Online Job Posts Contain Very Little Wage Information*

Honey Batra, Amanda Michaud, Simon Mongey

December 15, 2023

Abstract

We characterize the little wage information contained in online job posts. Wage information is rare: only 14% of posts contain any information. Of these, wage ranges are more common than point wages, and are wide on average, spanning 28% of the midpoint (e.g. $32,000-$42,000/yr). Posted wages are highly selected in low income occupations: 40% higher than wages of employed workers. High wage firms are more opaque, with more and wider ranges. We find zero correlation between wage information and local labor market tightness. We provide an example of bias in econometric inference that worsens as wage information falls.

Keywords: Labor, Wages, Search

JEL codes: J30, E24, D83

*Batra: Federal Reserve Bank of Minneapolis and University of Minnesota. Michaud: Federal Reserve Bank of Minneapolis. Mongey: Federal Reserve Bank of Minneapolis and NBER. Please contact amanda.michaud@mpls.frb.org.
For useful feedback on this project, we thank Benjamin Schoefer, Simon Quach, Sekyu Choi, Sydnee Caldwell, Marianna Kudlyak, Brenda Samaniego de la Parra, and especially Andrew Goodman-Bacon. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

Employers post tens of millions of jobs online each year.[1] Virtually all unemployed job seekers report performing job search online.[2] Wage information contained in job posts is obviously of use to these job seekers. The use of wages in job posts as a substitute for administrative data on workers' wages has also proliferated in economic research. Yet, a systematic study of the availability of wages in job posts in U.S. data has not, to date, been available. From these this paper answers the simple question *"How many job postings have wages?"*

Our main answer is "shockingly little." We study over 141 million posts scraped from over 45,000 different online sources between 2012 and 2017. The following summarizes our results:

1. **Job posts have little explicit wage information.** Only 13.5% have any wage information and just 5.8% state an exact, point wage. Private sector positions are less likely to have wage information: 10.7% with any and 4.9% with a point wage. The mean across postings disguises a skewed distribution and an even smaller median across firms. The postings-weighted median firm in the sample has any wage information in less than 1 percent of their posts. For example, 13 firms account for 4% of posts, but each have wage information in less than 0.5% of their posts. A good approximation is: there is no precise, direct, wage information.

2. **Ranges are more common than point wages.** The majority of posts with wages do not have a point wage. Instead, more than 57% feature a range. These ranges are wide on average: 28% of the mid-point. For example, $21.00-$27.50 per hour, or $50,000-$66,300 per year. This could indicate the wages of new hires are flexible, or at least designed to capture a broad set of applicants. Alternatively, there could be competitive incentives for firms to obfuscate wage information from competitors, applicants, and even current employees.

3. **Middle income jobs have more wage information overall.** Low and high wage occupations—measured using BLS OES data—have wage information in fewer than 10% of postings, while middle wage occupations have information in around 20% of postings.

4. **Ranges are more common in high wage jobs.** However, wages are presented more opaquely in high wage occupations. For this group, more than 70 percent of postings with wage information present that information as a range, while this is less than 30 percent for low wage

[1] The data we study contains the near universe of 141,816,864 job posts in the United States for the period 2012-2017, more than 20 million a year.

[2] Faberman et al. (2016) show this number rose from 25% to 76% from 2000 to 2011. A Pew Research Survey from 2015 finds online search was the most important job search activity for a third of workers Smith (2015).

occupations. Ranges are also slightly wider for high wage occupation postings. For empirical analysis, point wages in job postings cannot be treated as 'missing at random'.

5. **Wages in job postings systematically depart from survey data.** In low wage occupations, wages in job postings are 30 to 40 percent *higher* than observed in the BLS OES data, while they are 10 to 20 percent *lower* in high wage occupations. On the job wage growth would lead us to expect lower wages of new hires relative to incumbents. For low wage occupations, however, there is non-random and large positive selection into posting wage information.

6. **High wage employers present wage information less precisely.** Among all postings that do have wages—which is few—we find that within-occupation-year, firms that post higher wages are more likely to include a range instead of a point wage, and these ranges tend to be wider. These effects are significant, but small.

7. **Wage information is not correlated with local labor market conditions.** Consistent with Kuhn et al. (2021) we find that low unemployment labor markets are systematically tighter, but wage information is flat across markets.

8. **Lack of wage information is widespread in large private sector firms.** The most frequently posting firms are a broad, representative, cross-section of the U.S. economy: retail (Lowe's, Sears, Macy's, Dollar General), hospitality (Marriott), financial services (Wells Fargo, Accenture), health insurance (UnitedHealth Group, Blue Cross Blue Shield). Among the top 20, only the U.S. Air Force has wages in more than 1.5 percent of its postings. On the other hand, firms with high wage information is biased toward Federal and State Government departments, and firms in transportation and services that may pay piece rates.

Scarcity of wage information in job postings presents challenges to researchers looking to use job posting data as a stand-in for wage information of employed workers or even of new hires.[3] First, the above facts relate systematic bias in the availability of any wage data at all. Second, when wage data is available, it systematically departs from data on workers in jobs. As discussed, there is clear selection into posted wages in low wage occupations, where posted wages exceed the wages of employed workers. Third, inferring something about average wages at firms is challenging. With

[3] Our findings are specific to postings with wage information which contrasts with Hershbein and Kahn (2018) who find that overall postings are fairly representative. This agrees with our list of Top 20 posting firms in Table 2A, which is broadly representative of the U.S. economy.

so few postings per firm with wages, measures of the average wage at the firm level will be noisy, selected counterparts to administrative data.

A final exercise demonstrates that drawing inference about firms' wages from job postings data can lead to mistakes. Suppose we use job postings to measure the average wage at a firm. Lack of wage information means any estimate is a noisy proxy of the mean, which implies significant mean reversion. The estimated persistence of the log firm wage nearly halves when comparing firms with the most and least wage information.

Mean reversion also introduces bias in estimates of treatment effects using temporal variation. For example, a common practice in the minimum wage literature is to compare firms that are 'exposed' to a change in the minimum wage to those that are not. A common measure of exposure is the distance between the firm's average wage and the minimum wage. This boils down to a difference-in-difference design where the treated are pre-event low average wage firms, and the control are pre-event high average wage firms. Sparse wage information introduces noise into pre-event means such that pre-event low (high) wage firms mean revert upwards (downwards), generating a spurious effect. A placebo design picks up significant, three times larger wage increases when estimated using only firms with low wage information, relative to estimation on firms with high wage information. Econometric techniques may be sound when performed on high quality administrative data. Applying them to noisy wage data is not advised.

Our work is related to literatures studying the availability of price information in markets. In labor markets, Hall and Krueger (2012) surveyed 1,300 newly hired workers in 2008. They found that 23% of workers knew precise pay prior to interview and that workers with less education were more likely to have had this information. Our estimate of precise wage information in only 5% of posts suggests either (i) workers contact employers with requests for information, (ii) an intervening stage between viewing a job posting and interview when an applicant is informed of a wage, which could be interpreted as a first stage of bargaining. Recent surveys conducted by Caldwell et al. (2023) provide support for (ii). This suggests that what researchers may have inferred from Hall and Krueger (2012) as wage posting, may be a first round of wage bargaining.

