Supplemental Online Appendix For Monopsony and Employer Mis-optimization Explain Why Wages Bunch at Round Numbers

Arindrajit Dube, Alan Manning, and Suresh Naidu

April 18, 2025

Contents

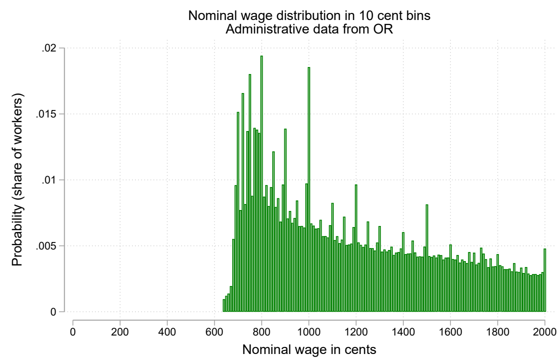
Online Appendix A
Additional Figures and Tables

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.

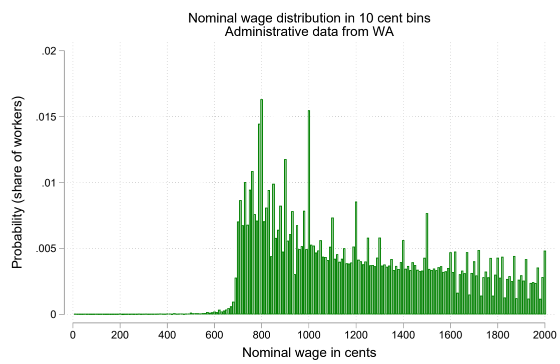
Figure A.1: **Histograms of Hourly Wages In Administrative Payroll Data from Minnesota, Oregon, and Washington, 2003-2007**

