Monopsony and Employer Mis-optimization Explain Why Wages Bunch at Round Numbers

By ARINDRAJIT DUBE, ALAN MANNING, AND SURESH NAIDU *

We show that administrative hourly wage data exhibits considerable bunching at round numbers. We run two experiments, randomizing wages around 10 cents and $1.00, to experimentally measure left-digit bias for identical tasks on Amazon Mechanical Turk, and fail to find any evidence of discontinuity in the labor supply function at round number, despite estimating a considerable degree of monopsony. We replicate these results in administrative worker-firm hourly wage data from Oregon. We can rule out inattention estimates found in the behavioral product market literature. We provide evidence that firms "misoptimize" wage-setting. More monopsony requires less employer misoptimization to explain bunching.
JEL: D03,D22,J3,J42

Behavioral economics has documented a wide variety of deviations from perfect rationality among consumers and workers. For example, in the product market, prices are more frequently observed to end in 99 cents than can be explained by chance, and a sizable literature has confirmed that this reflects firms taking advantage of customers' left digit bias (e.g Levy et al. 2011; Strulov-Shlain 2023). However, it is typically assumed that deviations from firm optimization are unlikely to survive, as competition among firms drives firms that fail to maximize profits out of business. Therefore, explanations for pricing and wage anomalies typically rely on human behavioral biases. In this paper, we show that when it comes to a quantitatively important anomaly in wage setting—bunching at round numbers—it is driven by choices made by firms (or other stakeholders in wage-setting) and not workers. In principle, the observed bunching can be rationalized either by "behavioral" firms forgoing significant profit in order to pay round numbers in relatively competitive labor markers, or by near-optimizing firms losing only small profits because employer monopsony power blunts the cost of mispricing labor. Empirically, a moderate amount of monopsony power, well within the range obtained in both our data and other recent high-quality U.S. evidence, is sufficient to explain the substantial bunching we document in this paper with limited loss of profits.

We begin by documenting bunching in the hourly wage distribution at "round" numbers, first noted by Jones (1896). Unlike previous studies, we use data from both administrative sources and an online labor market to confirm that there is true bunching of wages at round numbers—that it is not simply an artifact of survey reporting. We begin by providing the first (to our knowledge) credible evidence on the extent to which hourly wages are bunched at round numbers in high quality, representative data on hourly wages from Unemployment Insurance records from the three largest U.S. states (Minnesota, Washington, and Oregon) that collect information on hours.[1]

* Dube: University of Massachusetts Amherst, NBER, and IZA (adube@umass.edu); Manning: London School of Economics (a.manning@lse.ac.uk); Naidu: Columbia University and NBER (sn2430@columbia.edu). We thank Sandra Baquie, Clara Berestycki, Doruk Cengiz, Len Goff, Jeff Jacobs, and Jon Zytnick for excellent research assistance. We are grateful for helpful comments from David Autor, Matt Backus, Ellora Derenoncourt, David Green, Patrick Kline, Ilyana Kuziemko, Ioana Marinescu, Jim Rebitzer, Owen Zidar and from seminar participants at Microsoft Research, Columbia University, Boston University, Harvard University, Princeton, IIES, University of Toronto, UC Berkeley, MIT, and the 2018 Society of Labor Economists conference. Experiment 1 in this paper detailed in Section III was registered in the American Economic Association's registry for randomized controlled trials under ID AEARCTR-0001349. Replication material can be found at Dube et al. (2025)

[1]We were able to use individual-level matched worker-firm data from Oregon, and micro-aggregated data of job counts by detailed wage bins in Minnesota and Washington.

We further assess the extent of bunching in online labor markets, using a near universe of posted rewards on the online platform Amazon Mechanical Turk (MTurk).

We compare the size of the bunches in the administrative data to those in the Current Population Survey (CPS), where round number bunching is prevalent. For example, in the CPS data for 2016, a wage of $10.00 is about 50 times more likely to be observed than either $9.90 or $10.10. Figure 1 shows that the hourly wage distribution from the CPS outgoing rotation group (ORG) data between 2010 and 2016 has a visually striking modal spike at $10.00 (top panel). The middle panel of the figure shows that the share of wages ending in round numbers is remarkably stable over the past 35 years, between 30-40% of observations. The bottom panel of the figure also shows that between 2002 and 2016, the modal wage had been exactly $10.00 in at least 30 states, reaching a peak of 48 in 2008. As wages have increased, we see that this number drops off in recent years, but we see a commensurate rise in number of states with a modal wage of $15.00. Of course, survey data is likely to overstate the degree of round-number bunching. We also use a unique CPS supplement, which matches respondents' wage information with those from the employers to correct for reporting error in the CPS, and show that the measurement error-corrected CPS replicates the degree of bunching found in administrative data in the states where we observe both. It seems highly unlikely that such bunching at $10.00 is present in the distribution of underlying marginal products of workers.

We first attempt to account for bunching by building and experimentally testing an imperfectly competitive model where worker behavior exhibits discontinuities at round numbers, paralleling product price studies from the marketing literature that have documented pervasive "left-digit bias" where consumers ignore lower-order digits in price. These discontinuities could result either from individual cognitive biases or social norms held by workers. What is important is that it is individual worker-side behavior exhibits the discontinuity. We design and implement two experiments on an online platform (MTurk). In the first experiment (N=5,017), we randomly vary rewards above and below 10 cents for the same task to estimate the labor supply function facing an online employer. Our primary sample consists of workers who report being on the platform primarily for financial motivations, as their labor supply is more likely to respond to the rewards. Like the administrative data, the task reward distribution on MTurk exhibits considerable bunching. However, our experimentally estimated labor supply function shows no evidence of a discontinuity as would be predicted by worker left-digit bias. In a second follow-up experiment (N=2757), we vary rewards around $1.00, which also exhibits substantial bunching. Again, the estimated labor supply function exhibits no discontinuity around the $1.00 mode. Our results are precise enough to rule out levels of left-digit bias found in the product market literature.

Next, we use administrative matched worker-firm data on hourly wages from Oregon and a matched panel design to show that there is no discontinuity in the separation elasticity at $10.00 (the modal nominal wage during the study period) conditional on past worker wage history and other firm characteristics. These results suggest that our experimental findings on lack of left-digit bias are likely to have external validity beyond on-line labor markets. In addition, our analysis using the matched worker-firm data produces firm-level labor supply elasticities in the 2 to 3 range, which is consistent with estimates from other recent work (Sokolova and Sorensen (2021); Bassier et al. (2022)), and suggests a sizable amount of monopsony power in the U.S. low-wage labor market.

We then extend the model to allow imperfect firm optimization in the form of employer preferences for round wages, parameterized by the fraction of profits foregone in order to pay a round number. These could be administrative costs, cognitive biases of managers, public salience, or social norms that bind on the employer rather than workers: what is important is that they constrain employer wage-setting for reasons unrelated to worker behavior. We recover estimates of monopsony, employer misoptimization, and left-digit bias from the distribution of missing mass around $10.00 relative to a smooth latent density and conclude that employer misoptimization accounts for much of the observed bunching.[2]

In the administrative data from Oregon, we further show that small businesses, employers in the construction sector (with a prevalence of contractors), and firms with low adoption of frontier managerial practices, are most likely to pay round numbers, consistent with bunching being a result of potentially unsophisticated employer pay-setting. At the same time, we also find that bunching is the most prevalent among firms paying more labor market rent (in the sense of a high AKM firm fixed effects). These patterns are consistent with the theoretical prediction and empirical confirmation of greater monopsony power among firms at the top of the job ladder—thereby reducing the costs of mispricing. Using our estimates of monopsony power, we calculate that employer misoptimization in U.S. labor markets costs mis-optimizing employers between 1% and 5% of profits.

Our paper is related to a growing literature on behavioral firms, which documents a number of ways firms fail to maximize profits (DellaVigna and Gentzkow 2019; Goldfarb and Xiao 2011; Hortacsu and Puller 2008; Bloom and Van Reenen 2007; Cho and Rust 2010; Romer 2006; Zwick 2021). Our paper is the first to systematically quantify the extent of employer misoptimization in the labor market, which is an old and recurring theme in institutionalist labor economics[3], as well as the first to document the absence of left-digit bias among workers.[4] DellaVigna and Gentzkow (2019) examine two potential sources of the excess uniformity in prices they observe. The first is inertia, which include both agency as well as behavioral frictions within the firm that make it hard to set prices optimally. The second is reputational or brand concerns that may create long-term incentives to keep prices uniform despite opportunities for price discrimination. We think both of these are present in the labor market, but like Della Vigna and Gentzkow, we think the former dominates the latter in most cases. With the exception of very large public-facing brands that come under the eye of politicians and activists, most Americans have limited knowledge of what wages firms are paying, and thus public relations concerns for wage setting are likely low. Evidence that firm wage-setting is influenced these extra-labor market forces comes from the recent work on national wage setting by Hazell et al. (2022). They find that many national firms set wages in a manner that is highly uniform, not taking into account local labor market conditions. Further evidence comes from the literature on voluntary minimum wages (Derenoncourt and Weil, 2025), which are all at round numbers. While round-number bunching is a cognitively universal bias and thus visible in many diverse labor market contexts, it is symptomatic of pervasive mis-pricing of labor in the face of labor market frictions. The recent literature corroborates the view that firms both have considerable leeway in wage setting, and often engage in rule-of-thumb behavior such as use of uniform and round-number wages. Most closely related to this paper is Reyes (2024), who builds on our work to provide complementary evidence of bunching of wages using Brazilian data, along with firm-level correlates of such bunching.[5]

[2]While other configurations are logically possible, they do not easily explain why wages are bunched at round numbers. For example, if employers had a left-digit bias, any heaping would likely occur at \$9.99 and not at \$10.00, which is not true in reality. Similarly, if workers tended to round off wages to the nearest dollar, this would not encourage employers to set pay exactly at \$10.00. In contrast, both workers' left-digit bias and employers' tendency to round off wages provide possible explanations for a bunching at \$10.00/hour.

[3]Richard Lester, for example, noted that "Study of wage data and the actual processes of wage determination indicates that psychological, social, and historical factors ...are highly important influences in particular cases. There appears to be a fairly high degree of irrationality in the wage structure and in company wage policies, judged either by the market analysis of economists or job evaluation within and between plants."(Lester, 1946, pp. 158).

[4]A large literature has discussed cognitive biases in processing price information, but little of this has discussed applications to wage determination. Behavioral labor economics has extensively documented other deviations from the standard model (e.g. time-inconsistency and fairness, see Babcock et al. (2012) for an overview) for worker behavior. Worker behavioral phenomena have been replicated even in online spot labor markets (Chen and Horton, 2016; DellaVigna and Pope, 2018).

[5]Similar to us, Reyes (2024) finds bunching of wages are more prevalent at arguably "less sophisticated" firms, which is consistent with employer mis-optimization. In this paper, we additionally provide alternative theoretical accounts of bunching including worker left-digit bias and employer misoptimization; test these explanations including using experimental and observational evidence; and relate bunching to the presence of monopsony power in the labor market. On the other hand, Reyes (2024) extends our work by testing a hypothesis we develop below about how bunching relates to minimum wage spillovers.

In our account of wage-bunching, it is important to assume that firms have some labor market power. In this, we follow work in behavioral industrial organization that explores how firms choose prices when facing behavioral consumers in imperfectly competitive markets.[6] A recent and fast-growing literature has argued that monopsony is pervasive in modern, unregulated, labor markets (Manning 2011). While some of the recent literature on monopsony has focused on identifying effects of labor market concentration on wages (Berger et al. (2022); Azar et al. (2022)), more relevant for our paper is the literature that tries to directly estimate residual labor supply elasticities, either from firm level shocks to productivity, rents, or wage policies (Lamadon et al. (2022); Kline et al. (2019); Cho (2018); Derenoncourt and Weil (2025)), or from direct experimental manipulation of wages (Caldwell and Oehlsen (2018); Dube et al. (2020)). We show that moderate amounts of monopsony in the labor market can provide a parsimonious explanation of anomalies in the wage distribution, such as patterns of wage-bunching at arbitrary numbers. To demonstrate some additional economic implications of round number bunching, we show how inertia created by round-number bunching generates a new source of spillover from minimum wages and changes the interpretation of rent-sharing estimates and payroll tax incidence.

The plan of the paper is as follows. In section 2, we provide evidence on bunching at round numbers using administrative data as well as data from the CPS corrected for measurement error, and benchmark these against the raw CPS results. We recover the source of the bunched observations by comparing the observed distribution to an estimated smooth latent wage distribution. We show patterns in the degree of bunching by firm characteristics. In section 3, we develop a model of bunching with worker left-digit bias and develop testable hypotheses. Section 4 presents findings from the online experiment, administrative data, and stated-preference experiment–recovering both estimates of the degree of monopsony as well as the extent of worker left-digit bias. Section 5 extends the model to account for firm misoptimization, and presents estimates for the extent of mis-optimization required to rationalize the observed degree of bunching with our estimated labor supply elasticities. Section 6 concludes.

I. Bunching of wages at round numbers

There is little existing evidence on bunching of wages. One possible reason is that hourly wage data in the Current Population Survey comes from self-reported wage data, where it is impossible to distinguish the rounding of wages by respondents from true bunching of wages at round numbers. Documenting the existence of wage-bunching requires the use of other higher-quality data.

A. Administrative hourly wage data from selected states

Earnings data from administrative sources such as the Social Security Administration or Unemployment Insurance (UI) payroll tax records are of high quality, but most do not contain information about hours. However, 4 states (Minnesota, Washington, Oregon, and Rhode Island) have UI systems that collect detailed information on hours, allowing the measurement of hourly wages. We have obtained individual-level matched employer-employee data from Oregon, as well as micro-aggregated hourly wage data from Minnesota and Washington. The UI payroll records cover over 95% of all wage and salary civilian employment. Hourly wages are constructed by dividing quarterly earnings by total hours worked in the quarter. The micro-aggregated data are state-wide counts of employment (and hours) by nominal $0.05 bins between $0.05 and $35.00, along with a count of employment (and hours) above $35.00. The counts exclude NAICS 6241 and 814, home-health and household sectors, which were identified by the state data administrators as having substantial reporting errors.

[6]See Heidhues and Kőszegi (2018) for a survey and Gabaix and Laibson 2006 for an early example. Theoretical models to explain bunching in prices also assume firms have some market power: e.g., Basu (1997) has a single monopolist supplying each good, Basu (2006) has oligopolistic competition, and Heidhues and Kőszegi (2008) use a Salop differentiated products model.

Figure 2 shows the distribution of hourly wages in MN, OR and WA (we report the distributions separately in the Appendix). The histogram reports normalized counts in 10 cents (nominal) wage bins, averaged over 2003q1 to 2007q4. We focus on this period because in later years the nominal minimum wages often reach close to $10.00, making it difficult to reliably estimate a latent wage density in this range as we do in our analysis later in this paper. The counts in each bin are normalized by dividing by total employment. The wages are clearly bunched at round numbers, with the modal wage at the $10.00 bin representing more than 1.5 percent of overall employment. This suggests that observed wage bunching is not solely an artifact of measurement error, and is a feature of the "true" wage distribution. Further, the histogram reveals spikes at the MN, OR, and WA minimum wages in this period, suggesting that the hourly wage measure is accurate.

In Online Appendix D, we show very similar degree of bunching in a measurement-error corrected CPS, using the 1977 CPS Supplement that recorded wages from firms as well as workers. While the degree of bunching in the raw CPS falls with the measurement error correction, it remains substantial, and indeed comparable to the administrative data.

B. *Task rewards in an online market: Amazon Mechanical Turk*

Amazon MTurk is an online task market, where "requesters" (employers) post small online Human Intelligence Tasks (HITs) to be completed by "Turkers" (workers).[7] Psychologists, political scientists, and economists have used MTurk to implement surveys and survey experiments (e.g. Kuziemko et al. (2015)). Labor economists have used MTurk and other online labor markets to test theories of labor markets, and have managed to reproduce many behavioral properties in lab experiments on MTurk (Shaw et al. 2011).

We obtained the universe of MTurk requesters from Panos Ipeirotis at NYU. We then used the Application Programming Interface developed by Ipeirotis to download the near universe of HITs from MTurk from May 2014 to February 2016, resulting in a sample of over 350,000 HIT batches. We have data on reward, time allotted, description, requester ID, first time seen and last time seen (which we use to estimate duration of the HIT request before it is taken by a worker). The data are described more fully in Online Appendix F and in Dube et al. (2020).

Figure 3 shows that there is considerable bunching at round numbers in the MTurk reward distribution. The modal wage is 30 cents, with the next modes at 5 cents, 50 cents, 10 cents, 40 cents, and at $1.00. This is remarkable, as this is a spot labor market that has almost no regulations, suggesting the analogous bunching in real world labor markets is not driven by unobserved institutional constraints, including long-term implicit or explicit contracts. Nor is there an opportunity for bargaining, so the rewards posted are generally the rewards paid, although there may be unobserved bonus payments.

C. *Estimating Excess and Missing Mass Using a Bunching Estimator*

The excess mass in the wage distribution at a bunch that has been documented in the previous section must come from somewhere in the latent wage distribution that would result from the "nominal model" without any bunching (in the terminology of Chetty (2012)).[8] This section

[7] The sub header of MTurk is "Artificial Artificial Intelligence", and it owes its name to a 19th century "automated" chess playing machine that actually contained a "Turk" person in it.

[8] Following the literature, our procedure assumes that the missing mass is originating entirely from the surrounding bunching interval. In principle, it is possible that the missing mass is originating from latent non-employment—i.e., jobs that would not exist under the nominal model in the absence of bunching. However, the extent to which some of the excess jobs at $10 is coming from latent non-employment, one would need to assume either that (1) these jobs have latent productivity exactly at $10.00 so that employers are indifferent between entering and not entering, or (2) they have productivity greater than $10 but have a fixed cost of not paying exactly $10 that is independent of the size of the profits from paying different wages under the nominal model. Both of these assumptions strike us as implausible. As an empirical matter, if some of the excess mass at $10 are originating from latent non-employment, the estimated missing mass around $10 would be smaller in magnitude than the excess mass at $10. However, our estimated missing mass

describes how we estimate the origin of this "missing mass." We focus on the bunching at the most round number ($10.00 in the wage data, $1.00 in the MTurk rewards data). We ignore the secondary bunches; this will attenuate our estimate of the extent of bunching, as we will ignore the attraction that other round numbers exert on the distribution. However, we later assess robustness to controlling for secondary modes.

We assume that under the nominal model (i.e., without bunching at w_0), wages follow a latent wage distribution, $g(w)$. Following the standard approach in the bunching literature, we use a K^{th} order polynomial to model the latent wage distribution $g(w)$, and estimate this excluding the area affected by bunching (i.e., the "bunching interval") that ranges between $w_0 - \Delta w$ and $w_0 + \Delta w$ around the bunching point, w_0. Within the bunching interval, we model the bunching effects using 10 cent bins, indicated by j going up by 10 cent increments from $w_0 - \Delta$ to $w_0 + \Delta$.[9] (See Kleven 2016 for a discussion of this type of strategy used in the literature.)

We use bin-level counts of wages c_w in 10 cents bins, and define $p_w = \frac{c_w}{\sum_{j=0}^{\infty} c_j}$ as the normalized count or probability mass for each bin. We then estimate:

$$p_w = \sum_{j=w_0-\Delta w}^{w_0+\Delta w} \beta_j \mathbb{1}_{w=j} + \sum_{i=0}^{K} \alpha_i w^i + \epsilon_w \quad (1)$$

In this expression, j sums over 10 cent wage bins within the bunching interval, and $\sum_{i=0}^{K} \alpha_i w^i$ forms the K^{th}order polynomial approximating $g(w)$, while the β_j terms are coefficients on dummies for bins in the excluded range around w_0, between $w_L = w_0 - \Delta w$ and $w_H = w_0 + \Delta w$. β_{w_0} is the excess bunching (EB) at w_0. In addition, $\sum_{j=w_0-\Delta w}^{w_0-10} \beta_j$ is the missing mass strictly below w_0 (MMB), while $\sum_{j=w_0+10}^{w_0+\Delta w} \beta_j$ is the missing mass strictly above w_0 (MMA). However, it might be likely that the the number of workers is increasing in w; so our test for left-digit bias is whether $\sum_{j=w_0-\Delta w}^{w_0-10} \frac{\beta_j}{g(j)} = \sum_{j=w_0+10}^{w_0+\Delta w} \frac{\beta_j}{g(j)}$. In other words, we test whether the share of workers who would have been paid the latent wage j but are instead paid w_0 is different when considering bunchers from above versus below (averaged over all j within the bunching interval).

Since Δw is unknown, we use an iterative procedure similar to Kleven and Waseem (2013). Starting with $\Delta w = 10$, we estimate equation (1) and calculate the excess bunching EB and compare it with the missing mass $MM = MMA + MMB$. If the missing mass is smaller in magnitude than the excess mass, we increase Δw and re-estimate equation (1). We do this until we find a Δw such that the excess and missing masses are equalized. Since Δw is itself estimated, we estimate its standard error using a bootstrapping procedure suggested by Chetty (2012) and Kleven (2016). In particular, we resample (with replacement) the errors $\hat{\epsilon}_w$ from equation (1) and add these back to the fitted $\hat{p}_w$ to form a new distribution $\tilde{p}_w$, and estimate regression (1) using this new outcome. We repeat this 500 times to derive the standard error for Δw. The estimate of Δw and its standard error will be useful later for the estimation of other parameters of interest.

In Figure 4 we show the estimates for the administrative data from MN, OR, and WA, using polynomial order $K = 6$. For visual ease, we plot the kernel-smoothed $\hat{\beta}_j$ for the missing mass. Even leaving out the prominent spike at $10.00, the wage distribution is not smooth, and has relatively more mass at multiples of 5, 10 and 25 cents. For this reason, it is easier to detect the shape of the missing mass by looking at the kernel-smoothed $\hat{\beta}_j$. Moreover, we show the excess and missing mass relative to the counterfactual $\widehat{p_w^C} = \sum_{i=0}^{6} \alpha_i w^i$. There is clear bunching at $10.00

from the surrounding interval is, indeed, able to account for the size of the excess mass—which suggests that latent non-employment is unlikely to be an important contributor to the excess mass in our case.

[9]We use 1 cent bins in the MTurk data.

in the administrative data, consistent with evidence from the histogram above. We find that the excess bunching can be accounted for by missing mass spanning Δw = \$0.80; we can also divide Δw by w_0 and normalize the radius as $\omega = \frac{w_H - w_0}{w_0} = 0.08$. Visually, the missing mass is coming from both below and above \$10.00, which is relevant when considering alternative explanations.

These estimates are also reported in Table 1, column 1. The bunch at \$10.00 is statistically significant, with a coefficient of 0.010 and standard error of 0.002. In addition, the size of the missing mass from above and below w_0 are quantitatively very close, at -0.006 and -0.007 respectively. The t-statistic for the null hypothesis that the missing mass (relative to latent) are equal for bunchers from above and below is 0.338. This provides strong evidence against worker left-digit bias, which would have implied an asymmetry in the missing masses. The radius of the missing mass interval is $\omega = 0.08$, with a standard error of 0.027. In other words, employers who are bunching appear to be paying as much as 8% above or below the wage that maximizes profits under the nominal model.

In column 2, we use the CPS data limited to MN, OR, and WA only. We find a substantially larger estimate for the excess mass, around 0.043, consistent with rounding of wages by survey respondents. In column 3, we report estimates using the re-weighted CPS counts for MN, OR, and WA adjusted for rounding due to reporting error using the 1977 supplement (CPS-MEC). The CPS estimate of bunching adjusted for measurement error is much closer to the administrative data, with an estimated magnitude of 0.014; while it is still somewhat larger, we note that the estimate from the administrative data is within the 95 percent confidence interval of the CPS-MEC estimate. The use of the CPS supplement substantially reduces the discrepancy, which is re-assuring. At the same time, we note that the estimates for ω using the CPS (0.07) are remarkably close to those using the administrative data (0.08). The graphical analogue of column 3 is in Figure 5.

In Online Appendix B we probe the robustness of these bunching estimates to a variety of specifications of the latent distribution. We include indicators for 25 cent and 50 cent modes, higher degree polynomials, and Fourier bases. We also estimate specification that uses a polynomial in the real (rather than the nominal) wage distribution as the latent distribution; this is a "difference in bunching" specification which addresses issues raised about parametric assumptions in bunching estimators raised by Blomquist et al. (2021). The radius of the missing mass and ratio of the excess to missing mass are quite similar across these specifications.

The main conclusions from this section are that the missing mass seems to be drawn symmetrically from around the bunch and from quite a broad range. As the model below shows, these facts are informative about possible explanations for bunching and the nature of labor markets.