Brenčič (2012) provides the first analysis of wage information in job postings. Using U.S. data from "Monster.com" in 2006, she finds wages in 25% of posts, and that wages were more common in jobs where employers were less selective over skills or where skills were easier to measure. Using U.K. data from local employment agencies from 1988-1992, she finds wages in 86% of posts. This suggests the explosion of internet job search may have paradoxically reduced the amount of wage information available.

Why would employers choose to withhold pay information? One literature models firms' choice between wage posting and bargaining, contrasting commitment and flexibility (Flinn and Mullins, 2021; Doniger, 2023; Cheremukhin and Restrepo-Echavarria, 2020). A second literature on pay transparency within firms offers a different perspective: employers may want to hide from current employees the wages that it is offering to new hires.[4] A third literature on strategic price disclosure comes from an analysis of retail markets by Ellison and Ellison (2009). They show that easily available price comparisons made possible by the internet made consumers more price sensitive. In response, retailers strategically obfuscate price information. We think that our results would be useful to each of these literatures.

In recent work, Choi et al. (2020) use data from an online job board in Chile to study the cyclicality of new hire wages. Just 16% of job posts in these data contain explicit wage information, which suggests that the low prevalence of wage data in online job posts is not unique to our data nor to the United States. Banfi and Villena-Roldan (2019) use the same data to conclude that workers are not simply inferring the wage from the text of the job posting. They observe the 'target' wage for the position, which the employer must lodge with the platform but can choose not to advertise. Job postings with materially similar descriptions and target wages receive much higher applications when a wage is posted.

Recent papers have utilized the data used in this paper to make inferences about how labor markets function. Some focus on wages: Hazell and Taska (2020) (wage rigidity), Hazell et al. (2022) (national wage setting), Derenoncourt et al. (2021) (minimum wage spillovers). Our results suggest a challenge to such inference is (i) the overall lack of wages, (ii) bias in the availability of wages, (iii) systematic departure of the few posted wages that are available from wages in other data. Others focus on outcomes other than wages: Hershbein and Kahn (2018) and Modestino et al. (2015) (skill requirements), Braxton and Taska (2023) (technological change), Faberman et al. (2016) (advertising relative to open positions). Our study complements this second set.

[4] Cullen (2023) provides an overview of the evidence on the impacts of pay transparency. Cullen and Pakzad-Hurson (2023) builds a theoretical model of pay transparency and validates it with an event study analysis of state-level pay transparency laws.

1 Data

We use data from over 140 million online job postings over 2012-2017 provided by Burning Glass Technologies (Lightcast, 2013). Burning Glass Technologies (BGT) scraped, processed and encoded these data from over 45,000 different online sources including job boards, company websites, and others. The time frequency is daily.[5] Algorithmic methods are used to identify and delete copies of the same job posting across platforms. For example, removing an entry from a post on a job board that also appears on the company's corporate website.

BGT data is available to us from 2010 to 2023. We start in 2012 because BGT data throughout this period is harmonized to occupation codes that only are used in the OES from 2012 onwards. We end in 2017 because of massive breaks in 2018. From 2017 to 2018, (i) overall posts grow by 520%, (ii) the share of posts with wage information doubles, (iii) the share of posts with ranges jumps. This reflects major job boards adding their own imputed ranges of wages to posts, beginning in 2018.[6] As BGT does not identify these instances, this increases the fraction of posts coded with wages and the fraction with ranges.[7] These wages are not of interest to us. We believe the explosion in the sheer number of posts in BGT data reflects the de-duplication algorithm failing once the same post has different (imputed) wages on different job boards. Lafontaine et al. (2023) and Callaci et al. (2023) carefully remove posts with words like "estimated" indicating imputation. Lafontaine et al. (2023) Table 2 reports that 0.2% of posts in 2017 had such features, which jumps to 58.2% in 2018.

There are two main issues concerning whether or not the data are representative of job openings in the US. First, jobs employers post online may differ from those advertised by other means. Hershbein and Kahn (2018) provide a detailed comparison of BGT data to job openings measured in the Job Openings and Labor Turnover (JOLTS) and employment measured in OES data each from the BLS. They conclude BGT data are representative of many occupations and industries with a few exceptions such as over representing computer occupations and under representing food preparation and construction industries. Second, a single post may represent multiple posi-

[5] An observation in the data is created when the job is posted. If the job is still online after 60 days, a new observation is created. Posting duration data is available after 2018. When examining these data we see almost all postings appear for exactly 60 days (more than 50 percent expire on day 61), so differential durations of job postings do not skew the data. This also means, however, that we do not think that job post duration in the data is meaningful where it is available.

[6] For example, Indeed (link: May, 2018) and LinkedIn (link: February, 2018).

[7] The share of posts with wage information in the form of ranges jumps from 57 percent to more than 75 percent, reflecting that imputed wages are stated as ranges.

I. All sectors

	A. All job postings	B. Statistics by terciles of occupation income in OES			
		Low income occupations	Mid income occupations	High income occupations	No occ. information
	(1)	(2)	(3)	(4)	(5)
Fraction of jobs with any wage information	**13.5%**	9.7%	17.1%	11.6%	12.3%
Fraction of jobs with a point wage	**5.8%**	6.3%	8.2%	3.7%	5.7%
Fraction of jobs with wage info that post a range	**57.3%**	35.2%	52.3%	68.3%	53.3%
Average percent width of wage: $(\overline{w}-\underline{w})/(\underline{w}+\overline{w})/2$	**27.6%**	26.7%	25.8%	29.2%	28.8%
Average salary (BGT) [uses mid-point if range]	**\$62.411**	\$35.339	\$47.693	\$86.632	\$51.436
Average salary (OES)	**\$52.589**	\$25.577	\$43.346	\$96.705	-
Number of ads (million)	**141.8**	16.3	54.0	66.0	5.5
Fraction of all job postings	**100.0%**	11.5%	38.0%	46.5%	3.9%
Fraction of OES employment	**100.0%**	33.3%	33.3%	33.3%	-

II. Private sector

	A. All job postings	B. Statistics by terciles of occupation income in OES			
		Low income occupations	Mid income occupations	High income occupations	No occ. information
	(1)	(2)	(3)	(4)	(5)
Fraction of jobs with any wage information	**10.7%**	7.2%	14.0%	9.2%	8.6%
Fraction of jobs with a point wage	**4.9%**	4.6%	7.1%	3.2%	4.5%
Fraction of jobs with wage info that post a range	**54.6%**	36.1%	49.5%	65.7%	47.6%
Average percent width of wage: $(\overline{w}-\underline{w})/(\underline{w}+\overline{w})/2$	**26.7%**	26.9%	25.4%	27.8%	27.8%
Average salary (BGT) [uses mid-point if range]	**\$63.020**	\$35.255	\$49.563	\$86.755	\$52.417
Average salary (OES)	**\$52.589**	\$25.577	\$43.346	\$96.705	-
Number of ads (million)	**107.8**	13.9	40.8	49.4	3.8
Fraction of all job postings	**100.0%**	12.9%	37.8%	45.8%	3.5%
Fraction of OES employment	**100.0%**	33.3%	33.3%	33.3%	-

Table 1: Basic facts - BGT and OES data from 2012-2017 (salaries adjusted to 2020 dollars)

tions. Walmart may have a single post for cashiers when it seeks to fill a dozen positions. For this reason we do not make statements about quantities of positions, which cannot be inferred from the data.