Panel A: Minnesota

Panel B: Oregon

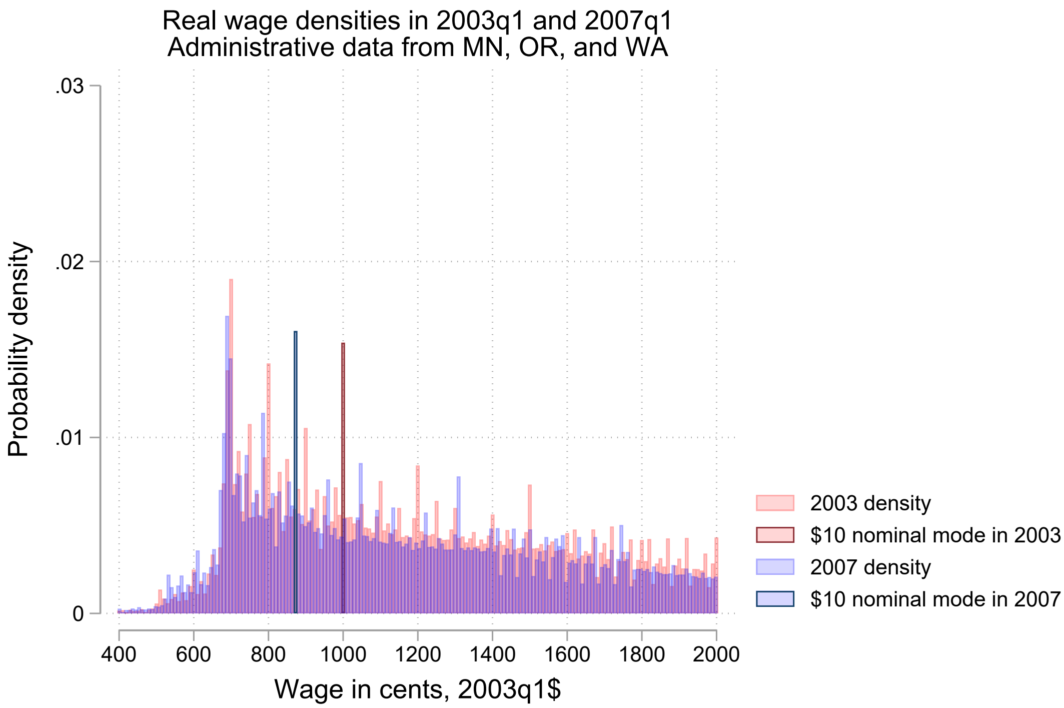


Panel C: Washington



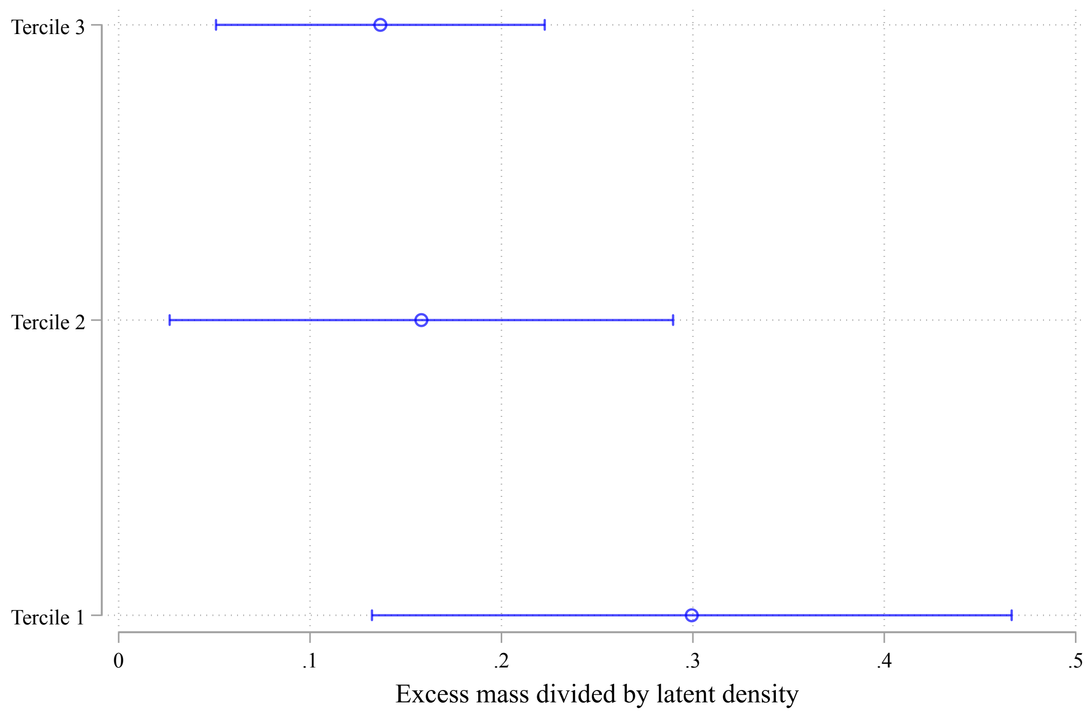
Notes. The figure shows histograms of hourly wages in 10 cents (nominal) wage bins, averaged over 2003q1 to 2007q4, using administrative Unemployment Insurance payroll records from the states of Minnesota (Panel A) , Oregon (Panel B), and Washington (Panel C). Hourly wages are constructed by dividing 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:

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

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:

$$z_0 = \frac{\delta + \gamma}{\eta\,(1+\eta)}, z_1 = \frac{\delta}{\eta\,(1+\eta)} \tag{2}$$

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 (2) 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 (5) implies that we will have, for $\omega < 0$, :

$$\phi(\omega_*) = 1 - \Lambda_0^z\left[\frac{\omega_*^2}{2}\right] \tag{3}$$

and for $\omega > 0$:

$$\phi(\omega^*) = 1 - \Lambda_1^z\left[\frac{\omega^{*2}}{2}\right] \tag{4}$$

In this paper, we empirically recover estimates of the left-hand sides of (3) and (4). 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:

$$\frac{2\delta}{\eta\,(1+\eta)} = \omega^2 \tag{5}$$

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). [1]

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

[1] 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).

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:

$$\phi(\omega) = [1 - \underline{G}]\, I\left[\omega^2 < \frac{2\delta^*}{\eta\,(1+\eta)}\right] \tag{6}$$

In this model, the share of jobs with a latent wage close to the bunch that continue to pay a non-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.[2]

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.[3] Finally, we continue to allow arbitrary heterogeneity

[2]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.

[3]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

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 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.

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).

