Willingness to Pay for Workplace Amenities*

Massimo Anelli[1] and Felix Koenig[2]

[1]Bocconi University, IZA, CESifo

[2]Carnegie Mellon University, CEP, IZA

May 5, 2025

Abstract

We develop a revealed preference approach to measure the value of workplace amenities by analyzing how variation in non-wage job attributes affects excess mass in the earnings distribution at budget discontinuities. The approach formalizes the idea that workers are less responsive to monetary incentives when amenities constitute a larger share of total compensation. Applying this method to workplace safety during COVID-19 waves, we find that workers are willing to sacrifice 9% of their earnings to reduce weekly fatality risks by one in 100,000. The findings suggest that conventional hedonic regressions substantially underestimate the value of workplace safety.

JEL-Codes: J17, J28, J32
Keywords: Non-wage Amenities, Labor Supply, Bunching, Workplace Safety, Value of Life, Job Satisfaction.

---

*Massimo Anelli, massimo.anelli@unibocconi.it, Bocconi University, Department of Social and Political Science. Felix Koenig, fkoenig@cmu.edu, Carnegie Mellon University, Heinz College. We are grateful to Elvin Bora and Akseli Palomaki for excellent research assistance. We thank Thomas Le Barbanchon, Giulia Giupponi, Hilary Hoynes, Brian Kovak, Matthew Notowidigdo, Lowell Taylor, Andrea Weber and seminar participants at NBER Summer Institute - Labor Studies, NBER COVID-19 and Health Outcomes Conference, Bocconi University, Boston College, Boston University, EEA-ESSEM, McGill, Barcelona summer forum, SOLE/MEA, North American Meetings of the Econometric Society, SOFI - Stockholm University for helpful comments. We gratefully acknowledge funding from the Block Center for Technology and Society at Carnegie Mellon University and support from the "Adopt a Paper mentoring program".

1 Introduction

Non-wage amenities significantly influence labor market dynamics, shaping job preferences and wage structures.[1] Workers value a wide range of amenities including workplace safety, meaningful work, and supportive work environments and often compel workers to make trade-offs between monetary and non-monetary compensation. The seminal literature on "compensating differentials" proposed a framework to quantify the value of amenities. The approach theorizes that the value of amenities is reflected in wage differences for otherwise similar jobs. However, estimating such compensating differentials in practice has proven difficult.[2]

This paper introduces a novel approach to measure the value of workplace amenities. Our strategy builds on the quasi-experimental approach developed in the bunching literature.[3] This literature leverages that budget discontinuities (e.g., earning limits for benefit eligibility) create financial incentives to work at or below a particular earning level. We argue that these incentives only apply to earnings (the monetary part of compensation) and not to non-monetary returns to work. The incentive effect of budget discontinuities is thus smaller if non-monetary compensation is a bigger share of total compensation. That means that workers are less likely to respond to a budget discontinuity when they derive large non-wage returns from working and the excess mass in the earnings distribution at the budget discontinuity will be smaller (i.e., there is less "bunching"). Conversely, when work produces substantial dis-amenities (e.g., health risks), the returns to work are mostly financial and budget discontinuities will

[1]This includes both classic Rosen 1986, 1974; Lucas 1977; Masters 1969 and recent studies Lavetti and Schmutte (2022); Lagos (2020); Roussille and Scuderi (2022); Lehmann (2022); Sockin (2022); Taber and Vejlin (2020); Morchio and Moser (2019); Goldin and Katz (2011, 2016); Card, Heining, and Kline (2013); Pierce (2001); Hamermesh (1999).

[2]Reviews of the empirical challenges estimates highlight issues from frictional wage setting and endogenous job mobility (Ashenfelter, 2006; Lavetti, 2023a).

[3]Card, Lee, Pei, and Weber (2017); Card, Johnston, Leung, Mas, and Pei (2015a); Card, Lee, Pei, and Weber (2015b); Kleven (2016a); Chetty, Friedman, and Saez (2013) and many more papers that leverage discontinuities in UI benefit schedules to study the effect of UI benefits on labor supply (reviewed in Schmieder and Von Wachter (2016)).

trigger larger responses and more excess mass in the earnings distribution around the relevant earning threshold. Changes in the relative importance of monetary vs. non-monetary compensation result in a change in excess mass. We show that this change in excess mass reveals the willingness to pay (WTP) for the change in amenities.

Our main empirical application implements this method to estimate the value of workplace safety during Covid-19 outbreaks. The identification strategy exploits earnings eligibility thresholds for partial unemployment insurance (UI). These thresholds are generated or amplified by the Federal Pandemic Unemployment Compensation (FPUC) scheme in March 2020, which provided workers with an additional $600 weekly UI payments. Unlike regular UI payments, FPUC payments were available in full below the threshold and not at all above the threshold, creating a "jump" (or notch) in the worker's budget constraint at the pre-existing partial UI threshold.[4] We use panel data for a within-worker analysis that compares behavior during high-risk weeks to low-risk ones. Findings show that workers are more likely to bunch when risks are higher. Complementary evidence shows similar results in cross-industry comparisons. Workers engaged in tasks with the greatest deterioration of health risk during the Covid-19 pandemic are more likely to bunch than workers in less affected tasks. We show that these differences are not explained by differences in economic shocks, differences in preferences, or differences in adjustment frictions across sectors. We also demonstrate that a placebo specification on workers ineligible for the UI benefit shows no evidence of spurious excess mass.

Our estimates imply that a one standard deviation increase in workplace risk is equivalent to a reduction in wage by around 30%. To put this magnitude into context, an increase in risk by one standard deviation is similar to switching from the safest to the riskiest occupation in normal times. Converting standard deviations into interpretable units, our estimates imply a willingness to pay of 9% of earnings to lower work

[4]Thresholds differ by state. Institutional details of the FPUC scheme are presented in Appendix C.1.

fatality risk by one in 100,000, a variation equivalent to the risk difference between a librarian and a roofer. While our findings align with the statistical value of life estimates (Viscusi, 2018), they markedly diverge from results obtained through canonical hedonic wage approaches. Using the same data and setting, the hedonic regression yields a willingness to pay of only 0.5% for one standard deviation of risk, which is nearly two orders of magnitudes smaller than our baseline 30% WTP estimate. Frictions in wage setting may explain why hedonic regressions show a much smaller WTP. If frictions prevent wages from adjusting, hedonic regressions will be biased downward. Relaxing the assumption of frictionless wage setting is one of the advantages of the approach developed in this paper.

The method developed in this work relies on two identification assumptions. The first is the standard "**smoothness assumption**" of the bunching literature, which assumes that confounding shocks (e.g. shocks to labor *demand*) are smooth at the threshold. This assumption guarantees that behavior around the budget discontinuity identifies labor *supply* choices. The second identification assumption requires that workers exposed to high and low amenity levels respond similarly to monetary incentives, which we refer to as the **preference-orthogonality assumption**. In a within-individual design, this assumption is inherently less problematic since we compare the same worker under different risk conditions, ensuring that their underlying responsiveness to financial incentives remains constant. However, when amenities vary between individuals rather than within, the assumption becomes more demanding. In this case, researchers must account for potential differences in labor supply elasticities across groups, as variation in responsiveness could confound estimates of willingness to pay for workplace amenities.

Several tests investigate the validity of these assumptions in our empirical application. We begin by addressing the "smoothness assumption." In a standard cross-sectional analysis, this requires that the counterfactual earnings distribution is smooth through the notch. There has been a debate whether it is feasible to identify a counterfactual distribution from a cross-sectional earnings distribution Blomquist et al. (2021). Our setting circumvents this problem, since we observe an untreated earnings distribution before the notch takes effect. The identification than relies on a slightly weaker assumption that *changes* in the earnings distribution are smooth through the notch. Any observed changes around the notch are therefore attributable to labor supply responses rather than spurious shocks. We confirm the validity of this assumption using a state-border design, a placebo sample and by directly controlling for potential spurious effects, from childcare constraints or local demand conditions. All of these tests support the smoothness assumption.

We probe the second "preference-orthogonality assumption" in two ways. First, we test whether labor supply elasticities are correlated with risk increases and find that the correlation of the two is extremely low. This suggests there is little potential for omitted variable bias. Second, we implement a within-worker design to probe whether the baseline results reflect differences in worker characteristics between workers exposed to higher or lower risks. This design compares the earnings of an individual in weeks when workplace risks are high with those in weeks when risks are lower. These within-worker results are similar to cross-industry results, where risk exposure varies across- rather than within-individuals. A more subtle possibility is that labor supply elasticities change over time for a given individual. For example, people's labor supply elasticity may vary with the economic cycle, which may invalidate even the within-individual design. We address this possibility by introducing controls for varying probabilities of bunching over the cycle. Such controls have little effect on the result.

A further potential identification threat in our empirical setting are adjustment frictions (e.g. inability of workers to set work-hours, search friction, etc.). Such frictions play an important role in canonical bunching estimates and we explore their impact for the WTP procedure developed in this paper. The theoretical section shows that the adjustment friction typically discussed in the bunching literature has no effect on the WTP estimator. The WTP estimator is a ratio of two bunching estimates, which cancels out the impact of any friction that affect both bunching estimates proportionally, as is the case with the optimization friction discussed in the bunching review by Kleven (2016b). We test that this theoretical construct applies to the data by showing that our results are robust to holding employers' schedule flexibility fixed using workplace fixed effects. A related but distinct concern about frictions is their impact on the location of excess mass. When workers cannot target specific earnings levels, excess mass may not spike exclusively at the notch earning, but could arise over a wider range of earnings. We adjust the empirical design to allow for excess mass over a wider interval.

Finally, we illustrate that the framework can be used to value broader bundles of amenities. Previous work on amenities estimates the value of *specific* amenities (e.g., workplace safety, parental leave, etc.) or the overall importance of amenities for compensation. Our approach can be flexibly used for either application, and we demonstrate this with a second empirical application aimed at estimating the value of broader amenity bundles. Specifically, we estimate the monetary value of a job with greater (self-reported) job satisfaction. The identification strategy leverages discontinuities in the lifetime budget constraint, generated by social security benefit rules. We find that both satisfied and less satisfied workers are more likely to retire at the threshold age, but this excess mass is bigger for individuals with lower job satisfaction. On average, less satisfied worker are willing to give up an extra two quarters of earnings to retire earlier and this difference cannot be explained by observable differences in education, health, industry, occupation, location. Through the lens of our model, this implies that holding a one likert point more satisfying job is equivalent to 12.5% higher earnings.

Related Literature – A large literature on hedonic regressions estimates how wages vary with amenities at different occupations (Lucas, 1977; Hwang, Reed, and Hubbard, 1992; Guardado and Ziebarth, 2019; Lee, 2022).[5] After controlling for selection effects with individual fixed effects, such studies consistently find that amenities have only

[5]Hedonic regressions are also popular for non-health related applications, e.g., Summers (1989); Gruber and Krueger (1991); Gruber (1994, 1997); Fishback and Kantor (1995); Stern (2004).

small effects on wages, leading to the impression that amenities account for a minor part of workers' compensation (Brown, 1980; Kniesner et al., 2012; Viscusi and Aldy, 2003).[6] A recent re-assessment of these estimates challenges this conclusion and finds a more prominent role for amenities. Such studies show that careful modeling of imperfect competition and/or endogenous job switching can reconcile large valuations of amenities with small hedonic regression results (Altonji and Paxson, 1992; Bonhomme and Jolivet, 2009; Ruppert, Stancanelli, and Wasmer, 2009; Lang and Majumdar, 2004; Lamadon, Mogstad, and Setzler, 2022; Bell, 2022; Lavetti and Schmutte, 2022).

Another complementary literature uses survey experiments (vignettes) to study the willingness to pay for workplace amenities and also provides evidence that amenities are more important than classic hedonic regressions suggest (Wiswall and Zafar, 2018; Mas and Pallais, 2017; Chen et al., 2021; Le Barbanchon, Rathelot, and Roulet, 2021; Dube, Naidu, and Reich, 2022; Einarsen et al., 2011; Maestas et al., 2018; Folke and Rickne, 2022; Tsao, 2024). However, concerns remain that stated preferences in survey experiments may not align with actual labor market behavior. While some studies attempt to mitigate this issue by embedding vignettes in hiring processes (Mas and Pallais, 2017), direct field experiments on workplace amenities are rare due to ethical and practical constraints. As a result, much of what is known about amenities relies on stated rather than revealed preferences.

Critics push back against both approaches, arguing that they require strong assumptions. The revised hedonic estimates only yield valid results if the equilibrium wage process is accurately modeled (accounting for search frictions, unions, efficiency wages, minimum wages, and other factors that influence equilibrium wages), while survey experiments assume that stated preferences match real-world decisions.[7]

Our results support the findings of these recent studies and show that non-wage

[6] Several studies leverage quasi-random variation in amenities to improve identification of the role of amenities (Lavetti, 2020; Gruber, 1997; Fishback and Kantor, 1995; Gruber and Krueger, 1991; Summers, 1989).

[7] Mas and Pallais (2017) address hypothetical bias by embedding a vignette study in the hiring process of call-center workers, making the responses incentive-compatible.

amenities play an important role in the labor market while sidestepping both criticisms. Our method uses a *revealed preferences* approach, avoiding potential biases from hypothetical survey responses. Moreover, the method shifts the focus away from equilibrium wages. Instead, we focus on an outcome—excess mass at benefit thresholds—that, by construction, is independent of the equilibrium wage-setting process. An additional advantage of our method is that the threshold-based research design is unaffected by any spurious variable with smooth effects through the cut-offs. This reduces the threat created by spurious labor market changes that coincide with changes in amenities and provides an identification strategy that estimates the willingness to pay for amenities without quasi-random variation in amenities.

Our work also relates to two studies that analyze how workplace amenities affect labor supply decisions (Sorkin, 2018; Powell, 2012). Powell (2012) shows that the presence of amenities reduces the tax elasticity of labor supply and Sorkin (2018) shows that one can rank the quality of employers by studying employer switching behavior. Unlike such an ordinal measure, our paper develops an estimator that provides cardinal metrics for the value of amenities. Furthermore, we extend the scope of amenity valuation from assessing the overall bundle of amenities offered by an employer to quantifying the value of *specific* amenities.

Finally, our estimate of a monetary value for avoiding workplace risk also relates to the large literature on health and safety at workplaces (for a summary, see Ruser and Butler, 2009). Recent contributions include studies of the (income or consumption) loss associated with falling ill (Dobkin et al., 2018), the causes for workplace illnesses and injuries (e.g., Pichler and Ziebarth, 2019; Johnson, 2020; Johnson, Levine, and Toffel, 2022), and a large literature on the "value of a statistical life" (prominent examples include Ashenfelter and Greenstone, 2004; Viscusi and Aldy, 2003).

2 Estimating the Value of Amenities from Budget Discontinuities

We present a framework to identify the willingness to pay (WTP) for workplace amenities that builds on the influential work on budget discontinuities. Responses around "kinks" and "notches" can be used to estimate labor supply elasticities – aka preferences over leisure and earnings (Card et al., 2015a,b, 2017; Kleven, 2016a). We extend this canonical two-good framework to a three-good framework with leisure, earnings, and amenities. We show that variation in workplace amenities affects the amount of excess/missing mass. How much the excess mass *varies* depends on the value of the change in amenities, so that the excess mass response can be used to identify the willingness to pay for amenities.