D. Which Employers Pay Round Numbers?

In his original paper, Jones (1896) explained the use of round numbers in wages as a result of difficulties and uncertainties in measuring productivity, the costs of calculating and administering wages, and the influence of social norms. [10] These factors are considered the costs of setting non-round wages for firms.

We begin by analyzing which types of firms are more likely to bunch wages exactly at \$10.00. If bunching is driven by managerial cognition or employer administrative costs, then firms that bunch should exhibit characteristics suggestive of "behavioral" factors, such as a lack of administrative sophistication or modern management practices. We use various firm characteristics from the matched employer-employee data from Oregon. Figure 6 presents the extent of bunching (ratio of excess mass to latent wage density) for different groups of firms, categorized by firm size, quintiles

[10] As Jones, (pp. 114) noted, "Round numbers frequently locate for us those fields in which propriety dictates that economic motives, which prompt exact measurements of value, shall be made subordinate to other considerations."

of hourly wage firm effects,[11] industry, the share of part-time workers (working less than 20 hours), and for starting wages.

The most significant pattern observed is that bunching is pervasive across firms of various industries, sizes, and wage policies. In all but one case, the extent of bunching is statistically significant at least at the 90% confidence level, and typically at the 95% level. The sole exception is the accommodation and food services sector, where Oregon's minimum wage (indexed to inflation) is highly binding.

At the same time, there are meaningful differences in bunching that are worth noting. We find that firms most likely to bunch are very small firms, particularly in the construction sector, which includes many small contractors. This is consistent with these firms being relatively "unsophisticated," more likely to pay in cash, and less likely to have standardized pay practices (e.g., automatic inflation escalation) that would remove employer discretion in wage setting. As a result, these firms are more prone to bunching wages at round numbers. The public sector also shows unusual levels of bunching, which could reflect low cost-minimization pressures in public wage-setting, or it may indicate a preference among public sector workers or unions for benefits over wages. Entry-level wages are also more likely to bunch, which aligns with the idea that firms may use other mechanisms for wage increases (e.g., applying a uniform percentage increase to all workers), which gradually erode the tendency to bunch wages at round numbers over time. Importantly, and consistent with Reyes (2024), we find that bunching is more prevalent among firms that exited prior to the end of the sample period; among firms with high turnover; and to an extent among firms that shrank in employment over the sample period. Overall, the variation in round-number wage use across firms supports the hypothesis that bunching is more prevalent among less sophisticated firms.

Although we lack extensive firm-level data on human resource practices to merge with the Oregon data, we incorporate the World Management Survey database from Bloom and Van Reenen (2007), aggregated to the industry level. Appendix Figure A.3 shows the prevalence of bunching by the degree of adoption of advanced human resource management practices. Overall, the matched subsample exhibits less bunching, as the management scores are available only for the manufacturing sector, where bunching is less common. However, within this sample, the extent of bunching decreases with managerial sophistication, with the lowest tercile showing the greatest bunching. This further supports the idea that less sophisticated firms are more likely to bunch.

Finally, we find that the firms with higher AKM effects are more likely to bunch; we return to this point later in the paper in discussing how monopsony power may affect bunching.

II. A model of round-number bunching in the labor market

As a first pass for explaining the bunching at round numbers, this section presents a model of the labor market where workers exhibit a left-digit bias, mirroring inattention in the product market. Suppose there is a mass 1 of workers differing in their marginal product p, assumed to have density $k(p)$ and CDF $K(p)$—assume labor is supplied inelastically to the market as a whole. We assume there is only one "round number" wage in the vicinity of the part of the productivity distribution we consider—denote this by w_0.

We model the left-digit bias of workers following the recent literature on left-digit bias in product markets. Starting from a discrete choice model of the labor market, a workers $i's$ utility at a job j is:

$$u_{ij} = \eta_1((1-\theta)w_i + \theta\lfloor w_i \rfloor) + \epsilon_{ij} \quad (2)$$

[11]These firm effects are estimated using the Abowd et al. (1999) decomposition. The Oregon-specific hourly wage decomposition results are detailed in Bassier et al. (2022).

Here ϵ_{ij} captures idiosyncratic valuation of the job that are independently drawn from a type I extreme value distribution. Importantly, here there can be a discontinuous jump in the utility at a round number $w_0 = \lfloor w_j \rfloor$ due to left-digit bias among workers (where $\lfloor . \rfloor$ is the "floor" or "greatest integer" function). left-digit bias has been documented in a wide variety of markets, used to explain prevalence of product prices that end in 9 or 99, and is a natural candidate explanation for bunching in the wage distribution. [12]

This model implies a logit labor-supply function that is linear in wage with a jump. When the number of jobs J is large, we get a log-linear labor supply curve given by:

$$\ln(l_j) = \underbrace{\eta_1(1-\theta)}_{\eta_0} w_j + \underbrace{\eta_1\theta}_{\gamma_0} \lfloor w_j \rfloor \tag{3}$$

We estimate variants of this specification in our empirical work below, and recover an estimate of $\theta = \frac{\gamma_0}{\gamma_0+\eta_0}$. For our theoretical results, in order to obtain tractable closed-form formulas for bunching, we will work with a constant elasticity specification that takes the standard log-log labor-supply specification and augments it with a wage-floor term. We also implement this specification empirically, and find very similar estimates of θ as the linear specification. We thus report results from both the linear specification above and log specification, given below as:

$$\ln(l_j) = \eta \ln(w_j) + \gamma_1 \lfloor w_j \rfloor \tag{4}$$

When a sample is restricted to the vicinity of a specific round number $w_o = \lfloor w_j \rfloor$ we can recover our pre-specified regression, where $\gamma = \gamma_1 \times w_0$:

$$\ln(l_j) = \eta \ln(w_j) + \gamma \cdot \mathbb{1}_{w_j \geq w_0} \tag{5}$$

It can be readily seen that $\theta \approx \frac{\gamma_1}{\gamma_1+\eta}$ with the γ and η from equation 5. [13]

When worker productivities p vary across firms, wages are dispersed and depend on p. Define $\rho(w,p) = (p-w)l^*(w,p)$. Here $\rho(w,p)$ is, in the language of Chetty (2012), the "nominal model" that parameterizes profits in the absence of left-digit bias. Optimizing wages in the nominal model would yield a smooth "primitive" profit function of productivity given by $\pi(p_j) = (\frac{p_j}{1+\eta})^{1+\eta}$, but the presence of worker biases induces discontinuities in true profits at round numbers. In deciding on the optimal wage for employers one simply needs to compare the profits to be made by maximizing the nominal model and paying the round number. Consider the wage that maximizes the nominal model. Given the isoelastic form of the labor supply curve to the individual firm, this can simply be shown to be:

[12]For example, Levy et al. (2011) show that 65% of prices in their sample of supermarket prices end in 9 (33.4% of internet prices), and prices ending in 9 are 24% less likely to change than prices ending in other numbers. Snir et al. (2012) also document asymmetries in price increases vs. price decreases in supermarket scanner data, consistent with consumer left-digit bias. A number of field and lab experiments document that randomizing prices ending in 9 results in higher product demand (Anderson and Simester 2003; Thomas and Morwitz 2005; Manning and Sprott 2009; Pope, Pope and Sydnor 2015). Pope, Pope and Sydnor (2015) show that final negotiated housing prices exhibit significant bunching at numbers divisible by $50,000, suggesting that round number focal points can matter even in high stakes environments. Lacetera, Pope and Sydnor (2012) show that car prices discontinuously fall when odometers go through round numbers such as 10,000. Allen et al. (2016) document bunching at round numbers in marathon times, and interpret this as reference-dependent utility. Backus, Blake and Tadelis (2019) show that posted prices ending in round numbers on eBay are also a signal of willingness to bargain down.

[13]Note that for w_j close to w_0 we have the elasticity at the round number w_0 given by $w_0 \frac{dlog(l_j)}{dw_j} = \frac{d(\eta_1 \log((1-\theta)w_j+\theta\lfloor w_j \rfloor))}{dw_j}|_{w_j=w_0} = \eta_1$ which would be given by $\eta+\gamma$ estimated from equation 5 and the elasticity below the round number given by $w_j \frac{d(\eta_1 \log((1-\theta)w_j+\theta\lfloor w_j \rfloor))}{dw_j}|_{w_j \neq w_0} = (1-\theta)\eta_1$, which would be estimated as η from equation 5. So θ and η_1 can be recovered from η and γ (and w_0).

$$w^*(p) = \left(\frac{\eta}{1+\eta}\right) p \tag{6}$$

where the size of the markdown on the marginal product is determined by the extent of competition in the labor market. If the labor market is perfectly competitive, $\eta \to \infty$, and wages are equal to the marginal product. We will refer to the wage that maximizes the nominal model as the latent wage. The firm with job productivity p will pay the round number wage as opposed to the latent wage if:

$$e^{\gamma \mathbb{1}_{w*(p)<w_0}} > \frac{\rho\left(w^*\left(p\right), p\right)}{\rho\left(w_0, p\right)} \tag{7}$$

Taking logs, we obtain that a firm will pay the round number if $p \in \left[p_*, \frac{(1+\eta)}{\eta} w_0\right]$, where p_* solves

$$\gamma = \ln \rho\left(w^*\left(p_*\right), p_*\right) - \ln \rho\left(w_0, p_*\right) \tag{8}$$

This implies that all jobs with latent wages between w_0 and $w^*(p_*)$ will pay w_0, and bunching will be larger the greater is γ . For the marginal bunching job, the proportionate increase in profits from paying a round number wage is γ. One immediate implication of this result is that the patterns of excess and missing mass documented in Section I.C are not consistent with left-digit bias. That is, even without direct estimates of γ and η the fact that the excess mass at 10.00 is accounted for by equal amounts of missing mass from both above and below 10.00 suggests that left-digit bias is not the immediate determinant of bunching at round numbers. We now turn towards directly identifying estimates of η and γ, and hence θ, from experimental and observational data.

III. Experimental and Observational Evidence on Worker Left-Digit Bias

A. Experimental Evidence From MTurk

While the observational and survey experiment evidence strongly points away from worker left-digit bias, one might still worry about omitted variables bias or survey responses not reflecting real stakes. Further, neither of the earlier experiments allow us to test whether productivity also responds to the wage. We therefore supplement these approaches by conducting two high-powered experiments on Amazon Mechanical Turk to test for left-digit bias. Online labor market experiments have a high degree of internal validity, though a disadvantage is that one is inevitably unsure about the external validity of the estimates. For example, one might expect that these "gig economy" labor markets are very competitive because they are lightly regulated and there are large numbers of workers and employers with little informational frictions or long-term contracting. However, in a companion paper, Dube, Jacobs, Naidu and Suri (2020) compile labor supply elasticities implicit in the results from a number of crowdsourcing compensation experiments on MTurk, finding a precision-weighted elasticity of 0.14 across 5 experiments (including the first one in this paper). This low estimate implies considerable market power in these types of "crowdsourcing" labor markets, widely used to obtain training data for machine learning applications (Kingsley, Gray and Suri (2015)).

In their original paper on labor economics on Amazon Turk, Horton et al. (2011) implement a variant of the experiments we conduct below, making take it or leave it offers to workers with random wages in order to trace out the labor supply curve. However, while they label this an estimate of labor supply to the *market*, it is in fact a labor supply to the requester that they are tracing out, as the MTurk worker has the full list of alternative MTurk jobs to choose from.

EXPERIMENT 1 (E1):

The first experiment, conducted in 2018 extends the Horton, Rand and Zeckhauser (2011) design to experimentally test for worker left-digit bias.[14] We randomize wages for a census image classification task to estimate discontinuous labor supply elasticities at 10 cents. We posted a total of 5,500 unique HITS on MTurk tasks for 10 cents that includes a brief survey and a screening task, where respondents view a digital image of a historical slave census schedule from 1850 or 1860, and answer whether they see markings in the "fugitives" column (for details on the 1850 slave census, see Dittmar and Naidu (2016)). The respondent is then offered a choice of completing an additional set of classification tasks for a specific wage. This is close to the maximum number of unique respondents obtainable on MTurk within a month-long experiment.

EXPERIMENT 2 (E2):

We conducted a second experiment in 2022 to measure left-digit bias at the more salient, and larger spike at $1.00. This experiment uses a "honeypot" design, where treated workers see the job among the jobs available and a survey-taking task centered around 1.00. We first pre-screened MTurk workers to as to exclude bots, and then randomized the wage offered to each one, for a job tailored to appear only in that worker's choice set of HITs while browsing MTurk. We posted 2826 HITs this way, where the task was taking a survey experiment on hypothetical job preferences over 30 pairs of jobs with random characteristics (Dube et al., 2020), along with writing essays. This survey was designed to take enough time so as to induce variation in take-up around 1.00, but had a smaller overall take-up rate.

We pool data from both experiments and present results for the sample of sophisticated users (with more than 10 hours and primarily motivated my money, as defined by Horton et al. (2011)) in order to increase the power to detect a significant labor-supply elasticity. A large sample of respondents who are more likely to be responsive to wage changes increases the precision with which we can test our hypothesis that wages differentially respond to one-cent increases at 9 or 99 cents. We also show results including amateurs, and disaggregating by experiment .

B. Specification

We begin by estimating experiment-specific regressions as follows:

$$Accept_i = \beta_{0A} + \eta_0 w_i + \gamma_0 \left(\mathbb{1} \left\{ w_i \geq w_0 \right\}_i \right) + \beta_1 T_i + \epsilon_i \tag{9}$$

This corresponds to the specification from equation 3, which is linear in the perceived wage. Recall that an estimate of θ can be calculated as $\frac{\gamma_0}{\gamma_0 + w_0 \eta_0}$. As discussed in section II, we also want to estimate a constant elasticity specification, which is given by:

$$Accept_i = \beta_{0A} + \eta_1 log(w_i) + \gamma_1 \left(\mathbb{1} \left\{ w_i \geq w_0 \right\}_i \right) + \beta_1 T_i + \epsilon_i \tag{10}$$

In both cases, $w_0 \in \{10, 100\}$ denotes the threshold depending on the experiment. T_i is within-experiment controls (6 or 12 images for the 2018 experiment, wave for the 2022 experiment). We also compute estimates of θ as $\frac{\gamma_1}{\gamma_1 + \eta_1}$.

[14] Pre-registered as AEA RCT ID AEARCTR-0001349.

Next, we pool our experiments, estimating a regression with a single log wage coefficient to gain precision, but otherwise allow separate coefficients on the jumps across the two experiments, resulting in our preferred specification:

$$\begin{aligned}(11)\quad & Accept_i = \beta_{0A} + \eta_1 log(w_i) + \gamma_1^{10}\Big(\mathbb{1}\{w_i \geq 10\}_i \times \mathbb{1}\{E1\}_i\Big) + \\ & \gamma_1^{100}\Big(\mathbb{1}\{w_i \geq 100\}_i \times \mathbb{1}\{E2\}_i\Big) + \beta_1 T_i + \beta_2 X_i + \beta_3 X_i \times \mathbb{1}\{E2\}_i + \epsilon_i\end{aligned}$$

In one variant, we additionally include worker-level controls,X_i, interacted with the experiment indicators, where experiments are denoted by $E1$ and $E2$. We estimate θ_{10} and θ_{100} analogously. When estimating this specification, we additionally calculate a precision weighted average of the two, and report this as a pooled θ.

While specifications restricted to the sophisticated agents are our preferred ones, we also show specifications including the "amateurs" who use Mechanical Turk less than ten hours a week, or whose primary motivation is not money. We can also include the data from Horton, Rand and Zeckhauser (2011) to improve the precision of the η_1 estimate, further refining our power to detect left-digit bias.

Finally, we can also pool fully across the experiments estimating a single γ_1 coefficient, along with a single semi-elasticity η_1, giving us the following specification:

$$\begin{aligned}(12)\quad & Accept_i = \beta_{0A} + \eta_1 log(w_i) + \gamma_1\Big(\mathbb{1}\{w_i \geq 10\}_i \times \mathbb{1}\{2018\}_i + \\ & \mathbb{1}\{w_i \geq 100\}_i \times \mathbb{1}\{2022\}_i\Big) + \beta_1 T_i + \beta_2 X_i + \beta_3 X_i \times \mathbb{1}\{2022\}_i + \epsilon_i\end{aligned}$$

This has the cost of imposing stronger assumptions with the benefit of greater precision. As a practical matter, as we shall see, the point estimates of γ_1^{10} and γ_1^{100} are numerically close, so the primary effect is on precision. Based on this specification, we calculate a fully pooled θ coefficient as $\frac{\gamma_1}{\gamma_1+\eta_1}$. Implicitly, this compares the average response to a 1-cent increase at each threshold to the sample-averaged estimate of the semi-elasticity of labor.

C. Results

We begin by presenting the raw data for sophisticated respondents in the scatterplots in Figure 7 shows the basic pattern. For both experiments, there is a clear positive slope. However, there are no discernible discontinuity at either 10 cents or $1.00.

Tables 2 and 3 below show the key experimental results. The former shows experiment-specific estimates based on equation (10), with the wage variable in levels (columns 1, 2) or log (columns 3, 4). The wage coefficients are positive and statistically significant in all samples and specifications. The implied labor supply elasticities are between 0.14 and 0.21. At the same time, these elasticities are small in magnitude indicating with a high degree of monopsony power, consistent with the MTurk evidence in Dube et al. (2020).

The second set of coefficients (labeled as"Jump") estimate γ_1. Contrary to the left-digit bias hypothesis, the estimates in all 4 columns are negative but not statistically different from zero. We also report estimates of θ as described above. All the point estimates of θ are negative in sign, and not distinguishable from zero. For experiment 1, the log-wage estimate rules out θ greater than 0.11 at the 5% level, while the level-wage estimate rules out θ larger than 0.24. Experiment 2 rules out θ larger than 0.35 in logs or 0.36 in levels. (Note that while the coefficient on log-wage are broadly similar across experiments (0.18 and 0.11), the level wage coefficients are of different orders of magnitude, which highlights why pooling across the two experiments is more feasible with the log-wage specification.)

Next, to gain precision, column 1 in Table 3 jointly combines data from both experiments and estimates a common log-wage coefficient using 11, while allowing each experiment to have a different jump at the relevant thresholds. The coefficient on log wage estimate of 0.14 is more precise, with smaller standard error than either individual experiment; at the same time, both jump coefficients continue to be small and statistically insignificant. Estimates here rule out θ_{10} greater than 0.21 and θ_{100} greater than 0.22. A more precise labor supply elasticity estimate increases our power to bound the extent of possible left-digit bias. In addition, we report a precision-weighted average θ across the two estimates: the confidence interval for the pooled θ rule out values larger than 0.10.

Columns 2 through 5 show robustness of the estimates to inclusion of amateurs, as well as adding data from the Horton et al. (2011) experiment. Estimates are mostly unchanged, though as expected, inclusion of amateurs reduce precision of θ because amateur labor supply is less elastic. At the same time, use of all 3 experiments and all workers (N-8,430) rules out a pooled θ greater than 0.14.

Columns 6 and 7 presents estimates from a specification with a single jump and a single log wage coefficient based on equation (12). Note that the the jump coefficients from the two experiments are not very different numerically, which suggest the main impact is on precision. Column 10 reports a θ of -0.22, ruling out values larger than 0.09. Finally, column 10 repeats this, but with the sample expanded maximally (including amateurs, and data from Horton et al. (2011)). The estimated θ is -0.14 and highly precise, again ruling out values greater than 0.09. These results are summarized graphically in Figure 9; we return to this figure below when we contextualize these estimate in relation to the extant literature.

We show additional specifications from the pre-analysis plan in Online Appendix F. Our pre-analysis plan also included a test for whether task quality could exhibit left-digit bias, and we report those in the Appendix; we do not find any evidence of a discontinuity in task quality at 10 cents either.

We interpret the evidence as strongly pointing away from a sizable left-digit bias on the workers' side. At the same time, we find considerable amount of wage-setting power in this online labor market: labor is fairly inelastically supplied to online employers, with an estimated elasticity η generally between 0.1 and 0.2.

D. Panel Evidence From Oregon Administrative Data

Using matched employer-employee data from Oregon, we estimate the responsiveness of employee separations to hourly wages, and test for a a discontinuity in separations at \$10.00/hour wage. The separation elasticity is widely used in the monopsony literature to measure the responsiveness of the supply of labor to a firm to the firm's posted wage (see Manning (2011)).

We use a matched design that compares separation responses of new hires based on their starting wage, conditioning on a rich set of controls. For the 2003-2007 period, we start with the universe of all newly hired workers who have at least one quarter of tenure with their current employer (whom we refer to as "full quarter hires").[15] With the hiring date in quarters denoted as t, our sample includes workers who were still employed at the beginning of quarter $t+1$. To focus on workers who were earning around \$10.00, we restrict our sample to those whose initial ("full quarter") hourly wage at date $t+1$ was between \$9.50 and \$10.59. We then calculate their four-quarter separation, $S_{i,t+4}$, which equals 1 if the worker leaves the firm by quarter $t+4$. Our analysis examines how this separation probability varies with the starting full-quarter wage $w_{i,t+1}$ for workers hired with an initial wage during the same period and with similar characteristics at their previous job (at time $t-1$).

[15]This restriction excludes workers with very marginal attachment to the workforce such as seasonal workers. Notably, the spike at \$10.00 persists in this sample.

A key concern in estimating this elasticity using observational data is the possibility that workers earning different wages may differ in other important dimensions. In our baseline specification, the vector $\mathbf{X_{it}}$ includes "workers controls": an indicator for calendar quarter, and the log of hourly wage (in 10 cents intervals) and log of weekly hours in their last full quarter at their previous job. In the most saturated specifications we also control for the value of the workers' (and previous employers') AKM effects.

Another concern arises from firm heterogeneity. As we saw in section I.D, firms engaging in greater round number bunching tend to be quite different. In the baseline specifications, "firm controls" include 1 digit NAICS sectors, and quintiles of share of jobs in the range at their current employer that pay exactly \$10.00, in order to control for firm heterogeneity. This is particularly important because some firms (about a third in our sample) have no jobs paying exactly \$10.00; and these firms are quite different than those bunching at \$10.00, including ones that have all jobs in this range bunched at \$10.00. In addition, in some specifications we trim the sample to exclude firms where a very high share (more than 95%) or very low share (less than 5%) of wages are bunched at \$10.00, to ensure the "common support" assumption holds. In our preferred specifications, we account for firm heterogeneity by interacting the quarter of hire with firm fixed effects.

Our most saturated specification considers workers who were hired within the same firm in the same quarter, and who had previously had earned very similar wages and worked similar hours at similar firms, but are now earning somewhat different wages at their current jobs, and considers their subsequent 4-quarter separation probability. This matched panel design is similar to the approach developed by Bassier et al. (2022).[16]

We estimate analogues of our experimental specifications above, regressing separations $S_{i,t+4}$ on log wages, and an indicator for earning more than \$10.00.

$$S_{i,t+4} = \beta_0 + \eta_1 log(w_{i,t+1}) + \gamma_{1A} \mathbb{1}\left\{w_{i,t+1} \geq 10.00\right\}_i + \Lambda \mathbf{X}_{i,t+1} + \epsilon_{i,t+1} \tag{13}$$

This specification tests for a jump in the labor supply at \$10, but constraining the slope to be the same on both sides. So left-digit bias is rejected if $\gamma_{1A} = 0$. We also calculate the left-digit bias parameter θ, which is approximaely equal to $\frac{\gamma_{1A}}{\gamma_{1A}+\eta_1}$.

The columns in Table 4 report the estimates from specifications with increasingly saturated controls. Using the dynamic monopsony steady state assumption, where firm labor supply elasticity is twice the negative of the separation elasticity, the labor supply elasticities, η, range between 1.7 and 2.6 and all are statistically significant at the 5 percent level. These indicate a fair degree of of monopsony power in the low-wage labor market in Oregon over this period, and are consistent with findings in Bassier et al. (2022).