Our analysis uses only the following variables: date; annual salary; occupation (6 digit SOC); location (Metropolitan Statistic Area); and employer name.[8] BGT provide an annualized salary measure for compensation listed at higher frequencies. We use the Consumer Price Index to express wages in 2020 dollars. We exclude potentially interesting variables due to the high incidence of missing data (e.g. education, part-time vs. full-time, and pay frequency).

We additionally use BLS data. First, employment and occupational wages at the 6-digit SOC level from the Occupational Employment and Wage Statistics (OES). Second, job opening (vacancies) data from JOLTS and unemployment data from the Current Population Survey (CPS).

[8] BGT uses textual algorithms to harmonize variations in employer names (e.g.: Hewlett Packard and HP would be the same employer) and to encode occupations based on job titles and the textual description. An audit by Hershbein and Kahn (2018) found occupation and industry codes to be 80% accurate.

2 Five Facts

Fact One: ***Wage information in job posts is scarce (less than 14% of posts), and when information is provided it is most often in the form of a range (more than 55%), and these ranges are wide (more than 25% of the midpoint).***

Table 1 provides basic summary statistics showing that wage information in job posts is scarce. We consider all sectors (Table 1.I) and only the private sector (Table 1.II). Unless otherwise stated we refer to all sectors.

Only 13.5% of all job postings include any type of wage information. Even when available, this wage information is often imprecise. Only 5.7% have a point figure and the rest post a range. In other words, of the 13.5% of posts with any information, the majority (57.3%) have a range. These ranges are usually sizeable. Define "*percentage width*" as the width of the range divided by the midpoint: $\frac{\overline{w}-\underline{w}}{0.5*\underline{w}+0.5*\overline{w}}$. By this metric, the average percentage width for posts with a range is 27.6%. To put this in context, examples are \$21-28/hr or \$32,000-\$42,000/yr. Looking across sectors, private sector jobs have less information than public sector jobs: only 10.7% have any wage information and only 4.9% have a point wage.

Concerning representativeness, the distribution of job posts differs from the distribution of workers currently employed. Table 1 shows that job posts are significantly skewed towards higher paying occupations than the distribution of occupations of U.S. workers measured in OES data. Columns (2), (3) and (4) use the OES to create employment weighted splits of occupations. If the distribution of postings and employment were the same, the row "Fraction of all posts" would read (approximately) one-third, one-third, one-third.[9]

Fact Two: ***There are systematic differences in the availability of wage information in job posts as we move across the wage distribution of occupations.***

Prior research found that posts for low wage or education jobs tend to have more wage information and that low wage workers are less likely to have bargained for their wage.[10] Since wage and education information is scare in online job posts, we explore this theme using occupational

[9] The skew in postings towards high wage occupations could be that on-line platforms are more used by workers and employers for these occupations. Low wage occupations could be filled by other means such as networks, referrals, or even physical "help wanted" signs. It could alternatively indicate differential turnover or vacancy duration. This would be consistent with high wage occupations having more turnover or, more likely, taking longer to fill a vacancy.

[10] Hall and Krueger (2012) find High School graduates were about 50% more likely than college graduates to know the pay for the job before they applied. See also Caldwell and Harmon (2019).

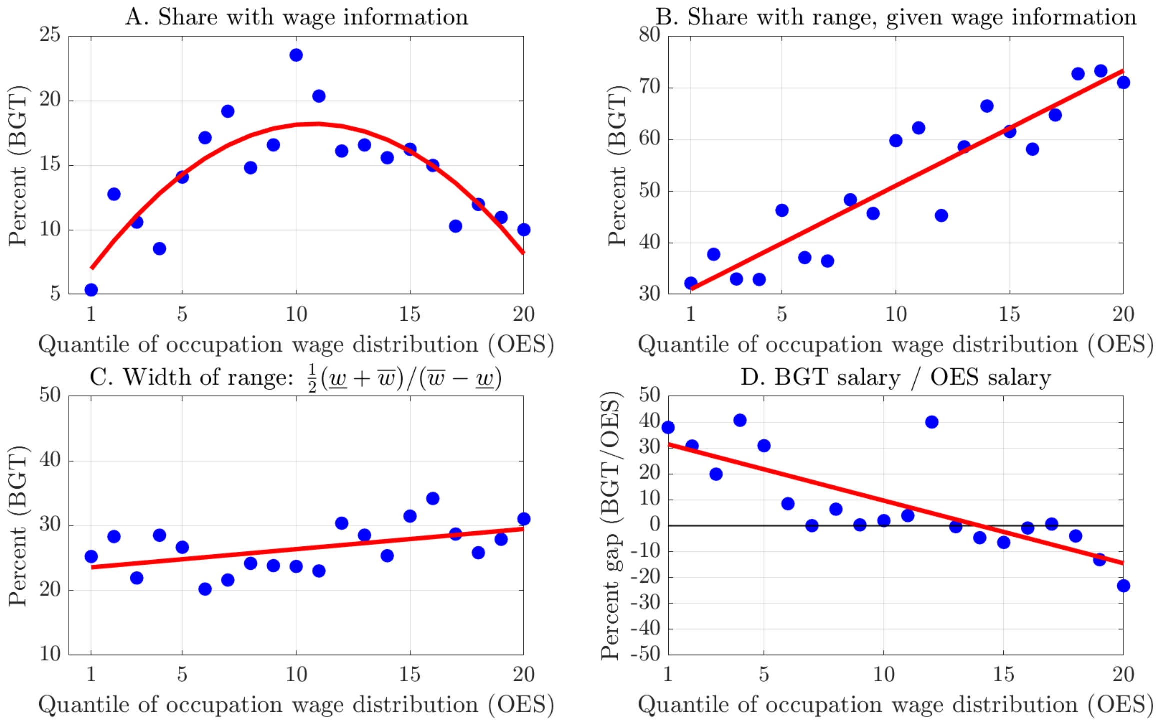

Figure 1: Comparing wage information across the wage distribution of occupations

wage information. Occupation is coded for 96% of posts in the data.[11] Since wage information in job postings is scarce, we rank occupations by average earnings in the OES. Our baseline results compare workers across twenty employment weighted bins.[12]

Job postings for middle income occupations are most likely to have wage information of any type but wage information is more precise in lower paying occupations. Figure 1 shows salary information available in job postings across the occupational wage spectrum.[13] Posts for occupations representing the lowest paid 5% of workers that contain a wage have a range only 30% of the time whereas the top 5% has a range around 70% of the time. The average width of salary ranges tends to be slightly higher in high wage occupations. Taken together, we conclude that posts for jobs in lower paying occupations tend to have more salary information.

Note that nowhere in the occupation wage distribution does the prevalence of any wage information exceed 25%. Thus the conclusion that pay information is scarce in online job posts is

[11]Table 1 column (5) shows that posts without occupation information tend to have even less wage information, hence focusing on postings with occupation information already assumes slightly more information.