Consider the standard notch case with an individual who obtains utility from after-tax earnings (or consumption) and pre-tax earnings (cost of effort). We augment this framework with a third good, so that the utility function becomes $U\left(T(m), \frac{m}{z}, a\right)$, with m pre-tax earnings, $T(m)$ after-tax earnings, z worker ability, and a workplace amenities. While the framework applies to a wide range of possible workplace amenities, we use the example of worker health to illustrate the approach. A worker is either healthy (a_0) or sick (a_1).[8] Heterogeneity in ability is captured by a distribution function $f(z)$. Assume that this ability distribution, the tax system and preferences are smooth so that the resulting earnings distribution is also smooth. Denote after-tax earnings by $T(m)$, the tax rate by t, benefits by $\mathcal{B}$ and the earnings eligibility threshold for accessing benefits by m^*. Individuals become ineligible for $\mathcal{B}$ when their pre-tax earnings exceeds m^*. The worker's budget constraint is therefore:

[8]The framework applies to all cases where the total (dis)amenity consumed grows with hours worked, and cases where amenities are a fixed part of work but the probability of experiencing/using the (dis)amenity depends on hours worked. The framework thus accommodates most amenities studied in the literature.

$$T(m) = \begin{cases} (1-t) * (m + \mathcal{B}) & m \leq m^* \\ (1-t) * m & m > m^* \end{cases} \tag{1}$$

This is the canonical case of a budget notch, where m^* is the notch point. The budget constraint "jumps" at m^*, as shown in Panel A of Figure 1. Panel B shows the resulting excess/missing mass in the earnings distribution. The notch creates an incentive to reduce earnings below m^* and generates excess mass (missing mass) in the earnings distribution below (above) m^*. We will focus on such budget notches in what follows, but the approach generalizes to budget kinks.

To add amenities to this framework, first consider cases where the probability of having a positive/negative experience at work increases the more time people spend at work (e.g. injuries, harassment, positive interactions, sense of achievement, etc.). We denote the probability that a relevant event occurs by $r(m)$. In the context of workplace safety, $r(m)$ is the risk of an illness or injury. This risk is the product of the per-period risk of an injury (θ) and the time at work. We let the risk increase with m instead of hours to simplify notation and link our theory closely to the canonical bunching literature: $r(m) \equiv m\theta$.[9] While the example is for an amenity that happens stochastically, like injuries, the same model applies to a broad range of amenities that become more valuable or more frequent the more time is spent at work.[10]

We can write the expected utility of a worker as:

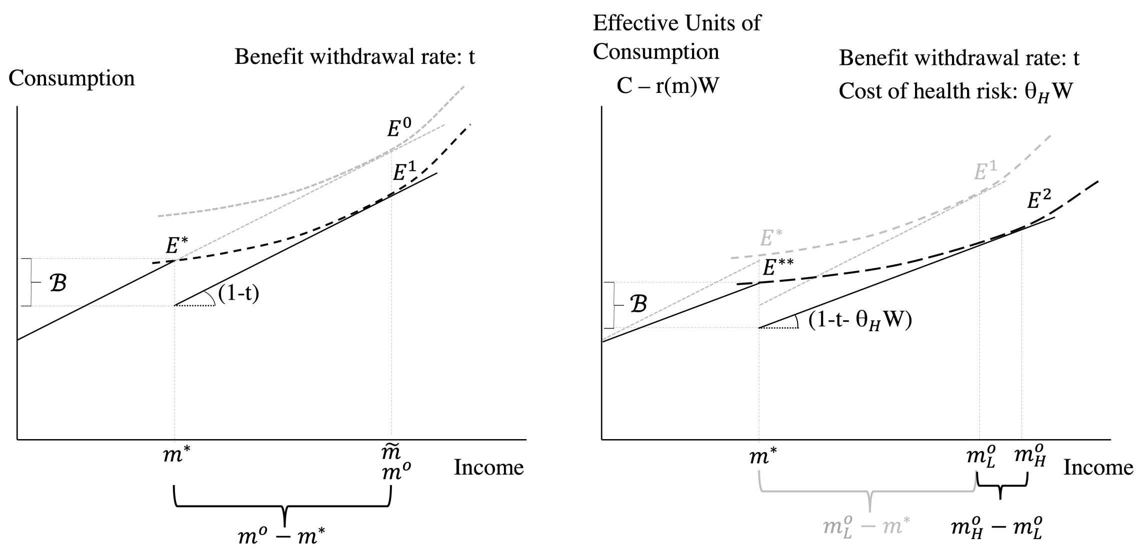
$$E\left(U(T(m), \frac{m}{z}, a)\right) = [1 - r(m)]\, U\left(T(m), \frac{m}{z}, a_0\right) + r(m) U\left(T(m), \frac{m}{z}, a_1\right) \tag{2}$$

Analogous to the iso-elastic quasi-linear assumption of the two-good bunching litera-

[9]Recall that hours and earnings are closely linked in this model with m equal to hours multiplied by productivity z.

[10]The model covers amenities that are a fixed part of work but are only valuable occasionally (e.g., sick leave, work time flexibility, etc.) or amenities that are a by-product of work (e.g., sense of making a difference, enjoyment of work or colleagues, etc.). The framework does not apply when work time has no effect on the value or frequency of the amenity (e.g., health care coverage).

Panel A: Labor Supply with Budget Notch



Panel B: Excess and Missing Mass in the Earnings Distribution

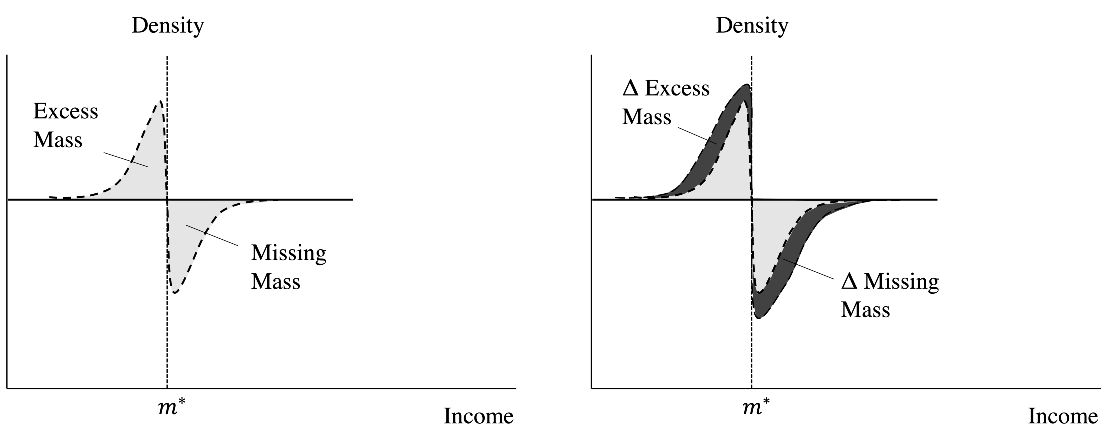


Figure 1: Worker Response to Budget Notch

Note: The left side of the figure shows the budget constraint from equation (1) and the indifference curve from equation (3) in Panel A and the resulting excess mass in Panel B for $\theta = \theta_L$. The right side shows the same for θ_L and θ_H. Panel A has total expected post-tax earnings (wage plus amenity) on the y-axis and labor supply (pre-tax earnings) on the x-axis. Panel B shows excess mass relative to the notch point m^*

ture, we assume that utility is separable and quasi-linear in earnings. This utility takes the form:

$$U\left(T(m), \frac{m}{z}, a\right) = T(m) - \frac{z}{1+1/e}\left(\frac{m}{z}\right)^{(1+1/e)} + a$$

where e is the labor supply elasticity.[11]

Next, we define the compensating variation W associated with an amenity. A worker is willing to pay W for the amenity: $U(T(m), \frac{m}{z}, a_1) = U(T(m) - W, \frac{m}{z}, a_0)$. In the context of health, a worker is willing to pay W to stay healthy (a_0). Using this definition of W in equation 2, expected utility becomes: $E\left(U\left(T(m), \frac{m}{z}, a\right)\right) = U\left(T(m), \frac{m}{z}, a_0\right) - r(m)W$ and normalizing $a_0 = 0$ we can express expected utility as:

$$E\left(U\left(T(m), \frac{m}{z}, a\right)\right) = T(m) - m\theta W - \frac{z}{1+1/e}\left(\frac{m}{z}\right)^{(1+1/e)}. \tag{3}$$

This expected utility nests the canonical bunching case when $W = 0$. In the general case with $W \neq 0$, the risk to health reduces the return to work. The health risk operates like an additional tax on working, with tax rate θW.[12] We can illustrate the impact of W in the standard leisure and consumption diagram (Figure 1). The "health tax" pivots the budget constraint downward like a standard fiscal tax.

We can identify the value of W by linking the expected utility above to the excess mass observed around the budget notch. The canonical bunching approach uses this idea to identify the parameter e. We will require an additional moment to identify W. To start, consider the canonical approach, which points out that the last person to bunch, the "marginal buncher," is indifferent between choosing the notch point m^* and an interior point $\tilde{m}$. The marginal buncher therefore has: $EU^* = E\tilde{U}$. The

[11] The linearity of utility in a is without loss of generality since a has no units and we can redefine any $f(\tilde{a}) \equiv a$. The assumption of an additive value of amenities is common in the literature (e.g., Morchio and Moser (2019)). For a more general utility function, see Appendix D.3.

[12] The implicit tax imposed by health costs is linear in this case, but the framework holds more broadly. The linear tax is an artifact of the functional form assumption on the utility, in the more general case WTP may vary with earnings and thus lead to a non-linear cost. As long as the optimization problem remains quasi-concave, the above linear framework still works and provides a local approximation that captures the WTP at earnings level m^*.

indifference curve for this worker is shown in Figure 1. At the interior point $\tilde{m}$ the first order condition from maximising (3) implies:

$$\tilde{m} = z(1 - t - \theta W)^e \tag{4}$$

and hence the indirect utility $E\tilde{U}$ is:

$$E\tilde{U} = \frac{z}{1+e}(1 - t - \theta W)^{(1+e)}. \tag{5}$$

At the notch point m^*, utility EU^* is:

$$EU^* = (1 - t - \theta W)m^* + (1 - t)\mathcal{B} - \frac{z}{1 + 1/e}\left[\frac{m^*}{z}\right]^{(1+1/e)} \tag{6}$$

Using $EU^* = E\tilde{U}$ together with equations (5), (6) and the fact that $z = m^o/(1 - t - \theta W)^e$ yields:

$$\frac{(1-t)\mathcal{B}}{(1 - t - \theta W)m^*} = \frac{m^o}{m^*}\gamma - 1 \tag{7}$$

where $\gamma = \frac{1}{1+e} + \left(\frac{m^o}{m^*}\right)^{-\frac{1+e}{e}} \frac{e}{1+e}$. When $W = 0$, this is the canonical bunching approach and identifies e using equation (7).

When $W \neq 0$, we can identify W by observing excess mass at the notch in a high and low-risk setting (θ_H, θ_L, respectively). We illustrate the impact of such changes on excess mass in the right panels of figure 1.[13] Excess mass increases when workplace risks worsen. Solving 7 for the parameters W and e follows the standard bunching approach and requires data on observed policy parameters $t, \mathcal{B}, \theta$, m^*; and on $\frac{m^o}{m^*}$. The canonical bunching literature (without frictions) identifies $\frac{m^o}{m^*}$ from the amount of excess mass at the budget discontinuity, denoted by η. The link between η and $(m^o - m^*)$

[13]Note that the ICs in the right picture of Panel A are not parallel upward shifts because E^1 and E^2 represent ICs of two marginal bunchers who are different individuals. Also note that the illustration in the standard notch figure is feasible because the worker utility is separable in health and consumption.

is: $\eta = \int_{m^*}^{m^o} d_0 = (m^o - m^*)d_0$, where d_0 is the baseline earnings distribution.[14] We generalise our model below to a case with frictions in hour adjustment.

To build intuition for the impact of W on excess mass, note that the left side of equation (7) is akin to the replacement rate. The numerator $[(1-t)\mathcal{B}]$ is the net-of-tax payoff from not working and the denominator $[(1-t-\theta W)m^*]$ the payoff from working net of both taxes and health risk. The right-hand side of equation (7) is a function of $\frac{m^o}{m^*}$ and parameters only and thus captures the amount of excess mass at the threshold. If work involves greater health risks ($\theta \uparrow$) the LHS increases and the excess mass at the notch must go up to increase the RHS. Health risk thus increases excess mass at the notch. This effect captures the intuition that working is less attractive at higher health risks and more workers are therefore willing to bunch when risks are higher.

To measure the value of amenities, we define a willingness to pay for amenities (denoted by WTP) as a percent of after-tax earnings: $WTP(r) \equiv \frac{r(m)W}{m^*(1-t)}$. Note that $WTP(r)$ is different from W in two ways. $WTP(r)$ is the cost of an increase in sickness risk by r and is expressed as a share of earnings, while W is the compensating variation for falling sick ($r = 1$) and expressed as an absolute dollar amount. $WTP(r)$ is thus independent of units and has a value between 0 and 1. We solve for $WTP(r)$ by evaluating (7) in θ_L and θ_H risk states and taking the ratio of the two. We use L and H subscripts to denote low and high risk states. Normalising $\theta_L = 0$ and re-arranging

[14] The last equality assumes d_0 is constant and simplifies the expression. The same approach, however, also works for cases with more flexible functions of d_0.

the ratio yields:

$$WTP(r) = 1 - \frac{\frac{m_L^o}{m^*}\gamma_L - 1}{\frac{m_H^o}{m^*}\gamma_H - 1} \tag{8a}$$

$$\simeq 1 - \frac{\frac{m_L^o}{m^*} - 1}{\frac{m_H^o}{m^*} - 1} \tag{8b}$$

$$= \frac{m_H^0 - m_L^0}{m_H^0 - m^*}. \tag{8c}$$

Equation (8a) shows that the WTP can be expressed in terms of the labor supply response to the notch in high and low-risk settings $(m_H^0/m^*\gamma_H, m_L^0/m^*\gamma_L)$.[15]

This expression simplifies in the case of regression kink designs (when $\gamma_L = \gamma_H = 1$) and similarly for notches when labor supply elasticity e is small (implying $\gamma_L, \gamma_H \to 1$). Most empirical estimates find small values of e, making small e a particularly relevant approximation. The simplified expression for WTP in equation (8c) is independent of e and is a simple ratio with the response in the high-risk state H $(m_H^0 - m^*)$ as the denominator and the additional response when risk increases from θ_L to θ_H $(m_H^0 - m_L^0)$ as the numerator. Simply put, we compare the magnitude of excess mass when workplace risks are high and low. If the excess mass is the same in both cases $(m_H^0 = m_L^0)$, then $WTP(r) = 0$. On the contrary, a large $WTP(r)$ implies that the excess mass increases sharply with risk $(m_H^0 > m_L^0)$.