For columns 2 and later, when we add an indicator for \$10/hour or more as a regressor, we do not see any indication that the separation rate changes discontinuously at \$10 in any specification, whether the sample includes all firms, is trimmed to exclude extremely bunched or extremely un-bunched firms, whether we control for firm fixed effects, and whether we control for previous job AKM values and values of worker fixed effects. In all cases, the jump coefficients are small, "wrong signed," and statistically indistinguishable from zero. We calculate the estimates for θ for each of these specifications. In the specifications with firm fixed effect (columns 4,5, and 6), the 95%

[16] Bassier et al. (2022) also examine the separation rates of newly hired workers as a function of their starting wage. One key difference is that they focus solely on wage policy differences across firms. Specifically, they regress the separation rate on the starting wage at time t, instrumented by the difference in average wages between the old and new firms, to isolate the firm-specific wage component. Their estimates of the labor supply elasticity for the 2003-2007 period, using a broader set of Oregon workers, is approximately 3.6, but with smaller estimates in the low-wage sector, which aligns with our findings. However, our main interest lies in detecting left-digit bias by testing for discontinuities in the worker-level separation elasticity around \$10.00, which requires using individual-level, rather than solely firm-level, wage variation. Moreover the bias from heterogeneity of firms paying exactly a round number is not a concern in their case, but it is for us here.

confidence intervals rule out estimates of θ greater than 0.17 in all cases. Finally, in Figure 8, we plot the coefficient on each 10 cent dummies between \$9.50 and \$10.40 associated with equation (13), with firm fixed effects and the full sample (i.e., column 4) of Table 4. Consistent with the findings in the table, there is a general negative relationship between wages and separation, with a similar implied separations elasticity as the baseline specification above. However, there is no indication of a discontinuous drop in the separation rate going from \$9.90 to \$10.00. Similarly to the online there seem to be no discontinuity in labor supply at \$10.00. And most estimates rule out a high degree of left-digit bias.

While our primary analysis focuses on starting wages for new hires, we also exploit variation in wages that comes from subsequent raises. For instance, consider two individuals who both were hired at the same firm at date t earning the same wage under \$10, but between $t+1$ and $t+2$ one person got a raise that placed her above \$10, while the other's raise was not sufficient to cross that threshold. We could ask what that does to the separation probability between dates $t+2$ and $t+3$.

In Appendix Table A.1 column 1, we begin by reproducing our starting wage analysis for comparability focusing on those with 1 quarter of tenure, using firm-by-quarter fixed effects. Then, in columns 2 and 3, we restrict to those with 3 quarters of tenure and further control for their log wage (and indicator for wage exceeding \$10.00) when they were at tenure of 1 quarter. Now the tenure of 2 quarter log wage and jump at \$10.00 is identified only from the raise that occurs between the first and second quarter at the firm. Column 3 further controls for previous job-based characteristics. The θ in both specifications are negative in sign, statistically indistinguishable from zero, and rule out values larger than 0.13 at the 95% level. Workers who started with similar wages under \$10 were no less likely to quit when raises placed them above the \$10 threshold.

E. *Comparison with product market estimates*

A natural question is whether our experiments allow us to reject levels of left-digit bias estimated in the product market literature. As discussed above, Strulov-Shlain (2023) parameterizes left-digit bias in demand with θ, the inattention parameter. He specifies the empirical demand function to be $log(Q+1) = \beta^0 p + \beta^1 \lfloor p \rfloor$, and recovers $\theta = \frac{\beta^1}{\beta^1+\beta_0}$.

Overall, Strulov-Shlain (2023) finds a θ of 0.22 among customers at supermarket in his data, indicating a substantial amount of left-digit bias. (Interestingly, he finds that supermarkets do not sufficiently bunch prices at \$0.99 given the extent of consumer left-digit bias.). List et al. (2023), estimating similar specifications for Lyft riders with experimental variation, finds a lowest θ of 0.44. A non-product market comparison comes from Lacetera et al. (2012) who evaluate change in used car prices around 10,000 thresholds in odometer readings; they find an overall estimate of θ equal to 0.31.

Figure 9 presents comparable estimates of θ from our analysis using both the experimental variation and observational estimates. The figure draws from a select set of experimental estimates presented in Tables 2 and 3.

The individual experiments rule out a high degree of left-digit bias, and in all cases rule out a θ as large as 0.36; in some cases they rule out any θ larger than 0.11. Pooling across experiments greatly increases our power to more sharply bound the extent of left-digit bias through a more precise estimate of the labor supply elasticity. Pooling across the experiments 1 and 2, and estimating a common log wage coefficient, the 95% confidence rule out θ_{10} above 0.21 and θ_{100} above 0.22—right around the Retail estimate. The confidence intervals for the precision-weighted pooled θ rules out anything greater than 0.10, well under both product market comparisons (Strulov-Shlain, 2023; List et al., 2023).

As the figure shows, most pooled estimates across a range of specifications and samples rule out the key product market estimates of θ. The most precise estimate comes from pooling all data including from Horton et al. (2011) while estimating a single jump parameter and a single log wage coefficient, ruling out any θ greater than 0.09. Overall the experimental estimates are inconsistent with left-digit bias of a substantial magnitude.

Turning to separation response estimates from the Oregon data, our preferred specifications (with firm FE controls) rule out θ larger than 0.17 at the 95% confidence level. The least saturated specifications and no trimming on share bunched are less precise (and also less reliable for reasons explained above), but do rule out θ greater than 0.34 at the 90% confidence level.

On the whole, the weight of evidence from both the experimental and observational data do not detect comparable levels of left-digit bias found in the product market. An interesting direction for future research is exploring why left-digit bias is so much larger in product prices than hourly wages; we conjecture it is because any given product is a small share of expenditure, while hourly wages are a large share of income for most wage workers, and so rational inattention could account for the differences.[17] In contrast, in the context of a job, the stakes are larger for workers—making mistakes more costly. For example, Caldwell et al. (2025) shows that when individual-level bargaining is feasible, workers often engage in many bargaining rounds before accepting or rejecting an offer—even when the wage gains resulting in such bargaining are relatively small. This is consistent with a high stakes environment where behavioral biases on worker side may be uncommon, whether wages are set by bargaining or wage posting.

IV. Incorporating employer optimization frictions

Given the lack of evidence for worker left-digit bias, we now consider the other possible explanation for bunching: employer misoptimization.We extend the model to allow employers to "benefit" by paying a round number, despite lowered profits.[18] While consistent with employers preferring to pay round numbers, it could reflect internal fairness constraints or administrative costs internal to the firm. These could be transactions costs involved in dealing with round numbers, cognitive costs of managers, or administrative costs facing a bureaucracy.

The heterogeneity in firm bunching presented so far already points towards some misoptimization being important. Section I.D showed that bunching was more prevalent in firms less likely to survive, as well as experiencing slower employer growth and higher turnover, just as Reyes (2024) showed in Brazil. Further, we showed that firms in industries with lower quality management were more likely to have jobs bunched at \$10. All of these suggest that bunching reflects a constraint on firm maximization, with the resulting prediction that firm-performance ought to be lower when there is a high degree of bunching.

We parameterize employer mis-optimization by δ, measuring the willingness to forego profit in order to pay a round number.The presence of δ modifies the profit function to be:

$$\pi(w, p) = (p - w)l(w, p)e^{\delta \mathbf{1}_{w=w_0}} \tag{14}$$

where δ is the percentage "gain" in profits from paying the round number.[19] This specification parallels that in Chetty (2012), who restricts optimization frictions to be constant fractions of optimal consumer expenditure (in the nominal model), except applied to the employer's choice of wage for a job rather than a consumer's choice of a consumption-leisure bundle. In the taxable income model, optimization frictions parameterize the lack of responsiveness to tax incentives, while in our model they parameterize the willingness to forgo profits in order to pay a round number.

[17] Matejka and McKay (2015) provide foundations for discrete choice that incorporates inattention, and see Gabaix (2019) for applications of inattention to a wide variety of behavioral phenomena, including left-digit bias.

[18] It would be equivalent to assume that firms suffer an effective loss from not paying a round number.

[19] While we do not microfound why employers may have preferences for paying a particular round number, this may reflect inattention among wage-setters. For example, Matějka (2015) shows that rationally inattentive monopoly sellers will choose a discrete number of prices even when the profit-maximizing price is continuous.

Given (5) and (14), profits from paying a wage w to a workers with marginal product p can be written as:

$$\pi(w,p) = (p-w)\frac{w^{\eta}}{C}e^{\gamma \mathbb{1}_{w \geq w_0}}e^{\delta \mathbf{1}_{w=w_0}}k\left(p\right) = (p-w)l^{*}(w,p)e^{\gamma \mathbb{1}_{w \geq w_0}}e^{\delta \mathbf{1}_{w=w_0}} \tag{15}$$

which can be written in terms of the nominal model ρ as:

$$\delta + \gamma \mathbb{1}_{w^{*}(p)<w_0} > \ln \rho\left(w^{*}\left(p\right),p\right) - \ln \rho\left(w_0,p\right) \tag{16}$$

This shows that bunching is more likely the greater is the left-digit bias of workers and the optimization cost for employers. The optimization bias is symmetric whether the latent wage is above or below the round number. But left-digit bias is asymmetric because it only has an impact if the latent wage is below the round number. The right-hand side of (16) can be approximated using the following second-order Taylor series expansion of $\rho\left(w_0,p\right)$ about $w^{*}\left(p\right)$[20]:

$$\ln \rho\left(w_0,p\right) \simeq \ln \rho\left(w^{*},p\right) + \frac{\partial \ln \rho\left(w^{*},p\right)}{\partial w}\left[w_0 - w^{*}\right] + \frac{1}{2}\frac{\partial^2 \ln \rho\left(w^{*},p\right)}{\partial w^2}\left[w_0 - w^{*}\right]^2 \tag{17}$$

The first-order term is zero by the definition of the latent wage and the envelope theorem (Akerlof and Yellen (1985) use this idea to explain price and wage rigidity). Using (6) the second-order term can be written as:

$$\frac{\partial^2 \ln \rho\left(w^{*},p\right)}{\partial w^2} = -\frac{\eta\left(1+\eta\right)}{w^{*2}} \tag{18}$$

Substituting (18) into (17) and then into (16) leads to the following expression for whether a firm pays the round number:

$$\frac{1}{2}\left[\frac{w_0 - w^{*}}{w^{*}}\right]^2 \equiv \frac{\omega^2}{2} \leq \frac{\delta + \gamma \mathbb{1}_{w^{*}<w_0}}{\eta\left(1+\eta\right)} \tag{19}$$

The left-hand side of (19) implies that the size of the loss in nominal profits from bunching is increasing in the square of the proportional distance of the latent wage from the round number (ω). The right-hand side tells us that, for a given latent wage, whether a firm will bunch depends on the extent of left-digit bias as measured by γ (only relevant for wages below the round number), the extent of optimization frictions as measured by δ and the degree of competition in the labor market as measured by η. The extent of labor market competition matters because the loss in profits from a sub-optimal wage are greater the more competitive is the labor market.

In Online Appendix D we formally show that with $\gamma = 0$, the missing mass is distributed symmetrically from above and below w_0, which is consistent with our empirical bunching estimates as well as our experimental evidence on left-digit bias. Taking $\gamma = 0$ as given, our estimates of the proportion of firms who bunch for each latent wage identifies the CDF of $z_1 = \frac{\delta}{\eta(1+\eta)}$, but does not allow us to identify the distributions of δ and η separately, since that requires additional assumptions about their joint distribution. This simplest place to start is to assume a single value of η and a two part distribution of δ, equal to δ^{*} for bunchers and 0 for others. This yields an expression for the $\delta^{*} - \eta$ locus for a given extent of bunching as measured by ω:

[20]One can use the actual profit function, instead of the approximation, but the difference is small for the parameters we use, and the approximation has a clearer intuition.

$$\frac{\omega^2}{2} = \frac{\delta^*}{\eta\,(1+\eta)} \tag{20}$$

Figure 10 plots the δ^*, η locus as expressed in equation (20), for the sample width of the normalized bunching interval ω based on the pooled administrative data, as well as for the 90 percent confidence interval around it. We can see visually that as we consider higher values of δ^*, the range of admissible $\eta's$ increases and becomes larger in value. However, even for sizable δ^*'s, the implied values of the labor supply elasticity are often modest, implying at least a moderate amount of monopsony power.

A basic consistency check is whether plausible magnitudes of δ^*, within "economic standard errors" (Chetty, 2012) of perfect profit maximization, imply estimates of η that are within the range of existing monopsony estimates. We calculate estimates of η assuming δ^* equal to 0.01, 0.05 or 0.1 in rows A, B, and C of Appendix Table C.2 respectively. The implied labor supply elasticity η varies between 1.3 and 5.1 when we vary δ^* between 0.01 and 0.1. These are well within the range of recent estimates of η for the low-wage U.S. labor market from the literature (Sokolova and Sorensen, 2021; Bassier et al., 2022).

We also have direct estimates of η from our experiments and administrative data. η in the matched worker-firm Oregon data is in the 2 to 3 range, and implies a δ of between 0.02 and 0.04. The amount of money being lost by firms mis-pricing labor seems similar or somewhat smaller than other estimates from the recent behavioral IO literature (DellaVigna and Gentzkow (2019) and Strulov-Shlain (2023) find 10.2-11.7% and 1-4% of profits are lost due to employer mispricing in retail, respectively). Overall, the profit losses implied by the optimization friction based explanation for round number bunching are consistent with the evidence from other types of firm mis-optimization.

Returning to the heterogeneity in Section I.D, one might wonder whether the variation across types of firms is explained more by η or by δ. As an example, we can use our preferred specification above (with firm fixed effects) to estimate separation response and η by management score tercile, and can confirm that there is little heterogeneity in the degree of monopsony-power across industries by this variable: η is around 0.970 for the bottom quartile versus 0.995 for the top quartile.

Similarly the much higher bunching among small firms cannot be explained by differences in the elasticity of labor supply faced by employers. For instance, while the degree of bunching among firms with 50 or fewer workers is substantially higher than in the full sample, the separation elasticities are quite similar. For firms with 50 or fewer workers, the separation elasticity using our baseline specification is -2.047 (SE=0.198), while for the full sample, it is -2.163 (SE=0.203).

At the same time, the Figure 6 also provides an indication that bunching is more prevalent among firms with high AKM firm effects than for firms with low firm effects. In dynamic monopsony models like Burdett and Mortensen (1998), labor supply elasticity is likely to be lowest at firms near the top of the (residual) wage distribution, reflecting the relative scarcity of better-paying competitors that can poach workers. We empirically confirm that separation elasticities are smaller in magnitude for firms in the top quintile of AKM firm effects, as compred to the bottom quintile.[21] Better paying firms facing limited competition bear fewer consequences for deviating from their profit-maximizing wage.

Mechanical Turk data shows a stark version of the trade-off. In Appendix Figure C.4 we plot the excess and missing mass from Mechanical Turk. Compared to the low labor supply elasticity recovered from our experiments above, the ω implied by this figure (0.18), can be rationalized by a

[21]While the separation elasticity using our baseline specification is 1.214(SE=0.436) for the top quintile, it is 2.212(SE=0.286) for the bottom quintile.

very small δ (0.003). The substantial monopsony power employers have on MTurk means that only a small degree of misoptimization is required to rationalize the large degree of observed bunching.

In Online Appendix D, we additionally show how flexible distributions of η, γ, and δ can be recovered from the amount and symmetry of the missing mass around a round number. These results suggest that in contexts where only granular histograms of wages are available (e.g. restricted use administrative data or historical payroll records), researchers may be willing to make assumptions about firm mis-optimization in order to recover bounds on the degree of employer market power.

In Appendix Table C.4 we explore robustness to different assumptions on heterogeneity in δ and η, including allowing for arbitrary heterogeneity in δ, or heterogeneity in η but not δ, and finally a semi-parametric approach where we impose a lognormal distrbution on δ but an arbitrary non-parametric distribution on η. We also examine robustness of the estimates to alternative specifications of the latent distribution of wages in Appendix Table **C.5**. This includes controlling for secondary modes at 50 and 25 cent bins, and varying the polynomial used to estimate the latent distribution.

V. Conclusion

Significantly more U.S. workers are paid exactly round numbers than would be predicted by a smooth distribution of marginal productivity and the neoclassical model. We document this fact in administrative data, mitigating any issues due to measurement error, and is present even in Amazon MTurk, an online spot labor market, where there are no regulatory constraints nor long-term contracts.

Unlike in the product market, where numerous researchers have rationalized prices ending in 99 cents with "left-digit bias", we rule out worker inattention with two high-powered experiments on Amazon Mechanical Turk as well as observational date on separation elasticities from Oregon. We instead find evidence that this is consistent with employer misoptimization: bunching firms are smaller, less likely to survive, and have worse management. We also quantify the combinations of monopsony power and employer misoptimization necessary to rationalize any degree of bunching: only moderate amounts of misoptimization are required to explain observed bunching given estimates of monopsony.

Finally, as we show in Online Appendix G, round number bunching of wages is relevant for a number of labor market policies. Inertia in monopsonistic wage-setting implies that minimum wage changes will have larger spillover effects when they cross a round number, and that small payroll tax increases will be borne solely by firms. Pervasive monopsony gives room for a variety of "satisficing" firm-level wage policies to persist in equilibrium, and our estimates suggest that firms may be leaving significant sums of money on the table.

REFERENCES

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

Akerlof, George A and Janet L Yellen, "A near-rational model of the business cycle, with wage and price inertia," *The Quarterly Journal of Economics*, 1985, *100* (Supplement), 823–838.

Allen, Eric J, Patricia M Dechow, Devin G Pope, and George Wu, "Reference-dependent preferences: Evidence from marathon runners," *Management Science*, 2016.

Anderson, Eric T and Duncan I Simester, "Effects of $9 price endings on retail sales: Evidence from field experiments," *Quantitative marketing and Economics*, 2003, *1* (1), 93–110.

Anderson, Patricia M and Bruce D Meyer, "The effects of the unemployment insurance payroll tax on wages, employment, claims and denials," *Journal of public Economics*, 2000, *78* (1-2), 81–106.

Andrews, Isaiah, Matthew Gentzkow, and Jesse M Shapiro, "Measuring the Sensitivity of Parameter Estimates to Estimation Moments," *The Quarterly Journal of Economics*, 2017, pp. 1553–1592.

Azar, José, Ioana Marinescu, and Marshall Steinbaum, "Labor market concentration," *Journal of Human Resources*, 2022, *57* (S), S167–S199.

Babcock, Linda, William J Congdon, Lawrence F Katz, and Sendhil Mullainathan, "Notes on behavioral economics and labor market policy," *IZA Journal of Labor Policy*, 2012, *1* (1), 2.

Backus, Matthew, Thomas Blake, and Steven Tadelis, "On the empirical content of cheap-talk signaling: An application to bargaining," *Journal of Political Economy*, 2019, *127* (4), 1599–1628.

Bassier, Ihsaan, Arindrajit Dube, and Suresh Naidu, "Monopsony in movers: The elasticity of labor supply to firm wage policies," *Journal of Human Resources*, 2022, *57* (S), S50–s86.

Basu, Kaushik, "Why are so many goods priced to end in nine? And why this practice hurts the producers," *Economics Letters*, 1997, *54* (1), 41–44.

_ , "Consumer cognition and pricing in the nines in oligopolistic markets," *Journal of Economics & Management Strategy*, 2006, *15* (1), 125–141.

Berger, David, Kyle Herkenhoff, and Simon Mongey, "Labor market power," *American Economic Review*, 2022, *112* (4), 1147–1193.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan, "Gender identity and relative income within households," *The Quarterly Journal of Economics*, 2015, *130* (2), 571–614.

Blomquist, Sören, Whitney K Newey, Anil Kumar, and Che-Yuan Liang, "On bunching and identification of the taxable income elasticity," *Journal of Political Economy*, 2021, *129* (8), 2320–2343.

Bloom, Nicholas and John Van Reenen, "Measuring and explaining management practices across firms and countries," *The Quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

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

Butcher, Tim, Richard Dickens, and Alan Manning, "Minimum wages and wage inequality: some theory and an application to the UK," Technical Report, Centre for Economic Performance, London School of Economics and Political 2012.

Caldwell, Sydnee and Emily Oehlsen, "Monopsony and the Gender Wage Gap: Experimental Evidence from the Gig Economy," *Massachusetts Institute of Technology Working Paper*, 2018.

__ **, Ingrid Haegele, and Jörg Heining**, "Bargaining and inequality in the labor market," Technical Report, National Bureau of Economic Research 2025.

Chen, Daniel L and John J Horton, "Are Online Labor Markets Spot Markets for Tasks? A Field Experiment on the Behavioral Response to Wage Cuts," *Information Systems Research*, 2016, *27* (2), 403–423.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins, "Double/debiased machine learning for treatment and structural parameters," *The Econometrics Journal*, 2017.

Chetty, Raj, "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply," *Econometrica*, 2012, *80* (3), 969–1018.

Cho, David, "The Labor Market Effects of Demand Shocks: Firm-Level Evidence from the Recovery Act," Technical Report, Princeton University 2018.

Cho, Sungjin and John Rust, "The flat rental puzzle," *The Review of Economic Studies*, 2010, *77* (2), 560–594.

Conlon, Christopher T and Nirupama L Rao, "Discrete prices and the incidence and efficiency of excise taxes," *American Economic Journal: Economic Policy*, 2020, *12* (4), 111–143.

Delaigle, A and Irène Gijbels, "Bootstrap bandwidth selection in kernel density estimation from a contaminated sample," *Annals of the Institute of Statistical Mathematics*, 2004, *56* (1), 19–47.

DellaVigna, Stefano and Devin Pope, "What motivates effort? Evidence and expert forecasts," *The Review of Economic Studies*, 2018, *85* (2), 1029–1069.

__ **and Matthew Gentzkow**, "Uniform pricing in us retail chains," *The Quarterly Journal of Economics*, 2019, *134* (4), 2011–2084.

__ **, John A List, and Ulrike Malmendier**, "Testing for altruism and social pressure in charitable giving," *The Quarterly Journal of economics*, 2012, *127* (1), 1–56.

Derenoncourt, Ellora and David Weil, "Voluntary Minimum Wages: The Local Labor Market Effects of National Retailer Policies," *The Quarterly Journal of Economics*, 2025, p. qjaf016.

Dittmar, Jeremiah and Suresh Naidu, "Contested Property: Fugitive Slaves in the Antebellum US South," Technical Report, Working paper 2016.

Dube, Arindrajit, Alan Manning, and Suresh Naidu, "Data and code for: Monopsony and employer mis-optimization explain why wages bunch at round numbers," 2025. Available at https://www.openicpsr.org/openicpsr/aea.

__ **, Jeff Jacobs, Suresh Naidu, and Siddharth Suri**, "Monopsony in online labor markets," *American Economic Review: Insights*, 2020, *2* (1), 33–46.

Gabaix, Xavier, "Behavioral inattention," in "Handbook of behavioral economics: Applications and foundations 1," Vol. 2, Elsevier, 2019, pp. 261–343.

__ **and David Laibson**, "Shrouded attributes, consumer myopia, and information suppression in competitive markets," *The Quarterly Journal of Economics*, 2006, *121* (2), 505–540.

Garin, Andrew and Filipe Silvério, "How responsive are wages to firm-specific changes in labour demand? Evidence from idiosyncratic export demand shocks," *Review of Economic Studies*, 2024, *91* (3), 1671–1710.

Goldfarb, Avi and Mo Xiao, "Who thinks about the competition? Managerial ability and strategic entry in US local telephone markets," *The American Economic Review*, 2011, *101* (7), 3130–3161.

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

Heidhues, Paul and Botond Kőszegi, "Competition and price variation when consumers are loss averse," *The American Economic Review*, 2008, *98* (4), 1245–1268.

— **and** — , "Behavioral industrial organization," *Handbook of Behavioral Economics: Applications and Foundations 1*, 2018, *1*, 517–612.

Hortacsu, Ali and Steven L Puller, "Understanding strategic bidding in multi-unit auctions: a case study of the Texas electricity spot market," *The RAND Journal of Economics*, 2008, *39* (1), 86–114.

Horton, John J, David G Rand, and Richard J Zeckhauser, "The online laboratory: Conducting experiments in a real labor market," *Experimental Economics*, 2011, *14* (3), 399–425.

Jones, Edward D, "Round Numbers in Wages and Prices," *Publications of the American Statistical Association*, 1896, *5* (35-36), 111–141.

Kingsley, Sara Constance, Mary L Gray, and Siddharth Suri, "Accounting for market frictions and power asymmetries in online labor markets," *Policy & Internet*, 2015, *7* (4), 383–400.

Kleven, Henrik J and Mazhar Waseem, "Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan," *The Quarterly Journal of Economics*, 2013, *128* (2), 669–723.

Kleven, Henrik Jacobsen, "Bunching," *Annual Review of Economics*, 2016, *8*, 435–464.

Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar, "Who profits from patents? rent-sharing at innovative firms," *The Quarterly Journal of Economics*, 2019, *134* (3), 1343–1404.

Kuziemko, Ilyana, Michael I Norton, Emmanuel Saez, and Stefanie Stantcheva, "How elastic are preferences for redistribution? Evidence from randomized survey experiments," *The American Economic Review*, 2015, *105* (4), 1478–1508.

Lacetera, Nicola, Devin G Pope, and Justin R Sydnor, "Heuristic thinking and limited attention in the car market," *American Economic Review*, 2012, *102* (5), 2206–2236.

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.

Lester, Richard A, "Wage diversity and its theoretical implications," *The Review of Economics and Statistics*, 1946, *28* (3), 152–159.