[12]OES provides employment and earnings at the 6 digit level. We pool data over our sample period and employment weight to obtain a measure of the average wage at the 6 digit level. We then order occupations by the average wage and create twenty bins with equal 2012-2017 employment.

[13]The qualitative patterns we find hold when dividing the occupations into 100 wage bins (Figure A1) and when the sample is limited to only job posts from firms with more than 100 postings.

not driven by the composition of online job posts being tilted towards higher paying occupations where information is scarce. This practice is broadly common across the distribution.

Fact Three: ***In low wage occupations, posted wages are 40% higher than average occupation wages from BLS data, while in high wage occupations, posted wages are 20% lower.***

Can workers and researchers fill in missing wage information by inferring wages of posts without wage information using posts with wage information? We find that this is unlikely to be true.

Wages in job posts are lower in high wage occupations and higher in low wage occupations relative to the OES data on all employees. This is clear in the lower right-hand panel of Figure 1. Average wages are 40% higher for low wage occupations in job postings than for OES employed workers and 20% lower for high wage occupations.

That wages of new hires are lower than incumbents may not be surprising given on the job wage growth. What is remarkable is the higher wages in posts for new hires than current employees in low wage occupations. If a low wage job seeker draws a random job posting with a wage offer, the likelihood is that the wage is far above the occupation mean. This requires a degree of selection towards high wage jobs or a bias in information posted large enough to undo the forces of wage growth on the job.[14]

Fact Four: ***Within an occupation category, employers with job posts with (relatively) higher wages have a (relatively) higher tendency toward posting a range and wider percentage widths when they do.***

What roles do firms or employers play in determining the amount of wage information in job posts? If we rank firms by the fraction of their posts that contain *any* wage information, we find that the postings weighted median firm has information in less than 1 percent of their postings (Figure 2A).[15] It is hard to find many firms with a lot of wage information: less than 14 percent of posts belong to firms that have wage information in more than even 5 percent of their postings. Note that the distribution in Figure 2A integrates up to the average value of 13.5% in Table 1. A fat right-tail of firms with wage information in more than 20 percent of their postings skews the means in Table 1.

[14] Hazell and Taska (2020) Section 2.1 conducts an exercise that compares 2-digit occupation-quarter averages of BGT wages and wages of new hires in the CPS. Including time fixed effects, the elasticity of BGT occupation wages with respect to CPS wages is 1.34, with an R^2 of 0.07 (Table 2, column 3).

[15] Since many firms only ever have one post in the sample, presenting the unweighted distribution would make this even more extreme.

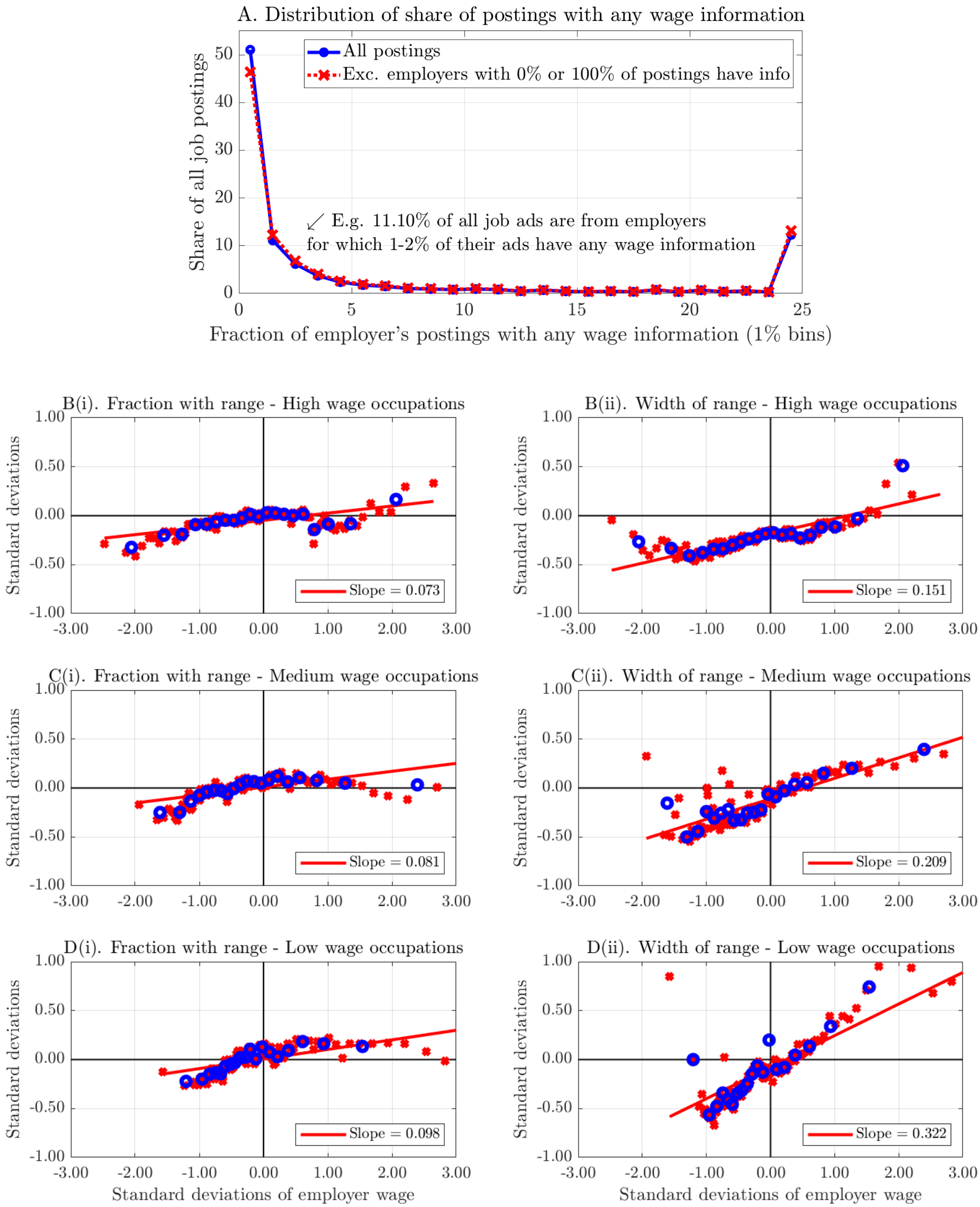

Figure 2: Within-'occupation-wage-category'-year, Across-employer relationship between employer wage and (i) fraction of postings with a range, (ii) average width of ranges

Notes: (i) Blue circles show averages within the x- and y-variables for 20 bins of the x-variables, weighted by observations, (ii) Red crosses show averages within the x- and y-variables for 100 bins of the x-variables, weighted by observations.

In the remaining panels of Figure 2, we first drop all posts without wage or employer information. Using the remaining posts, a 'firm-occupation' is constructed as the interaction of a firm and low, middle and high wage occupations (Table 1). Within a firm-occupation-year cell we compute the average wage in all postings. We compute the relative wage by subtracting the within-occupation-year average and dividing by the within-occupation-year standard deviation. We repeat this for *fraction of posts with a range* and *average width of range*.