Using the approximation in (8c) instead of the structural equation in (8a) has several advantages. First, it yields a simple expression that depends only on two behavioral responses which can be transparently estimated using familiar quasi-experimental tools. Second, the expression is independent of the value of e and accommodates a wider range of functional form assumptions and/or adjustment frictions. In addition, using the approximation comes at a relatively small cost under a wide range of plausible parameter values. The approximation always provides a lower bound, and the bound

[15] When implementing this approach empirically, one also has to account for potential frictions in work-hour choices. We will address this issue in the online Appendix section D.2.

remains close to the truth for a wide range of plausible parameter values (see Appendix D.1).

The WTP approach shares several of the advantages of canonical budget discontinuity designs. The behavioral responses can be estimated non-parametrically using transparent quasi-experimental tools. The theoretical framework translates these estimates into structural parameters that hold validity beyond the specific estimation context, enabling the study of policy counterfactuals.

The above framework presents a frictionless labor market and the bunching literature has shown that traditional bunching methods are biased by optimisation frictions (adjustment costs, inability to chose hours, etc.). As a solution, the literature shows that the model can be extended and accomodate frictions explicitly (Kleven and Waseem, 2013). We show that such extensions are less important in our case since the WTP calculation is mostly unafffected by the presence of optimisation frictions. In fact, when a fixed share of workers faces optimisation frictions, the impact of frictions cancels out exactly (see Appendix D.2) and does not affect the WTP at all. This sharp contrast with the canonical bunching literature arises because the canonical literature interprets the *level* of bunching, while the WTP approach studies the *percent change* in bunching (aka a ratio) and frictions that affect the denominator and numerator proportionally cancel out. The frictions discussed in Kleven (2016b) are one example of such frictions and therefore do not affect the WTP estimates. Our approach of course, shares the drawback of the bunching literature that more complex types of friction could lead to biased estimates.

In the online Appendix, we extend this framework in several dimensions: we consider different functional form assumptions (D.3), and the role of income effects (D.4).

3 Data and Sample

Our main application studies WTP for workplace safety from Covid health risks and leverages a unique dataset on workers labor supply during the Covid pandemic. The data is provided by Homebase, a private company used by small businesses to track the hours and earnings of their workers.[16] The data mainly covers sectors with hourly and frontline workers (such as those in the restaurant, food and beverage, retail, health and beauty, and healthcare industries), the type of worker who faced the decision whether to reduce their work hours to diminish the risk of contracting Covid-19.[17]

Studying partial UI programs is challenging due to eligibility criteria tied to weekly earnings, whereas most datasets report earnings and hours on a monthly, quarterly, or annual basis. The dataset of this study reports daily earnings and hours, which we use to compute weekly records.

A second important advantage of the data is the third party reporting of hours through an app. This reduces the well-known issue of noise in self-reported work hours data (c.f., classic work by Bollinger, 1998; Bound and Krueger, 1991). The accuracy of hours data is the core product feature of Homebase: workers use a mobile phone app to clock in and out of work, and the phone's geo-location tracking ensures accurate clocking.[18] A third advantage of the data is its coverage. Typical administrative UI records cover only a single state and/or are available with a substantial time lag. Our sample includes data from 21 states and is available in near-real time. The study of multiple states simultaneously offers additional sources of institutional variation. In our application, each state has its own eligibility threshold for partial UI, making for a stronger identification strategy. Furthermore, it allows us to use border designs

[16]The data is provided and licensed by Homebase (joinhomebase.com).

[17]In Appendix B, we compare the Homebase data with nationally representative data. Our sample's weekly earnings, hourly wages and hours worked are similar to the average hourly worker in small firms in the 21 states under analysis.

[18]When the app recognizes that workers get to or leave the workplace, it sends a check-in/out notification as shown from the app screenshot in Appendix figure A1.

and compare neighboring counties with similar characteristics but different eligibility thresholds for UI.[19]

A drawback of this type of private-sector data is that it lacks information on individuals who exit the sample. When individuals are not observed, they could have either changed employers or left the labor force entirely. This is a lesser concern in our setting since the theoretical framework focuses on intensive margin changes and we exclude weeks with zero earnings from the main analysis. In Appendix E.6, we show that the results are robust to including workers that leave the homebase sample.

We impose four restrictions on our sample along the following dimensions: time period, eligibility for partial UI benefits, geography and work-spell length.

First, we restrict data to the time window between October 1, 2019, and July 31, 2020 (the end of the FPUC program) – five months before and five months after the onset of Covid-19 pandemic in the U.S. in March 2020.

Second, we restrict the sample to individuals eligible for the partial UI policy we study.[20] The eligibility criteria are similar to regular UI payments. Instead of requiring that people are out of work, partial UI requires that earnings fall below a threshold level. In principle, the relevant threshold varies across workers based on their past earnings. However, for workers that qualify for *maximum* weekly benefits (MWB) the same threshold applies within each U.S. state. We focus on these workers eligible for MWB and thus study one eligibility threshold per state.[21] Homebase does not directly report whether workers are eligible, but we can infer eligibility based on retrospective work histories and state-specific eligibility rules.[22] While for most workers we observe

[19]Strategic misreporting of work hours is also less of a concern with Homebase data since these records are not used to administer UI benefits.

[20]Partial Unemployment Insurance schemes are not specific to the Coronavirus Aid, Relief, and Economic Security (CARES) Act and were already in place before the start of the Pandemic. Several works have studied these schemes even before the onset of the Pandemic. Notable examples are Lee et al. (2021); Boeri and Cahuc (2023); Le Barbanchon (2020).

[21]We would only be able to approximate the partial UI threshold for workers eligible for benefits below the MWB level.

[22]We rely on information collected by the Department of Labor to reconstruct eligibility rules. In Appendix C.2 we report all sources and details for our calculations.

the full earnings history required to determine benefit eligibility, for workers who have only a partial history, or might have a second job that is not in the Homebase system, we calculate theoretical quarterly earnings based on their hourly wage multiplied by 40 hours and 13 weeks. We then estimate whether they would be eligible for maximum weekly benefits based on these theoretical quarterly earnings. Given the greater uncertainty implied by this estimate, we down-weight these observations based on the ratio of observed quarterly earnings over theoretical quarterly earnings, so that workers with shorter earnings histories have a smaller weight in the analysis.

Third, data availability limits the analysis to a subsample of U.S. states. We exclude states where Homebase is not active and states where only few workers in our sample earn enough to meet the high earnings requirements to qualify for MWB.[23] The resulting sample covers 21 U.S. states.

Fourth, the baseline analysis also excludes the least attached workers who only work in the period before the onset of the pandemic or only after it. Relaxing this restriction has minimal impact on the result (see E.6). Finally, the baseline sample is "balanced" in the sense that each worker is in the sample for the same number of weeks before and after the onset of the pandemic, so that each worker contributes equally to the pre- and post-Covid-19 earnings distributions. While this is not strictly necessary, it alleviates concerns about selection effects and makes it easier to interpret excess and missing mass as changes among the same pool of workers.[24] We again show that relaxing this restriction has minimal effects on the estimates (see Appendix E.6).

Summary statistics for the sample are reported in Table 1. Panel A reports worker information. The sample includes 9,063 workers and 169,450 worker-week observations. On average, they work 36 hours per week and earn \$660. The median hourly wage is \$16 and does not vary much (the 25^{th} percentile is \$14, and the 75^{th} is \$20). Panel B

[23]For these states, we would therefore have insufficient workers to study behavior at the threshold.

[24]Take the specific example of a worker that worked continuously but had a two-week temporary absence (e.g., sickness or holidays) before the start of Covid-19 pandemic. We include all the active weeks before Covid-19 and trim the last two active weeks in the post-Covid-19 period for this worker to maintain a balanced number of work-week observations before and after Covid-19.

of Table 1 shows summary statistics for the 3,500 small businesses in our sample. On average, they have 1.1 branches and 13.26 employees, of which 97% are hourly waged workers in the median firm. 32% of all firms operate in the Food and Drink sector, with Retail, Health Care, and Professional Services being the next most represented sectors in the data.

Table 1: Descriptive Statistics

	Mean	S.D.	p50	p25	p75
Panel A: Workers					
Weekly earnings	660.04	345.06	617.63	449.59	813.02
Weekly hours	36.49	12.78	38.58	29.17	44.47
Hourly wage	18.35	8.15	16.00	14.00	20.00
Number of weeks in data per worker	27.33	10.07	30.00	18.00	36.00
Worker-week observations	169,450				
Number of workers	9,063				
Panel B: Firms					
Size	13.26	20.05	8.03	4.25	15.52
Share of salaried workers	0.10	0.17	0.03	0.01	0.13
Number of Branches	1.14	0.64	1.00	1.00	1.00
Food and Drink	0.32	0.47	0.00	0.00	1.00
Health Care	0.18	0.39	0.00	0.00	0.00
Professional Services	0.04	0.21	0.00	0.00	0.00
Retail	0.03	0.18	0.00	0.00	0.00
Number of firms	3,500				

Note: Homebase data between October 1, 2019, and July 31, 2020. Sample of hourly workers with sufficient past earnings to qualify for MWB payments in their home state.

4 Willingness To Pay for Workplace Safety

4.1 Design: Response to partial UI Budget Notches

The main implementation of our WTP methodology first estimates the magnitude of the excess mass at a budget notch threshold and then examines how this excess mass varies with workplace risk.

The research design is based on a budget notch created by the Federal Pandemic Unemployment Compensation (FPUC program)[25], which introduced a lump-sum $600 expansion of UI weekly benefits and was approved as part of the Coronavirus Aid, Relief, and Economic Security (CARES) Act, enacted on March 27, 2020, and ended on July 31, 2020.[26] Importantly, workers can qualify for FPUC while working, as long as they are eligible for partial unemployment insurance, which requires their earnings to be below a threshold level (the "earnings test").[27] Above the threshold, workers become ineligible for FPUC. This creates incentives to reduce labor supply and decrease earnings below the threshold, potentially resulting in excess and missing mass in the earnings distribution around the threshold (for a theoretical illustration, see Fig 1).[28] Consistent with this, we do see workers excess mass in the earnings distribution at the state-specific eligibility thresholds.

While FPUC was introduced uniformly in all US states, the administration of the benefit was left to the states and therefore depended on pre-Covid state-specific eligibility thresholds. We have calculated these thresholds for our sample of workers eligible for maximum UI benefits using information collected by the Department of

[25]The FPUC benefit played a central role in U.S. mitigation policies during Covid-19. Numerous papers have measured the take-up of this program and studied its impact on labor supply (some notable examples are Marinescu, Skandalis, and Zhao, 2021; Gallant et al., 2020; Holzer, Hubbard, and Strain, 2024; Ganong, Noel, and Vavra, 2020; Ganong et al., 2022; Forsythe, 2023; Forsythe et al., 2020; Gupta et al., 2023).

[26]No FPUC benefits were payable between July 31, 2020, and December 26, 2020. The FPUC was re-established by the Continued Assistance Act as a $300 per-week supplement to unemployment benefits from December 26, 2020, to March 14, 2021. Please consult Online Appendix C.1 for more details on FPUC and subsequent programs.

[27]Formally, FPUC is paid to all individuals on UI and on partial UI benefits. The qualifying criteria for these benefits vary by state and for our sample states these criteria always include an earnings test.

[28]Standard additional rules aimed at mitigating moral hazard are also in place, e.g. UI recipients are not allowed to refuse job offers, and job loss or hours reduction should not in principle be due to the fault of the worker. These rules are notoriously difficult to enforce and a large literature on UI benefits studies the moral hazard problems that may prevail despite these rules. Monitoring was especially weak during the first weeks of the pandemic when unemployment offices prioritized processing the major inflow of claims and executing a variety of new programs. Authorities also had an incentive to allow people to stay home to reduce the spread of infections.

Labor. Details of the calculations are reported in Appendix C.2. Figure 2a shows the variation of the Partial UI eligibility threshold across states. A worker earning \$500 a week would be eligible for benefits in California and Pennsylvania, but not in Arizona or Florida.

To identify the baseline labor supply response to the budget notch (the first step of our implementation), we stack these different thresholds to combine 21 difference-in-differences (DiD) analyses across the sample states. Each DiD compares workers in a window below and above the state-specific threshold before and after the onset of the Covid-19 pandemic to estimate the excess mass below the threshold.

Before comparing these DiD estimates in low- vs. high-risk settings to capture how the excess mass changes with varying risk and back-out the WTP for workplace safety, we discuss the identification details for the baseline labor supply response.

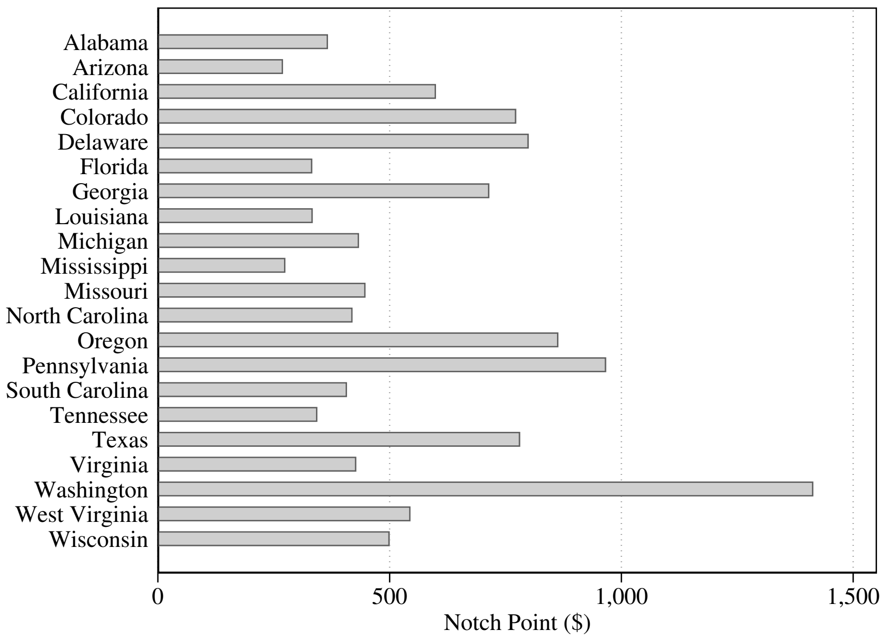
We estimate excess and missing mass with the following DiD specification:

$$E_{wmtk} = \pi^{mt} + \sum_{k=-650}^{1300} \beta^k \cdot I_k + \sum_{k=-650}^{1300} \eta^k \cdot I_k \cdot C_t + \varepsilon_{wmtk} \tag{9}$$