Levy, Daniel, Dongwon Lee, Haipeng Chen, Robert J Kauffman, and Mark Bergen, "Price points and price rigidity," *Review of Economics and Statistics*, 2011, *93* (4), 1417–1431.

List, John A, Ian Muir, Devin Pope, and Gregory Sun, "Left-digit bias at lyft," *Review of Economic Studies*, 2023, *90* (6), 3186–3237.

Manning, Alan, “Imperfect competition in the labor market,” *Handbook of labor economics*, 2011, *4*, 973–1041.

Manning, Kenneth C and David E Sprott, “Price endings, left-digit effects, and choice,” *Journal of Consumer Research*, 2009, *36* (2), 328–335.

Matějka, Filip, “Rationally inattentive seller: Sales and discrete pricing,” *The Review of Economic Studies*, 2015, *83* (3), 1125–1155.

Matejka, Filip and Alisdair McKay, “Rational inattention to discrete choices: A new foundation for the multinomial logit model,” *American Economic Review*, 2015, *105* (1), 272–98.

Pope, Devin G, Jaren C Pope, and Justin R Sydnor, “Focal points and bargaining in housing markets,” *Games and Economic Behavior*, 2015, *93*, 89–107.

Reyes, Germán, “Coarse Wage-Setting and Behavioral Firms,” *Review of Economics and Statistics*, 2024, pp. 1–41.

Risch, Max, “Does taxing business owners affect employees? Evidence from a change in the top marginal tax rate,” *The Quarterly Journal of Economics*, 2024, *139* (1), 637–692.

Romer, David, “Do firms maximize? Evidence from professional football,” *Journal of Political Economy*, 2006, *114* (2), 340–365.

Saez, Emmanuel, “Do taxpayers bunch at kink points?,” *American Economic Journal: Economic Policy*, 2010, *2* (3), 180–212.

Shaw, Aaron D, John J Horton, and Daniel L Chen, “Designing incentives for inexpert human raters,” in “Proceedings of the ACM 2011 conference on Computer supported cooperative work” ACM 2011, pp. 275–284.

Snir, Avichai, Daniel Levy, Alex Gotler, and Haipeng Allan Chen, “Not all price endings are created equal: price points and asymmetric price rigidity,” 2012.

Sokolova, Anna and Todd Sorensen, “Monopsony in labor markets: A meta-analysis,” *ILR Review*, 2021, *74* (1), 27–55.

Stefanski, Leonard A and Raymond J Carroll, “Deconvolving kernel density estimators,” *Statistics*, 1990, *21* (2), 169–184.

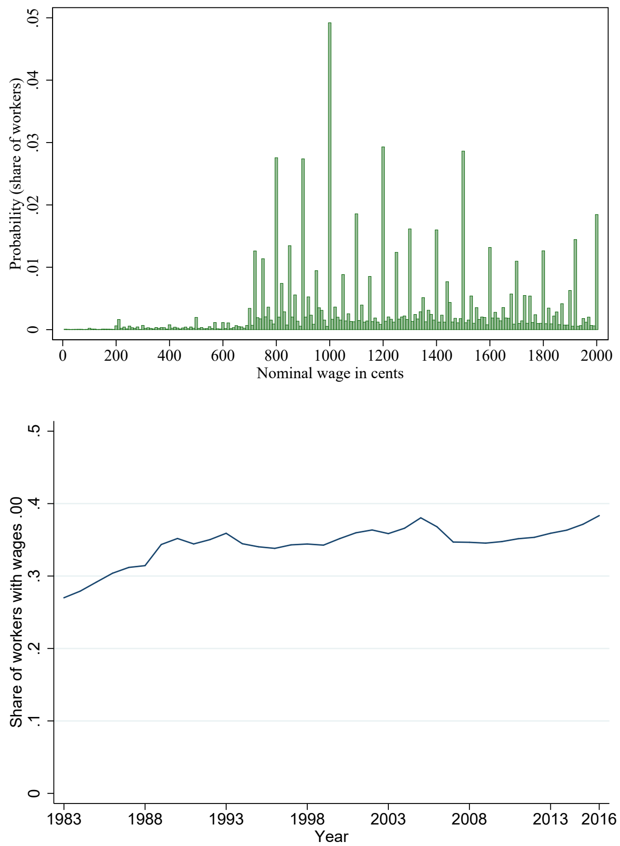
Strulov-Shlain, Avner, “More than a penny’s worth: Left-digit bias and firm pricing,” *Review of Economic Studies*, 2023, *90* (5), 2612–2645.

Thomas, Manoj and Vicki Morwitz, “Penny wise and pound foolish: the left-digit effect in price cognition,” *Journal of Consumer Research*, 2005, *32* (1), 54–64.

Wang, Xiao-Feng and Bin Wang, “Deconvolution Estimation in Measurement Error Models: The R Package Decon,” *Journal of Statistical Software*, 2011, *39* (10), i10.

Webber, Douglas A, “Firm-Level Monopsony and the Gender Pay Gap,” *Industrial Relations: A Journal of Economy and Society*, 2016, *55* (2), 323–345.

Zwick, Eric, “The costs of corporate tax complexity,” *American Economic Journal: Economic Policy*, 2021, *13* (2), 467–500.



figures/NumberStates10_and_15_Mode_to2023.pdf

FIGURE 1. PREVALENCE OF ROUND NOMINAL WAGES IN THE CPS

Notes: The top figure shows the CPS hourly nominal wage distribution, pooled between 2010 and 2016, in 10 cent bins. The middle figure shows the fraction of hourly wages in the CPS that end in .00 from 2003 through 2016. The bottom figure shows the fraction of states with $10.00 and $15.00 modal wages in the CPS. We exclude imputed wages.

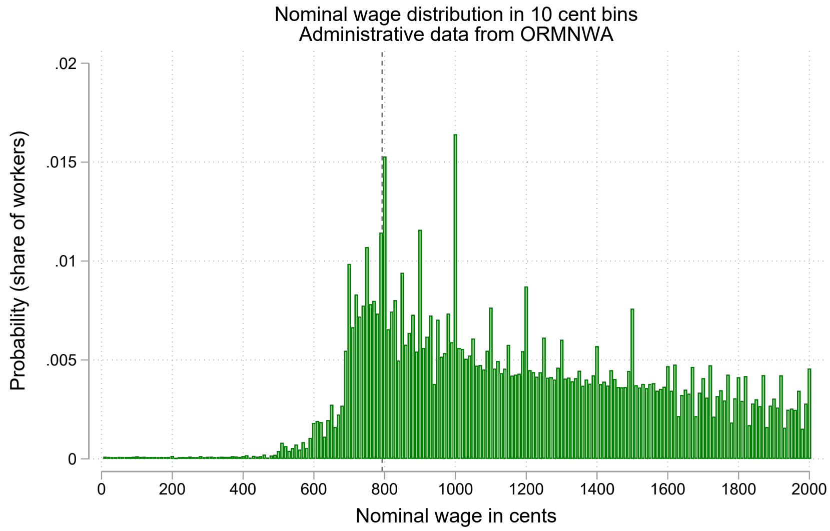


FIGURE 2. HISTOGRAM OF HOURLY WAGES IN POOLED ADMINISTRATIVE PAYROLL DATA FROM MINNESOTA, OREGON, AND WASHINGTON, 2003-2007

Notes: The figure shows a histogram of hourly wages in 10 cents (nominal) wage bins, averaged over 2003q1 to 2007q4, using pooled administrative Unemployment Insurance payroll records from the states of Oregon, Minnesota and Washington. Hourly wages are constructed by dividing quarterly earnings by the total number of hours worked in the quarter. The counts in each bin are normalized by dividing by total employment in that state, averaged over the sample period. The UI payroll records cover over 95% of all wage and salary civilian employment in the states. The vertical line is the highest minimum wage in the sample (7.93 in WA 2007). The counts here exclude NAICS 6241 and 814, home-health and household sectors, which were identified by the state data administrators as having substantial reporting errors.

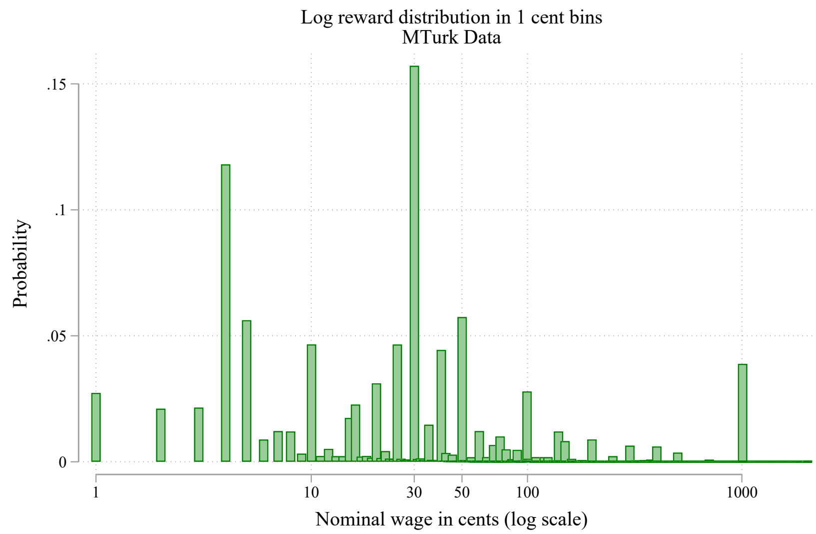


FIGURE 3. BUNCHING IN TASK REWARDS IN ONLINE LABOR MARKETS - MTURK

Notes: The figure shows a histogram of posted rewards by $0.01 (nominal) bins scraped from MTurk. The x-axis is in log scale for ease of visualization. The sample represents all posted rewards on MTurk between May 01, 2014 and September 3, 2016.

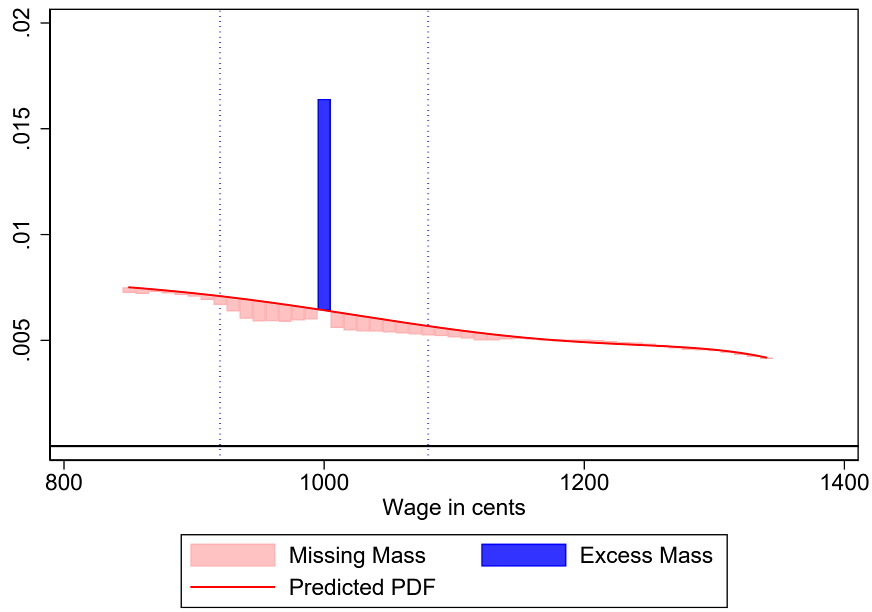


FIGURE 4. EXCESS BUNCHING AND MISSING MASS AROUND $10.00 USING ADMINISTRATIVE DATA ON HOURLY WAGES (MN, OR, WA)

Notes: The reported estimates of excess bunching at $10.00, and missing mass in the interval around $10.00 as compared to the smoothed predicted probability density function, using administrative hourly wage counts from OR, MN and WA, aggregated by 10 cents bins, over the 2003q1-2007q4 period. The darker shaded blue bar at $10.00 represents the excess mass, while the lighter red shaded region represents the missing mass. The dotted lines represent the estimated interval from which the missing mass is drawn. The predicted PDF is estimated using a sixth order polynomial, with dummies for each 10 cents bin in the interval from which the missing mass is drawn. The width of the interval is chosen by iteratively expanding the interval until the missing and excess masses are equal, as described in the text.

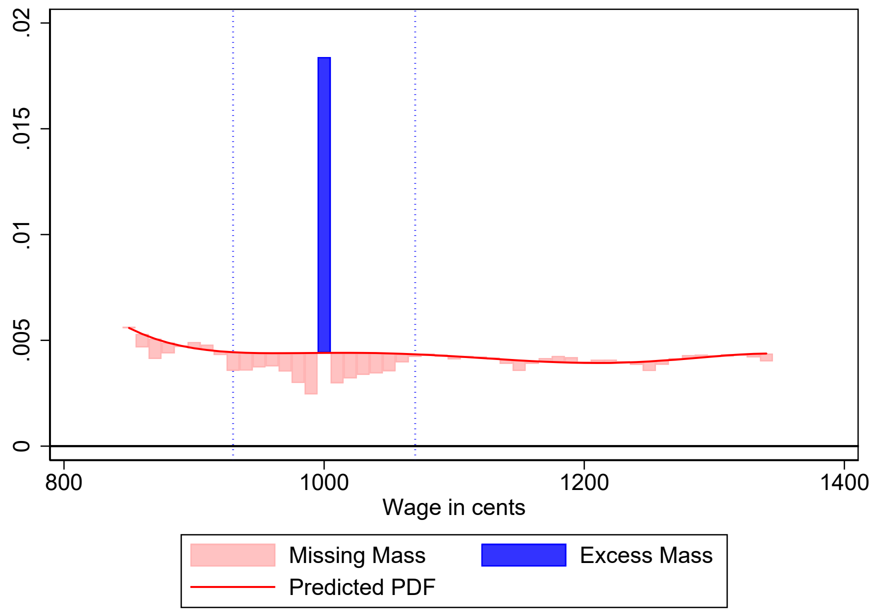


FIGURE 5. EXCESS BUNCHING AND MISSING MASS AROUND $10.00 USING MEASUREMENT ERROR CORRECTED CPS DATA

Notes: The reported estimates of excess bunching at $10.00, and missing mass in the interval around $10.00 as compared to the smoothed predicted probability density function, using CPS data for MN, OR, and WA, corrected for measurement error using the 1977 administrative supplement. The darker shaded blue bar at $10.00 represents the excess mass, while the lighter red shaded region represents the missing mass. The dotted lines represent the estimated interval from which the missing mass is drawn. The predicted PDF is estimated using a sixth order polynomial, with dummies for each 10 cents bin in the interval from which the missing mass is drawn. The width of the interval is chosen by iteratively expanding the interval until the missing and excess masses are equal, as described in the text.

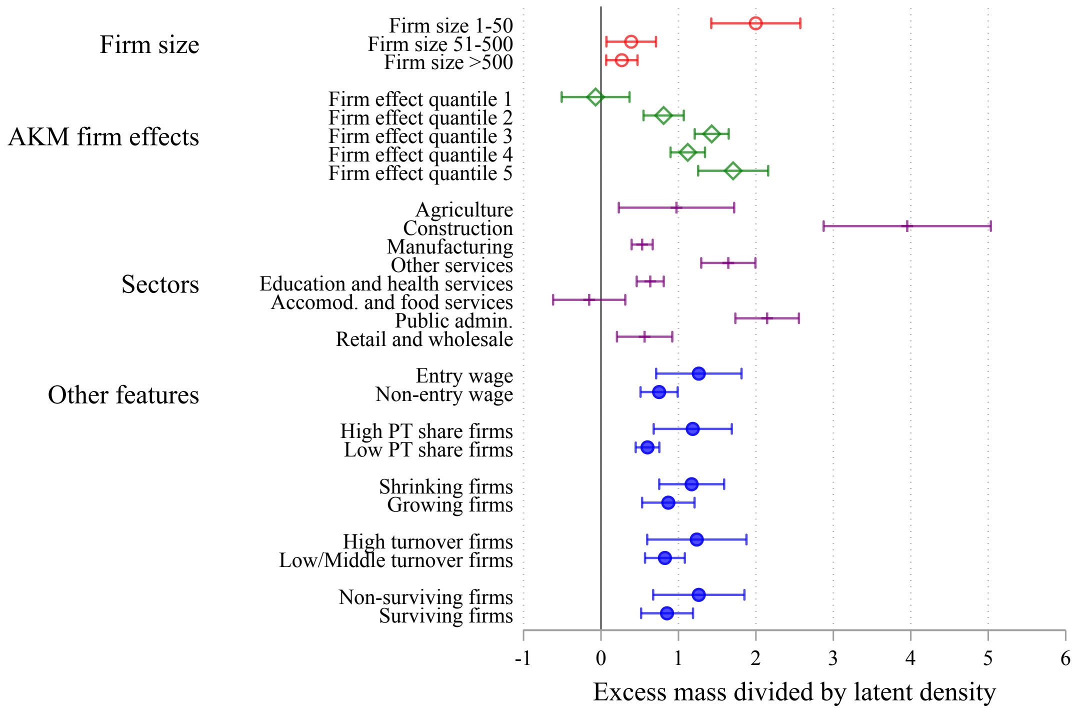
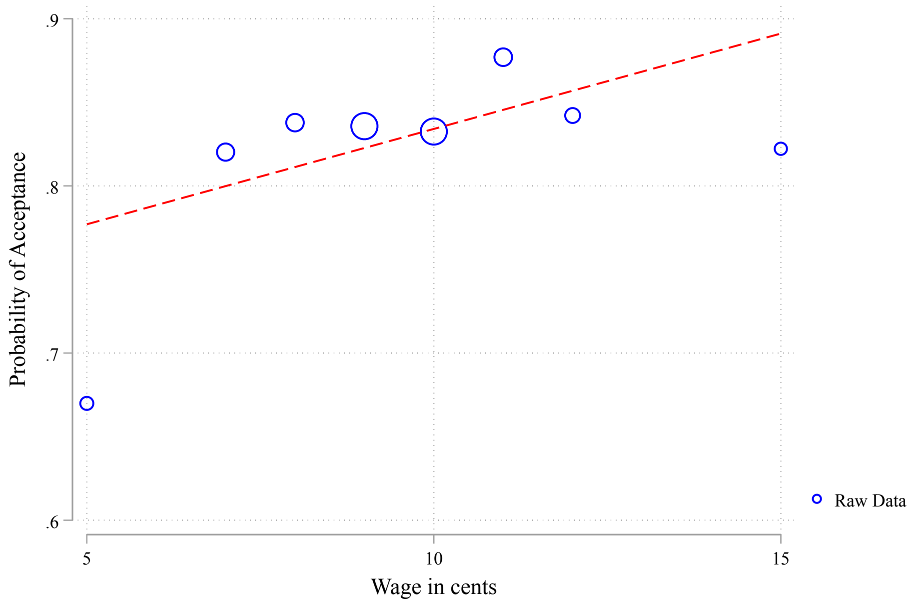


FIGURE 6. HETEROGENEITY IN BUNCHING BY OREGON FIRM CHARACTERISTICS

Notes: The figure plots the extent of bunching (excess mass at $10 divided by latent density at that wage) for different subgroups. An estimate greater than 0 indicates bunching. AKM firm effects refer to deciles of firm effects estimated via Abowd et al. (1999). PT refers to part-time workers (less than 20 hours); high PT share firms are those in the top quartile. High turnover firms have separation rates in the top quartile. Shrinking firms saw employment fall over the sample (2003-2007), while growing firms saw employment weakly increase. Non-surviving firms are those that were no longer in the dataset by 2007q4. 95 percent confidence intervals are based on standard errors clustered at the wage bin level.



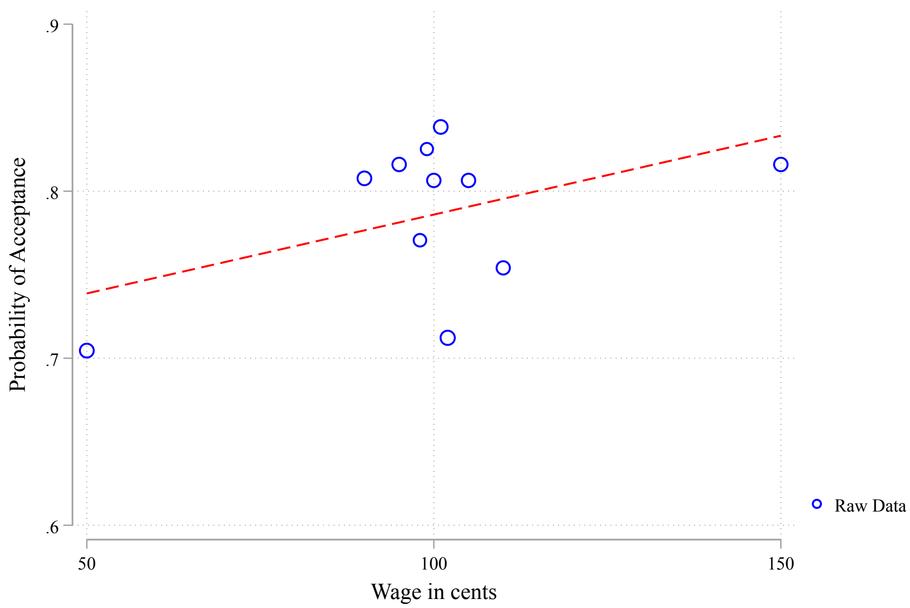


FIGURE 7. DISTRIBUTION OF RANDOMIZED REWARDS IN THE MTURK EXPERIMENT, AND RESULTING PROBABILITY OF TASK ACCEPTANCE

Notes: The figure shows the percentage change in probability of accepting the bonus task as a function of the wage for both experiments, with the 2018 experiment on top and the 2022 experiment below. Slope of the fitted line is estimated from the pooled data, with a jump at the discontinuity, with no controls. Dots are scaled proportionally to the number of observations.

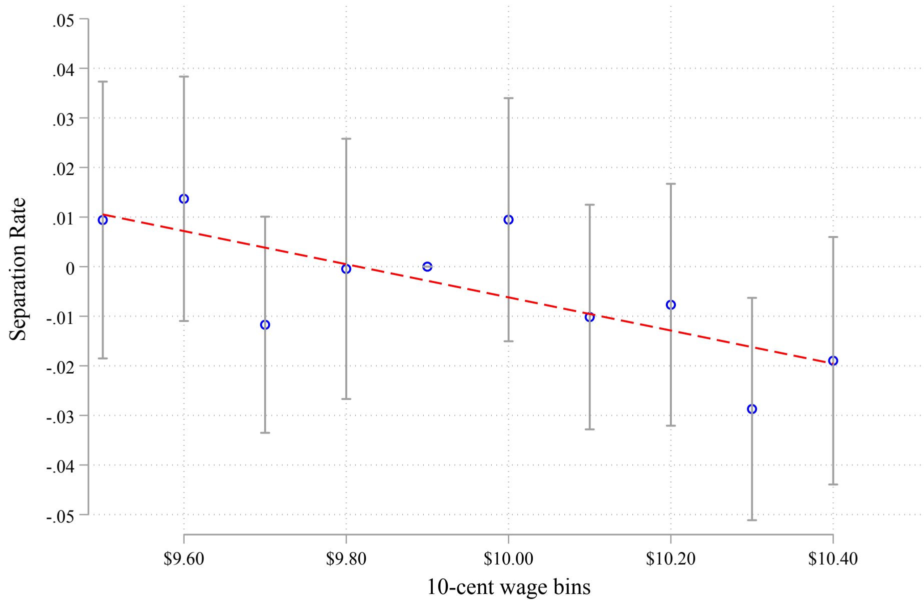


FIGURE 8. OREGON SEPARATION RESPONSE BY WAGE BIN

Notes: Estimated from quarterly matched worker-establishment data from Oregon 2003-2007. Sample includes all "full quarter hires'–i.e., with 1 quarter tenure at their new job–who were employed at a different company 2 quarters prior, whose hourly wages fall between $9.50 and $10.59, and whose first quarter hours exceeded the 25th percentile (184 hours). The outcome is 4-quarter separation which takes on 1 if the person leaves their job within 4 quarters of start date. 4-quarter separation rate with blue circles are coefficient on 10 cents wage bin dummies relative to $9.90 (the bin right below $10.00), and control for fully saturated interactions of quarter of hire and firm fixed effects, as well as: log hours and log wage at the last job 2 quarters prior. Robust standard errors are clustered at the 1-cent wage bin level.