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:

$$phi\,(\omega) = 1 - \hat{G}(\eta(1+\eta)\frac{\omega^2}{2}) \tag{7}$$

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

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

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 (8) 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:

$$\hat{H}(\eta) = \frac{\hat{\phi}\left(\sqrt{\left(\frac{8\delta^*}{(2\eta+1)^2 - 1}\right)}\right)}{1 - \underline{G}} \tag{9}$$

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

$$E(\hat{\eta}) = \int_0^{\infty} \eta d\hat{H}(\eta) \tag{10}$$

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 5 (again maintaining that $\gamma = 0$) we get

$$2\ln(\omega) = ln(2) - \ln(\eta(1+\eta)) + \ln(\delta) = ln(2) - \ln(\eta(1+\eta)) + E[\ln(\delta)\,|\delta > 0] + \ln(\delta_{res}) \quad (11)$$

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 $U = \ln(\delta_{res})$.)

$$W = X + U \quad (12)$$

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:

$$\widehat{f_X} = \frac{1}{2\pi}\int e^{-itx} K(th) \frac{\sum_{j=1}^{N} \frac{e^{itW_j}}{N}}{\phi(t)} dt \tag{13}$$

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.

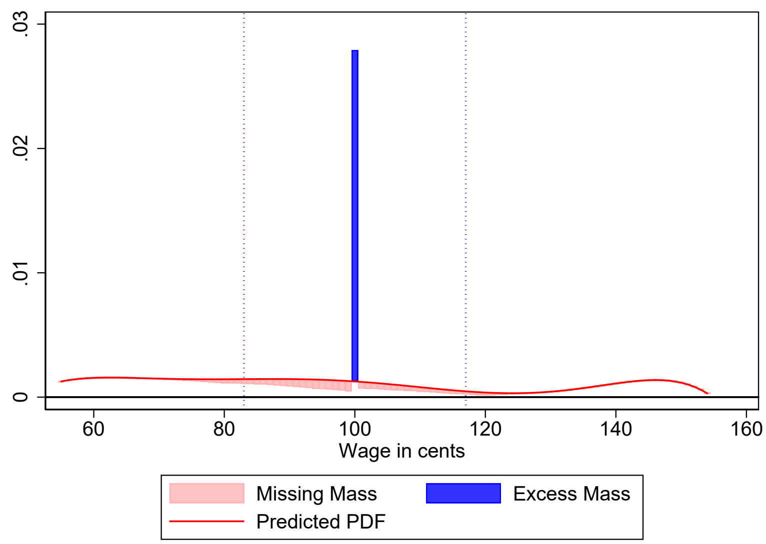
For normal $U = \ln(\delta_{res})$, Delaigle and Gijbels (2004) suggest a kernel of the form:

$$K(x) = 48\frac{cos(x)}{\pi x^4}(1 - \frac{15}{x^2}) - 144\frac{sin(x)}{\pi x^5}(1 - \frac{5}{x^2}) \tag{14}$$

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.

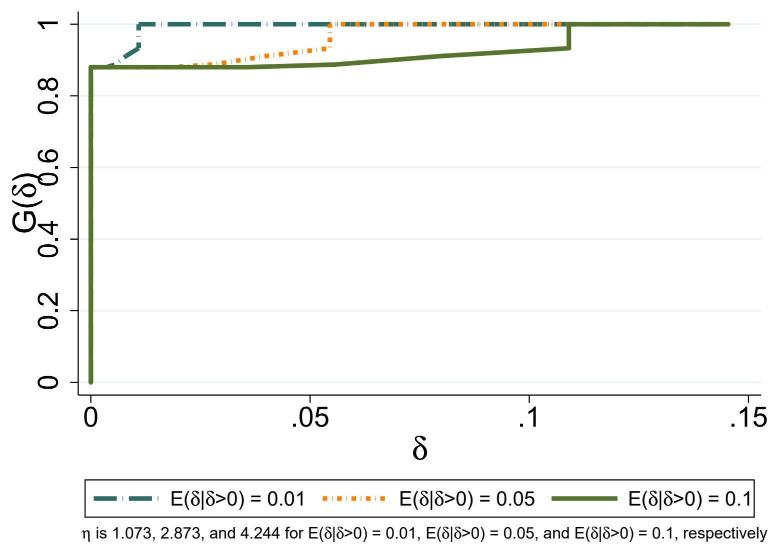
In Figure C.7, we show the distribution of η using the deconvolution estimator, assuming a lognormal distribution of δ. In the first panel, we estimate $H(\eta)$ assuming the standard deviation $\sigma_{ln(\delta)} = 0.1$, which is highly concentrated around the mean. In the second panel, we instead 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**