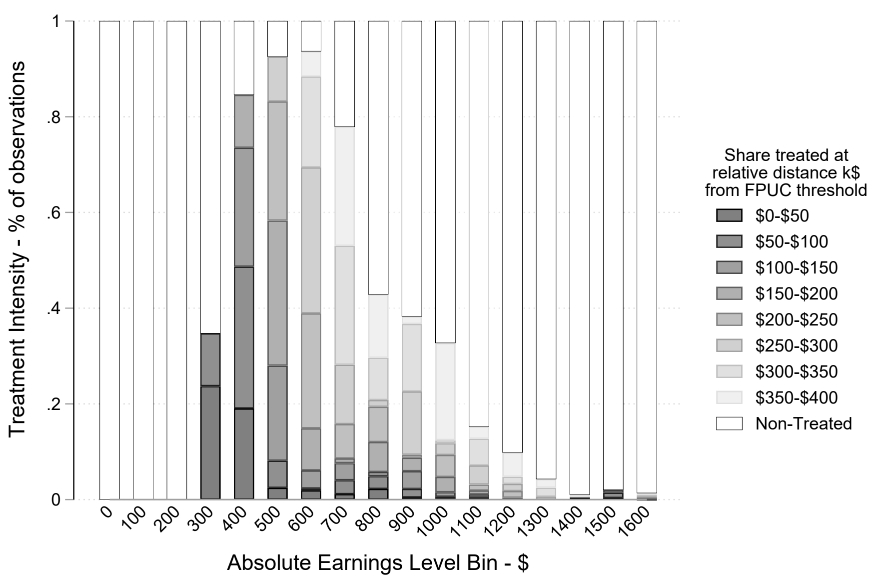
where E_{wmtk} is a dummy with value 1 if a workers' (w) earnings are in range m, in week t, \$$k$ away from the UI eligibility threshold, and C_t is an indicator with value 1 after the onset of Covid-19 pandemic. The coefficients are η^k, β^k, π^{mt}. β^k captures the excess or missing mass around the eligibility threshold *before* Covid, and η^k captures the change in the mass *after* the onset of the pandemic and is the main parameter of interest. π^{mt} are fixed effects that capture changes in the aggregate earnings distribution and vary by \$100 bins of earnings (m) and before/after the onset of the Covid-19 pandemic (t). The remaining identifying variation comes from whether earnings fall to the left or the right of the local FPUC eligibility threshold. Instead of a single eligibility indicator, we use finer dummies that capture the distance to the eligibility threshold (k). Theory would predict that responses are starkest close to the eligibility threshold and weaker further away from the threshold. I_k is an indicator that takes value 1 if a workers'

Figure 2: Identifying Variation

(a) Notch Points by State



(b) Treatment Intensity by Earnings Level



Note: Panel A shows maximum allowable earnings while receiving FPUC payments for maximum weekly benefit (MWB) recipients across US states. Panel B shows the variation in the treatment exposure for individuals with the same earnings. The x-axis shows absolute earning levels in $100 bins, the level of our $\pi_m t$ fixed effects. Within each bin, the share of workers in the $400 treatment window above the threshold and therefore subject to the financial incentive (treated) is shown in gray colors and the share of those out of the treatment window (non-treated) is shown in white. The treatment is disaggregated by treatment intensity, with workers closer to the relevant state-specific FPUC eligibility earnings threshold in darker gray, as they are exposed to a stronger incentive.

earnings are in a \$50-wide bin $\$k$ away from the UI eligibility threshold. Given that π^{mt} controls for before/after Covid-19 changes in the earnings distribution, η_k captures differences in the behavior of individuals with identical earnings, say \$300, but who fall on different sides of the eligibility thresholds. Figure 2b provides graphical intuition for the variation leveraged in the analysis: it plots the share of treated workers in each \$100 bin of earnings and dis-aggregates "treatment" into intensity bins I_{wtk}, that capture the distance (\$50, \$100 … or \$400) from the state-specific earnings threshold. The figure illustrates that there is much variation in treatment status for individuals with identical absolute earnings, which allows for an identification strategy that compares the behavior of individuals with identical earnings facing different financial incentives.

There are at least two empirical challenges with canonical bunching estimates. A first challenge is that workers typically have constraints on hour choices (i.e. adjustment frictions) that reduce their ability to bunch around budget notches (Kleven and Waseem (2013)). Section 2 considered the case where a fraction of workers are unable to chose their workhours. Such constraints have important effects on the traditional bunching estimates, however, the impact of such frictions cancels out in the WTP estimator developed in this paper. Our approach is thus unaffected by the presence of such frictions. Aside from this widely studied binary distinction of whether workers can or cannot adjust their hours, workers may face more complicated restrictions on their hours choices. For instance, workers may choose from a set of indivisible shifts or full workdays rather than adjusting their hours incrementally, leading to earnings changes that occur in discrete steps rather. Such a constraint prevents workers from precisely targeting an earnings threshold and the response to a threshold then is less concentrated in a single point of the earnings distribution but rather more spread out over a broader earnings ranges. A spread-out response is difficult to identify in traditional cross-sectional bunching designs but our setting that leverages panel data allows us to make progress. Comparing changes in the pre-period to the post FPUC period allows us to identify "anomalies" in the earnings distribution over broader ranges. We allow for a bunching interval rather than a single bunching point and thus accommodate the possibility of indivisible shifts in our empirical set-up.

A second identification challenge—emphasized by Blomquist et al. 2021—is the credible estimation of the counterfactual earnings distribution. In the canonical cross-sectional bunching framework, identification hinges on extrapolating from regions of the earnings distribution that are assumed to be unaffected by the notch, a strategy that rests on strong and contentious assumptions. This concern is specific to cross-sectional designs and is less relevant to our setting, which leverages a Difference-in-Differences-style identification strategy. The Difference-in-Differences design identifies the effect of the notch through *changes* in the earnings distribution over time, eliminating the need to impose functional form assumptions on the untreated portion of the distribution for counterfactual inference.

The main identifying assumption of our approach is that the pre-notch earnings distribution provides a valid counterfactual for the post-notch distribution. In a setting with a single notch, this implies that spurious shocks have smooth effects through the threshold. Two features of our setting allow us to further relax and probe this assumption: First, the notch does not affect some groups of workers (creating a placebo test[29] and second, the notches happens at different income levels across states. We use the differences across states to implement a border design and flexibly control for potential changes in the aggregate distribution. The baseline specification 9 allows for flexible changes in every $100 earning bin, captured by $\pi_{m,t}$.[30]

Finally, we examine the raw data before proceeding to the regression analysis. Appendix Figure E.1 presents the unadjusted earnings distribution, allowing us to assess excess and missing mass without the influence of control variables. The figure provides visual evidence that excess mass emerges around the threshold following the

[29]We implement a placebo test with workers who are eligible for FPUC at different thresholds than those on which we focus. See the next section for further details on our placebo sample.

[30]Further checks additionally allow for heterogeneity in income changes among specific demographic sub-groups. The identifying variation in these specifications comes from the variation in the threshold level of the notch across states.

introduction of FPUC, consistent with a behavioral response to the policy. In the pre-FPUC period there is no such excess mass, which serves as a placebo check and shores up the credibility of the identification strategy.[31]

4.2 Results: Response to partial UI Budget Notches

We now turn to the empirical results, leveraging the Difference-in-Differences (DiD) specification in equation 9 to estimate behavioral responses to partial UI budget notches. This specification exploits the fact that the eligibility threshold varies across states, leading us to expect excess mass at different points in the earnings distribution in Alabama vs. Georgia, etc. A key advantage over simply plotting the raw data is that this approach accounts for potentially large, time-varying shifts in the earnings distribution due to the macroeconomic volatility induced by the COVID-19 pandemic. Specifically, we include π^{mt} fixed effects, which flexibly absorb shocks specific to each $100 earnings bin, ensuring that observed bunching patterns are not driven by broader shifts in the labor market. These fixed effects control for abrupt changes in employment patterns, sectoral demand shifts, and other effects that could otherwise confound our estimates of policy-induced responses.

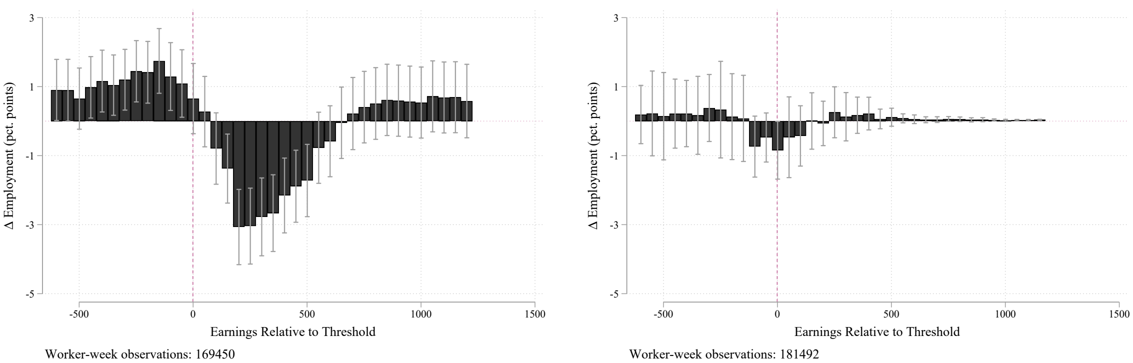
Figure 3 presents the η_k coefficients, which measure excess mass in $50 bins around state-specific eligibility thresholds after the onset of the pandemic. Earnings bins are defined relative to the applicable threshold in each state, with the threshold bin normalized to zero. Positive values indicate earnings above the state-specific threshold, while negative values represent earnings below it. Panel A shows that the introduction of FPUC created strong incentives for workers to adjust earnings downward to remain below the eligibility threshold. Our analysis reveals a substantial reduction in the frequency of earnings just above the FPUC threshold, accompanied by a corresponding

[31] There is no bunching at the earnings threshold without FPUC. Note that without FPUC there is still withdrawal of UI payments below the threshold and since payments run out at the threshold, such withdrawals stop at the threshold. This could in theory create a kink at the threshold, even in the absence of FPUC, however it does not generate bunching behavior.

increase in excess mass below it. The share of workers in bins above the threshold declines by 3 percentage points at the bin with the largest drop (i.e., “threshold+$250”), representing a 33% decrease in frequency relative to a baseline frequency of 9% in that bin.

Figure 3: Excess and Missing Mass around the Partial UI Notch

(a) FPUC Eligible Workers **(b)** Placebo Sample: Ineligible Workers



Note: The figure shows η_k coefficients from equation (9). Standard errors are clustered at the state, earnings bin, and week level, and 95 percent confidence intervals are reported. The sample in panel (a) is hourly workers with sufficient past earnings to qualify for MWB payments in their home state (169,450 worker-week observations). The baseline mass in the most affected bin is around 9%. The sample in panel (b) is hourly workers “ineligible for MWB” as defined in the text (181,492 worker-week observations). These workers do not face an eligibility threshold at 0. Source: Homebase.

The pattern around the threshold largely aligns with standard labor supply intuition: the frequency of earnings observations drops most sharply just above the threshold, reflecting that workers near the eligibility cutoff have the strongest incentive to adjust their earnings. In contrast, workers who would need to reduce earnings substantially to qualify for benefits are less likely to respond, leading to a smaller shift in frequency further up the earnings distribution.

As expected, this shift in frequency is accompanied by excess mass below the threshold, but notably, the excess mass is not concentrated at a single point exactly at the threshold. Instead, it spreads out over a broader earnings range, with the peak appearing slightly further below rather than precisely at the threshold. At first glance, this dispersion might seem surprising, as standard bunching models would predict a sharper concentration of observations just below the threshold. However, this broader spread can be largely attributed to adjustment frictions, particularly for workers with indivisible shifts that prevent precise hour reductions. When workers must reduce earnings by dropping entire shifts (e.g., giving up an 8-hour work shift), their responses will be more dispersed rather than clustering tightly around the threshold.[32]

We formally show that such scheduling restrictions lead to a weaker bunching pattern by splitting the sample into industries with more or less flexible scheduling policies. Data on scheduling flexibility comes from the American Time Use Survey (ATUS) Leave Module questions on workers' ability to choose their shifts' start and end times. The data confirms that the non-sharp nature of the excess mass is partly due to the effective ability of workers to adjust hours. Excess mass is sharper and much larger for industries that have more flexible schedule policies allowing for a more granular targeting of worked hours (Figure 4). Additional potential factors that may explain any remaining non-sharp excess mass in industries with high hour flexibility are measurement error in how we calculate the partial UI eligibility earnings threshold.[33]

Next, we conduct a placebo test to further examine potential spurious changes in the earnings distribution around eligibility thresholds. The placebo sample consists of workers who, based on their prior earnings records, are ineligible for FPUC at the thresholds analyzed in our study. This group includes workers with insufficient past earnings history to qualify for Unemployment Benefits, as well as workers eligible for less-than-maximum benefits whose relevant thresholds fall below the Maximum Benefit threshold in their state – the threshold on which our analysis focuses. These workers share the same labor market shocks, but they do not have incentives to respond to

[32]An additional pattern of interest is the small, though statistically insignificant, excess mass beginning around $600 above the threshold. One possible explanation is within-firm labor market adjustments: to accommodate coworkers reducing hours and dropping shifts, some workers who are far enough above the threshold—and thus not directly affected by the UI policy—may need to increase their hours, leading to a small upward shift in the frequency of earnings observations in that range.

[33]When determining the partial UI threshold, differences in observed earnings and UI relevant earnings arise in some jurisdictions from allowances for families, special circumstances, or multiple jobholders.

Figure 4: Excess and Missing Mass by Hour Flexibility

(a) Least Flexible Quartile **(b)** Most Flexible Quartile

Note: The figure estimates excess mass patterns in industries where workers have more or less flexibility in choosing work hours. Information on flexibility comes from the 2017-2018 ATUS data. We calculate the average ability to frequently adjust work hours at 2-digit NAICS industry-level. Panel A shows the bottom 25% and Panel B the top 25% of the distribution of schedule flexibility. The sample covers hourly workers with sufficient past earnings to qualify for MWB payments in their home state. Source: Homebase.

the benefit eligibility thresholds we study. By comparing this placebo group to the main treatment sample, we assess whether observed excess mass patterns are driven by genuine behavioral responses to FPUC or by broader, unrelated shifts in the earnings distribution. The results are shown in Figure 3 (b), which plots the behavioral response around the eligibility threshold for ineligible workers. The effects are insignificant and of small magnitude, which confirms that there are no spurious shocks. Indeed, this test rules out many alternative explanations for the observed bunching patterns – the pattern exists only for workers who face the eligibility threshold.

4.3 Results: Response of Excess Mass to Workplace Risk

The second step of the WTP approach estimates how excess mass responds to variation in workplace risk. To achieve this, we implement a "shift-share"-like design that captures the intuition that COVID-19 risk increases over time with local outbreaks (the "shift") and does so proportionally more for more vulnerable industries (with the industry-specific risk weights representing the "shares" in this analogy).

We measure industry vulnerability based on the nature of job-specific tasks, specifically the extent to which tasks require workers to engage in interpersonal interactions. Hairdressers, for example, experienced larger shocks to workplace risk than workers with less interpersonal contact, such as landscape gardeners. The industry risk scores are computed by combining pre-COVID-19 information on tasks' risks developed by Basso et al. (2021)[34] with American Community Survey data on the distribution of occupations and tasks across 3-digit industries. Our industry vulnerability index is therefore the product of each task's riskiness and the frequency of the task in the industry.[35] The advantage of this risk measure is that it is based exclusively on pre-pandemic data and rules out a reverse causality issue where the (lack of) fear of COVID-19 drives local infection rates.

We then interact the time-invariant, pre-determined industry risk score with time-varying data on local outbreaks. Denote the time-invariant risk score for industry i by P_i and denote the local fatality rate in the c' neighbor counties of c by $R_{tc'}$.[36] The product of these two components yields a time-varying, county- and industry-specific workplace risk measure: $\theta_{tci} = R_{tc'} \cdot P_i$. If there are no local outbreaks, $\theta_{tci} = 0$ for all industries in county c. As local outbreaks increase, θ_{tci} rises, with a proportionally greater impact on more vulnerable industries. $R_{tc'}$ proxies local outbreaks with neighboring counties' fatality rates (c') to avoid potential reverse causality issues that could arise from mass outbreaks at local employers[37]. Since θ_{tci} has no natural units, we normalize this variable to start at 0 and have a standard deviation of 1, so that treatment can be read in terms of standard deviations.