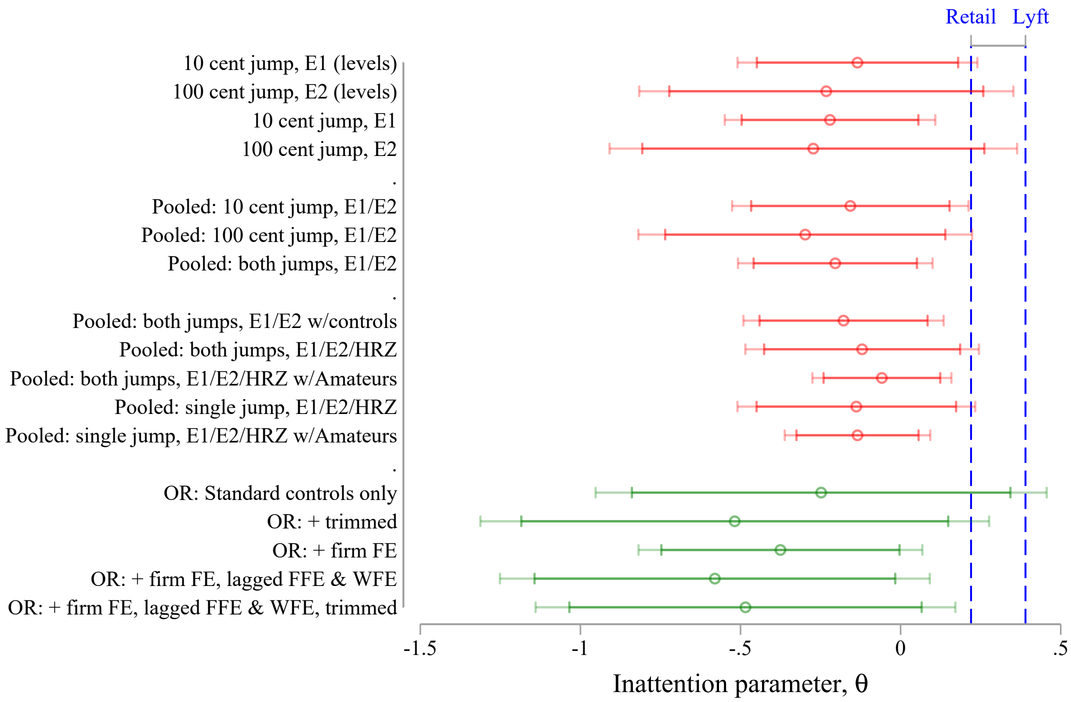


FIGURE 9. COMPARISON WITH PRODUCT MARKET INATTENTION PARAMETERS

Notes: shows estimates and confidence intervals (95% and 90%) of the attention parameter θ. The red points correspond to specifications estimated on the Mturk experiments from Table 2. E1 indicates the 2018 experiment testing the discontinuity at 10 cents, while E2 indicates the 2022 experiment testing the discontinuity at $1.00. HRZ denotes that data from the MTurk experiment conducted in Horton et al. (2011) is included. Unless âAmateurs" is specified all specifications are estimated on the sample of "sophisticates": Turkers who respond that they work more than 10 hours a week and their primary motivation is money. The âpooled, both jumps" specifications estimate separate jump coefficients at 10 cents and $1.00 discontinuities, and then take a precision-weighted average of the two estimates of θ. The âpooled, single jump" specifications impose a single coefficient on both discontinuities, and calculate θ under the assumption of constant η and constant γ across both experiments. The green points show estimates from the administrative matched worker-firm data in Oregon, under progressively finer specifications as in Table 4 The vertical lines correspond to estimates of the inattention parameter from retail in Strulov-Shlain (2023) and from Lyft in List et al. (2023).

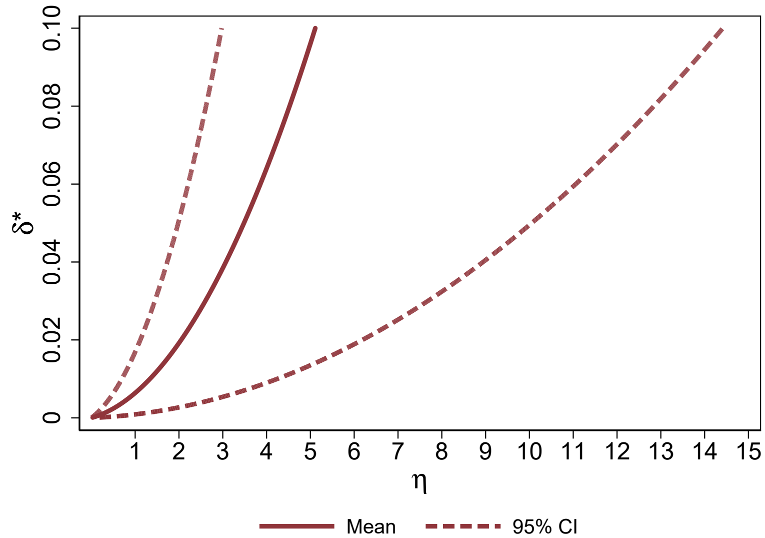


FIGURE 10. RELATIONSHIP BETWEEN LABOR SUPPLY ELASTICITY (η) , OPTIMIZATION FRICTIONS (δ) AND SIZE OF BUNCHING (ω): ADMINISTRATIVE HOURLY WAGE DATA FROM MN, OR, AND WA

Notes: The solid, red, upward sloping line shows the locus of the labor supply elasticity η and optimization frictions $\delta^* = E[\delta|\delta > 0]$ consistent with the extent of bunching ω estimated using the administrative hourly wage data from MN, OR, and WA between 2003q1-2007q4, as described in equation 25 in the paper. The dashed lines are the 95 percent confidence intervals estimated using 500 bootstrap replicates.

TABLE 1—ESTIMATES FOR EXCESS BUNCHING, MISSING MASS, AND INTERVAL AROUND THRESHOLD

	(1)	(2)	(3)	(4)
Value of w_0	\$10.00	\$10.00	\$10.00	\$10.00
Excess mass at w_0	0.010	0.034	0.014	0.041
	(0.002)	(0.008)	(0.004)	(0.009)
Total missing mass	-0.013	-0.043	-0.016	-0.033
	(0.009)	(0.034)	(0.017)	(0.047)
Missing mass below	-0.007	-0.024	-0.009	-0.019
	(0.007)	(0.023)	(0.010)	(0.029)
Missing mass above	-0.006	-0.019	-0.007	-0.014
	(0.008)	(0.029)	(0.013)	(0.036)
Test of equality of missing shares of latent $<$ and $>$ 0: t-statistic	-0.338	-0.097	-0.065	-0.151
Bunching = $\frac{Actual\ mass}{Latent\ density}$	2.555	6.547	4.172	8.394
	(0.343)	(6.476)	(1.772)	(7.405)
w_L	\$9.20	\$9.30	\$9.30	\$9.30
w_H	\$10.80	\$10.70	\$10.70	\$10.70
$\omega = \frac{(w_H - w_0)}{w_0}$	0.080	0.070	0.070	0.070
	(0.026)	(0.033)	(0.034)	(0.033)
Data:	Admin OR & MN & WA	CPS-Raw OR & MN & WA	CPS-MEC OR & MN & WA	CPS-Raw

Notes: The table reports estimates of excess bunching at threshold w_0, missing mass in the interval around w_0 as compared to the smoothed predicted probability density function, and the interval (ω_L, ω_H) from which the missing mass is drawn. It also reports the t-statistic for the null hypothesis that the size of the missing mass to the left of w_0 is equal to the size of the missing mass to the right. The predicted PDF is estimated using a sixth order polynomial, with dummies for each bin in the interval from which the missing mass is drawn. The width of the interval is chosen by iteratively expanding the interval until the missing and excess masses are equal, as described in the text. In columns 1-3, estimates are shown for bunching at \$10.00 from pooled MN, OR, and WA using the administrative hourly wage counts, the raw CPS data, and measurement error corrected CPS (CPS-MEC) over the 2003q1-2007q4 period. Bootstrap standard errors based on 500 draws are in parentheses.

TABLE 2—TASK ACCEPTANCE PROBABILITY BY OFFERED TASK REWARD ON MTURK - INDIVIDUAL EXPERIMENTS

	(1)	(2)	(3)	(4)
Log Wage	0.178	0.111		
	(0.0633)	(0.0557)		
Wage			0.0147	0.00113
			(0.00688)	(0.000619)
Jump at 10 cents	-0.0321		-0.0174	
	(0.0270)		(0.0272)	
Jump at 1.00		-0.0237		-0.0214
		(0.0263)		(0.0266)
DV Mean	0.828	0.786	0.828	0.786
$\theta(10)$	-0.220		-0.135	
SE	(0.168)		(0.191)	
$\theta(100)$		-0.273		-0.232
SE		(0.325)		(0.298)
N	1702	1363	1702	1363
Spec.	E1	E2	E1-levels	E2-levels
Sample	Soph.	Soph.	Soph.	Soph.
Controls	N	N	N	N

Notes: E1 indicates the 2018 experiment testing the discontinuity at 10 cents, while E2 indicates the 2022 experiment testing the discontinuity at $1.00. For columns 1 and 2, the wage variable is in logs, while for columns 3 and 4, it is in levels. The sample of sophisticates is restricted to Turkers who respond that they work more than 10 hours a week and their primary motivation is money. Robust standard errors in parentheses

TABLE 3—TASK ACCEPTANCE PROBABILITY BY OFFERED TASK REWARD ON MTURK - POOLED EXPERIMENTS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Wage	0.140	0.132	0.100	0.0703	0.108	0.141	0.108
	(0.0419)	(0.0411)	(0.0246)	(0.0273)	(0.0151)	(0.0417)	(0.0151)
Jump at 10 cents	-0.0190	-0.0143	-0.00526	-0.0102	-0.0231		
	(0.0225)	(0.0225)	(0.0198)	(0.0145)	(0.0124)		
Jump at 1.00	-0.0321	-0.0300	-0.0207	0.0163	0.00562		
	(0.0246)	(0.0236)	(0.0230)	(0.0175)	(0.0164)		
Jump (Pooled)						-0.0252	-0.0128
						(0.0186)	(0.0104)
DV Mean	0.809	0.811	0.804	0.777	0.774	0.809	0.774
θ	-0.204	-0.182	-0.112	0.0462	-0.0777	-0.217	-0.135
SE	(0.155)	(0.159)	(0.185)	(0.171)	(0.110)	(0.156)	(0.116)
$\theta(10)$	-0.157	-0.122	-0.0555	-0.170	-0.271		
SE	(0.188)	(0.194)	(0.215)	(0.242)	(0.171)		
$\theta(100)$	-0.298	-0.295	-0.260	0.188	0.0493		
SE	(0.266)	(0.268)	(0.351)	(0.196)	(0.139)		
N	3065	3018	3226	8010	8430	3065	8430
Spec.	E1/E2	E1/E2	E1/E2/HRZ	E1/E2	E1/E2/HRZ	E1/E2	E1/E2/HRZ
Sample	Soph.	Soph.	Soph.	All	All	Soph.	All
Controls	N	Y	N	N	N	N	N

Notes: E1 indicates the 2018 experiment testing the discontinuity at 10 cents, while E2 indicates the 2022 experiment testing the discontinuity at $1.00. HRZ indicates the MTurk experiment conducted by Horton et al. (2011). The sample of sophisticates is restricted to Turkers who respond that they work more than 10 hours a week and their primary motivation is money. Pooled estimates of θ are precision-weighted averages of jump-specific θ, except in columns 6 and 7, where a common jump coefficient is imposed. Robust standard errors in parentheses

TABLE 4—OREGON MATCHED PANEL ESTIMATES: SEPARATION RESPONSE TO STARTING WAGE AROUND $10.00/HOUR JOBS

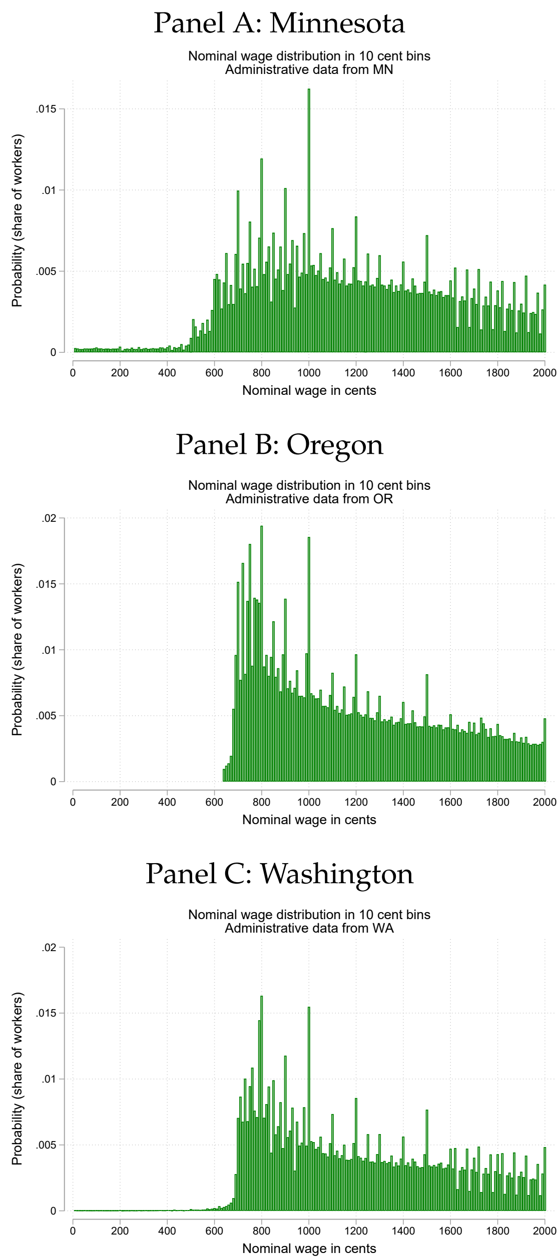
	(1)	(2)	(3)	(4)	(5)	(6)
Log wage	-0.381	-0.455	-0.615	-0.564	-0.425	-0.426
	(0.087)	(0.208)	(0.293)	(0.170)	(0.171)	(0.179)
Jump at 10.00		0.009	0.021	0.015	0.016	0.014
		(0.014)	(0.020)	(0.010)	(0.011)	(0.011)
θ		-0.248	-0.518	-0.375	-0.580	-0.484
		(0.359)	(0.405)	(0.226)	(0.342)	(0.334)
η	1.650	1.971	2.606	2.440	1.838	1.797
	(0.374)	(0.897)	(1.225)	(0.736)	(0.738)	(0.755)
Obs	107631	107623	71362	63785	63765	53163
Worker controls		Y	Y	Y	Y	Y
Firm controls		Y	Y			
Trimmed sample (round share)			Y			Y
Firm FE				Y	Y	Y
Lag WFE FFE value					Y	Y

Notes: Sample includes all âfull quarter hiresââi.e., with 1 quarter tenure at their new jobâwho were employed at a different company 2 quarters prior, whose hourly wages fall between $9.50 and $10.59, and whose first quarter hours exceeded the 25th percentile (184 hours). The outcome is 4-quarter separation which takes on 1 if the person leaves their job within 4 quarters of the start date. The mean 4-quarter separation rate in this sample is 0.49. Standard controls include: quarter fixed effects, log hours and log wage at the last job 2 quarters prior, quintiles in the share of jobs bunched at exactly $10.00 at the current employer, and 1-digit NAICS industry dummy. Firm FE indicates a fixed effect for the current employer, interacted with quarter dummies. The trimmed sample excludes workers at firms where very few (less than 5%) or very high (greater than 95%) of jobs in this range are bunched at $10.00. η is the elasticity of labor supply facing the firm. θ is the inattention parameter capturing left-digit bias. Robust standard errors are clustered at the 1-cent wage bin level.

Online Appendix A.
Additional Figures and Tables

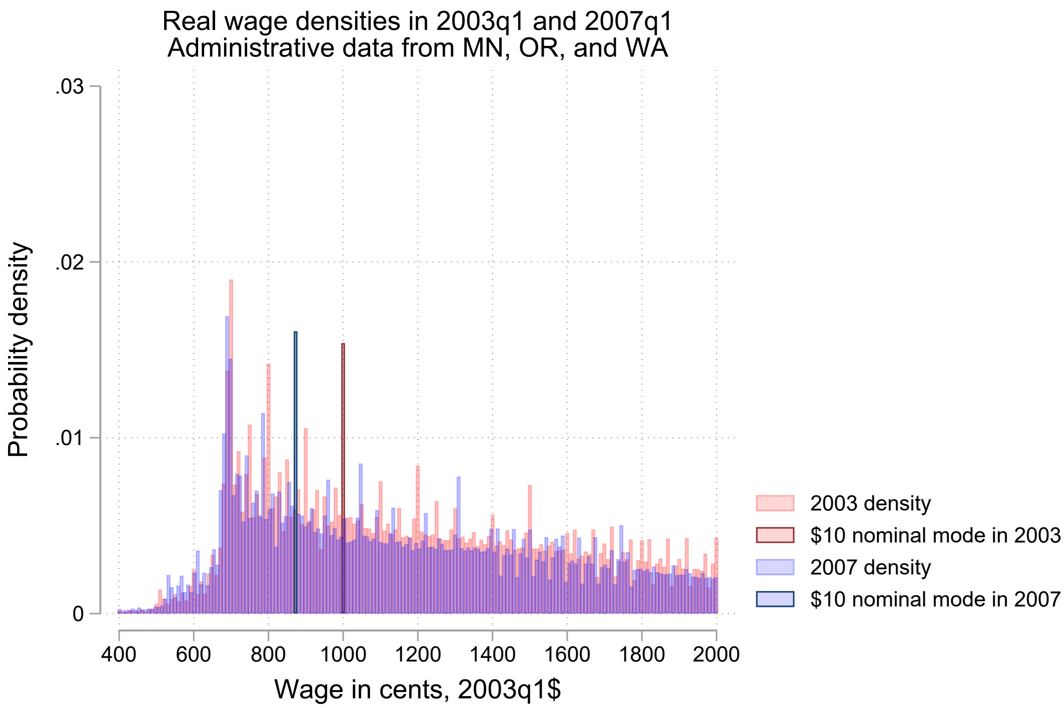
Appendix Figure A.1 plots the histograms of hourly wages in (nominal) 10 cents bins using administrative data separately for the states of Minnesota (panel A), Oregon (Panel B) and Washington (panel C). All are based on hourly wage data from UI records from 2003-2007. Hourly wages are constructed by dividing quarterly earnings by the total number of hours worked in the quarter. The counts are normalized by dividing by total employment in that state, averaged over the sample period. The figure shows very clear bunching at multiples of $1 in both states, especially at $10. Appendix Figure A.2 plots the overlaid histograms of hourly wages, pooled across MN, OR, and WA, in real 10 cents bins from 2003q4 and 2007q4, and shows that the nominal bunching at $10.00 occurs at different places in the real wage distribution in 2003 and 2007.

FIGURE A.1. **HISTOGRAMS OF HOURLY WAGES IN ADMINISTRATIVE PAYROLL DATA FROM MINNESOTA, OREGON, AND WASHINGTON, 2003-2007**



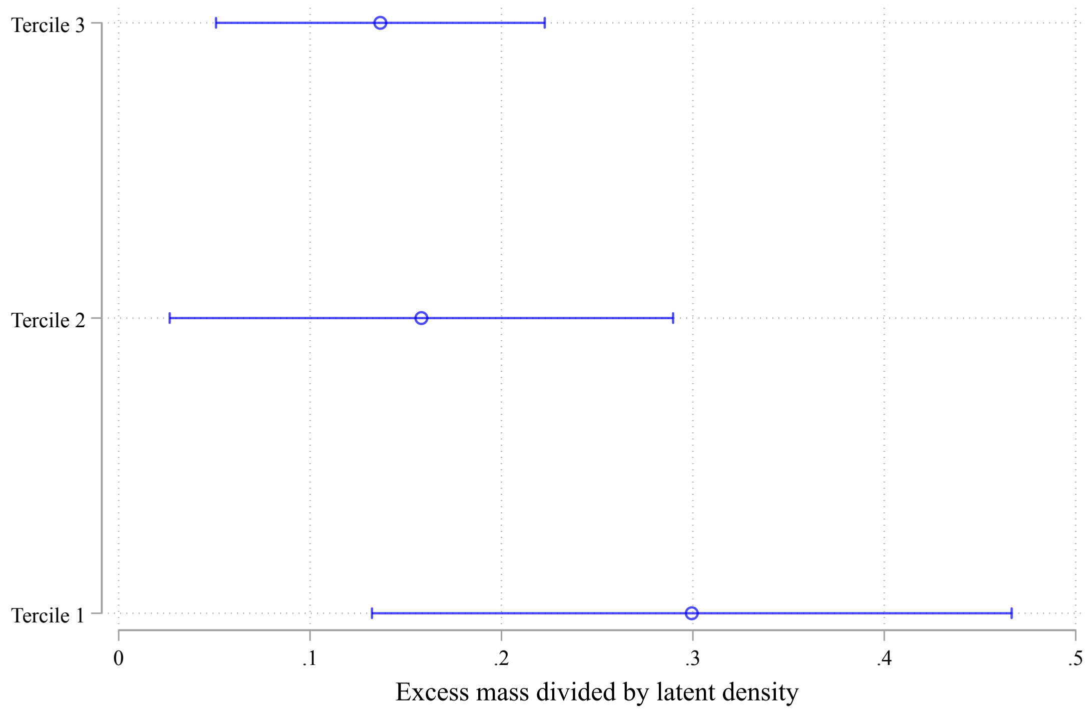
Notes. The figure shows histograms of hourly wages in 10 cents (nominal) wage bins, averaged over 2003q1 to 2007q4, using administrative Unemployment Insurance payroll records from the states of Minnesota (Panel A) , Oregon (Panel B), and Washington (Panel C). Hourly wages are constructed by dividing quarterly earnings by the total number of hours worked in the quarter. The counts in each bin are normalized by dividing by total employment in that state, averaged over the sample period. The UI payroll records cover over 95% of all wage and salary civilian employment in the states. The counts here exclude NAICS 6241 and 814, home-health and household sectors, which were identified by the state data administrators as having substantial reporting errors.

FIGURE A.2. **HISTOGRAMS OF REAL HOURLY WAGES IN ADMINISTRATIVE PAYROLL DATA FROM MINNESOTA, OREGON, AND WASHINGTON, 2003-2007**



Notes. The figure shows a histogram of hourly wages in 10 cents real wage bins (2003q1 dollars) for 2003q1 and 2007q1, using pooled administrative Unemployment Insurance payroll records from the states of Minnesota and Washington. The nominal $10 bin is outlined in dark for each year—highlighting the fact that this nominal mode is at substantially different part of the real wage distributions in these two periods. Hourly wages are constructed by dividing quarterly earnings by the total number of hours worked in the quarter. The counts in each bin are normalized by dividing by total employment in that state for that quarter. The UI payroll records cover over 95% of all wage and salary civilian employment in the states. The counts here exclude NAICS 6241 and 814, home-health and household sectors, which were identified by the state data administrators as having substantial reporting errors.

FIGURE A.3. HETEROGENEITY IN BUNCHING BY OREGON FIRM HUMAN RESOURCE MANAGEMENT SCORE TERCILES



Notes: The figure plots the extent of bunching (excess mass at $10 divided by latent density at that wage) for terciles of Human Resource Management scores matched by industry to the World Management Survey data from Bloom and Van Reenen (2007). An estimate greater than 0 indicates bunching. 95 percent confidence intervals are based on standard errors clustered at the wage bin level.

TABLE A.1— OREGON MATCHED PANEL ESTIMATES: SEPARATION RESPONSE TO RAISES AROUND $10.00/HOUR JOBS

	(1)	(2)	(3)
Log wage (at start)	-0.564		
	(0.170)		
Jump at 10.00 (at start)	0.015		
	(0.010)		
Log wage (after raise)		-0.856	-0.754
		(0.253)	(0.252)
Jump at 10.00 (after raise)		0.017	0.016
		(0.013)	(0.013)
θ (at start)	-0.375		
	(0.226)		
θ (after raise)		-0.239	-0.280
		(0.177)	(0.207)
η	2.440	4.225	3.719
	(0.736)	(1.250)	(1.245)
Obs	63785	21162	21155
Standard controls	Y	Y	Y
Firm FE control	Y	Y	Y
Lag WFE, FFE value controls			Y
Log wage at tenure=1q control		Y	Y
Jump at 10.00 at tenure=1q control		Y	Y
Tenure at job	1 qtr	2 qtr	2 qtr

Notes: Sample in column 1 includes all "full quarter hires'—i.e., with 1 quarter tenure at their new job—who were employed at a different company 2 quarters prior, and whose hourly wages fall between $9.50 and $10.59, and whose first quarter hours exceeded the 25th percentile (184 hours). The outcome is 4-quarter separation which takes on 1 if the person leaves their job within 4 quarters of start date Columns 2 and 3 restrict the sample further to those with 2 quarters of tenure, and control for tenure=1 log wage, and whether the wage exceeded $10.00. Standard controls includes include: log hours and log wage at the last job, and quintiles in the share of jobs bunched at exactly 10.00 at the current employer. Firm FE indicates a fixed effect for the current employer, interacted with quarter dummies. η is the elasticity of labor supply facing the firm. θ is the inattention parameter capturing left-digit bias. Robust standard errors are clustered at the 1-cent wage bin level.