Firms with higher posted wages—conditional on occupation-year—have less precise wage information: more frequent ranges, and wider ranges. Figure 2 plots these relationships using 20 (blue circles) and 100 (red crosses) bins of relative wages. For example, from Figure 2 panels D(i) and D(ii), a two standard deviation increase in the relative wage of the firm within the low wage occupation market increases the incidence of ranges by half a standard deviation and the width of ranges by one standard deviation. The gradient is steeper in low wage occupations.

Less precise wage information is not just associated with high paying occupations (Figure 1) but also with high paying firms within low paying occupations. In other words, not only the occupation of the worker matters but also the pay-rank of the firm. We again moderate *'matters'* with the 1,000 foot view that hardly any of these postings contain wage information at all. Results in this subsection suggest that it would be difficult to characterize differences in firm wage policies using job posting data.

Fact Five: ***There is no systematic relationship between wage information and various measures of local labor market tightness.***

While few postings have wage information, it might be that tighter labor markets have more wage information. On the one hand, Flinn and Mullins (2021), show that under standard theories of the labor market, posting is more likely when workers' bargaining power is stronger. A common understanding would be that workers' bargaining power is higher in tight labor markets. On the other hand, in tight labor markets employers may view the ability to negotiate after matching as a useful margin of recruiting, leading to less precise wage information in posts.

We take the lead of Kuhn et al. (2021) who show, geographically, that low unemployment markets are persistently tighter. Consistent with Kuhn et al. (2021) we first document systematic, persistent, differences across U.S. cities in labor market tightness. To measure market tightness we compare the number of job openings V to the number of workers searching for work S.[16] Different

[16]JOLTS provides estimates of total job openings across 18 MSA's, and We measure S using either (i) total un-

measures of tightness all yield the expected relationship: cities with twice as high unemployment have labor markets that are slacker by four standard deviations (Appendix Figure A2).

Despite the tight relationship between unemployment and the prevalence of open jobs across cities, we find startlingly little relationship between measures of wage information and market tightness. This null relationship is important for researchers. It suggests little hope in disentangling theories by appealing to the prevalence of wage posting across markets, differential in their tightness.

3 Which firms post online and which include wage information?

To place the lack of wage information in context and further shed light on the information that job seekers have when searching, we compare the identities of firms that have many postings to those that have many postings with wage information.

Most postings. Predominant employers in online job postings are consistent with what we would expect in the U.S. economy. Table 2A shows the employers with the most job posts, regardless of whether they contain salary information, from 2012-2017. These are mostly Fortune 500 companies with a broad national presence. Among the top 20 employers, only one—the U.S. Air Force—exceeds the national average in terms of the share of postings with any wage information. Consistent with Figure 2A, wage information is contained in less than 1 percent of postings at 15 of the top 20 posting firms.

Most wage information. If a worker is limiting themselves to applying to jobs that feature wage information, the set of firms they would apply to would be quite different to the U.S. economy. Of the top 20 firms in terms of total postings with wage information (Table 2B), four sectors account for 17 slots: (i) Federal, State and Local government (Dept. Veterans Affairs, US Army, US Airforce, Army National Guard, US Dept. Defense, State of LA, SC Dept. Public Safety), (ii) logistics / trucking (Roehl, J.B. Hunt, Schneider, Centerline, Enterprise), (iii) security services (G4S, Allied Barton), (iv) health (Mayo Foundation, Intermountain Healthcare, Onward Health). In government, laws or other institutional guidelines often mandate wage posting. Outside government, piece-rate work (e.g. truck driving) and part-time work (e.g. security subcontracting) appears to

employed workers or (ii) total unemployed workers plus a fraction 0.15 of employed workers (from estimated employed search efficiency in Bilal et al., 2022).

Firm		Total postings (1)	Share any info. (2)	Share range if info (3)
1.	Lowe's Companies, Inc	768.648	0.4%	31.7%
2.	Sears	670.538	2.0%	46.1%
3.	Macy's	541.567	0.2%	64.3%
4.	Hospital Corporation of America	524.694	0.8%	28.5%
5.	Anthem Blue Cross	518.325	0.2%	25.8%
6.	Wells Fargo	461.091	1.5%	16.7%
7.	Dollar General	443.853	0.1%	42.1%
8.	Marriott International Incorporated	441.537	1.4%	19.6%
9.	The Home Depot Incorporated	407.645	0.4%	42.3%
10.	Accenture	400.598	0.2%	40.8%
11.	Best Buy	367.859	0.1%	23.0%
12.	US Air Force	367.165	19.1%	59.7%
13.	Deloitte	366.764	1.2%	4.2%
14.	CVS Health	365.362	0.2%	45.0%
15.	JP Morgan Chase Company	346.337	0.2%	26.9%
16.	UnitedHealth Group	337.378	0.9%	34.7%
17.	Bank of America	333.762	0.3%	33.5%
18.	CACI	333.270	0.1%	8.6%
19.	Pizza Hut	306.177	0.2%	43.3%
20.	Hy-Vee	288.234	0.0%	9.8%

A. Top 20 firms by *Total postings*

Firm		Total postings (1)	Share any info. (2)	Share range if info (3)
1.	Department of Veterans Affairs	266.446	62.8%	96.7%
2.	US Army	220.536	45.1%	66.1%
3.	Roehl Transport	138.825	51.4%	10.4%
4.	US Air Force	367.165	19.1%	59.7%
5.	Schneider National Incorporated	81.217	75.2%	6.2%
6.	Army National Guard	135.136	38.2%	99.4%
7.	Centerline	53.140	89.7%	92.3%
8.	G4S	77.737	57.3%	23.1%
9.	Werner Enterprises	86.697	41.6%	94.3%
10.	Intermountain Healthcare	68.370	52.4%	1.2%
11.	Enterprise Rent-A-Car	114.984	29.8%	4.2%
12.	US Department of Defense	48.820	69.4%	94.5%
13.	State of Louisiana	41.913	79.7%	72.8%
14.	J.B. Hunt Transport, Inc.	66.123	48.7%	18.6%
15.	Bridgestone / Firestone	146.664	21.7%	94.9%
16.	Mayo Foundation for Medical Education and Research	77.732	40.7%	0.9%
17.	Onward Health	36.447	85.3%	99.9%
18.	YMCA	69.775	42.7%	66.7%
19.	AlliedBarton Security Services	215.526	13.1%	8.8%
20.	Sc Department Public Safety	31.995	86.5%	95.1%