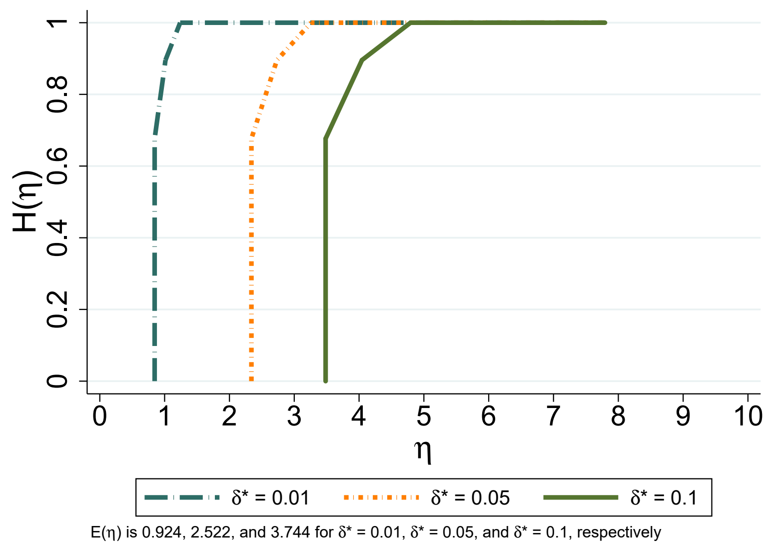
Notes. This figure plots the excess and missing mass around 1.00 on Amazon Mechanical Turk. The latent distribution is modelled with a 6th degree polynomial.

Figure C.5: **Implied Distribution of δ Under Constant η**



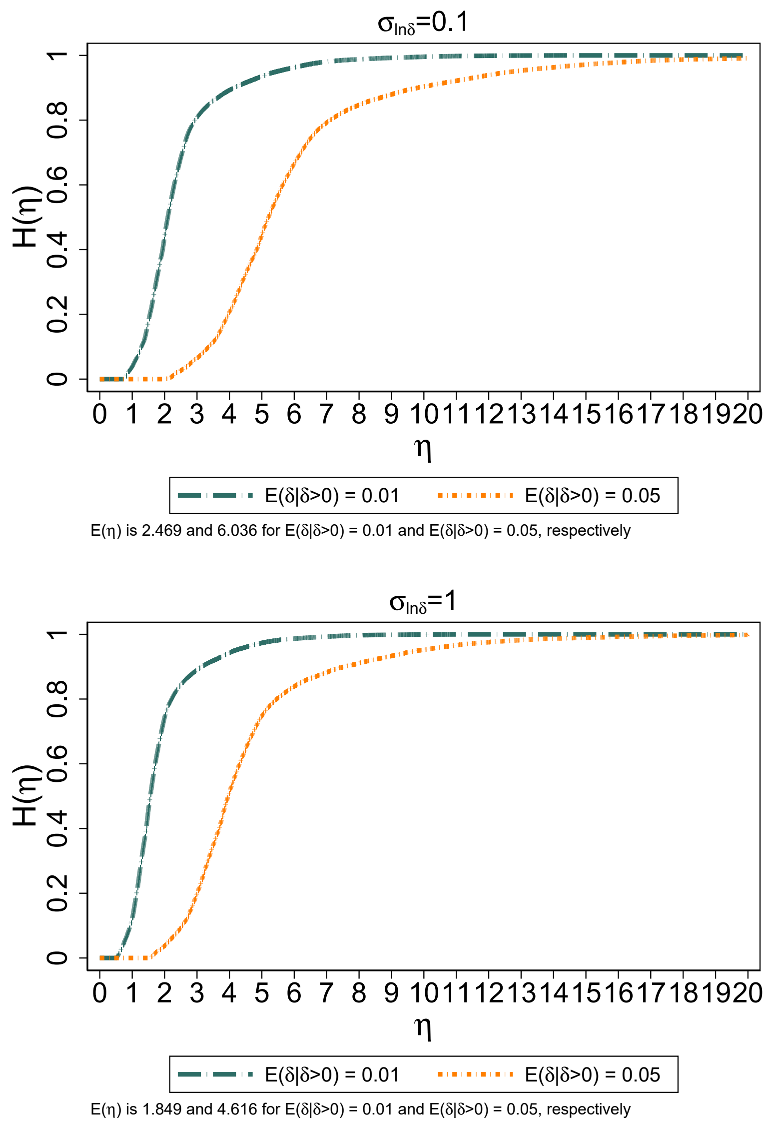
Notes. The figure plots the cumulative distributions $G(\delta)$ based on equation 8, for alternative values of $E(\delta|\delta > 0)$. The elasticity η is assumed to be a constant. The estimates use administrative hourly wage data from MN, OR, and WA.