[34]Basso et al. (2021) use O*NET data to compute task-specific risk measures based on proximity to others at work and the possibility of working remotely.

[35]The risk scores in Basso et al. (2021) are reported at the occupation level, so that we compute industry averages for the lowest-digit industries available in the American Community Survey (mostly 3 and 4 digit) by taking an employment-weighted average of occupational risks in each industry. We compute the riskiness at the industry level rather than the occupation level because our worker data only includes industry information.

[36]Note that we focus on fatality rates – rather than infection rates – to measure risks because of a lack of reliable infection data during the first months of the pandemic.

[37]See Appendix E.5 for alternative results obtained using fatality rates in the same county

To provide transparency on the variation that drives our results, we first present preliminary cross-industry evidence on how the response to FPUC differs by industry vulnerability score P_i (the "shares" component of our measure). In practice, we estimate excess mass separately for workers engaged in more or less vulnerable tasks and show that industries with tasks particularly vulnerable to Covid-19 exhibit more excess mass at the FPUC threshold. Figure 5 plots the industry-specific excess mass against the workplace risk measure.[38] This summarises the excess mass shown in Figure 3 at the industry level and measure the average excess/missing mass within a $400 treatment window around the threshold.[39] The Figure shows a strong correlation between excess mass and the risk exposure of the industry. These results are highly significant, a standard deviation in risk increases excess mass by 0.51 percentage points. This is consistent with the prediction that workers are willing to leave more money on the table to bunch at the threshold when there is a greater health risk at work. It is also noteworthy that the data show that workplace risk is a first order predictor of excess mass. Most observations are close to the regression line and this single variable explains half of the variation in excess mass at the FPUC threshold across industries (the R^2 of the regression is 0.56).

A causal interpretation of these cross-industry results requires that no omitted variables are simultaneously correlated with both COVID-19 risk and bunching at the threshold. This means identification can be achieved under at least one of two conditions: (i) any omitted variable is uncorrelated with the regressor of interest (COVID-19 risk), or (ii) any omitted variable is uncorrelated with the outcome of interest (excess mass around UI eligibility thresholds). Hedonic regressions typically rely on condition (i), seeking variation in amenities that is "as good as randomly assigned." However, the approach in this paper shifts the focus to condition (ii), requiring only that potential confounders do not systematically influence the distribution of earnings around UI

[38] The omitted industry is real estate services.

[39] Results with alternative treatment windows are reported in Appendix A2.

thresholds.[40]

In our empirical application, workplace safety from COVID-19 outbreaks is almost certainly correlated with other labor market changes, potentially violating condition (i). However, as long as these labor market changes affect earnings in a smooth manner around the FPUC threshold, they satisfy condition (ii) and do not threaten identification. By construction, this approach is robust to any confounding variables that induce smooth changes in the earnings distribution, ensuring that the estimates are not biased by standard endogeneity concerns, such as workers' exposure to simultaneous economic shocks when workplace amenities change.

Despite these advantages, an important identification concern remains: cross-industry differences in bunching may not solely reflect worker preferences for amenities but could instead be driven by differences in labor supply elasticities among more and less exposed workers. In the cross-industry setting, the concern is that industries affected by larger risk shocks are also industries with higher labor supply elasticities, leading to spurious differences in bunching. We investigate this possibility in two ways. First, we estimate labor supply elasticities by industry and check if they are correlated with task risks. Elasticities are estimated using an IV approach that exploits firm-wide wage changes. This design takes advantage of the fact that firms typically update all worker wages simultaneously, isolating wage changes that are plausibly exogenous to shifts in individual workers' labor supply (see Appendix E.3 for details).[41] Results show that elasticities are only very weakly correlated with task risks – the correlation coefficient is 0.02. As a result, controlling for elasticities in the main analysis has virtually no effect on the results.[42] Second, we leverage the panel dimension of our data to ex-

[40]Meeting condition (i) is notoriously difficult, as amenities are rarely randomly assigned and are often linked to promotions or labor market factors that affect both wages and working conditions. The approach in this paper sidesteps this issue by exploiting condition (ii), allowing identification even when changes in amenities correlate with broader labor market dynamics.

[41]We use leave one out wage changes that use wage updates of all workers excluding the focal worker.

[42]A figure that controls for greater excess mass for more elastic industries shows near identical results to figure 5.

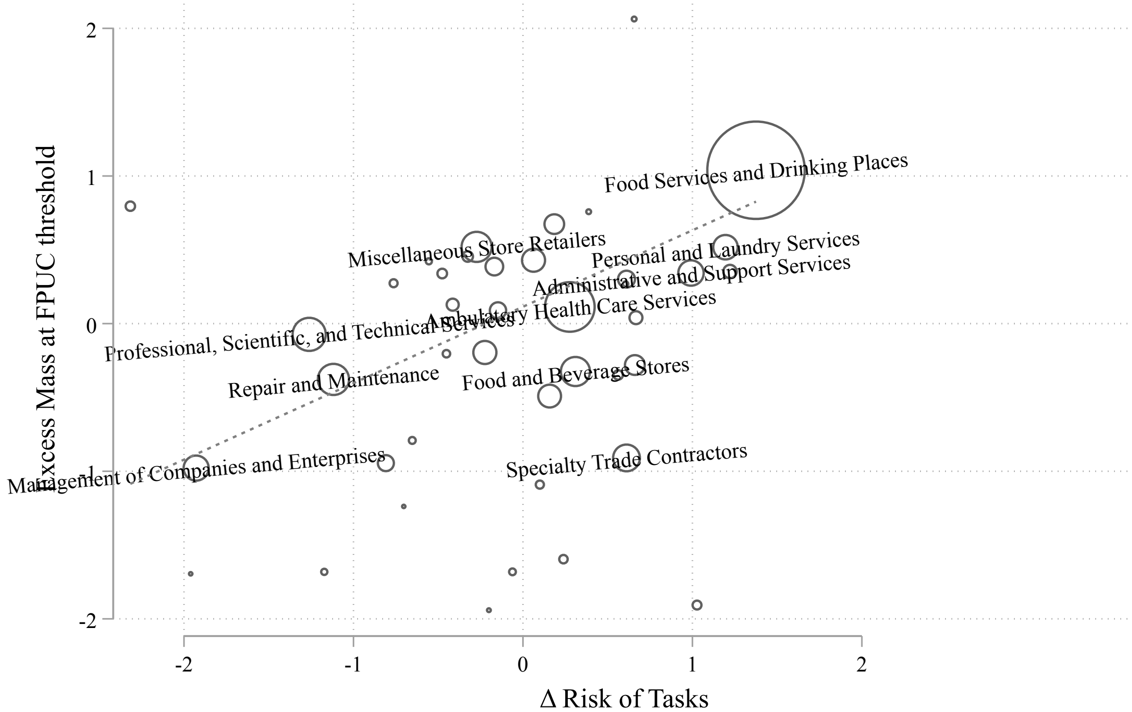
amine whether an individual's bunching behavior varies between weeks of lower and higher workplace risk. The key concern is that some individuals may simply be more responsive to financial incentives, regardless of workplace conditions. To address this, we exploit time variation in county- and industry-specific workplace risk (θ_{tci}) and implement a specification that interacts worker fixed effects with a time dummy for the introduction of FPUC. This specification controls for each worker's baseline responsiveness to the threshold and isolate the effect of changes in workplace risk on excess mass. Results, presented in the next section, show that a given individual is more likely to bunch near the eligibility threshold in weeks when Covid risks are high compared to weeks when risks are low and we find substantially more excess mass in weeks with higher risk. This confirms the preliminary cross-industry findings and shows that differences in time invariant worker characteristics are not driving the results.

Our main analysis is based on the variation of workplace risk in θ_{tci} that combines variation across industry with county and time variation. We estimate the excess mass at the threshold separately for the highest and lowest health risk quintiles. This replicates Figure 3 for the two different risk groups. Figure 6 (a) shows the response in the lowest risk quintile in grey and the highest risk quintile in black. Both the excess mass and the missing mass are more pronounced in the high-risk settings. And higher risk raises the excess mass at earnings levels just below the threshold and generates extra missing mass in the bins above the threshold, in line with our previous results and theoretical predictions. Repeating the earlier placebo test with ineligible workers again confirms that there are no spurious shocks that may confound our results (Figure 6b).

4.4 Implementing the WTP Estimator

We now move from the graphical evidence to estimating point estimates necessary for the WTP calculation in equation 8c. The estimator is constructed as the ratio of

Figure 5: Effect of Workplace Safety on Labor Supply – Task Risk Proxy



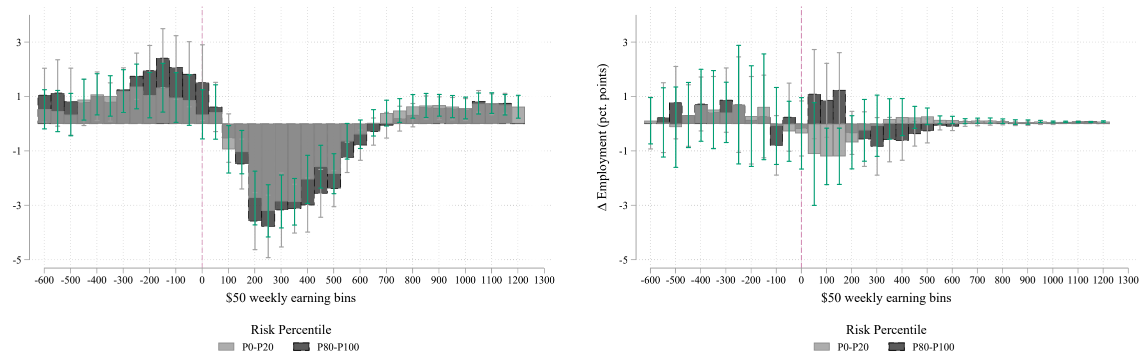
Note: The figure shows the amount of excess mass at the FPUC threshold for 3-digit NAICS industries, relative to the omitted industry (real estate services, NAICS 531). The y-axis shows the excess mass generated by the FPUC eligibility threshold in industry i relative to the omitted industry. The x-axis is based on the riskiness of tasks used in industry i, using the data on Covid-19 risk in specific tasks from Basso et al. (2021) standardized to have a standard deviation of 1. Industry titles are shown for the ten largest industries and for display purposes we only show industries with at least 1,000 observations. The size of the markers corresponds to the number of observations in the industry and regressions are weighted by this number. The fitted line has a slope coefficient of 0.51 and an $R^2 = 0.56$ Source: Homebase.

baseline excess mass (Figure 3) to the change in excess mass in response to variation in workplace risk (Figure 6). In practice, we implement this by interacting a treatment window indicator, T_k[43], which captures a ±$400 earnings range around the eligibility threshold, with the continuous risk measure θ_{tci} defined in Section 4.3 and estimate the

[43]We replace the granular bin dummies I_k in equation (9) with a categorical variable (T_k) that takes value 0 outside the ±$400 treatment window, and inside the window takes value 1 below the threshold (excess mass), and value -1 above the threshold (missing mass). We use the term *excess mass* for the sake of simplicity, however the coefficient of T_k captures both excess and missing mass effects.

Figure 6: Excess and Missing Mass around the Partial UI Notch for Low and High-Risk Settings

(a) FPUC Eligible Workers **(b)** Placebo Sample: Ineligible Workers



Note: The figure shows η_k coefficients from equation (9) for the highest and lowest quintiles of Covid-19 risk (θ_{tci}). The gray bars represent the response in the lowest quintile and the black bars in the highest quintile. The sample in panel (a) covers hourly workers with sufficient past earnings to qualify for MWB payments in their home state and is based on 169,450 work-week spells. The sample in panel (b) has 181,492 work-week spells and covers hourly workers with insufficient past earnings to qualify for MWB payments in their home state. For these workers, the threshold should not be relevant. Source: Homebase.

following triple interaction specification:

$$E_{wmtkci} = \pi^{mt} + \delta \cdot T_k \cdot C_t \cdot \theta_{tci} + \boldsymbol{X\beta} + \varepsilon_{wmtkci} \tag{10}$$

where the coefficient of interest δ measures how excess mass at the threshold responds to changes in Covid-19 risks. C_t and T_k are the Covid-19 period dummy and the $400 treatment window[44] and $\boldsymbol{X}$ is a vector of pairwise interactions and single variable entries of T_k, C_t, and θ respectively, while $\boldsymbol{\beta}$ is the associated vector of coefficients.

The results from this regression are shown in Table 2. Panel A shows results for the denominator of equation (8c) (baseline excess mass), while Panel B shows the numerator (changes in excess mass when workplace risks increase). Panel A estimates excess mass at average risk levels and shows that FPUC creates an excess mass of around 0.86 percentage points in earnings bins around the threshold. Panel B shows

[44] In Figure A2 we test the sensitivity of our DiD estimate to changing treatment windows around the threshold. Our estimate is statistically significant if we consider window equal or larger than $150 around the threshold. We identify only a subset of the response if we focus on a narrow window: once the window is $250 or bigger, the effect is very stable.

Table 2: Willingness To Pay for Workplace Safety

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Baseline Excess Mass					
FPUC	0.858	0.858	0.858	0.858	0.858	0.858
	(0.096)	(0.010)	(0.010)	(0.096)	(0.010)	(0.096)
	Panel B: Additional Excess Mass					
FPUC × Workplace Risk	0.260	0.234	0.232	0.254	0.230	0.260
	(0.053)	(0.052)	(0.052)	(0.052)	(0.052)	(0.053)
Panel C: WTP (% of weekly income) for reduced Workplace Risk by:						
1 std. dev. - Red. Form	30.3	27.3	27.0	29.6	26.8	30.3
1 s.d. - Structural - $e = 0.25$	33.0	29.9	29.5	32.3	29.3	33.0
1 s.d. - Structural - $e = 1.5$	34.8	31.5	31.2	34.0	31.0	34.8
1 death per 100,000	9.0	8.1	8.0	8.8	8.0	9.0
	Panel D:		Value of Statistical Life (million $)			
VSL (perfect information)	$ 5.56	$ 5.01	$ 4.95	$ 5.43	$4.91	$5.56
VSL (worker beliefs)	$ 7.97	$ 7.18	$ 7.10	$ 7.78	$7.05	$7.97
Income FE x Covid FE (π^{mt})	yes	yes	yes	yes	yes	yes
Covid FE x Income x FE for		state	county	industry	worker	-
Labor Supply Elasticity						yes

Note: The Table shows how Covid-19 risk affects excess mass at the FPUC eligibility threshold. Panel A shows excess mass around the FPUC threshold for average risk. Panel B shows δ estimates from equation (10) and captures how excess mass changes with fatality rates. Willingness to pay in Panel C is based on equation (8c), and is the ratio of panel B and panel A estimates. The structural estimation rows additionally use an estimate of labor supply elasticity of $e = 0.25$ or $e = 1.5$, a marginal tax rate $t = 0.12$ and the average FPUC eligibility threshold $m^* = 409$. Panel D computes $VSL = \frac{WTP*m}{\Delta fatality}$, where m is median earnings (m = \$617), and $\Delta fatality$ is one standard deviation of workplace risk increases fatality rates by 3.365 cases (perfect information), or 2.346 cases (worker beliefs) per 100,000 workers. Controls are state, county, and two-digit NAICS fixed effects, interacted with a dummy for the Covid-19 period and a continuous earnings variable. The Labor Supply Elasticity control, interacts elasticity estimates from Appendix E.3 with $T_k \cdot C_t$ to allow for greater response among more elastic workers. The results are based on 169,450 worker-week spells. Source: Homebase, Chen et al. (2021).

that this excess mass increases by 0.26 percentage points for a standard deviation increase in risk. Combining these results ($\frac{0.26}{0.86}$) implies the willingness to pay for a standard deviation of risk is 30.3% of weekly earnings (Panel C), or around $187 in weekly earnings for the median earner in our sample.