Online Appendix B. Bunching Robustness Specifications

Since the counterfactual involves fitting a smooth distribution using a polynomial in the estimation range, in Table C.3 we assess the robustness of our estimates to alternative polynomial orders between 4 and 7. Both the size of the bunch, and the radius of the interval with missing mass, ω, are highly robust to the choice of polynomials. For example, using the pooled administrative data, the bunching β_0 is always 0.01, and ω is always 0.08 for all polynomial orders K.

One concern with bunching methods in cross sectional data is that the estimation of missing mass requires parametric extrapolation of the wage distribution around \$10. In our case, however, the bunching is at a nominal number (\$10) that sits on a different part of the real wage distribution in each of the 20 quarters of our sample. As an alternative, instead of collapsing the data into a single cross section, we use quarterly cross sectional data and fit a polynomial in the real wage $w_r = w/P_t$ where P_t is the price index in year t relative to 2003. Defining p_{w_r} as the probability mass for a real wage bin w_r, we specify the regression equation as:

$$p_{w_r} = \sum_{j=w_0-\Delta w}^{w_0+\Delta w} \beta_j \mathbb{1}_{w_r \times P_t = j} + \sum_{i=0}^{K} \alpha_i w_r^i + \epsilon_{w_r} \tag{21}$$

We again iterate estimating this equation until $MM = MMA + MMB$ to recover Δw. If the real wage distribution is assumed to be stable during this period (i.e. the α_i are constant over time), then in principle the latent wage distribution within the bunching interval can be identified nonparametrically, because each w_r bin falls outside of the bunching interval in at least some periods. More precisely, suppose there were only two periods, and $(w_0 - \Delta w)/P_{T_1} \geq (w_0 + \Delta w)/P_{T_0}$, for some T_1 and T_0. In this case β_j is identified from the mass at $w_r \times P_{T_1}$ controlling for a flexible function of w_r which is effectively identified from the real wage distribution in T_0 as well as the mass at $w_r \times P_{T_0}$ conditional on the real wage density in T_1. This specification is an example of a "difference in bunching" approach that compares the same part of the real wage distribution across years (Kleven (2016)), and addresses criticisms of bunching estimators being dependent on parametric assumptions about the shape of the latent distribution (Blomquist et al., 2021). To show that this assumption of non-overlapping bunching intervals is satisfied for at least some portion of our data, Appendix Figure A.2 shows that the bunching interval around the nominal \$10.00 mode in 2007 does not overlap with that from the 2003 real wage distribution, allowing for estimation of the latent (real) density around the nominal \$10.00 mode using variation in the price level over time. In column (8) of Table C.3 we show that estimates with the repeated cross section and real wage polynomials are virtually identical to our baseline estimates, providing reassurance that our estimates are not being driven by parametric assumptions about the latent distribution within the bunching interval.

Online Appendix C.
Recovering the joint distribution of η and δ, allowing for heterogeneity

We begin by clarifying what is identifiable from the empirical estimates of bunching. Define:

$$z_0 = \frac{\delta + \gamma}{\eta\,(1+\eta)}, z_1 = \frac{\delta}{\eta\,(1+\eta)} \tag{22}$$

Assume, for the moment, that there is some potential variation in (δ, γ, η) across firms which is independent of the latent wage and leads to a CDF for z_0 of $\Lambda_0^z(z)$ and a CDF for z_1 of $\Lambda_1^z(z)$. From (22) it must be the case that $\Lambda_0^z(z) \leq \Lambda_1^z(z)$ with equality if there is no left-digit bias. The way in which we use this is the following—suppose the fraction of firms who bunch from above w_0 is denoted by $\phi(\omega^*) = \phi\left(\frac{w_H^* - w_0}{w_H^*}\right)$, where ω^* is the proportionate gap between the latent optimal wage under the nominal model,w^*, the round number w_0, w_H^* as the optimal wage under the nominal model for the marginal buncher from above. Similarly, $\phi(\omega_*)$ is defined as $\phi(\frac{w_0 - w_L^*}{w_L^*})$. Then (25) implies that we will have, for $\omega < 0$, :

$$\phi(\omega_*) = 1 - \Lambda_0^z\left[\frac{\omega_*^2}{2}\right] \tag{23}$$

and for $\omega > 0$:

$$\phi(\omega^*) = 1 - \Lambda_1^z\left[\frac{\omega^{*2}}{2}\right] \tag{24}$$

In this paper, we empirically recover estimates of the left-hand sides of (23) and (24). The results in this Appendix imply that these estimates of the source of the missing mass in the wage distribution can be used to nonparametrically identify the distributions of z_0 and z_1, Λ_0 and Λ_1. However, these estimates alone do not allow us to nonparametrically identify the distribution of (δ, γ, η), the underlying economic parameters of interest, and below we estimate the degree of monopsony and employer misoptimization under a variety of parametric assumptions.

The first result of our framework above is that worker left-digit bias implies that the degree of bunching is asymmetric, in that missing mass will come more from below the round number than above. Thus, finding symmetry in the origin of the missing mass implies that we can approximate ω^* and ω_* with the harmonic mean of the two, which we denote $\omega \equiv \left|\frac{w - w_0}{w_0}\right|$, and is exactly the proportional radius of the bunching interval in Table 1. This further implies that $\Lambda_0 = \Lambda_1$ and allows us to accept the hypothesis that $\gamma = 0$. The intuition for this is that left-digit bias implies that firms with a latent wage 5 cents below the round number have a higher incentive to bunch than those with a latent wage 5 cents above. We fail to reject symmetry of the missing mass in Table 1 and so we proceed holding $\gamma = 0$.

Under the assumption that bunchers have latent wages near the round number, the presence of missing mass greater than w_0 also rules out a number of explanations that do not require monopsony in the labor market. If the labor market were perfectly competitive, then no worker could be *underpaid*, even though misoptimizing firms could still *overpay* workers. Explanations involving product market rents or other sources of profit for firms cannot explain why firms systematically can pay below the marginal product of workers; only labor market power can account for this. Similarly, however, the presence of missing mass below w_0 rules out pure employer collusion around a focal wage of w_0, as the pure collusion case would imply that all the missing mass was coming from *above* w_0.

Taking $\gamma = 0$ as given, our estimates of the proportion of jobs that are bunched at w_0 for each latent wage identifies the CDF of $z_1 = z_0 = \frac{\delta}{\eta(1+\eta)}$, but does not allow us to identify the distributions of δ and η separately without further assumptions. First, note that if there is perfect competition in labor markets ($\eta = \infty$) or no optimization frictions ($\delta = 0$), we have that $z_1 = 0$ in which case there would be no bunches in the wage distribution. The existence of bunches implies that we can reject the joint hypothesis of perfect competition for all jobs and no optimization frictions for all firms. But there is a trade-off between the extent of labor market competition and optimization friction that can be used to rationalize the data on bunches. To see this note that if the labor market is more competitive i.e. η is higher, a higher degree of optimization friction is required to explain a given level of bunching. Similarly, if optimization frictions are higher i.e. a higher δ, then a higher degree of labor market competition is required to explain a given level of bunching.

To estimate η and δ separately from $\phi(\omega)$, we need to make assumptions about the joint distribution. A natural first place to start is to assume a single value of η and a single value of δ. In this case, the missing mass is a constant proportion within the bunching interval around the whole number bunch with all latent wages inside the interval and none outside, an extremely sharp spike with no jobs nearby in the wage distribution. Therefore, ω, η and δ must satisfy:

$$\frac{2\delta}{\eta\,(1+\eta)} = \omega^2 \tag{25}$$

This expression shows that, armed with an empirical estimate of ω, the size of the interval of wages bunched at w_0, we can trace out a δ-η locus, showing the values of δ and η that can together rationalize a given ω. For a given size of the bunching interval ω, a higher value of optimization frictions (higher δ) implies a more competitive labor market (a higher η). [22]

Our estimates of the "missing mass" do not suggest a bunching interval with such a stark spike where all jobs within the interval pay w_0 . Instead, at all latent wages, there seem to be some jobs whiche are bunched and others which are not. To rationalize this requires a non-degenerate distribution of δ and/or η . We make a variety of different assumptions on these distributions in order to investigate the robustness of our results.

We always assume that the distributions of η and δ are independent with cumulative distributions $H(\eta)$ and $G(\delta)$. At least one of these distributions must be non-degenerate because, by the argument above, if they both have a single value for all firms one would observe an area around the bunch where all jobs bunch so the missing mass would be 100% - this is not what the data look like. Our estimates imply that there are always some jobs which are not bunched at w_0, however close is their latent wage to the bunch. We rationalize this as being some fraction of jobs who are always optimized i.e. have $\delta = 0$.

We first make the simplest parametric assumptions that are consistent with the data: we assume that η is constant, and δ has a 2-point distribution with δ=0 with probability $\underline{G}$ and $\delta = \delta^*$ with probability $1 - \underline{G}$, so that $E[\delta|\delta > 0] = \delta^*$. Below, we will extend this formulation to consider other possible shapes for the distribution $G(\delta|\delta > 0)$, keeping a mass point at $G(0) = \underline{G}$.

This then implies the missing mass at w is given by:

$$\phi(\omega) = [1 - \underline{G}]\, I\left[\omega^2 < \frac{2\delta^*}{\eta\,(1+\eta)}\right] \tag{26}$$

In this model, the share of jobs with a latent wage close to the bunch that continue to pay a non-

[22] Andrews, Gentzkow and Shapiro (2017) make a similar point in a different context, arguing that differing percentages of people with optimization frictions can substantively affect other parameter estimates using the example of DellaVigna, List and Malmendier (2012).

round w identifies $\underline{G}$, and the radius of the bunching interval (ω) in the distribution identifies $\frac{\delta^*}{\eta(1+\eta)}$. The width of the interval was estimated, together with its standard error, in the estimation of the missing mass where, relative to the bunch, it was denoted by $\frac{\Delta w}{w_0}$. Under assumptions about δ^* we can recover a corresponding estimate of η and vice versa.

Alternative assumptions on heterogeneity

While assuming a single value of non-zero δ and a constant elasticity η may seem restrictive, it is a restriction partially made for empirical reasons as our estimate of the missing mass at each latent wage is not very precise and we will also be unable to distinguish heterogeneous elasticities in our experimental design. Nonetheless, there is a concern that different assumptions about the distribution of δ and η might be observationally indistinguishable but have very different implications for the extent of optimization frictions and monopsony power in the data. This section briefly describes a number of robustness exercises that vary the possible heterogeneity in δ and η, with details relegated to the next subsection Online Appendix C.[23]

While it is not possible to identify arbitrary nonparametric distributions of δ and η, as robustness checks we consider polar cases allowing each to be unrestricted one at a time, and then finally a semi-parametric deconvolution approach that allows for an unrestricted, non-parametric distribution $H(\eta)$, along with a flexible, parametric distribution $G(\delta)$. First, we continue to assume a constant η but allow δ to be have an arbitrary distribution $G(\delta|\delta > 0)$ while continuing to fix the probability that $\delta = 0$ at $\underline{G}$. Second, at the opposite pole, we allow each job to have its own labor supply elasticity η, which is either mis-optimized by a fixed fraction $\delta^*$of profits or not at all.[24] Finally, we continue to allow arbitrary heterogeneity in η but only restrict $G(\delta)$ to have a continuous lognormal distribution, with prespecified variances of .1 and 1. We present detailed derivations of the estimators for this in Online Appendix C, below.

We quantitatively show robustness of our main estimates to these four alternate specifications in Table **C.4.** Column 1 shows the implied $E[\delta|\delta > 0]$ and $\bar{\delta}$ when an arbitrary distribution of δ is allowed. The implied η for $E[\delta|\delta > 0] = 0.01$ is 1.67 instead of 1.33 in the baseline estimates from Table C.2. Similarly, in column 2 we see the estimates under the 2-point distribution for δ and an arbitrary distribution for η. The mean η of 1.56 in this case is quite close to column 1. The implied bounds are somewhat larger, with a 1% loss in profits for those bunching (i.e., $E(\delta|\delta > 0) = 0.01$) generating 95% confidence intervals that rule out estimates of 5.4 or greater. Under 5% loss in profits, we get elasticities in columns 1 and 2 that are close to 4 , somewhat larger than the comparable baseline estimate of 3.5, but with similarly close to 20 percent wage markdown. Therefore, allowing for heterogeneity in either δ or η only modestly increases the estimated mean η as compared to our baseline estimates.

In columns 3 and 4 we report our estimates allowing for an arbitrary distribution for η, along with a lognormal conditional distribution for δ. These estimates are obtained using a deconvolution estimator to recover the distribution of a difference in random variables, described in more detail below. As in columns 1 and 2, we consider the case where $E(\delta|\delta > 0) = 0.01$ or 0.05, but now allow the standard deviation σ_δ to vary. In column 3 we take the case where δ is fairly concentrated around the mean with $\sigma_\delta = 0.1$. Here the estimated $E(\eta)$ is equal to 2.5, which is larger than the analogous baseline estimates in columns 1 and 2 allowing for an arbitrary distributions for δ and

[23]In Appendix Table D.6 we examine heterogeneity in η by worker characteristics, holding fixed δ and using measurement error corrected CPS data. The estimates are consistent with plausible heterogeneity in residual labor supply elasticities: women have lower estimated η while new workers have higher values, but the extent of heterogeneity is generally limited.

[24]This exercise is in the spirit of Saez (2010) who estimates taxable income elasticities using bunching in income at kinks and thresholds in the tax code (Kleven 2016). Kleven and Waseem (2013) use incomplete bunching to estimate optimization frictions, similar to our exercise in this paper; however, in our case optimization frictions produce bunching while in Kleven and Waseem (2013) they prevent it. This has been applied to estimating the implicit welfare losses due to various non-tax kinks, such as gender norms of relative male earnings (Bertrand, Kamenica and Pan 2015) as well as biases due to behavioral constraints (Allen et al. 2016).

η, respectively. In column 4, we allow δ to be much more dispersed, with $\sigma_\delta = 1$. In this case the estimated $E(\eta)$ falls somewhat to 2. With $E(\delta|\delta > 0) = 0.05$, we get $E[\eta] = 6$ and 4.6 under $\sigma_\delta = 0.1$ and $\sigma_\delta = 1$, respectively, and we are able to rule out markdowns less than 5 percent easily. Encouragingly, for a given mean value of optimization friction, $E[\delta|\delta > 0]$, allowing for heterogeneity in δ and η together only modestly affects the estimated mean η as compared to our baseline estimates.

Overall, a wide range of assumptions made about the distribution of δ and η continue to suggest that the degree of bunching observed in the data is consistent with a moderate degree of monopsony along with a modest reduction in profits from optimization errors; and that an assumption of a more competitive labor market implies larger profit loss from mispricing.

Detailed Derivations

In this Appendix, we provide details on the derivations of the robustness checks in section Online Appendix C. We also show the estimated CDFs for the distributions of δ and η under the different distributional assumptions.

For the first exercise, we continue to assume a constant η but allow δ to have an arbitrary distribution $G(\delta|\delta > 0)$ while continuing to fix the probability that $\delta = 0$ at $\underline{G}$. In this case, for a given value of η the non-missing mass at ω would equal:

$$phi\,(\omega) = 1 - \hat{G}(\eta(1+\eta)\frac{\omega^2}{2}) \tag{27}$$

This expression implicitly defines a distribution $\hat{G}(\delta)$:

$$\hat{G}(\delta) = 1 - \phi\left(\sqrt{\frac{2\delta}{\eta(1+\eta)}}\right) \tag{28}$$

Note that this implies that $\delta \in [0, \delta_{max}]$ where $\delta_{max} = \frac{\omega^2}{2}\eta(1+\eta)$ where ω is the radius of the bunching interval. We then fix $E(\delta|\delta > 0)$ at a particular value, similar to what we do with δ^*, and then can identify both an arbitrary shape of $\hat{G}(\delta)$ as well as η. Figure C.5 shows the distribution along with values of η from equation (28) in the MN-OR-WA administrative data. As can be seen, a higher η implies a first-order stochastic dominating distribution of δ; thus average δ is higher for higher η. This CDF also suggests our 2-point distribution is not too extreme an assumption: the non-zero δ are confined to about 20% of the distribution, and are bounded above by 0.11, suggesting that most firms are not foregoing more than 10% of profits in order to pay a round number.

A natural question is how our estimates could differ if, instead of a constant η and flexibly heterogeneous δ, we assume a heterogeneous η with an arbitrary distribution $H(\eta)$, along with some specified distribution $G(\delta)$. The simplest variant of this is to consider a two-point distribution (where δ is either 0 or δ^*) as in our baseline case above. In this variant of the model each firm is allowed to have its own labor supply elasticity, and each firm either misoptimizes profits by a fixed fraction δ^* or not at all. In this case, solving for the positive value of η, the missing mass at ω should be equal to:

$$\phi\,(\omega) = [1 - \underline{G}]\,H\left(\frac{1}{2}\left(\sqrt{1 + \frac{8\delta^*}{\omega^2}} - 1\right)\right)$$

Since we can identify $\underline{G} = G(0) = 1 - \lim_{\omega \to 0^+} \hat{\phi}(\omega)$, for a particular δ^* we can empirically estimate the distribution of labor supply elasticities as follows:

$$\hat{H}(\eta) = \frac{\hat{\phi}\left(\sqrt{\left(\frac{8\delta^*}{(2\eta+1)^2-1}\right)}\right)}{1-\underline{G}} \tag{29}$$

We can use $\hat{H}(\eta)$ to estimate the mean $E\hat{(}\eta)$ for a given value of δ^*:

$$E\hat{(}\eta) = \int_0^\infty \eta d\hat{H}(\eta) \tag{30}$$

Note that under these assumptions, η is bounded from below at $\eta_{min} = \frac{1}{2}\sqrt{1+\frac{8\delta^*}{\omega^2}} - 1$. In other words, the lower bound of η from the third method is equal to the constant estimate of η from our baseline, both of which come from the marginal bunching condition at the boundary of the interval ω. While we can only recover the distribution of η conditional on $\delta > 0$ (i.e. those that choose to bunch), we can make some additional observations about the parameters for non-bunchers. In particular, we can rule out the possibility that some of the the non-bunchers have $\delta > 0$ while being in a perfectly competitive labor market with $\eta = \infty$. This is because in our model those firms would be unable to attract workers from those firms with $\delta = 0$ and $\eta = \infty$. The gradual reduction in the missing mass $\phi(\omega)$ that occurs from moving away from $\omega = 0$ is entirely due to heterogeneity in $\eta's$. It rules out, for instance, that such a gradual reduction is generated by heterogeneity in $\delta's$ in contrast to the second method. As a result, the third method is likely to provide the largest estimates of the labor supply elasticity.

In parallel fashion to the previous case, we graphically show the implied distribution of η with a 2-point distribution for δ in Figure C.6. This figure shows the distribution of η implied by different values of δ from the MN-OR-WA administrative data. As can be seen, a higher η implies a first-order stochastic dominating distribution of η, thus average η is higher for higher δ.

Finally, we can extend this framework to allow for $G(\delta)$ to have a more flexible parametric form (with known parameters) than the 2-point distribution. We rely on recently developed methods in non-parametric deconvolution of densities to estimate the implicit $H(\eta)$. If we condition on $\delta > 0$,we can take logs of equation 25 (again maintaining that $\gamma = 0$) we get

$$2\ln(\omega) = ln(2) - \ln(\eta(1+\eta)) + \ln(\delta) = ln(2) - \ln(\eta(1+\eta)) + E[\ln(\delta)\,|\delta > 0] + \ln(\delta_{res}) \tag{31}$$

Here $\ln(\delta_{res}) \sim N(0, \sigma_\delta^2)$, and we fix $E[\ln(\delta)\,|\delta > 0] = \ln(E(\delta|\delta > 0)) + \frac{1}{2}\sigma_\delta^2$. We can use the fact that the cumulative distribution function of $2\ln(\omega)$ is given by $1 - \phi\left(\frac{1}{2}\exp(2\ln(\omega))\right)$ to numerically obtain a density for $2\ln(\omega)$. This then becomes a well-known deconvolution problem, as the density of $-\ln(\eta(1+\eta))$ is the deconvolution of the density of $2\ln(\omega)$ by the Normal density we have imposed on $\ln(\delta_{res})$. We can then recover the distribution of η,$H(\eta)$, from the estimated density of $-\ln(\eta(1+\eta)) + E[\ln(\delta)\,|\delta > 0]$.

We now illustrate how Fourier transforms recover the distribution $H(\eta)$. Consider the general case of when the observed signal (W) is the sum of the true signal (X) and noise (U). (In our case $W = 2\ln(\omega) - E[\ln(\delta)\,|\delta > 0]$ and $U = \ln(\delta_{res})$.)

$$W = X + U \tag{32}$$

Manipulation of characteristic functions implies that the density of W is $f_W(x) = (f_X * f_U)(x) = \int f_X(x-y) f_U(y) dy$ where $*$ is the convolution operator. Let W_j be the observed sample from W.

Taking the Fourier transform (denoted by $\sim$) , we get that $\tilde{f_W} = \int f_W(x)e^{itx}dx = \tilde{f_X} \times \tilde{f_U}$. To recover the distribution of X, in principle it is enough to take the inverse Fourier transform of $\frac{\tilde{f_W}}{\tilde{f_U}}$. This produces a "naive" estimator $\widehat{f_X} = \frac{1}{2\pi} \int e^{-itx} \frac{\sum_{j=1}^{N} \frac{e^{itWj}}{N}}{\phi(t)} dt$, but unfortunately this is not guaranteed to converge to a well-behaved density function. To obtain such a density, some smoothing is needed, suggesting the following deconvolution estimator:

$$\widehat{f_X} = \frac{1}{2\pi} \int e^{-itx} K(th) \frac{\sum_{j=1}^{N} \frac{e^{itW_j}}{N}}{\phi(t)} dt \tag{33}$$

where K is a suitably chosen kernel function (whose Fourier transform is bounded and compactly supported). The finite sample properties of this estimator depend on the choice of f_U. If $\tilde{f_U}$ decays quickly (exponentially) with t (e.g. U is normal), then convergence occurs much more slowly than if $\tilde{f_U}$ decays slowly (i.e. polynomially) with t (e.g. U is Laplacian). Note that once we recover the density for $X = \ln(\eta(1+\eta))$, we can easily recover the density for η.

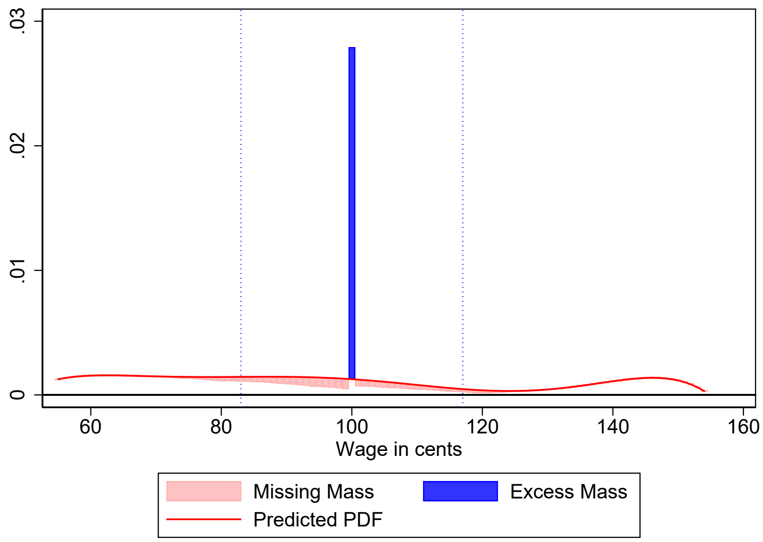
For normal $U = \ln(\delta_{res})$, Delaigle and Gijbels (2004) suggest a kernel of the form:

$$K(x) = 48 \frac{cos(x)}{\pi x^4}(1 - \frac{15}{x^2}) - 144 \frac{sin(x)}{\pi x^5}(1 - \frac{5}{x^2}) \tag{34}$$

We use the Stefanski and Carroll (1990) deconvolution kernel estimator. This estimator also requires a choice of bandwidth which is a function of sample size. Delaigle and Gijbels (2004) suggest a bootstrap-based bandwidth that minimizes the mean-integral squared error, which is implemented by Wang and Wang (2011) in the R package `decon`, and we use that method here, taking the bandwidth that minimizes the mean-squared error over 1,000 bootstrap samples.