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 9, 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 11, 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 $\sigma_\delta = 0.1$; in contrast, the bottom panel assumes a rather dispersed distribution with $\sigma_\delta = 1$. 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. $\delta^* = 0.01$				
$\overline{\delta}$	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. $\delta^* = 0.05$				
$\overline{\delta}$	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. $\delta^* = 0.1$				
$\overline{\delta}$	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 5 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. $E(\delta \mid \delta > 0)$= 0.01				
$\bar{\delta}$	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. $E(\delta \mid \delta > 0)$= 0.05				
$\bar{\delta}$	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 8 and 9, 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 11. 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. $\delta^* = 0.01$							
$\overline{\delta}$	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. $\delta^* = 0.05$							
$\overline{\delta}$	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. $\delta^* = 0.1$							
$\overline{\delta}$	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 5 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

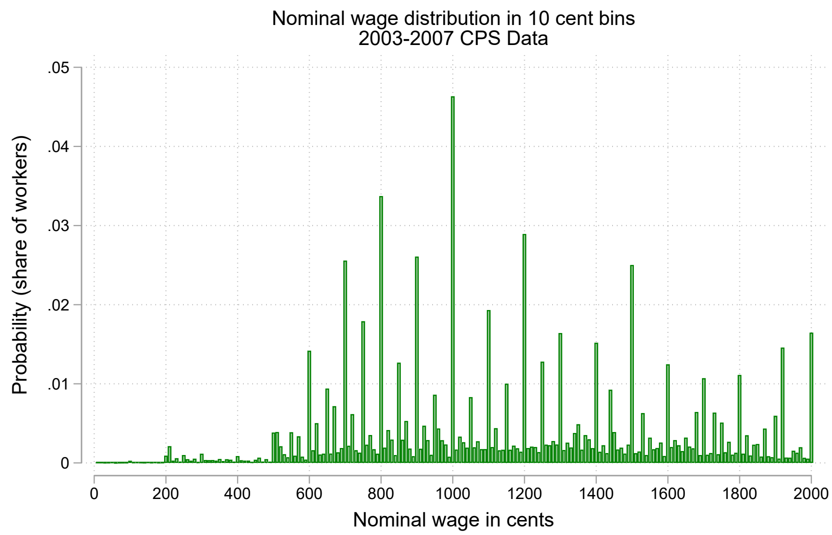
In this appendix, we show the degree of bunching in hourly nominal wage data using the national CPS data. In Figure D.8, we plot the nominal wage distribution in U.S. in 2003 to 2007 in 10 cents bins. There are notable spikes in the wage distribution at $10, $7.20 (the bin with the federal minimum wage), $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

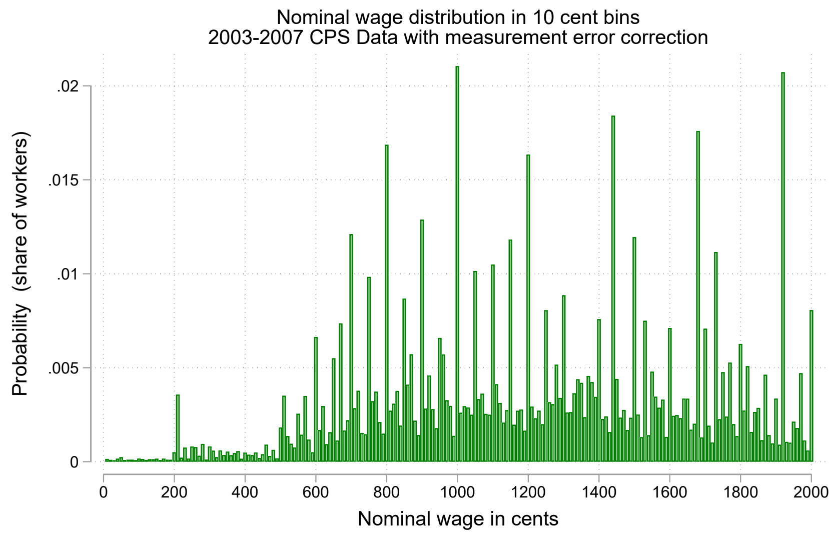
In Appendix Table D.6 we estimate the implied η for different δ^* under our baseline 2-point model across subgroups of the measurement corrected CPS data, as we do not have worker-level covariates for the administrative data. We examine young and old workers, as well as male and female separately. Consistent 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 $\delta = 0.01$; whereas it is $\delta = 0.05$ 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 5 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:

$$\ln(duration_h) - \ln(\widehat{duration_h}) = \eta \times (\ln(wage_h) - \ln(\widehat{wage_h})) + \gamma \mathbf{1_{w>w_0}} + \epsilon \tag{15}$$

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:

$$\begin{aligned} Accept_i = \beta_{0B} + \eta_{1B} log(w_i) + \gamma_{2B} \times (log(w_i) - log(0.09)) \times \mathbb{1}\{w_i \geq 0.09\}_i \\ + \gamma_{3B} \times (log(w_i) - log(0.10)) \times \mathbb{1}\{w_i \geq 0.1\}_i + \beta_{2B} T_i + \beta_2 X_i + \epsilon_i \end{aligned} \quad (16)$$

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:

$$Accept_i = \sum_{k \in S} \delta_k \mathbb{1}\{w_i = k\}_i + \gamma \beta_{3B} T + \beta_2 X_i + \epsilon_i \quad (17)$$

And then calculates the following statistics:

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

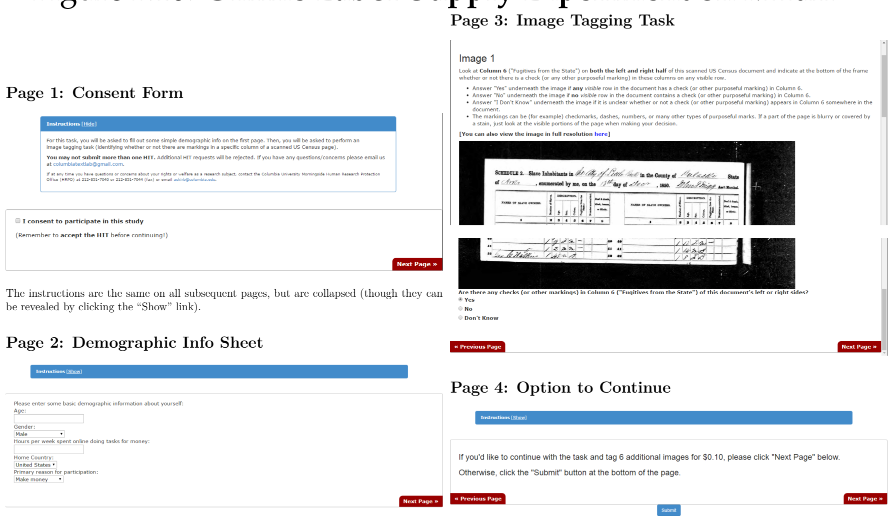
$$\beta_{local} = (\delta_{0.1} - \delta_{0.09}) - \frac{\left(\sum_{k=.08, k \neq 0.1}^{0.12} \delta_k - \delta_{k-0.01}\right)}{4}$$

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

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.[4] 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).

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

(18)

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"

[4] 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.

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).[5]

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

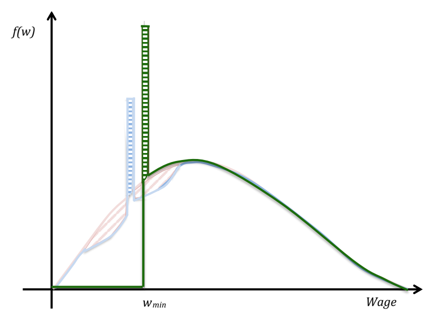
[5]While a bit further away from our baseline model, of interest is the case where there is another mode, $w_1 > w_0$, for example $w_1 = \$11$. If some firms are particularly prone to paying whole numbered wages so that $\pi(w_0) > \pi(w_1) > (1-\delta)\pi(w)$, then when the minimum wage crosses w_0, 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.

behavioral responses.[6] 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.").[7] 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.

[6]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.

[7]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}.

References

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

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.

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.

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.

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.

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, 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.

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

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.

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.

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

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.

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

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

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.