B. Top 20 firms by *Total postings with any wage information*

Firm		Total postings (1)	Share any info. (2)	Share range if info (3)
1.	Home'n'Happy	1.615	100.0%	100.0%
2.	The Army Civilian Service	2.017	100.0%	94.4%
3.	Rehabilitation Correction	1.075	100.0%	98.9%
4.	Summerford Truck Line Inc	1.179	100.0%	100.0%
5.	CRST Atlanta Expedited	1.309	100.0%	0.0%
6.	Study Smart Tutors	1.503	99.9%	0.0%
7.	California Department Social Services	1.209	99.9%	99.9%
8.	Maricopa County Attorneys Office	1.611	99.9%	92.0%
9.	Symmetry Business Group	2.818	99.8%	95.2%
10.	Human Services Department	1.847	99.8%	99.7%
11.	Topgear Transportation	1.189	99.7%	99.7%
12.	New Mexico Behavioral Health Institute	1.745	99.7%	99.9%
13.	Natural Resources Department	1.556	99.7%	93.1%
14.	Minnesota State Colleges and Universities	1.451	99.7%	99.7%
15.	Riverside County Regional Medical Center	2.007	99.7%	96.5%
16.	North Carolina Department Of Health And Human Services	3.457	99.6%	96.4%
17.	J Iverson Riddle Developmental Center	6.201	99.6%	97.4%
18.	California Division of Correctional Rehabilitation	12.952	99.5%	96.4%
19.	North Carolina Department Of Justice	6.991	99.5%	97.2%
20.	Office Employment Training	2.083	99.5%	99.7%

C. Top 20 firms by *Share of postings with any information (firms with more than 1,000 postings)*

Table 2: Ranking of top 20 firms by different metrics, pooled over 2012-2017

be associated with more wage information. Similar industries is observed if we rank all firms with more than 1,000 postings by the prevalence of wage information in their posts (Table 2C). Virtually all include a wage range rather than a point wage (Column 3). Firms that have a lot of information have relatively imprecise information.

4 Implications for data use in research

4.1 Using job ad wages as a proxy for administrative or survey data.

We close by arguing one strong prescriptive conclusion: wages in job posts should not be used to study prevalent wages of current or new employees. As we have seen, the most common outcome is that wage information is missing, and the incidence of this is biased across occupations, sectors and firms. Missing data will bias any sample unless they are Missing Completely at Random (MCAR); ie: the data that do appear are a simple random sample of all potentially observed values. This bias can be ignored only if it is orthogonal to the outcome one would like to estimate. Our analysis clearly suggests that online ad wage data to not satisfy MCAR. One might use ranges of wages and take the mid-point to expand the data available. However, (i) the wide width of ranges is a cause for concern, (ii) our analysis shows that the choice of posting a range is also not random, and hence dropping posts with ranges also introduces bias.

One might be hopeful that recent and proposed laws mandating wage information in job posts would facilitate imputation. We are skeptical. Our preliminary analysis of Colorado after it's wage transparency laws were enacted in shows compliance is far from complete.[17] The incidence of any wage information has increased, but via more frequent and wider ranges. Posting a range satisfies the law, but keeps wages opaque for workers and researchers.

4.2 An implication of noise for inference - Using measures of firm average wages

We provide a simple example of how using wages from job posts can lead to mistaken inference. Suppose we want to compare how wages of low versus high wage firms respond to a particular economic shock. The usual approach in such a setting is a difference-in-difference design, comparing the response of wages at the treated (low wage) versus control (high wage) firms. Such a design arises in minimum wage studies. Low wage firms are *exposed*, while high wage firms are

[17] Preliminary work by Arnold et al. (2022) estimates that this law increased the number of job postings with wage information by 25 percentage points.

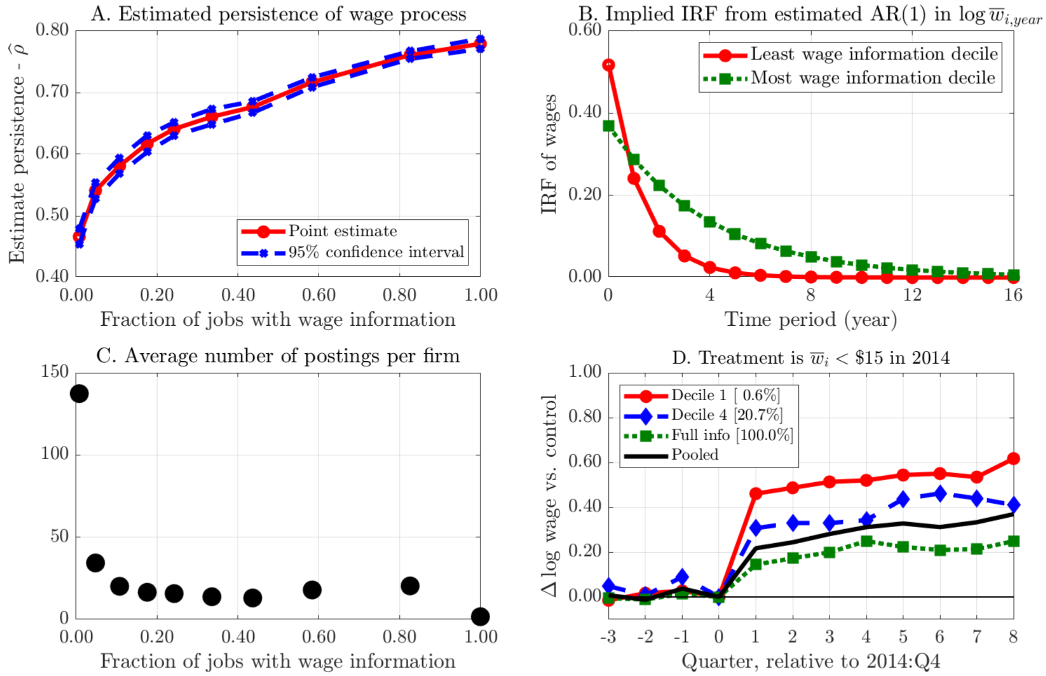

Figure 3: Mean reversion and inference with noisy measures of firm wages

Notes: Panel D plots the point estimates of difference in difference specifications, where the differences are across time (the period shown compared to 2014:Q4) and across treatment (firms with an average posted wage below \$15 in 2014) and control (all other firms).

less affected by the change in policy and hence control for coincident common factors.

When the variable used to classify firms in the pre-shock sample is the same variable studied post-shock, noisy measurement variable will lead to spurious results. Let $\overline{w}_{it}$ be the measure of the average wage of firm i in period t. If $\overline{w}_{it}$ is assembled with few observations from firm i, there is a chance that the pre-shock classification of the firm is mistakenly to the low-wage group, while its true average wage is high. Subsequent reversion to the mean, will push misclassified low-wage firms' wages up, and misclassified high-wage firms' wages down subsequent to any classification. A researcher will measure this as a positive effect of the shock on low-wage firms' wages. This issue will be *worse* the *less wage information* is used to assemble $\overline{w}_{it}$.

Figure 3 shows how these issues surface in practice when using job posting data. At random, we pick a three year time frame: 2014-2016. We split firms by the fraction of wage information in all of their postings in 2014, and conduct our analysis separately for each (unweighted) decile of firms by this measure. Figure 3A shows that if we run a simple OLS of quarterly $\log \overline{w}_{it}$ on its lag, estimated persistence drops off in deciles of firms with less wage information. Lack of information generates a noisy mean and high mean reversion, picked up as less persistence. Panel C shows this is not due to less posts per firm in this group, but by less wage information per post.

