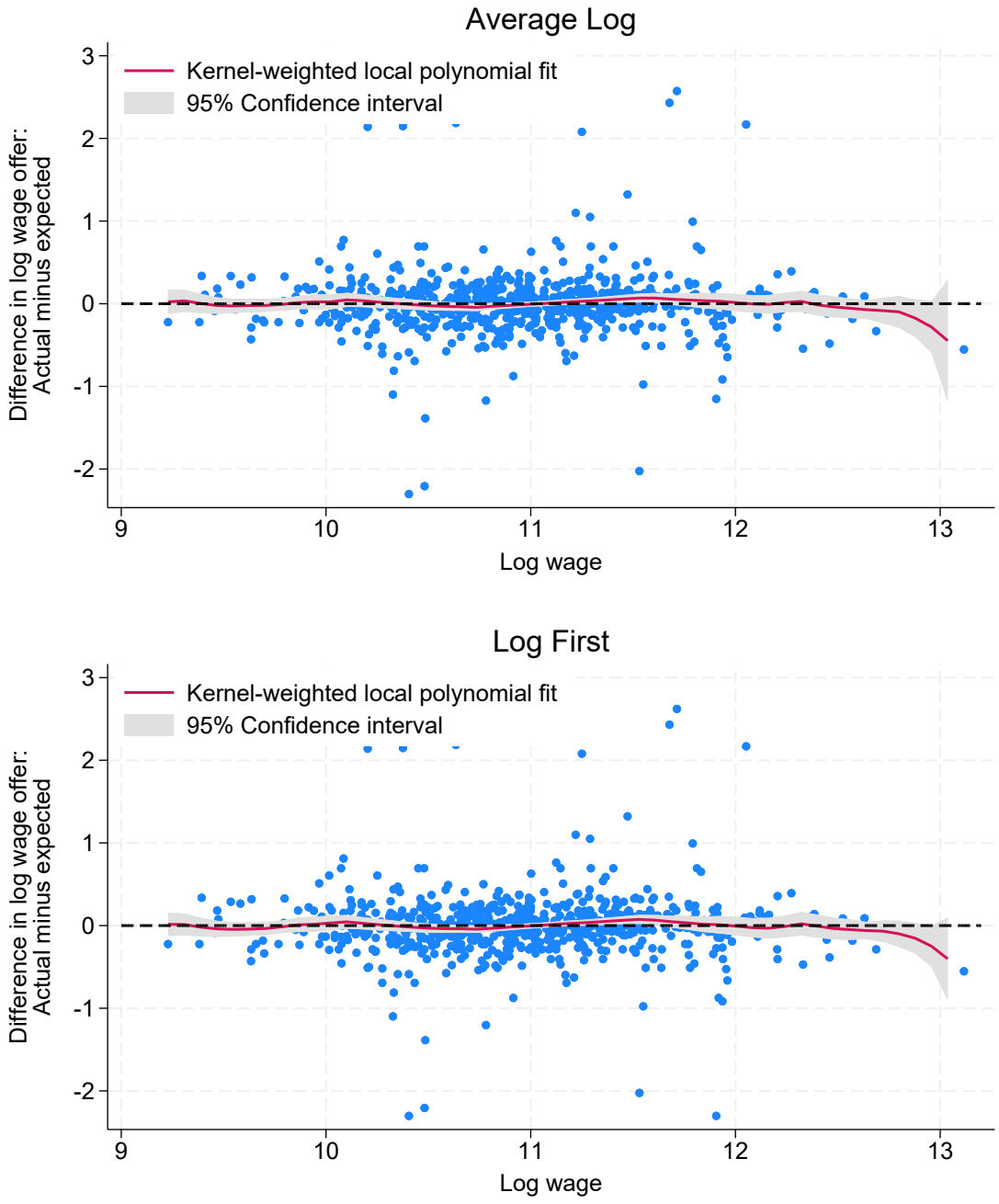


Figure A2: Alternative Measures of Biases in Worker Beliefs About Outside Offers vs Current Wages

Notes: The two panels are similar to Figure 4 in the paper. The only difference is the variable on the vertical axis, which is $B^{AL}_{i,t,t+4}$ for the top panel, $B^{q_1}_{i,t,t+4}$ for the bottom panel, and $B^{LA}_{i,t,t+4}$ for Figure 4. See the main text and the note of Figure 4 for the definitions of the three variables and other details.