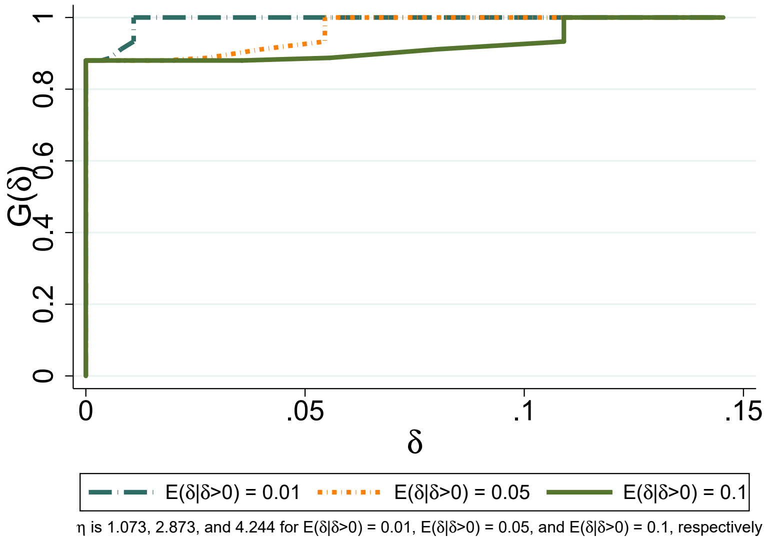
In Figure C.7, we show the distribution of η using the deconvolution estimator, assuming a lognormal distribution of δ. In the first panel, we estimate $H(\eta)$ assuming the standard deviation $\sigma_{ln(\delta)} = 0.1$, which is highly concentrated around the mean. In the second panel, we instead assume $\sigma_{ln(\delta)} = 1$. This is quite dispersed: among those with a non-zero optimization friction, δ around 16% have a value of δ exceeding 1, and around 31% have a value exceeding 0.5. As a result, we think the range between 0.1 and 1 to represent a plausible bound for the dispersion in δ. As before, we see a higher $E[\delta|\delta > 0]$ leads to first-order stochastic dominance of $H(\eta)$. For both cases with high- and low-dispersion of δ, the distribution $H(\eta)$ is fairly similar, though increase in $\sigma_{ln(\delta)}$ tends to shift $H(\eta)$ up somewhat, producing a smaller $E(\eta)$.

FIGURE C.4. **BUNCHING AT 1.00 ON AMAZON MECHANICAL TURK**



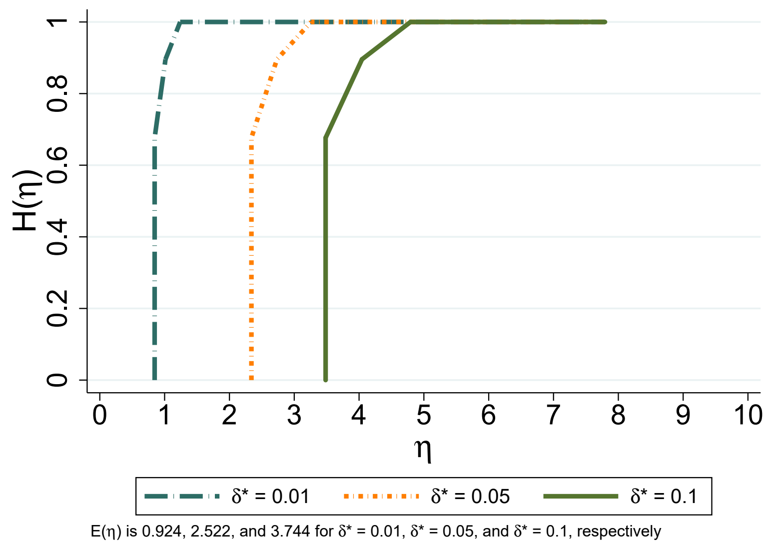
Notes. This figure plots the excess and missing mass around 1.00 on Amazon Mechanical Turk. The latent distribution is modelled with a 6th degree polynomial.

FIGURE C.5. **IMPLIED DISTRIBUTION OF δ UNDER CONSTANT η**



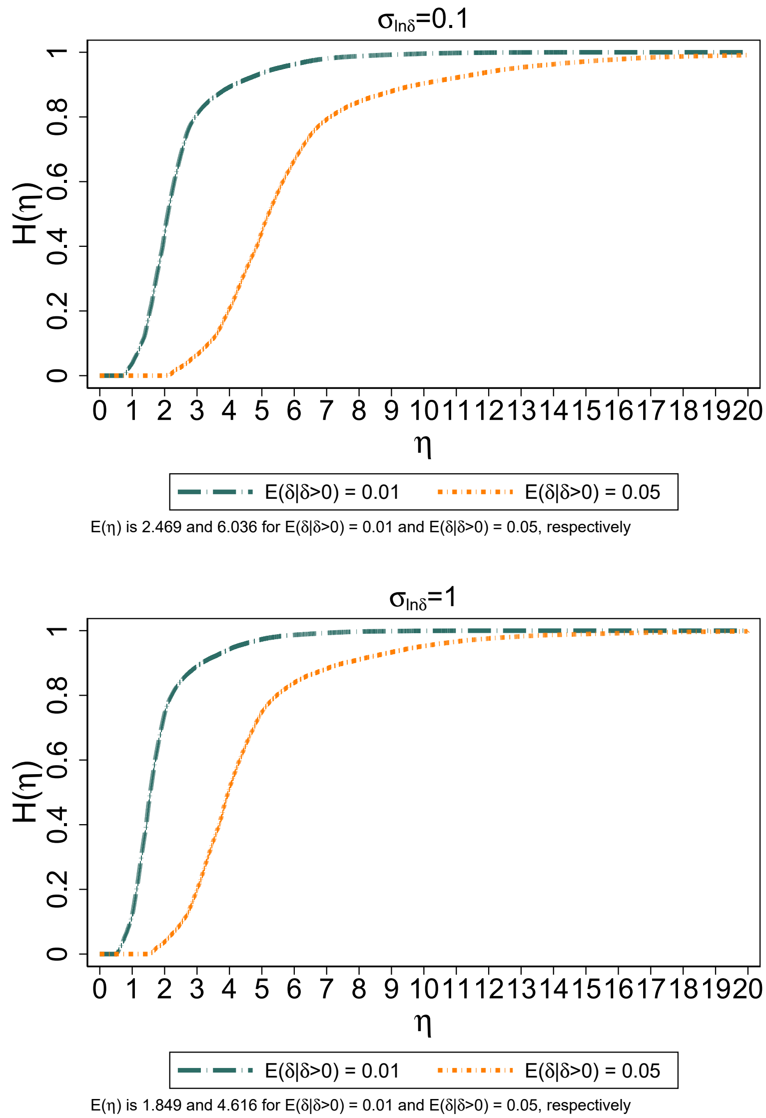
Notes. The figure plots the cumulative distributions $G(\delta)$ based on equation 28, for alternative values of $E(\delta|\delta > 0)$. The elasticity η is assumed to be a constant. The estimates use administrative hourly wage data from MN, OR, and WA.

FIGURE C.6. **IMPLIED DISTRIBUTION OF η WITH A 2-POINT DISTRIBUTION OF δ**



Notes. The figure plots the cumulative distributions $H(\eta)$ based on equation 29, for alternative values of $\delta^* = E(\delta|\delta > 0)$. δ is assumed to follow a 2-point distribution with $\delta = 0$ with probability $\underline{G}$ and $\delta = \delta^*$ with probability $1 - \underline{G}$. The estimates use administrative hourly wage data from MN, OR, and WA.

FIGURE C.7. **IMPLIED DISTRIBUTION OF η USING A DECONVOLUTION ESTIMATOR WHERE δ HAS A CONDITIONAL LOGNORMAL DISTRIBUTION**



Notes. The figure plots the cumulative distributions $H(\eta)$ using a deconvolution estimator based on equation 31, for alternative values of $E(\delta|\delta > 0)$. The procedure allows for an arbitrary smooth distribution of η, while assuming δ is lognormally distributed (conditional on being non-zero) with a standard deviation σ_δ. The top panel assumes a relatively concentrated distribution of δ with σ_δ = 0.1; in contrast, the bottom panel assumes a rather dispersed distribution with σ_δ = 1. The estimates use administrative hourly wage data from MN, OR, and WA.

TABLE C.2—**BOUNDS FOR LABOR SUPPLY ELASTICITY IN ADMINISTRATIVE DATA**

	(1)	(2)	(3)	(4)
A. $\delta^* = 0.01$				
$\bar{\delta}$	0.001	0.004	0.002	0.003
η	1.337	1.581	1.581	1.581
90% CI	[0.417, 2.050]	[0.417, 4.525]	[0.472, 9.512]	[0.417, 4.525]
95% CI	[0.417, 2.871]	[0.417, 4.525]	[0.417, 9.512]	[0.417, 4.525]
B. $\delta^* = 0.05$				
$\bar{\delta}$	0.005	0.020	0.011	0.017
η	3.484	4.045	4.045	4.045
90% CI	[1.291, 5.112]	[1.291, 10.692]	[1.429, 21.866]	[1.291, 10.692]
95% CI	[1.291, 6.970]	[1.291, 10.692]	[1.291, 21.866]	[1.291, 10.692]
C. $\delta^* = 0.1$				
$\bar{\delta}$	0.010	0.041	0.022	0.034
η	5.112	5.908	5.908	5.908
90% CI	[1.983, 7.421]	[1.983, 15.319]	[2.182, 31.127]	[1.983, 15.319]
95% CI	[1.983, 10.053]	[1.983, 15.319]	[1.983, 31.127]	[1.983, 15.319]
$G(0) = \underline{G}$	0.896	0.592	0.785	0.662
Data:	Admin OR & MN & WA	CPS-Raw OR & MN & WA	CPS-MEC OR & MN & WA	CPS-Raw

Notes. The table reports point estimates and associated 95 percent confidence intervals for labor supply elasticities, η, and markdown values associated with different values of optimization friction δ. All columns use the pooled MN, OR, and WA administrative hourly wage data. In columns 1, 2 and 3, we use hypothesized values of δ of 0.01, 0.05 and 0.1 respectively. The labor supply elasticity, η, and the markdown are estimated using the estimated extent of bunching, ω, and the hypothesized δ, using equations 25 and 6 in the paper. The 95 percent confidence intervals in square brackets are estimated using 500 boostrap draws.

TABLE C.3—ROBUSTNESS OF ESTIMATES FOR EXCESS BUNCHING, MISSING MASS, AND INTERVAL AROUND THRESHOLD

	Dum. for \$0.5 (1)	Dum. for \$0.25 & \$0.5 (2)	Poly. of degree 4 (3)	Poly. of degree 7 (4)	Fourier, degree 3 (5)	Fourier, degree 6 (6)	Real wage poly. (7)
Value of w_0	\$10.00	\$10.00	\$10.00	\$10.00	\$10.00	\$10.00	\$10.00
Excess mass at w_0	0.010	0.010	0.010	0.010	0.010	0.009	0.010
	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Total missing mass	-0.010	-0.011	-0.012	-0.013	-0.009	-0.017	-0.009
	(0.005)	(0.005)	(0.007)	(0.011)	(0.004)	(0.006)	(0.003)
Missing mass below	-0.007	-0.007	-0.006	-0.006	-0.007	-0.007	-0.004
	(0.005)	(0.005)	(0.005)	(0.007)	(0.004)	(0.005)	(0.003)
Missing mass above	-0.004	-0.004	-0.006	-0.006	-0.002	-0.009	-0.005
	(0.004)	(0.004)	(0.004)	(0.008)	(0.004)	(0.004)	(0.003)
Test of equality of missing shares of latent $<$ and $>$ w_0:							
t-statistic	-0.325	-0.341	-0.812	-0.595	-0.875	-0.830	-1.216
Bunching = $\frac{Actual\ mass}{Latent\ density}$	2.658	2.625	2.594	2.566	2.643	2.233	2.664
	(0.297)	(0.298)	(0.293)	(0.342)	(0.238)	(0.285)	(0.238)
w_L	\$9.40	\$9.40	\$9.20	\$9.20	\$9.30	\$9.40	\$9.20
w_H	\$10.60	\$10.60	\$10.80	\$10.80	\$10.70	\$10.60	\$10.80
$\omega = \frac{(w_H - w_0)}{w_0}$	0.060	0.060	0.080	0.080	0.070	0.060	0.080
	(0.023)	(0.023)	(0.025)	(0.026)	(0.038)	(0.025)	(0.026)

Notes. The table reports estimates of excess bunching at the threshold w_0 as compared to a smoothed predicted probability density function, and the interval (ω_L, ω_H) from which the missing mass is drawn. All columns use the pooled MN, OR, and WA administrative hourly wage data. The predicted PDF is estimated using a K-th order polynomial or values of K between 2 and 6 as indicated, with dummies for each bin in the interval from which the missing mass is drawn. The width of the interval is chosen by iteratively expanding the interval until the missing and excess masses are equal, as described in the text. Columns 1 and 2 include indicator variables for wages that are divisible by 50 cents and 25 cents, respectively. Columns 3 and 4 vary the order of the polynomial used to estimate the latent wage. Columns 5 and 6 represent the latent wage with a 3 and 6 degree Fourier polynomial, respectively. Column 7 estimates the predicted PDF using a sixth order polynomial of real wage bins, as opposed to the nominal ones. Bootstrap standard errors based on 500 draws are in parentheses.

TABLE C.4—**BOUNDS FOR LABOR SUPPLY ELASTICITY IN OFFLINE LABOR MARKET - HETEROGENEOUS δ AND η**

	Heterogeneous δ	Heterogeneous η	Heterogeneous δ & η, $\sigma_\delta = 0.1$	Heterogeneous δ & η, $\sigma_\delta = 1$
A. $E(\delta\|\delta > 0)$= 0.01				
$\overline{\delta}$	0.001	0.001	0.001	0.001
η	1.668	1.559	2.469	1.849
90% CI	[0.969, 4.394]	[0.917, 4.650]	[1.056, 5.446]	[0.758, 4.163]
95% CI	[0.845, 4.816]	[0.823, 5.328]	[0.905, 6.589]	[0.649, 5.062]
Markdown	0.375	0.391	0.288	0.351
90% CI	[0.185, 0.508]	[0.177, 0.522]	[0.155, 0.486]	[0.194, 0.569]
95% CI	[0.172, 0.542]	[0.158, 0.548]	[0.132, 0.525]	[0.165, 0.606]
B. $E(\delta\|\delta > 0)$= 0.05				
$\overline{\delta}$	0.006	0.006	0.006	0.006
η	4.244	3.991	6.036	4.616
90% CI	[2.629, 10.397]	[2.503, 10.965]	[2.808, 12.739]	[2.108, 9.833]
95% CI	[2.337, 11.346]	[2.284, 12.453]	[2.469, 15.445]	[1.837, 11.894]
Markdown	0.191	0.200	0.142	0.178
90% CI	[0.088, 0.276]	[0.084, 0.285]	[0.073, 0.263]	[0.092, 0.322]
95% CI	[0.081, 0.300]	[0.074, 0.304]	[0.061, 0.288]	[0.078, 0.352]
$G(0)$= $\underline{G}$	0.880	0.880	0.880	0.880
Data:	Admin OR & MN & WA	Admin OR & MN & WA	Admin OR & MN & WA	Admin OR & MN & WA

Notes. The table reports point estimates and associated 95 percent confidence intervals for labor supply elasticities, η, and markdown values associated with hypothesized δ=0.01 and δ=0.05. All columns use the pooled MN, OR, and WA administrative hourly wage counts. Heterogeneous δ and η are allowed in columns 1 and 2, using equations 28 and 29, respectively. Columns 3 and 4 allow heterogeneous δ and η, and assume a conditional lognormal distribution of δ, using a deconvolution estimator based on equation 31. The third column assumes a relatively concentrated distribution of δ ($\sigma_\delta = 0.1$); whereas the fourth column assumes a rather dispersed distribution ($\sigma_\delta = 1$). In row A, we hypothesize $\delta = 0.01$; whereas it is $\delta = 0.05$ in row B. The 90 and 95 percent confidence intervals in square brackets in columns 1 and 2 (3 and 4) are estimated using 500 (1000) boostrap draws.

TABLE C.5—**BOUNDS FOR LABOR SUPPLY ELASTICITY IN ADMINISTRATIVE DATA — ROBUSTNESS TO SPECIFICATIONS OF LATENT WAGE**

	Dum. for \$0.5 (1)	Dum. for \$0.25 & \$0.5 (2)	Poly. of degree 4 (3)	Poly. of degree 7 (4)	Fourier, degree 3 (5)	Fourier, degree 6 (6)	Real wage poly. (7)
A. $\delta^* = 0.01$							
$\bar{\delta}$	0.001	0.001	0.001	0.001	0.001	0.002	0.001
η	1.909	1.909	1.337	1.337	1.581	1.909	1.337
90% CI	[0.472, 2.050]	[0.472, 2.050]	[0.417, 2.050]	[0.417, 2.050]	[0.300, 2.871]	[0.538, 2.871]	[0.417, 2.050]
95% CI	[0.417, 2.050]	[0.417, 2.050]	[0.417, 2.050]	[0.417, 2.871]	[0.247, 2.871]	[0.417, 4.525]	[0.417, 2.871]
B. $\delta^* = 0.05$							
$\bar{\delta}$	0.006	0.006	0.005	0.005	0.005	0.008	0.004
η	4.794	4.794	3.484	3.484	4.045	4.794	3.484
90% CI	[1.429, 5.112]	[1.429, 5.112]	[1.291, 5.112]	[1.291, 5.112]	[0.984, 6.970]	[1.593, 6.970]	[1.291, 5.112]
95% CI	[1.291, 5.112]	[1.291, 5.112]	[1.291, 5.112]	[1.291, 6.970]	[0.839, 6.970]	[1.291, 10.692]	[1.291, 6.970]
C. $\delta^* = 0.1$ $\bar{\delta}$	0.889	0.883	0.906	0.899	0.908	0.835	0.919

Notes. The table reports point estimates and associated 95 percent confidence intervals for labor supply elasticities, η, and markdown values associated with hypothesized $\delta = 0.01$ and $\delta = 0.05$. All columns use the pooled MN, OR and WA administrative hourly wage counts. The first two columns control for bunching at wage levels whose modulus with respect to \$1 is \$0.5, and \$0.5 or \$0.25, respectively. Column 3 uses a quadratic polynomial to estimate the wage distribution, whereas column 4 uses a quartic. In columns 5 and 6, instead of polynomials, Fourier transformations of degree 3 and 6 are employed. Column 7 estimates the predicted PDF using a sixth order polynomial of real wage bins, as opposed to the nominal ones. In row A, we hypothesize $\delta = 0.01$; whereas it is $\delta = 0.05$ in row B. The labor supply elasticity, η, and markdown values are estimated using the estimated extent of bunching, ω, and the hypothesized δ, using equation 25 and 6 in the paper. The 95 percent confidence intervals in square brackets are estimated using 500 boostrap draws.

Online Appendix D.
Bunching in Hourly Wage Data from Current Population Survey and Supplement

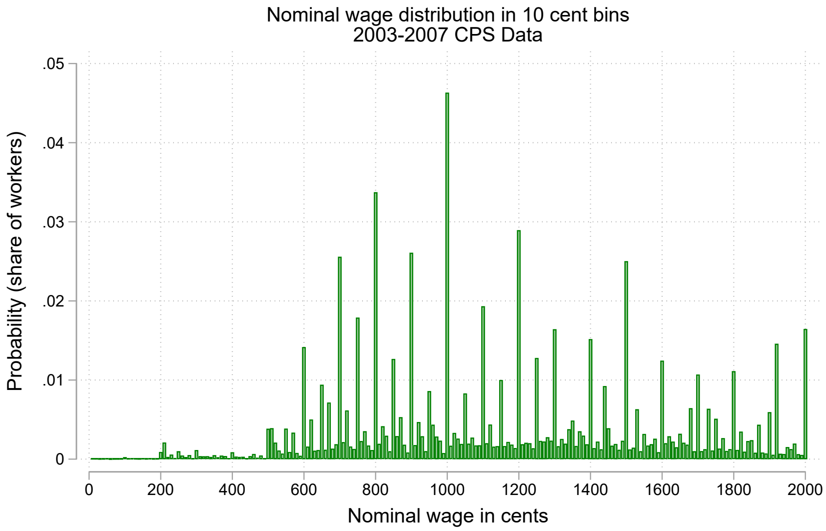
In this appendix, we show the degree of bunching in hourly nominal wage data using the national CPS data. In Figure D.8, we plot the nominal wage distribution in U.S. in 2003 to 2007 in 10 cents bins. There are notable spikes in the wage distribution at $10, $7.20 (the bin with the federal minimum wage), $12, $15, along with other whole numbers. At the same time, the spike at $10.00 is substantially larger than in the administrative data (exceeding 0.045), indicating rounding error in reporting may be a serious issue in using the CPS to accurately characterize the size of the bunching.

We also use a 1977 CPS supplement, which matches employer and employee reported hourly wages, to correct for possible reporting errors in the CPS data. We re-weight wages by the relative incidence of employer versus employee reporting, based on the two ending digits in cents (e.g., 01, 02, ... , 98, 99). As can be seen in Figure D.9, the measurement error correction produces some reduction in the extent of visible bunching, which nonetheless continues to be substantial. For comparison, the probability mass at $10.00 is around 0.02, which is closer to the mass in the administrative data than in the raw CPS. This is re-assuring as it suggests that a variety of ways of correcting for respondent rounding produce estimates suggesting a similar and substantial amount of bunching in the wage distribution.

Heterogeneous η by Worker Characteristics-CPS

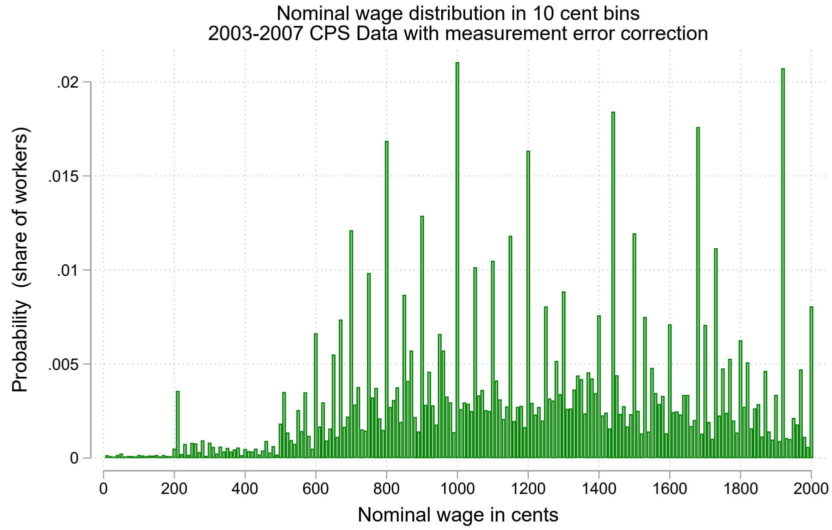
In Appendix Table D.6 we estimate the implied η for different δ^* under our baseline 2-point model across subgroups of the measurement corrected CPS data, as we do not have worker-level covariates for the administrative data. We examine young and old workers, as well as male and female separately. Consistent with other work suggesting that women are less mobile than men (Webber (2016); Manning (2011)), the estimated η for women is somewhat lower than that for men. We do not find any differences between older and younger workers. However, the extent of bunching is substantially larger for new hires consistent with bunching being a feature of initial wages posted, while workers with some degree of tenure are likelier to have heterogeneous raises that reduce the likelihood of being paid a round number. We find that among new hires the estimated η is somewhat higher than non-new hires. However, even for new hires—who arguably correspond most closely to the wage posting model—the implied η is only 1.58 if employers who are bunching are assumed to be losing 1% of profits from doing so, increasing to 4 when firms are allowed to lose up to 5% in profits.

FIGURE D.8. **HISTOGRAM OF HOURLY WAGES IN NATIONAL CPS DATA, 2003-2007**



Notes. The figure shows a histogram of hourly wages by 10 cents (nominal) wage bins, averaged over 2003q1 and 2007q4, using CPS MORG files. Hourly wages are constructed by average weekly earnings by usual hours worked. The sample is restricted to those without imputed earnings. The counts here exclude NAICS 6241 and 814, home-health and household sectors. The histogram reports normalized counts in 10 cents (nominal) wage bins, averaged over 2003q1 and 2007q4. The counts in each bin are normalized by dividing by total employment, averaged over the sample period.