Figure 3D shows the mistakes this generate, by estimating a placebo difference-in-difference specification around the non-event of 2015:Q1. Firms are allocated as *low wage* (treatment) if their average wage in postings in 2014 was less than \$15. When estimated on treatment and control firms with low wage information, estimated effects are large as $\overline{w}_{it}$ is a noisy measure with mean reversion generating a spurious effect. When estimated on firms with substantial wage information, there is still a spurious effect, but it is much smaller.

Generally, mean reversion in such exercises is always a cause for concern. Even in administrative data, the average firm is around 20 to 25 workers, making any average at the firm level subject to mean reverting shocks and measurement error. However, when the underlying data provides an especially noisy estimate of the classifying and outcome variable due to the lack of information—here, the wage—inference becomes even more complicated.[18] The lack of wage information at some of the largest employers in the U.S. (Table 2A) precludes using these techniques in this way. When employers are large, however, and administrative data is used, such issues will disappear.

5 Conclusion

Wage information in job posts is exceedingly scarce, both to job searchers and researchers. This finding is robust. Lack of wage information is pervasive in job posts across occupations, sectors, and locations. Nonetheless, there are systematic differences across these in terms of the availability of point wages, and the imprecision of ranges of wages when point wages are not given.

Our findings raise concern about research taking wages in online job posts as a proxy for wage offers to new hires or wages paid to continuing employees. We recognize some advantages of these data over administrative measures of wages. They are high frequency, at the firm and establishment level, spatially precise, and contain occupation data. The high prevalence of missing wages—which our findings show is not at random—and the wide ranges associated with most wage data in job posts are likely to bias results when testing a wide range of hypotheses economists are interested in. As an example, we show the bias this can cause in a common empirical design that requires identifying a treated group by classifying firms as 'low wage'.

[18] Related, in an epidemiological context, Daw and Hatfield (2018) show that matching units on pre-shock outcome levels produces biased estimates and this bias increases when the pre-period difference is larger or the serial correlation in the outcome is smaller. See also Chabé-Ferret (2015).

Our findings provide fertile ground for researchers trying to answer open questions in labor markets. First, why do employers choose to post little wage information? There is a growing body of work in this area (Cheremukhin and Restrepo-Echavarria, 2020; Flinn and Mullins, 2021; Doniger, 2023). Theories proposed should be consistent with the facts presented here: many of the top employers present essentially zero wage information, there is systematic variation across occupations, and no variation across locations. As new theories are developed, one can return to our results to test implications. Second, does scarce wage information inhibit matching or dissuade search? Some research suggests that workers get wage information elsewhere, whether in additional text in the job post or through social networks. Still, employers are systematically making the choice to withhold wages. Third, does the lack wage information in job posts indicate a higher than previously thought incidence of bargaining? Detailed surveys on the stages of bargaining like Caldwell et al. (2023) will be useful.[19]

[19] Preliminary results from this work are consistent with our view that bargaining is wide spread. Consider a worker filling out a job application that did not state a point wage (95% of posts), where the application asks for a expected salary range, and the firm replies with a wage ahead of the interview. Answering the survey of Hall and Krueger (2012), the individual *did know the precise wage when interviewed*, but the process represents bargaining. Caldwell et al. (2023) finds that the vast majority of application processes elicit a candidate's wage expectations. Agan et al. (2021) finds that employee responses to salary expectations questions are interpreted by firms as indicative of outside options and worker productivity.

References

Agan, A. Y., B. Cowgill, and L. K. Gee (2021): "Salary History and Employer Demand: Evidence from a Two-Sided Audit," NBER Working Papers 29460, National Bureau of Economic Research, Inc.

Arnold, D., S. Quach, and B. Taska (2022): "The Impact of Pay Transparency in Job Postings on the Labor Market," *Available at SSRN 4186234*.

Banfi, S. and B. Villena-Roldan (2019): "Do high-wage jobs attract more applicants? Directed search evidence from the online labor market," *Journal of Labor Economics*, 37, 715–746.

Bilal, A., N. Engbom, S. Mongey, and G. L. Violante (2022): "Firm and Worker Dynamics in a Frictional Labor Market," *Econometrica*, 90, 1425–1462.

Braxton, J. C. and B. Taska (2023): "Technological Change and the Consequences of Job Loss," *American Economic Review*, 113, 279–316.

Brenčič, V. (2012): "Wage posting: evidence from job ads," *Canadian Journal of Economics/Revue canadienne d'économique*, 45, 1529–1559.

Caldwell, S., I. Haegele, and J. Heining (2023): "Bargaining and Inequality in the Labor Market," Tech. rep., Working Paper.

Caldwell, S. and N. Harmon (2019): "Outside options, bargaining, and wages: Evidence from coworker networks," .

Callaci, B., M. Gibson, S. Pinto, M. Steinbaum, and M. Walsh (2023): "The Effect of Franchise No-poaching Restrictions on Worker Earnings," Tech. rep.

Chabé-Ferret, S. (2015): "Analysis of the bias of matching and difference-in-difference under alternative earnings and selection processes," *Journal of Econometrics*, 185, 110–123.

Cheremukhin, A. A. and P. Restrepo-Echavarria (2020): "Wage Setting Under Targeted Search," Working Papers 2020-041, Federal Reserve Bank of St. Louis.

Choi, S., N. Figueroa, and B. Villena-Roldan (2020): "Wage Cyclicality Revisited: The Role of Hiring Standards," Tech. rep., University Library of Munich, Germany.

CULLEN, Z. B. (2023): "Is Pay Transparency Good?" Tech. rep., National Bureau of Economic Research.

CULLEN, Z. B. AND B. PAKZAD-HURSON (2023): "Equilibrium effects of pay transparency," *Econometrica*, 91, 765–802.

DAW, J. R. AND L. A. HATFIELD (2018): "Matching and regression to the mean in difference-in-differences analysis," *Health services research*, 53, 4138–4156.

DERENONCOURT, E., C. NOELKE, D. WEIL, AND B. TASKA (2021): "Spillover effects from voluntary employer minimum wages," Tech. rep., National Bureau of Economic Research.

DONIGER, C. (2023): "Wage Dispersion with Heterogeneous Wage Contracts," *Review of Economic Dynamics (forthcoming)*.

ELLISON, G. AND S. F. ELLISON (2009): "Search, obfuscation, and price elasticities on the internet," *Econometrica*, 77, 427–452.

FABERMAN, R. J., M. KUDLYAK, ET AL. (2016): "What does online job search tell us about the labor market," *Economic perspectives*, 40, 1–15.

FLINN, C. AND J. MULLINS (2021): "Firms' Choices of Wage-Setting Protocols," Tech. rep., Discussion paper, New York University.

HALL, R. E. AND A. B. KRUEGER (2012): "Evidence on the incidence of wage posting, wage bargaining, and on-the-job search," *American Economic Journal: Macroeconomics*, 4, 56–67.

HAZELL, J., C. PATTERSON, H. SARSONS, AND B. TASKA (2022): "National wage setting," Tech. rep., National Bureau of Economic Research.

HAZELL, J. AND B. TASKA (2020): "Downward Rigidity in the Wage for New Hires," Discussion Papers 2028, Centre for Macroeconomics (CFM).