Variation in workplace risk during Covid-19 pandemic was large compared to the magnitude of job hazards that workers face in normal times. A standard deviation of θ_{tci} in our data corresponds to an increase in fatality rates by 3.37 fatalities per 100,000 workers.[45] The most deadly occupation in normal times is fishing and hunting, with a fatality rate of 2.9 cases per 100,000 workers per week.[46] A one standard deviation increase in our Covid risk measure (3.37 fatalities per 100,000 workers) is thus comparable to moving from a zero-risk occupation to one of the riskiest occupations in non-Covid times. Alternatively, our estimate implies that workers are willing to forgo around 9% ($=\frac{30.3}{3.365}$) of their earnings to reduce weekly fatality risks by one in a 100,000 (Panel C), a risk variation comparable to the difference in job hazards between a librarian and a roofer in non-pandemic times.

These WTP calculation use the reduced-form approximation in equation (8c). This is a lower bound to the true WTP. We also estimate the corresponding structural parameter to assess whether the lower bound is close to the true parameter. This exercise requires estimates of the labor supply elasticity e, the marginal tax rate t and the threshold value m^*. We use the entry level marginal tax rate $t = 0.12$ and the average threshold in our setting $m^* = \$409$. For e we choose two values (0.25 and 1.5) that represent the lower and the upper bound estimated in the literature.[47] Using indirect inference, we find that our WTP approximation of 30.3 corresponds to

[45] Data on weekly local industry-specific death rates are not available. Therefore, we rely on county/week death counts (D_{tc}) and compute the death counts in each industry by apportioning the deaths to industries based on time-invariant fatality rates in industries and the employment share of the industry. Data on industry-specific fatality rates (ρ_i) are only available for California and we use the data published by Chen et al. (2021). Industry by county employment counts (l_{ci}) come from the ACS 2014-2018. We apportion county-week fatalities to industries as follows: $R_{tci} = D_{tc}\frac{l_{ci}\cdot\rho_i}{\sum_i l_{ci}\cdot\rho_i}$

[46] Source: BLS Census of Fatal Occupational Injuries (CFOI).

[47] $e = 0.25$ is based on the meta-study by (Chetty, 2012). As upper bound we use estimates from the literature on short-run labor supply of stadium vendors, bicycle messengers, and taxi drivers that typically find larger labor supply elasticities. Specifically, we use an upper range estimate in this literature from Fehr and Goette (2007), who estimate $e = 1.5$ for bicycle messengers in Zurich. We follow these studies and interpret such estimates as structural labor supply elasticity. Powell (2012) points out that reduced-form elasticities can represent a combination of structural labor supply elasticities and values of amenities. One could follow this idea and recover structural elasticities from the above papers by solving two equations in two unknowns. We do not do this here and instead take the results from the previous articles at face value.

a structural WTP of 33.0 (Panel C) for an elasticity of 0.25 and 34.8 for $e = 1.5$. The reduced-form lower bound is thus between 2.7 and 4.5 percentage points below the structural parameter and illustrates that the approximation provides a relatively tight bound for the true structural parameter.

We probe the "smoothness assumption" and the "preference-orthogonality assumption" in several ways. The first assumption requires that the threshold responses are not generated by spurious shocks. The placebo test in the previous section already provided evidence in favor of this assumption. We now check further whether this assumption holds among the workers of the treatment group, by introducing flexible controls. First, we probe for potential spurious effects from location-specific policies, such as local lock-downs or school closures by interacting the semi-parametric baseline control for demand shocks π^{mt} (i.e. the Covid-19 period dummy interacted with income level) with either state (column 2) or county fixed effects (column 3). These controls absorb state- or county-wide pre- vs. post-Covid19 changes in the earning distribution. The remaining identifying variation in θ_{tci} comes from cross-industry over time heterogeneity in risk *within* the local area. The results are similar to our baseline results. In column 4 we absorb cross-industry differences in the response to the pandemic, leveraging the interaction of the Covid-19 period dummy with income level *and* industry fixed effects. The remaining identifying variation in θ_{tci} comes from over time heterogeneity in risk *within* industry and net of industry-specific average pre- vs post-Covid-19 changes in the earning distribution. This specification therefore controls for the possibility that workers of some industries adjust their hours on average more at the onset of the pandemic and the FPUC introduction. [48] We again find similar results, confirming that other shocks are orthogonal to our threshold design.

The second "preference-orthogonality" assumption requires that individual labor supply elasticities are uncorrelated with the risk variable θ_{tci}. To address this concern,

[48]This specification therefore captures complementary variation relative to the preliminary cross-industry evidence of Figure 5

in column 5 of Table 2, we introduce an interaction of the Covid-19 period dummy with income level *and* individual fixed effects. Such specification compares an individual's earnings in weeks of high workplace risk to those in weeks of lower risk, while controlling for their individual pre- vs. post-Covid 19 labor supply response. Although these controls should capture most cross-individual differences in preferences, we further ensure robustness by explicitly controlling for industry-level labor supply elasticities, estimated using pre-Covid-19 data, as detailed in Appendix E.3. In column 6, we interact these labor supply elasticities with the treatment window and Covid-19 dummy ($T_k \cdot C_t$). The results remain virtually unchanged results, confirming that heterogeneity in labor supply elasticity is not a major source of bias in our study.

4.5 Additional Robustness Tests

We address lingering concerns about the impact of spurious effects of deteriorating economic conditions. Because states use different eligibility thresholds, we can implement a border design. This design narrows in on the counties at state borders, where different partial UI thresholds apply but arguably demand conditions are similar (see Appendix E.2). We also examine whether rising excess mass could be explained by employers becoming more willing to let workers adjust their hours when demand softens. We add controls for demand variation at the local level and allow these to have different effects around our thresholds.[49] The results remain virtually unchanged (see Appendix E.4). Finally, we consider the possibility that the labor supply reaction is driven by the increased childcare responsibility rather than by health risk. Controlling for local school closures (Parolin and Lee, 2021a,b) again has little effect on the results (see Appendix E.4). All these checks confirm our claim that other shocks are orthogonal to our threshold design.

In Appendix E.6 we also discuss the robustness of our estimates to different sam-

[49]Controls include employment, business revenues, and the number of open businesses at the week and county level from https://tracktherecovery.org/ by Chetty et al. (2020a,b).

ple selection strategies and to the inclusion of an extensive labor supply margin. We relax the work-week restriction and extend the analysis to less-attached workers. The resulting estimates for the willingness to pay remain very close to the baseline estimate (29% of weekly earnings instead of 30%). Finally, estimates obtained using alternative approaches to consider extensive margin responses in the analysis range between 23% and 26% of weekly earnings, slightly smaller, but in the ballpark of our baseline estimates.

5 Discussion

5.1 Comparison with Hedonic Wage Regressions

We compare these findings to a conventional hedonic wage approach, by regressing hourly wages on the measure of workplace risk ($\theta_{c,t,i}$) using the same data. Individual fixed effects control for time-invariant worker ability and ensure that selection effects do not bias these results. The hedonic regression shows that wages are broadly unchanged by workplace risk, and the point estimate is insignificant. The estimate is also quantitatively small and suggests that wages increased by 11 cents with one standard deviation of θ_{tci}, which corresponds to a 0.5% wage increase (results are available upon request). Interpreted through the lens of a hedonic regression, these estimates would lead us to conclude that workers attach next to no value to workplace safety. However, another explanation for the small coefficient is that wages are slow to adjust for the small businesses we analyze, which did not implement Covid-19 hazard pay as some notable large companies did. Wages are thus unlikely to fully price in changes in workplace risk at least in the short-run. This echos findings in the literature that also report no impact of workplace risks on wages in ongoing spells (Brown, 1980; Kniesner et al., 2012; Viscusi and Aldy, 2003). Such results in the prior literature sparked a debate how one can recover underlying WTP parameters using a more structural approach that models adjustment frictions. Different from such approaches, our approach does not make assumptions about the wage-setting process. Our results show that workers respond substantially to workplace risks, and their behavior around notches suggests that the WTP is two orders of magnitude greater than the hedonic estimate. This aligns with recent studies arguing that amenities are more valuable than traditional hedonic regressions would suggest (e.g., Lavetti, 2023b; Lamadon, Mogstad, and Setzler, 2022; Maestas et al., 2023).

Another important consideration is heterogeneity in WTP across workers. The parameter of interest may depend on the context but is typically the average value of an amenity among workers with access to the amenity. The compensating differential approach identifies the WTP of the marginal worker indifferent between choosing the amenity job or an alternative higher-paid job. This is typically a lower bound for the parameter of interest. Different from this, our approach estimates an average WTP, taking the average of WTPs for workers near the threshold used for the identification. This is an average local treatment effect (LATE) for the population at the threshold and when the threshold is independent of workers' WTP, this will correspond to the parameter of interest. Another useful feature of our set-up is that it enables the researcher to study heterogeneity in WTP directly by analyzing excess mass changes in different demographic sub-groups of the population. This has been challenging in hedonic regressions framework, since there is only one market clearing wage and compensating differentials can thus only be estimated for one worker. Providing estimates for the heterogeneity in WTP across sub-groups is particularly useful for typical policy settings that seek to expand access to amenities (like workplace safety) to subgroups that would benefit most from them.

5.2 Value of a Statistical Life

A popular approach for quantifying responses to health risks is to compute a "value of a statistical life" (VSL), which infers the implicit value of life from observed responses to risks. Such estimates typically assume that individuals know and understand their risk exposure and that the fear of dying is the sole driver of the observed behavior. Since higher fatality rates are typically accompanied by unpopular safety measures and by risks of non-fatal injuries, this assumption effectively imposes that workers attach zero value to such non-fatal aspects. Under these assumptions common to the VSL literature, we can compute VSL as the ratio of WTP (in absolute dollars) to the change in fatality risk: $VSL = \frac{WTP*m}{\Delta fatality}$ with WTP being our main estimate from Table 2 column 1 and m the median earnings in our sample. Using our estimates, we find $VSL = \frac{0.303*\$617}{3.37/100{,}000} = \$5.56$ million (Panel D of Table 2). A value of \$5.56 million broadly aligns with the literature, a recent meta-study by Viscusi (2018) concludes that VSL is somewhere between \$3 and \$13 million (in 2020 USD). Our results align with these findings and lean towards the lower side of this range.

The main purpose of this exercise is to benchmark our WTP method and illustrate that it produces reasonable results. However, when generalizing our estimates to non-Covid-19 workplace risk, we need to consider the level of information people have about the workplace risk and the transmissible nature of the risk under study.

First, our empirical context offers a unique opportunity to assess the importance of the perfect information assumption. Ideally, researchers would relax the perfect information assumption and compute $VSL = \frac{WTP}{E[\Delta fatality]}$, where $E[\Delta fatality]$ is the workers' perception of fatality risk. Since these perceptions are not usually observed, studies instead use the statistical fatality rates as a proxy for perception, thereby imposing perfect information and rational expectations assumptions.[50] During the Covid-19 outbreak, beliefs about fatality risks were collected as part of the Understanding America

[50]Frequent violations of these assumptions are famously documented in Kahneman and Tversky (1979).

Study (UAS), which allows us to relax the perfect information assumption.[51] The estimate is an instrumental variable approach that instruments fatality beliefs with our risk measure. Using this strategy to adjust the VSL estimate for (mis)perception of risk ($E[\Delta fatality]$), the VSL value increases to $7.97 million (Panel D of Table 2). Accounting for imperfect information thus increases the VSL estimate by nearly 50%, highlighting the importance of the popular assumptions underpinning VSL calculation.[52]

Second consideration, our WTP estimate could partly reflect workers internalizing the risk of Covid-19 transmission to others. The WTP for a non-transmissible illness or injury might therefore be lower. The higher the weight workers place on others' health in their utility function, the more likely our estimate represents an upper bound for non-transmissible workplace risk. Conversely, in the canonical case of self-interested individuals, who only care about their own utility, the WTP for transmittable and non-transmittable health risks coincide.[53] To get a sense of the importance of the pro-social feature in our WTP estimate, we perform a back-of-the-envelope calculation for a worker who cares about the well-being of other household members in Appendix F. The exercise suggests that pro-social concerns make up less than 1% of the estimated WTP and the concern for one's own health is the main component of the WTP estimate.

[51]The data covers a representative sample of the US population and uses weekly rounds of interviews. Individuals were asked about their probability of contracting Covid-19 and conditional on this, their probability of dying. We use this data to compute expectations at the week-state-industry level and then use these to impute expectations for our sample. The expectation measure thus undoubtedly includes measurement error.

[52]It is unclear whether individuals are particularly poorly informed in our context. On the one hand, we study an event with enormous press coverage that was almost certainly salient to everyone. On the other hand, there was substantial uncertainty around the risks of Covid-19.

[53]The prior literature almost exclusively considers self-interested agents when interpreting risky behavior of individuals.

5.3 Workplace Safety Policy

Our results suggest that workers value workplace safety highly and that more stringent safety regulations provide substantial gains to workers.[54] To illustrate this point, we perform three back-of-the-envelope calculations. The first quantifies the hazard pay required during Covid-19 to make workers indifferent between working when exposed to Covid risk vs when not. Our results imply that the utility cost of working under an increase of Covid risk by one standard deviation would require an offsetting hourly wage increase of \$4.8. This is larger than the wage change we observe in practice (about 11 cents) and also larger than reported hazard rates at large retailers which top out between \$2 and \$4. A substantial part of the increased cost induced by Covid-19 workplace risk was thus not priced into wages. Second, we turn to the construction industry, one of the largest industries with substantial workplace risk. Weekly fatality rates in this industry in the US are 0.3 workers per 100,000 full-time employees per week, while comparable estimates for Germany and the UK are respectively 0.04 and 0.07 weekly deaths per 100,000 workers.[55] Our estimates imply that reducing US fatality rates to the level seen in the UK or Germany would provide substantial gains to workers, valued equivalently to a wage increase of 2.2%. Such gains happen to be similar in magnitude to the average wage gains from the introduction of a \$15 minimum wage in the industry, a popular labor market intervention proposal.[56] Finally, we consider the gains implied by switching between industries with different risk levels. Such an exercise helps to evaluate the potential of compensating differentials to explain the dispersion of wages. The gains from greater safety by changing from the construction sector to the safer accommodation and food services sector are worth around 2.5% of

[54]This result is consistent with the literature on the value of disability insurance, which finds sizable welfare gains from disability benefit payments (Cabral and Cullen, 2019; Deshpande and Lockwood, 2022)

[55]ILO data is converted to weekly deaths per 100.000 workers for comparison. Annual fatality rates are 16 per 100.000 workers in 2018. Source: ILOSTAT, series "INJ FATL ECO RT A" 2018.