FIGURE D.9. **WAGE BUNCHING IN CPS DATA, 2003-2007, CORRECTED FOR REPORTING ERROR USING 1977 CPS SUPPLEMENT**



Notes. The figure shows a histogram of hourly wages by 10 cents (nominal) wage bins, averaged over 2003q1 to 2007q4, using CPS MORG files, where individual observations were re-weighted to correct for overreporting of wages ending in particular two-digit cents using the 1977 CPS supplement. Hourly wages are constructed by dividing average weekly earnings by usual hours worked. The sample is restricted to those without imputed earnings. The counts here exclude NAICS 6241 and 814, home-health and household sectors. The histogram reports normalized counts in 10 cents (nominal) wage bins, averaged over 2003q1 and 2007q4. The counts in each bin are normalized by dividing by total employment, averaged over the sample period.

TABLE D.6—**BOUNDS FOR LABOR SUPPLY ELASTICITY IN U.S. LABOR MARKET — HETEROGENEITY BY DEMOGRAPHIC GROUPS**

	Male	Female	Age<30	Age≥30	Same job as last month	Different job from last month
Excess mass at w_0	0.018	0.015	0.030	0.012	0.014	0.028
	(0.003)	(0.004)	(0.006)	(0.003)	(0.003)	(0.006)
Total missing mass	-0.011	-0.012	-0.042	-0.012	-0.015	-0.023
	(0.009)	(0.007)	(0.013)	(0.006)	(0.007)	(0.012)
Bunching = $\frac{Actual\ mass}{Latent\ density}$	5.906	3.890	4.923	3.907	4.137	6.347
	(2.034)	(0.989)	(1.634)	(1.033)	(1.122)	(2.273)
A. $\delta^* = 0.01$						
$\overline{\delta}$	0.002	0.001	0.003	0.001	0.002	0.003
η	1.581	1.337	1.337	1.337	1.337	1.581
90% CI	[0.538, 4.525]	[0.618, 9.512]	[0.538, 4.525]	[0.618, 9.512]	[0.576, 9.512]	[0.538, 4.525]
95% CI	[0.472, 9.512]	[0.538, 9.512]	[0.472, 4.525]	[0.472, 9.512]	[0.472, 9.512]	[0.472, 4.525]
B. $\delta^* = 0.05$						
$\overline{\delta}$	0.009	0.005	0.014	0.007	0.008	0.013
η	4.045	3.484	3.484	3.484	3.484	4.045
90% CI	[1.593, 10.692]	[1.791, 21.866]	[1.593, 10.692]	[1.791, 21.866]	[1.687, 21.866]	[1.593, 10.692]
95% CI	[1.429, 21.866]	[1.593, 21.866]	[1.429, 10.692]	[1.429, 21.866]	[1.429, 21.866]	[1.429, 10.692]
$G(0)= \underline{G}$	0.820	0.895	0.713	0.863	0.834	0.750
Data:	CPS-MEC	CPS-MEC	CPS-MEC	CPS-MEC	CPS-MEC	CPS-MEC

Notes. The table reports point estimates and associated 95 percent confidence intervals for labor supply elasticities, η, and markdown values associated with hypothesized δ=0.01 and δ=0.05. All columns use the national measurement error corrected CPS data. The first two columns analyze by gender, the third and fourth by age, and the columns 5 and 6 by incumbency. In row A, we hypothesize δ = 0.01; whereas it is δ = 0.05 in row B. The labor supply elasticity, η, and markdown are estimated using the estimated extent of bunching, ω, and the hypothesized δ, using equations 25 and and 6 in the paper. The 95 percent confidence intervals in square brackets are estimated using 500 boostrap draws.

Online Appendix E.
Testing Discontinuous Labor Supply on Amazon Mechanical Turk Observational Data

Our Amazon Mechanical Turk experiment focused on discontinuities at 10 cents, while our bunching estimator used the excess mass at $1.00. In this appendix we present evidence from observational data scraped from Amazon Mechanical Turk to show that there is also no evidence of a discontinuity in worker response to rewards at $1.00. Our primary source of data was collected by Panos Ipseiros between January 2014 and February 2016, and, in principle, kept track of all HITs posted in this period.

We keep the discussion of the data and estimation details brief, as interested readers can see details in Dube et al. (2020). Dube et al. (2020) combines a meta-analysis of experimental estimates of the elasticity of labor supply facing requesters on Amazon Mechanical Turk with Double-ML estimators applied to observational data.. That paper does not look at discontinuities in the labor supply at round numbers.

Following Dube et al. (2020) we use the observed duration of a batch posting as a measure of how attractive a given task is as a function of observed rewards and observed characteristics. We calculate the duration of the task as the difference between the first time it appears and the last time it appears, treating those that are present for the whole period as missing values. We convert the reward into cents. We are interested in the labor supply curve facing a requester. Unfortunately, we do not see individual Turkers in this data. Instead we calculate the time until the task disappears from our sample as a function of the wage. Tasks disappear once they are accepted. While tasks may disappear due to requesters canceling them rather than being filled, this is rare. Therefore, we take the time until the task disappears to be the duration of the posting—i.e., the time it takes for the task to be accepted by a Turker. The elasticity of this duration with respect to the wage will be equivalent to the elasticity of labor supply when offer arrival rates are constant and reservation wages have an exponential (constant hazard) distribution.

In order to handle unobserved heterogeneity, Dube et al. (2020) implement a double-machine-learning estimator proposed by Chernozhukov et al. (2017), which uses machine learning (we used random forests) to form predictions of log duration and log wage (using one half of the data), denoted $\ln(\widehat{duration_h})$ and $\ln(\widehat{wage_h})$,and then subtracts them from the actual variable values in the other sample, leaving residualized versions of both variables. The predictions use a large number of variables constructed from the metadata and textual descriptions of each task, and have high out-of-sample predictive power, and so the residuals are likely to reflect variation that, if not exogenous, are at least orthogonal to a very flexible and predictive function of all the other observable characteristics of a task. See Dube et al. (2020) for further details on implementation and estimation.

We then estimate regressions of the form:

$$\ln(duration_h) - \ln(\widehat{duration_h}) = \eta \times (\ln(wage_h) - \ln(\widehat{wage_h})) + \gamma \mathbf{1}_{\mathbf{w}>\mathbf{w_0}} + \epsilon \tag{35}$$

Results are shown in Table E.7. We restrict attention to windows of wages around our two most salient round numbers, 10 cents, where the window is 6 to 14 cents, and $1.00, where the window is $0.80 to $1.20. Across specifications, there is a clear negative relationship between wages/rewards and duration, with a coefficient on η similar in magnitude to the - 0.11 estimate obtained on the whole sample in Dube et al. (2020), and close to the experimental estimates reported there. We also show analogues of our experimental specifications from our pre-analysis plan. The first approach tests for a discontinuity by adding an indicator for rewards greater than or equal to 10 or 100 ("Jump at 10/100"). This level discontinuity is tested in specifications 3 and 4, and there is no evidence of log durations becoming discontinuously larger above either 10 cents or $1.00. The second approach tests for a slope break at $1.00 by estimating a knotted spline that allows the elasticity to vary between 6 and 9 cents, 9 and 10 cents, and then greater than 10 cents, or 81 and 95 cents, 95 cents and $1.00, and then greater than $1.00 up to $1.20. The slope break specification is tested in specifications 5 and 6, where we report the change in slopes at 10 cents and $1.00 ("Spline"). Again, there is no evidence of a change in the relationship between log duration and log reward between 9 and 10 cents, vs greater than 10 cents, or $0.95 and $1.00 versus greater than $1.00.

TABLE E.7—**DURATION OF TASK POSTING BY LOG REWARD AND JUMP AT $1.00**

	(1)	(2)	(3)	(4)	(5)	(6)
Log Wage	-0.089***	-0.066***	-0.089***	-0.069***	-0.090***	-0.070***
	(0.024)	(0.014)	(0.024)	(0.015)	(0.025)	(0.015)
GEQ 10			0.014			
			(0.018)			
GEQ 100				0.027		
				(0.026)		
Spline 10					0.084	
					(0.225)	
Spline 100						0.693
						(0.700)
Double-ML	Y	Y	Y	Y	Y	Y
Window	6-14	80-120	6-14	80-120	6-14	80-120
Sample size	59,654	39,442	59,654	39,442	59,654	39,442

Notes. Sample is restricted to HIT batches with rewards between 80 and 120 cents. Columns 3, 4 and 8 estimate a specification testing for a discontinuity in the duration at $1.00, as in our pre-analysis plan, while columns 5 and 6 estimate the spline specification testing for a change in the slope of the log duration log reward relationship at $1.00, also from the pre-analysis plan. Significance levels are * 0.10, ** 0.05, *** 0.01.

Online Appendix F.
Additional Experimental Details and Specifications from Pre-analysis Plan

Additional specifications allow for heterogeneous slopes in labor supply above and below 10 cents using a knotted spline, where the knots are at $0.09 and 10 cents:

$$\begin{aligned} Accept_i &= \beta_{0B} + \eta_{1B} log(w_i) + \gamma_{2B} \times (log(w_i) - log(0.09)) \times \mathbb{1}\{w_i \geq 0.09\}_i \\ &+ \gamma_{3B} \times (log(w_i) - log(0.10)) \times \mathbb{1}\{w_i \geq 0.1\}_i + \beta_{2B} T_i + \beta_2 X_i + \epsilon_i \end{aligned} \tag{36}$$

Our main test here is that the slope between $0.09 and 10 cents (i.e., $\eta_{1B} + \gamma_{2B}$) is greater than the average of the slopes below $0.09 and above 10 cents $\left(\frac{1}{2} \times \eta_{1B} + \frac{1}{2} \times (\eta_{1B} + \gamma_{2B} + \gamma_{3B})\right)$; or equivalently to test: $\gamma_{2B} - \gamma_{3B} > 0$. Note that $(\gamma_{2B} - \gamma_{3B})$ is analogous to γ_{1A} in the spline specification, and measures the jump at 10 cents.

Finally, our most flexible specification estimates:

$$Accept_i = \sum_{k \in S} \delta_k \mathbb{1}\{w_i = k\}_i + \gamma \beta_{3B} T + \beta_2 X_i + \epsilon_i \tag{37}$$

And then calculates the following statistics:

$$\delta_{jump} = (\delta_{0.1} - \delta_{0.09})$$

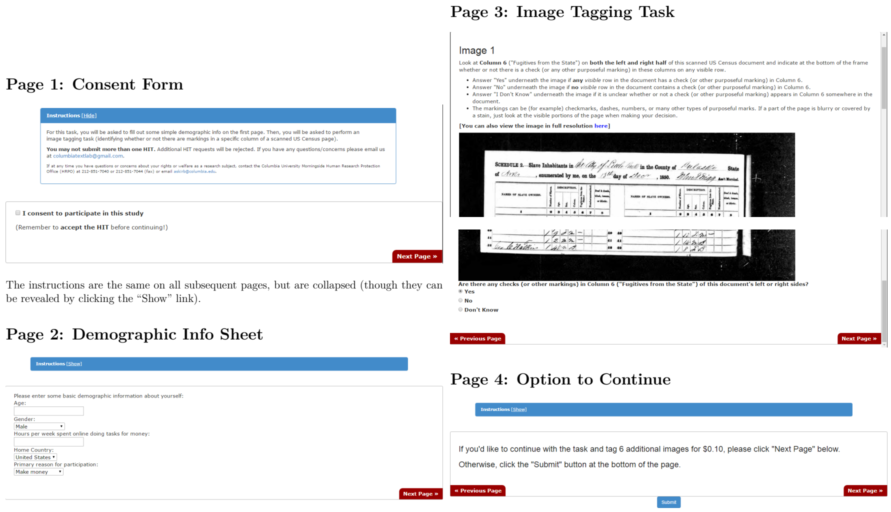
$$\beta_{local} = (\delta_{0.1} - \delta_{0.09}) - \frac{\left(\sum_{k=.08, k \neq 0.1}^{0.12} \delta_k - \delta_{k-0.01}\right)}{4}$$

$$\beta_{global} = (\delta_{0.1} - \delta_{0.09}) - \frac{1}{10}(\delta_{0.15} - \delta_{0.05})$$

The β_{local} estimate provides us with a comparison of the jump between $0.09 and 10 cents to other localized changes in acceptance probability from $0.01 increases. In contrast, β_{global} provides us with a comparison of the jump with the full global (linear) average labor supply response from varying the wage between $0.05 and $0.15.

Figure F.10 shows screenshots from the experimental layout facing MTurk subjects. Table F.8 shows the pre-analysis plan specifications for the accept decision for the first experiment. Across the specifications described above, we see no significant effect of left-digit bias at 10 cents. The pre-analysis plan had levels of wages on the right-hand side, and did not include the log specification shown in the main text, but elasticities are quantitatively extremely close and there is no evidence of left-digit bias in any specification. Table F.9 shows the pre-specified regression with the Any Correct variable as the outcome, to measure possible efficiency wage effects. This table shows no left-digit bias, but also no significant effect of the wage on effort or skill.

FIGURE F.10. **ONLINE LABOR SUPPLY EXPERIMENT ON MTURK**



Notes. The figure shows the screen shots for the consent form and tasks associated with the online labor supply experiment on MTurk.

TABLE F.8—**PREANALYSIS SPECIFICATIONS: TASK ACCEPTANCE PROBABILITY BY OFFERED TASK REWARD ON MTURK**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Wage	0.005	0.012	0.007		0.002	0.003	-0.002		0.009	0.021	0.015	
	[0.003]	[0.006]	[0.004]		[0.004]	[0.008]	[0.005]		[0.004]	[0.009]	[0.006]	
Jump at 10			-0.007				0.019				-0.035	
			[0.017]				[0.022]				[0.026]	
Spline		-0.067				0.188				-0.328		
		[0.161]				[0.213]				[0.242]		
Local				0.006				0.001				0.013
				[0.011]				[0.014]				[0.017]
Global				-0.004				0.012				-0.019
				[0.017]				[0.022]				[0.025]
η	0.065	0.139	0.081		0.022	0.040	-0.017		0.116	0.262	0.193	
	[0.033]	[0.077]	[0.050]		[0.042]	[0.093]	[0.064]		[0.052]	[0.127]	[0.078]	
Sample	Pooled	Pooled	Pooled	Pooled	6 HITs	6 HITs	6 HITs	6 HITs	12 HITs	12 HITs	12 HITs	12 HITs
Sample Size	5184	5184	5184	5184	2683	2683	2683	2683	2501	2501	2501	2501

Notes. The reported estimates are linear regressions of task acceptance probabilities on log wages, controlling for number of images. Column 1 reports specification 1 that estimates the labor-supply elasticity, without a discontinuity. Column 2 estimates specification 2, which tests for a jump in the probability of acceptance at 10 cents. Column 3 estimates a knotted spline in log wages, with a knot at 10 cents, and reports the difference in elasticities above and below 10 cents. Column 4 estimates specification 4, including indicator variables for every wage and testing whether the difference in acceptance probabilities between 10 and 9 cents is different from the average difference between 12 and 8 (local) or the average difference between 5 and 15 (global). Columns 5-8 repeat 1-4, but restrict the sample to "sophisticates": Turkers who respond that they work more than 10 hours a week and their primary motivation is money. Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$

TABLE F.9—**PREANALYSIS SPECIFICATIONS: TASK CORRECT PROBABILITY BY OFFERED TASK REWARD ON MTURK**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Wage	-0.001	0.001	-0.001		0.001	0.005	0.001		-0.003	-0.004	-0.002	
	[0.001]	[0.003]	[0.002]		[0.002]	[0.004]	[0.003]		[0.002]	[0.004]	[0.002]	
Jump at 10			-0.001				0.000				-0.003	
			[0.007]				[0.011]				[0.009]	
Spline		-0.013				0.006				-0.030		
		[0.070]				[0.105]				[0.089]		
Local				0.001				0.003				-0.002
				[0.005]				[0.006]				[0.007]
Global				-0.004				-0.008				0.001
				[0.007]				[0.010]				[0.010]
η	-0.008	0.012	-0.005		0.008	0.047	0.008		-0.026	-0.035	-0.021	
	[0.013]	[0.026]	[0.019]		[0.019]	[0.039]	[0.029]		[0.015]	[0.031]	[0.024]	
Sample	Pooled	Pooled	Pooled	Pooled	6 HITs	6 HITs	6 HITs	6 HITs	12 HITs	12 HITs	12 HITs	12 HITs
Sample Size	5184	5184	5184	5184	2683	2683	2683	2683	2501	2501	2501	2501

Notes. The reported estimates are linear regressions of task acceptance probabilities on log wages, controlling for number of images. Column 1 reports specification 1 that estimates the labor-supply elasticity, without a discontinuity. Column 2 estimates specification 2, which tests for a jump in the probability of acceptance at 10 cents. Column 3 estimates a knotted spline in log wages, with a knot at 10 cents, and reports the difference in elasticities above and below 10 cents. Column 4 estimates specification 4, including indicator variables for every wage and testing whether the difference in acceptance probabilities between 10 and 9 cents is different from the average difference between 12 and 8 (local) or the average difference between 5 and 15 (global). Columns 5-12 repeat 1-4, but restrict attention to subsamples based on the number of images given (randomized to be either 6 or 12). Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$

Online Appendix G. Implications for wage dynamics

The presence of round-number bunching has economically important implications in understanding how wages respond to various shocks.[25] In this section we discuss two such examples relevant to recent literatures. First, we argue how presence of round-number bunching creates a novel source of wage spillovers from minimum wages higher up in the distribution. Second, we discuss how bunching can also imply wage responses to productivity or payroll tax shocks can be both nonlinear and heterogeneous by types of firms.

Wage spillovers from minimum wages

If employers are mispricing, then minimum wage changes can have heterogeneous effects depending on whether they cross a round number. Minimum wages that pass through a round number will induce additional spillovers distinct from those that do not. To see this, consider a small increase in the minimum wage when it is initially equal to the round number, w_0. In this case, there will be a mass in jobs that pay $w_{min} = w_0$ that is the composed of two sets of firms: those firms that are bound by the mandated wage ("bound by minimum wage") and those that are misoptimizing and paying a round wage ("bunchers from above"). Note here that since mis-optimizing bunchers from below are still bound by the minimum wage, only $\delta/2$ of the firms are bunching down to the minimum (for simplicity, here we are assuming the distribution of firms around w_0 is symmetric).

$$g(w_{min}) = \underbrace{\frac{\delta}{2}\int_{\frac{w_0}{\mu}}^{\frac{w_u}{\mu}} l(w_{min})f(p)dp}_{\text{bunchers from above}} + \underbrace{(1-\frac{\delta}{2})\int_{w_{min}}^{\frac{w_{min}}{\mu}} l(w_{min})f(p)dp}_{\text{bound by minimum wage}} \quad (38)$$

In the Butcher et al. (2012) version of the monopsonistic competition model, the minimum wage has 2 effects: it forces exit of low productivity firms, but forces higher productivity firms to raise their wages to the minimum. With full employment, workers who lose their job are reallocated to higher paying jobs, so there are increases in employment at wages above the new minimum. With bunching a third force is added: the effect of increasing w_{min} on the distribution $g(w)$ will depend on where w_{min} sits relative to w_0 and the extent of bunching. The effect of increasing the minimum wage slightly from w_0 to w'_{min} eliminates both sources of the mass point at w_0, but the "bunchers from above" set wages according to their latent wage w, while those who are bound by the minimum wage (and do not exit) set wages at the new minimum w'_{min}. Once the round number w_0 is unavailable, wages of those bunching from above jump up to the latent wage which exceeds the new w'_{min} for a small increase. Relative to a minimum wage increase that does not begin at w_0 (or cross w_0), this results in larger increase in jobs paying between the new minimum w'_{min} and w_u than at all other wages $w > w_u$. This is an entirely new reason for spillovers than has been considered in the literature; moreover, it suggests that minimum wage spillovers are likely to be particularly large when the minimum wage crosses an important round number mode in the distribution (e.g., \$10 or \$15).[26]

[25]Note that the direct efficiency and distributional implications of employer misoptimization in an imperfect competition context are small. In our baseline estimates a δ^* of 5% implies only a deadweight loss and decrease in labor share of 0.2% , because bunchers from below reduce the monopsony distortion by overpaying relative to the monopsonist wage even as bunchers from above exacerbate it by underpaying.

[26]While a bit further away from our baseline model, of interest is the case where there is another mode, $w_1 > w_0$, for example

Passthrough, Rent-Sharing, and Payroll Tax Incidence

A second example concerns how the presence of round number bunching affects passthrough of productivity or taxes to wages. Intuitively, when some employers are forgoing monopsony profits by paying round number wages, small changes in the marginal product of labor, say due to productivity or tax shocks, will not translate into a realized wage increase. The resulting estimates of structural parameters (like degree of market power) can therefore be biased, just as in the aggregate labor supply elasticity literature (Chetty, 2012).

In our model with constant elasticity, passthrough rates would be $\frac{\Delta w}{\Delta p} = (1-\delta)\mu + \delta\mu\mathbf{1}(\frac{\frac{\Delta\pi}{\Delta p}}{\pi(p)} > \delta^*) \approx (1-\delta)\mu + \delta\mu\mathbf{1}(\frac{\Delta p}{p}(1+\eta) > \delta^*)$, where $\mu = \frac{\eta}{\eta+1}$. In other words, large percentage changes in p will result in passthrough estimates of μ but small percentage changes in p will result in passthrough estimates of $(1-\delta)\mu$. Taking an estimate of $\eta = 3$ and fixing $\delta^* = .05$, it would imply that increases in value-added per worker less than 2.5% would recover estimates of passthrough roughly 10% smaller than those estimated from larger increases.

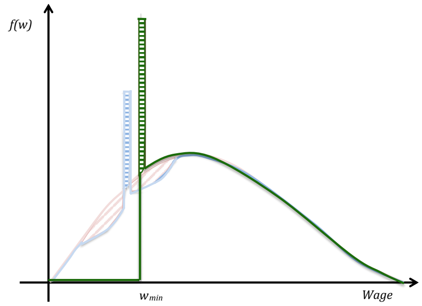
Recent rent-sharing estimates provide some evidence of this nonlinearity. For example, Figure 2 of Kline et al. (2019), shows a clear non-linearity in the response of the wage bill to surplus per worker: patents that create a large percentage change to firm surplus per worker, also generate a positive percentage increase in the wage bill, but much smaller percentage increases in firm surplus per worker do not.. Similarly Garin and Silvério (2024) report concave effects of levels of rent on wages (Table A.5), consistent with larger percentage changes in rent having larger effects on wages. Further, the nonlinearity in passthrough implies that small payroll tax changes are completely borne by misoptimizing employers who continue to pay w_0, suggesting that small tax changes will underestimate behavioral responses.[27] In particular, as we find low-wage employers are particularly likely to misprice labor and use round numbered wages, we expect that the wage response to a (large) revenue shock is likely to be more pronounced among small firms, as a large enough shock would lead these firms to "over-adjust" (just like a small shock would lead these firms to "under-adjust.").[28] While there are numerous possible reasons for why rent sharing elasticities may vary by the size of change in value added, we see round-number bunching due to misoptimization as an additional source of nonlinearities and heterogeneity in passthrough worth exploring in future research.

w_1 = \$11. If some firms are particularly prone to paying whole numbered wages so that $\pi(w_0) > \pi(w_1) > (1-\delta)\pi(w)$, then when the minimum wage crosses w_0, it might lead these firms (initially paying w_0) to jump to w_1, creating a sizable spillover and a larger spike at the next round number. This is consistent with recent findings in Derenoncourt and Weil 2025 , who study Burning Glass wage postings in counties with Amazon distribution centers after Amazon raised its entry wage to \$15, and find that the mass of postings at exactly \$20.00 and \$25.00 increased more than any other wage greater than \$16. This is consistent with some non-Amazon employers (initially paying more than Amazon) responding to Amazon's \$15 minimum wage by raising their wage to the next round number , e.g., \$20 or \$25.

[27] Note that under monopsony employers already bear a significant share of the payroll tax incidence, and do not shift it all to workers, consistent with Anderson and Meyer (2000). Conlon and Rao (2020), who show lumpy 1\$ units of price adjustment of liquor stores to excise taxes, similarly show undershifting of small taxes.

[28] Some suggestive evidence comes from Risch (2024) who find that small firms (under 100 workers), were much more likely to pass through tax cuts (on owners' income) to workers' wages than middle or large sized firms.

FIGURE G.11. **SPILLOVERS FROM MINIMUM WAGES IN THE PRESENCE OF ROUND NUMBER BUNCHING**



Notes. The red line represents the latent wage distribution in the absence of a minimum wage. The blue line represents the wage distribution in the presence of employer misoptimization where there is bunch at w_0, but still without a minimum wage. The solid green line represents the wage distribution in the presence of a minimum wage that exceeds w_0, but without any maket-level reallocation effects (i.e., lost jobs below w_{min}getting reallocated to firms paying at or above w_{min}.