HERSHBEIN, B. AND L. B. KAHN (2018): "Do recessions accelerate routine-biased technological change? Evidence from vacancy postings," *American Economic Review*, 108, 1737–1772.

KUHN, M., I. MANOVSKII, AND X. QIU (2021): "The geography of job creation and job destruction," Tech. rep., National Bureau of Economic Research.

LAFONTAINE, F., S. SAATVIC, AND M. SLADE (2023): "No–Poaching Clauses in Franchise Contracts: Anticompetitive or Efficiency Enhancing?" Tech. Rep. 440415, SSRN Workng Paper.

LIGHTCAST (2013): "Job Postings Data," Tech. rep.

MODESTINO, A., D. SHOAG, AND J. BALLANCE (2015): "Upskilling: do employers demand greater skill when skilled workers are plentiful?" *Federal Reserve Bank of Boston, Working Paper*.

SMITH, A. (2015): "Searching for Work in a Digitial Era," Tech. rep., Pew Research Center.

ONLINE APPENIX - Additional figures

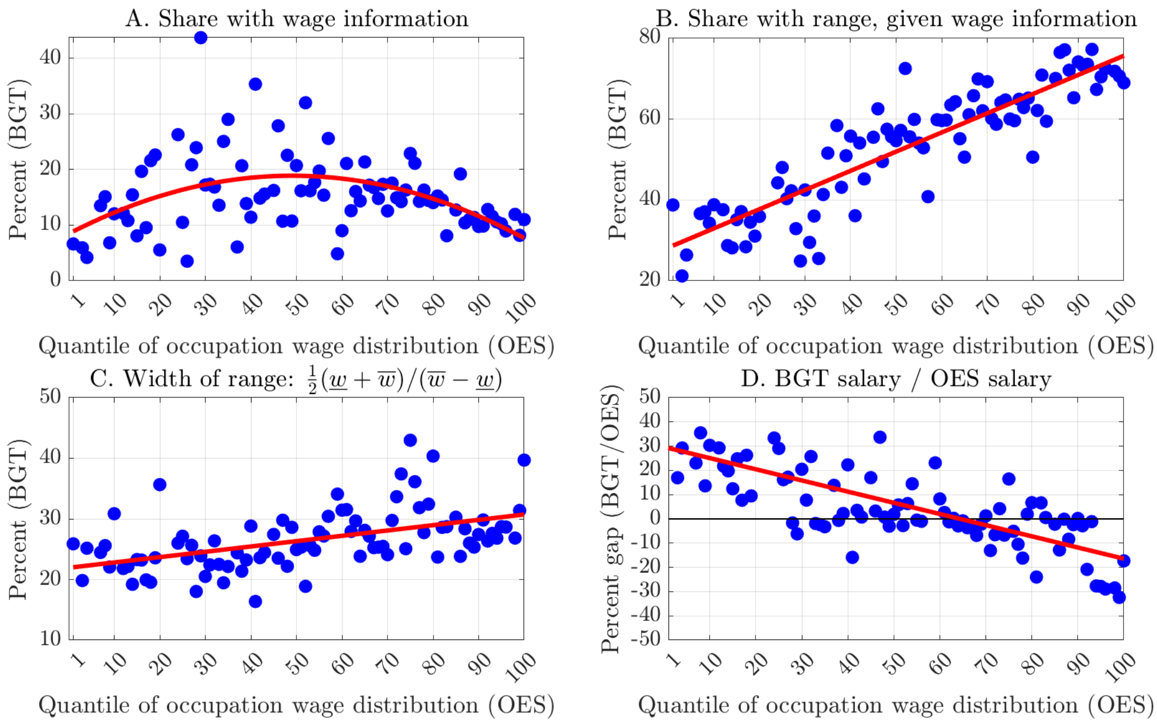

Figure A1: Within-'occupation-wage-category'-year, Across-employer relationship between employer wage and (i) fraction of postings with a range, (ii) average width of ranges

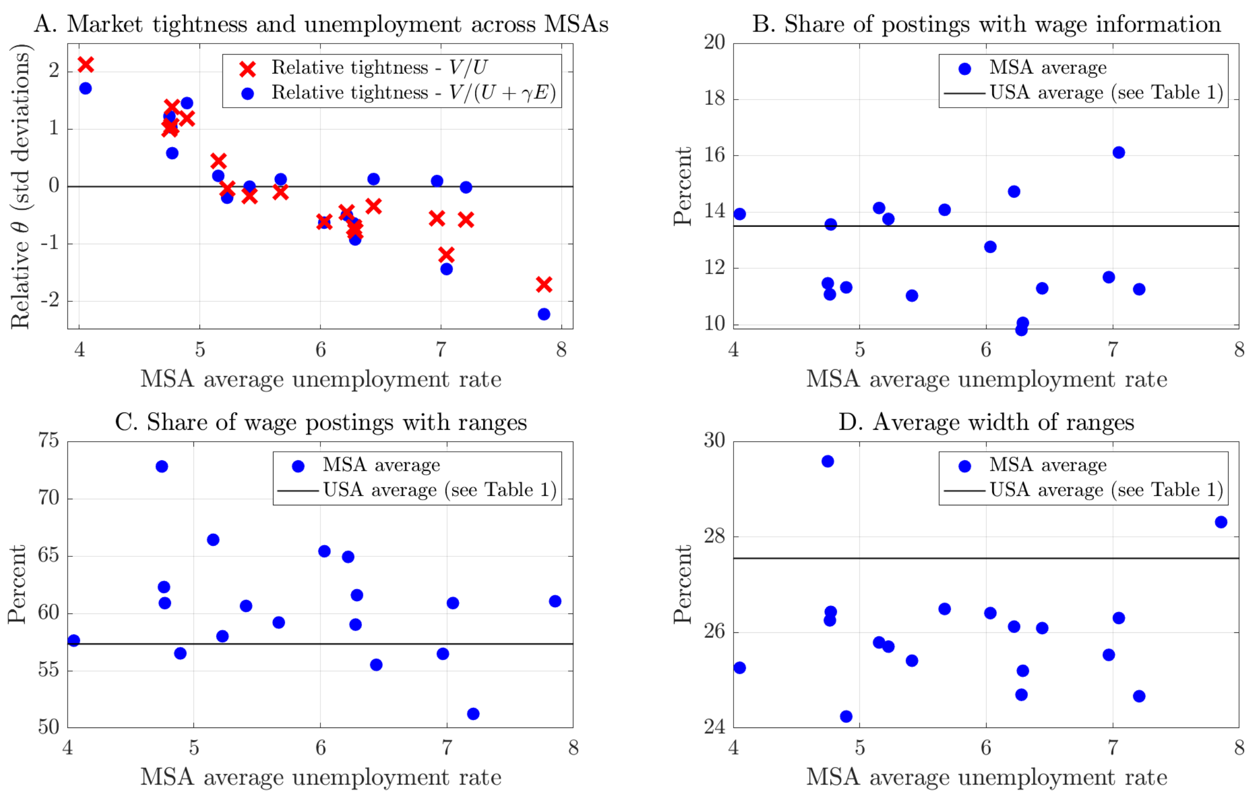

Figure A2: Wage information across labor markets