[56]The minimum wage calculation computes the wage floor that is equivalent to a 2.2% mean wage increase (assuming no employment loss). The data source is the 2019 and 2020 CPS ASEC data.

earnings, while moving to the riskier agricultural sector is equivalent to a wage loss of 8%. The magnitude of these gains is comparable to the value of other work amenities analyzed by Maestas et al. (2018), who find values ranging from 2% to 16%.[57]

5.4 Valuing Bundles of Amenities

Finally, we provide an entirely different application of the WTP method in appendix G to illustrate that the approach applies more widely and can be used for different types of amenities. This application focuses on the monetary value of enjoyable work and provides a money metric for job satisfaction scores that are widely collected in labor market surveys. Estimating the value of "good" and "bad" jobs has been central in labor economics (for an overview, see Lavetti (2023a)). Work enjoyment captures an aggregate (net) value of several amenities in a given job and the application also illustrates that the approach can be used to identify the value of broader bundles of amenities. For this exercise, we use data from the US Health and Retirement Survey (HRS) and analyze bunching around the U.S. early retirement age threshold. Workers accumulate social security entitlements for each quarter worked and once they reach age 62 the marginal value of additional quarters changes, creating a kink in the lifetime budget constraint. We study how bunching at the 62 age threshold differs for workers in high and low-enjoyment jobs. We find that workers are willing to take a 12.5% pay-cut to move from an average satisfying job to a highly satisfying job.

6 Conclusions

This paper presents a new revealed-preference method to estimate the value of non-wage amenities based on bunching in the earnings distribution around budget discontinuities in response to varying amenities. The approach formalizes the idea that workers

[57]Maestas et al. (2018) study the value of schedule autonomy, telecommuting, physical activity, sitting, relaxed work environment, work autonomy, PTO, teamwork, training, and opportunity to serve.

will be less responsive to financial incentives when non-wage amenities make up a larger part of workers' compensation. We apply this method to measure the value workers attach to safe workplaces.

Our identification leverages a budget constraint notch created by the launch in the United States of an extra-ordinary UI benefit program in March 2020 (the so-called "FPUC"). We find substantial baseline workers' reaction to this notch and show that these labor supply responses increase during periods of heightened Covid-19 risks, creating magnified excess mass. The estimates imply that workers are willing to sacrifice 30% of their weekly earnings to decrease their risk by one standard deviation. This is equivalent to giving up 9% of earnings to avoid a 1 in 100,000 risk of dying, a variation equivalent to the different in risk between a librarian and a roofer. These estimates are two orders of magnitude larger than canonical hedonic wage regressions. A difference that is likely driven by frictions in wage setting, as discussed by the recent literature on hedonic regressions. Our novel framework is designed to provide unbiased estimates of the value of workplace amenities even if the perfectly competitive wage-setting assumption of the canonical compensating differential models fails.

The revealed-preference framework introduced in this paper for estimating the value of non-wage amenities offers distinct advantages and limitations compared to other methods, such as those based on stated preferences. While it relies on specific empirical conditions—namely, a budget discontinuity and variations in amenities—that may not be universally present, it has the benefit of leveraging existing surveys or administrative data, eliminating the need for custom survey experiments. The approach can also be flexibly applied to estimate WTP for specific or bundled amenities. We hope this novel method will expand the set of empirical tools available to researchers interested in estimating non-wage amenities, which constitute a large and increasing part of workers' compensation, as signaled by the prominent role these amenities are playing in the discussions around the changing nature of work, from the gig economy to work-from-home and the "Great Resignation."

References

Altonji, Joseph G. and Christina H. Paxson. 1992. "Labor Supply, Hours Constraints, and Job Mobility." *The Journal of Human Resources* 27 (2):256–278.

Ashenfelter, Orley. 2006. "Measuring the Value of a Statistical Life: Problems and Prospects." *Economic Journal* 116 (510).

Ashenfelter, Orley and Michael Greenstone. 2004. "Using Mandated Speed Limits to Measure the Value of a Statistical Life." *Journal of Political Economy* 112 (S1):S226–S267.

Basso, Gaetano, Tito Boeri, Alessandro Caiumi, and Marco Paccagnella. 2021. "Unsafe Jobs, Labour Market Risk and Social Protection." Economic Policy.

Bell, Alex. 2022. "Job Amenities and Earnings Inequality." *SSRN Electronic Journal* .

Blomquist, Sören, Whitney K. Newey, Anil Kumar, and Che-Yuan Liang. 2021. "On Bunching and Identification of the Taxable Income Elasticity." *Journal of Political Economy* 129 (8):2320–2343.

Boeri, Tito and Pierre Cahuc. 2023. "Labor Market Insurance Policies in the Twenty-First Century." *Annual Review of Economics* 15.

Bollinger, Christopher R. 1998. "Measurement Error in the Current Population Survey: A Nonparametric Look." *Journal of Labor Economics* 16 (3):576–594.

Bonhomme, Stéphane and Grégory Jolivet. 2009. "The Pervasive Absence of Compensating Differentials." *Journal of Applied Econometrics* 24 (5):763–795.

Bound, John and Alan B. Krueger. 1991. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics* 9 (1):1–24.

Brown, Charles. 1980. "Equalizing Differences in the Labor Market." *The Quarterly Journal of Economics* 94 (1):113–134.

Cabral, Marika and Mark R Cullen. 2019. "Estimating the value of public insurance using complementary private insurance." *American Economic Journal: Economic Policy* 11 (3):88–129.

Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *The Quarterly Journal of Economics* 128 (3):967–1015.

Card, David, Andrew Johnston, Pauline Leung, Alexandre Mas, and Zhuan Pei. 2015a. "The Effect of Unemployment Benefits on the Duration of Unemployment Insurance Receipt: New Evidence from a Regression Kink Design in Missouri, 2003–2013." *American Economic Review: Papers & Proceedings* 105 (5):126–130.

Card, David, David S. Lee, Zhuan Pei, and Andrea Weber. 2015b. "Inference on Causal Effects in a Generalized Regression Kink Design." *Econometrica* 83 (6):2453–2483.

———. 2017. "Regression Kink Design: Theory and Practice." In *Regression Discontinuity Designs (Advances in Econometrics, Vol. 38)*, edited by Matias D. Cattaneo and Juan Carlos Escanciano. Emerald Publishing Limited, 341–382.

Chen, Yea Hung, Maria Glymour, Alicia Riley, John Balmes, Kate Duchowny, Robert Harrison, Ellicott Matthay, and Kirsten Bibbins-Domingo. 2021. "Excess Mortality Associated with the COVID-19 Pandemic Among Californians 18-65 Years of Age, by Occupational Sector and Occupation: March through November 2020." *PLoS ONE* 16 (6):e0252454.

Chetty, Raj. 2012. "Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." *Econometrica* 80 (3):969–1018.

Chetty, Raj, John N. Friedman, Nathaniel Hendren, Michael Stepner, and The Opportunity Insights Team. 2020a. "Data for: The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data." URL `https://tracktherecovery.org/`.

———. 2020b. "The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data." NBER Working Paper 27431.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 2011. "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." *The Quarterly Journal of Economics* 126 (2):749–804.

Chetty, Raj, John N. Friedman, and Emmanuel Saez. 2013. "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings." *American Economic Review* 103 (7):2683–2721.

Deshpande, Manasi and Lee M. Lockwood. 2022. "Beyond Health: Nonhealth Risk and the Value of Disability Insurance." *Econometrica* 90 (4):1781–1810. URL `https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA19668`.

Dobkin, Carlos, Amy Finkelstein, Raymond Kluender, and Matthew J. Notowidigdo. 2018. "The Economic Consequences of Hospital Admissions." *American Economic Review* 108 (2):308–352.

Dube, Arindrajit, Suresh Naidu, and Adam Reich. 2022. "Power and Dignity in the Low-Wage Labor Market: Theory and Evidence from Walmart Workers." Working paper.

Einarsen, Ståle Valvatne, Helge Hoel, Dieter Zapf, and Cary L. Cooper. 2011. *Bullying and Harassment in the Workplace: Developments in Theory, Research, and Practice*. CRC Press, 2 ed.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf. 2017. "Bunching at the Kink: Implications for Spending Responses to Health Insurance Contracts." *Journal of Public Economics* 146:27–40.

Fehr, Ernst and Lorenz Goette. 2007. "Do workers work more if wages are high? Evidence from a randomized field experiment." *American Economic Review* 97 (1):298–317.

Fishback, Price V. and Shawn Everett Kantor. 1995. "Did Workers Pay for the Passage of Workers' Compensation Laws?" *The Quarterly Journal of Economics* 110 (3):713–742.

Folke, Olle and Johanna Rickne. 2022. "Sexual harassment and gender inequality in the labor market." *The Quarterly Journal of Economics* 137 (4):2163–2212.

Forsythe, Eliza. 2023. "Unemployment Insurance Recipiency During the Covid-19 Pandemic." *National Tax Journal* 76 (2):000–000.

Forsythe, Eliza, Lisa B Kahn, Fabian Lange, and David Wiczer. 2020. "Labor demand in the time of COVID-19: Evidence from vacancy postings and UI claims." *Journal of public economics* 189:104238.

Gallant, Jessica, Kory Kroft, Fabian Lange, and Matthew J Notowidigdo. 2020. "Temporary Unemployment and Labor Market Dynamics During the COVID-19 Recession." Working Paper 27924, National Bureau of Economic Research. URL `http://www.nber.org/papers/w27924`.

Ganong, Peter, Fiona E Greig, Pascal J Noel, Daniel M Sullivan, and Joseph S Vavra. 2022. "Spending and Job-Finding Impacts of Expanded Unemployment Benefits: Evidence from Administrative Micro Data." Working Paper 30315, National Bureau of Economic Research. URL `http://www.nber.org/papers/w30315`.

Ganong, Peter, Pascal Noel, and Joseph Vavra. 2020. "US unemployment insurance replacement rates during the pandemic." *Journal of public economics* 191:104273.

Goldin, Claudia and Lawrence F. Katz. 2011. "The Cost of Workplace Flexibility for High-Powered Professionals." *The Annals of the American Academy of Political and Social Science* 638 (1):45–67.

———. 2016. “A Most Egalitarian Profession: Pharmacy and the Evolution of a Family-Friendly Occupation.” *Journal of Labor Economics* 34 (3):705–746.

Gruber, Jonathan. 1994. “The Incidence of Mandated Maternity Benefits.” *American Economic Review* 84 (3):622–641.

———. 1997. “The Incidence of Payroll Taxation: Evidence from Chile.” *Journal of Labor Economics* 15 (S3):S72–S101.

Gruber, Jonathan and Alan B. Krueger. 1991. “The Incidence of Mandated Employer-Provided Insurance: Lessons from Workers’ Compensation Insurance.” In *Tax Policy and the Economy*, vol. 5, edited by David Bradford. The MIT Press, 111–143.

Guardado, José R. and Nicolas R. Ziebarth. 2019. “Worker Investments in Safety, Workplace Accidents, and Compensating Wage Differentials.” *International Economic Review* 60 (1):133–155.

Gupta, Sumedha, Laura Montenovo, Thuy Nguyen, Felipe Lozano-Rojas, Ian Schmutte, Kosali Simon, Bruce A Weinberg, and Coady Wing. 2023. “Effects of social distancing policy on labor market outcomes.” *Contemporary economic policy* 41 (1):166–193.

Hamermesh, Daniel S. 1999. “Changing Inequality in Markets for Workpalce Amenities.” *Quarterly Journal of Economics* CVIV (4):1085–1123.

Holzer, Harry J, Glenn Hubbard, and Michael R Strain. 2024. “Did pandemic unemployment benefits increase unemployment? Evidence from early state-level expirations.” *Economic Inquiry* 62 (1):24–38.

Hwang, Hae-shin, W. Robert Reed, and Carlton Hubbard. 1992. “Compensating Wage Differentials and Unobserved Productivity.” *Journal of Political Economy* 100 (4):835–858.

Johnson, Matthew S. 2020. “Regulation by shaming: Deterrence effects of publicizing violations of workplace safety and health laws.” *American Economic Review* 110 (6):1866–1904.

Johnson, Matthew S., David Levine, and Maichael Toffel. 2022. “Improving Regulatory Effectiveness through Better Targeting: Evidence from OSHA.” *Working Paper* .

Kahneman, Daniel and Amos Tversky. 1979. “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica* 47 (2):263–292.

Kleven, Henrik J. and Mazhar Waseem. 2013. “Using Notches To Uncover Optimization Frictions And Structural Elasticities: Theory and Evidence from Pakistan.” *The Quarterly Journal of Economics* 128 (2):669–723.

Kleven, Henrik Jacobsen. 2016a. “Bunching.” *Annual Review of Economics* 8 (1):435–464.

———. 2016b. “Bunching.” *Annual Review of Economics* 8:435–464.

Kniesner, Thomas J., W. Kip Viscusi, Christopher Woock, and James P. Ziliak. 2012. “The value of a statistical life: Evidence from panel data.” *Review of Economics and Statistics* 94 (1):74–87.

Lagos, Lorenzo. 2020. “Labor Market Institutions and the Composition of Firm Compensation: Evidence from Brazilian Collective Bargaining.” :1–96.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler. 2022. “Imperfect Competition, Compensating Differentials and Rent Sharing in the U.S. Labor Market.” *American Economic Review* 112 (1).

Lang, Kevin and Sumon Majumdar. 2004. “The pricing of job characteristics when markets do not clear: Theory and policy implications.” *International Economic Review* 45 (4):1111–1128.

Lavetti, Kurt. 2020. “The Estimation of Compensating Wage Differentials: Lessons From the Deadliest Catch.” *Journal of Business & Economic Statistics* 38 (1):165–182.

———. 2023a. “Compensating Wage Differentials in Labor Markets: Empirical Challenges and Applications.” 37 (3):189–212.

———. 2023b. “Compensating wage differentials in labor markets: Empirical challenges and applications.” *Journal of Economic Perspectives* 37 (3):189–212.

Lavetti, Kurt and Ian M Schmutte. 2022. “Estimating Compensating Wage Differentials with Endogenous Job Mobility.” *working paper* URL `http://digitalcommons.ilr.cornell.edu/ldi/46`.

Le Barbanchon, Thomas. 2020. “Taxes today, benefits tomorrow.” Tech. rep., Working Paper.

Le Barbanchon, Thomas, Roland Rathelot, and Alexandra Roulet. 2021. “Gender Differences in Job Search: Trading off Commute against Wage.” *The Quarterly Journal of Economics* 136 (1):381–426.

Lee, David S, Pauline Leung, Christopher J O’Leary, Zhuan Pei, and Simon Quach. 2021. “Are sufficient statistics necessary? nonparametric measurement of deadweight loss from unemployment insurance.” *Journal of Labor Economics* 39 (S2):S455–S506.

Lee, Jungho. 2022. “Start-up firms and corporate culture: Evidence from advertised corporate culture.” .

Lehmann, Tobias. 2022. "Non-Wage Job Values and Implications for Inequality." (November).

Lucas, Robert E. B. 1977. "Hedonic Wage Equations and Psychic Wages in the Returns to Schooling." *American Economic Review* 67 (4):549–558.

Maestas, Nicole, Kathleen J. Mullen, David Powell, Till von Wachter, and Jeffrey B. Wenger. 2018. "The Value of Working Conditions in the United States and Implications for the Structure of Wages." NBER Working Paper 25204.

Maestas, Nicole, Kathleen J Mullen, David Powell, Till Von Wachter, and Jeffrey B Wenger. 2023. "The value of working conditions in the United States and implications for the structure of wages." *American Economic Review* 113 (7):2007–2047.

Marinescu, Ioana, Daphne Skandalis, and Daniel Zhao. 2021. "The impact of the federal pandemic unemployment compensation on job search and vacancy creation." *Journal of Public Economics* 200:104471.

Mas, Alexandre and Amanda Pallais. 2017. "Valuing Alternative Work Arrangements." *American Economic Review* 107 (12):3722–3759.

Masters, Stanley H. 1969. "An Interindustry Analysis of Wages and Plant Size." *The Review of Economics and Statistics* 51 (3):341–345.

Morchio, Iacopo and Christian Moser. 2019. "The Gender Gap: Micro Sources and Macro Consequences." Available at SSRN.

Parolin, Zachary and Emma Lee. 2021a. "U.S. School Closure & Distance Learning Database." Data retrieved from OSF.

Parolin, Zachary and Emma K. Lee. 2021b. "Large Socio-economic, Geographic and Demographic Disparities Exist in Exposure to School Closures." *Nature Human Behaviour* 5:522–528.

Pichler, Stefan and Nicolas R. Ziebarth. 2019. "Reprint of: The pros and cons of sick pay schemes: Testing for contagious presenteeism and noncontagious absenteeism behavior." *Journal of Public Economics* 171 (December 2017):86–104. URL `https://doi.org/10.1016/j.jpubeco.2019.03.005`.

Pierce, Brooks. 2001. "Compensation inequality." *Quarterly Journal of Economics* 116 (4):1493–1525.

Powell, David. 2012. "Compensating differentials and income taxes: Are the wages of dangerous jobs more responsive to tax changes than the wages of safe jobs?" *Journal of Human Resources* 47 (4):1023–1054.

Rosen, Sherwin. 1974. “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition.” *Journal of Political Economy* 82 (1):34–55.

———. 1986. “The Theory of Equalizing Differences.” In *Handbook of Labor Economics*, vol. 1, edited by Orley C. Ashenfelter and Richard Layard. Elsevier, 641–692.

Roussille, Nina and Benjamin Scuderi. 2022. “Bidding for Talent: Equilibrium Wage Dispersion on a High-Wage Online Job Board.” Working paper.

Ruppert, Peter, Elena Stancanelli, and Etienne Wasmer. 2009. “Commuting, Wages and Bargaining Power.” *Annals of Economics and Statistics* (95/96):201–220.

Ruser, John and Richard Butler. 2009. “The economics of occupational safety and health.” *Foundations and Trends in Microeconomics* 5 (5):301–354.

Schmieder, Johannes F and Till Von Wachter. 2016. “The effects of unemployment insurance benefits: New evidence and interpretation.” *Annual Review of Economics* 8:547–581.

Smith, Adam. 1776. *An Inquiry into the Nature and Causes of The Wealth of Nations*. Chicago: The University of Chicago Press.

Sockin, Jason. 2022. “Show Me the Amenity: Are Higher-Paying Firms Better All Around?” *SSRN Electronic Journal* .

Sorkin, Isaac. 2018. “Ranking Firms Using Revealed Preference.” *The Quarterly Journal of Economics* 133 (3):1331–1393.

Stern, Scott. 2004. “Do Scientist Pay to Be Scientist?” *Management Science* 50 (6):835–853.

Summers, Lawrence H. 1989. “Some Simple Economics of Mandated Benefits.” *American Economic Review, Papers and Proceedings of the Hundred and First Annual Meeting of the American Economic Association* 79 (2):177–183.

Taber, Christopher and Rune Vejlin. 2020. “Estimation of a Roy/Search/Compensating Differential Model of the Labor Market.” *Econometrica* 88 (3):1031–1069.

Tsao, Carolyn. 2024. “It’s Not (Just) About the Money: The Value of Working Conditions in Teaching.” .

U.S. Bureau of Labor Statistics. 2021. “Survey of Occupational Injuries and Illnesses"; “Census of Fatal Occupational Injuries"; “Occupational Employment and Wage Statistics.”

Viscusi, W. Kip. 2018. “Best Estimate Selection Bias in the Value of a Statistical Life.” *Journal of Benefit-Cost Analysis* 9 (2):205–246.

Viscusi, W. Kip and Joseph E. Aldy. 2003. “The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World.” *Journal of Risk and Uncertainty* 27 (1):5–76.

Wiswall, Matthew and Basit Zafar. 2018. “Preference for the Workplace, Investment in Human Capital, and Gender.” *The Quarterly Journal of Economics* 133 (1):457–507.

Online Appendix for the manuscript "Willingness to Pay for Workplace Amenities"

Massimo Anelli[1] and Felix Koenig[2]

[1]Bocconi University, IZA, CESifo

[2]Carnegie Mellon University, CEP, IZA

May 5, 2025

A Online Appendix Figures

B Homebase Data Benchmarking

We compare the Homebase data and our analysis sample to the characteristics of the labor force from the Annual Social and Economic (ASEC) supplement to the Current Population Survey (CPS) and the Quarterly Workforce Indicators (QWI). Table A1

*Massimo Anelli, massimo.anelli@unibocconi.it, Bocconi University, Department of Social and Political Science. Felix Koenig, fkoenig@cmu.edu, Carnegie Mellon University, Heinz College. We are grateful to Elvin Bora and Akseli Palomaki for excellent research assistance. We thank Thomas Le Barbanchon, Giulia Giupponi, Hilary Hoynes, Brian Kovak, Matthew Notowidigdo, Lowell Taylor, Andrea Weber and seminar participants at NBER Summer Institute - Labor Studies, NBER COVID-19 and Health Outcomes Conference, Bocconi University, Boston College, Boston University, EEA-ESSEM, McGill, Barcelona summer forum, SOLE/MEA, North American Meetings of the Econometric Society, SOFI - Stockholm University for helpful comments. We gratefully acknowledge funding from the Block Center for Technology and Society at Carnegie Mellon University and support from the "Adopt a Paper mentoring program".

Figure A1: Scheduling App Screenshot

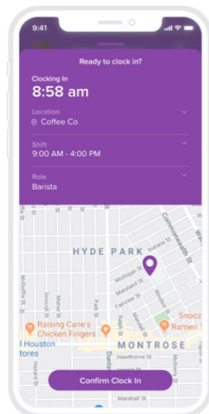
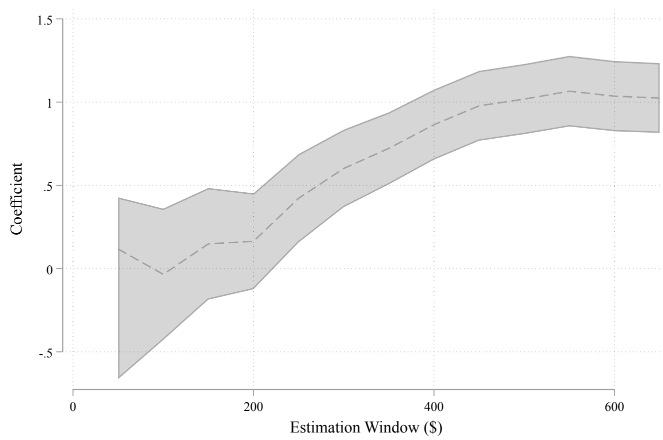


Figure A2: Effect of FPUC with Alternative Treatment Windows



Note: This figure shows results for equation (10) with alternative treatment windows T_k. The horizontal axis refers to the width of T_k to the left and right of the threshold.

presents summary statistics for wages, weekly earnings, and hours worked and Table A2 lists the distribution of observations by 2-digit NAICS sectors in these data.

Homebase provides 6 digit NAICS codes but ASEC does not provide an industry classification that uses NAICS. Therefore, to allow for comparability of ASEC to the Homebase sample, the industry classification in ASEC is first crosswalked to NAICS using the crosswalk provided by IPUMS.[1] Next, the ASEC sample is restricted to

[1]See "IND AND INDNAICS: CODES FOR INDUSTRY (IND) AND NAICS INDUSTRY (INDNAICS) IN THE 2000 CENSUS AND THE ACS/PRCS SAMPLES FROM 2000 ONWARD"

Homebase NAICS codes in a step-by-step manner: if an ASEC industry is linked to a 6-digit NAICS code, it is classified as being in the Homebase sample only if it matches a 6-digit Homebase code, and it is classified as not in the sample if it does not match any 6-digit Homebase code. Next, if an ASEC industry is linked to a 5-digit NAICS code, it is classified as in the Homebase sample if it matches the first 5 digits of a 6-digit Homebase NAICS code. This process is repeated until all ASEC NAICS codes are classified, and the resulting crosswalk is used to restrict ASEC in column (4).

Table A1: Summary Statistics: Hourly Wages, Weekly Earnings, and Hours Worked

	(1) ASEC Full	(2) HB Full	(3) ASEC HB States	(4) ASEC Sample	(5) HB Sample	(6) QWI Sample
Hourly wage	18.69	12.37	18.46	16.68	18.35	
	(10.84)	(4.858)	(10.66)	(9.101)	(8.146)	
Weekly earnings	1016.7	381.8	999.6	631.8	660.0	805.5
	(724.4)	(245.5)	(716.1)	(432.3)	(345.1)	(328.4)
Hours usually worked per week at all jobs	39.25		39.32	35.78		
	(11.30)		(11.11)	(11.00)		
Hours usually worked per week at main job	38.55	30.03	38.66	35.13	36.49	
	(10.84)	(13.26)	(10.69)	(10.60)	(12.78)	
Hours worked last week	38.45		38.49	34.84		
	(12.80)		(12.63)	(12.06)		

Note: Mean coefficients and standard errors are in parentheses. ASEC and HB Full data include 2019 and 2020. QWI data is from 2019 only. Column (3) ASEC is restricted to the 21 HB states. Column (4) ASEC sample is restricted to hourly workers, who are not self-employed, working in small businesses ($<$ 25 employees) in a HB state and industry. Column (5) HB sample is restricted to individuals eligible for full UI benefits (defined as meeting state-specific earnings requirements in previous quarters) with a balanced number of week spells before and after the onset of the Covid-19 pandemic. Column (6) QWI sample is restricted to privately owned small firms ($<$ 20 employees) in HB states. Weekly earnings are calculated from the average monthly earnings (divided by 4.345) of the beginning-of-quarter employment. Source: Homebase, ASEC, QWI.

https://usa.ipums.org/usa/volii/indtoindnaics18.shtml and "ATTACHMENT 9: INDUSTRY CLASSIFICATION: Industry Classification Codes for Detailed Industry (4 digit) (Starting January 2020)" https://www2.census.gov/programs-surveys/cps/methodology/Industry%20Codes.pdf.

Table A2: Distribution of Observations by NAICS 2

	(1) ASEC Full %	(2) HB Full %	(3) ASEC HB States %	(4) ASEC Sample %	(5) HB Sample %	(6) QWI Sample %
11 Agriculture	1.52	0.33	1.60	2.29	0.30	2.20
21 Mining	0.50	0.00	0.55	0.11		0.32
22 Utilities	0.84	0.00	0.86	0.41	0.00	0.11
23 Construction	7.06	1.45	7.40	15.41	4.32	11.05
31–33 Manufacturing	9.82	0.72	9.67	2.86	1.40	4.88
42 Wholesale Trade	2.16	0.00	2.21	0.60		4.58
44–45 Retail Trade	10.43	13.53	10.66	10.74	16.06	10.20
48–49 Transportation	4.82	1.04	4.92	4.90	1.87	2.60
51 Information	1.82	0.43	1.78	0.80	0.22	1.16
52 Finance & Insurance	4.73	0.18	4.34	1.38	0.34	2.95
53 Real Estate	2.04	0.31	2.15	1.47	1.15	3.17
54 Professional Services	8.00	2.20	8.04	5.98	4.68	10.59
55 Management	0.09	1.34	0.10		3.14	0.26
56 Admin. & Support	4.31	1.02	4.61	6.93	2.79	5.50
61 Education. Services	9.25	1.50	8.78	4.16	1.80	1.54
62 Health Care	13.60	4.75	13.09	14.45	13.10	16.18
71 Arts, Entertainment	2.30	3.78	2.27	3.15	3.58	2.08
72 Accommodation & Food	7.29	62.20	7.50	16.13	37.90	11.74
81 Other Services	4.78	5.14	4.88	7.10	7.13	8.90
92 Public Administration	4.64	0.04	4.56	1.13	0.21	0.00

Note: ASEC and HB Full data include 2019 and 2020. QWI data is from 2019 only. Column (3) ASEC is restricted to the 21 HB states. Column (4) ASEC sample is restricted to hourly workers, who are not self-employed, working in small businesses (< 25 employees) in a HB state and industry. Column (5) HB sample is restricted to individuals eligible for full UI benefits (defined as meeting state-specific earnings requirements in previous quarters) with a balanced number of week spells before and after the onset of Covid-19 pandemic. Column (6) QWI sample is restricted to privately owned small firms (< 20 employees) in HB states, and only beginning-of-quarter employment.

C Institutional Details

C.1 FPUC

Federal Pandemic Unemployment Compensation (FPUC), the weekly $600 supplement to unemployment benefits, was introduced by the CARES act enacted on March 27, 2020, and ended on July 31, 2020.[2] No FPUC benefits were payable between July 31,

[2] U.S. Department of Labor news release dated April 4, 2020.

2020, and December 26, 2020. FPUC was re-established by the Continued Assistance Act as a $300 per week supplement to unemployment benefits from December 26, 2020, to March 14, 2021.[3] American Rescue Plan Act extended FPUC through September 6, 2021.[4] Any individual eligible to receive at least $1 of state unemployment benefits was also eligible to receive federally-funded FPUC for that week. Individuals who were working part-time and fulfilled state eligibility requirements for partial UI benefits were also eligible to receive FPUC payments.[5]

During the gap in FPUC payments, from August 1, 2020, Lost Wages Assistance (LWA) program was funded through Federal Emergency Management Agency (FEMA). States had the option of choosing between two weekly benefit amounts, $300 or $400, with different cost-sharing requirements.[6]

FPUC and LWA together supplemented weekly unemployment benefits in the following periods depending on eligibility: $600 (FPUC) from March 28, 2020, through July 31, 2020; $300 (LWA) or $400 from August 1, 2020, through the week ending September 5, 2020 (week ending August 22, 2020, in Florida); gap between September 5, 2020 and December 26, 2020; and $300 (FPUC) from December 26, 2020 through September 6, 2021, with some states ending the program early.[7]

C.2 Eligibility for Partial Unemployment Insurance

The $600 FPUC benefit was received by all workers on Unemployment Insurance or Partial Unemployment Insurance (i.e. who reduced their hours worked or are working a

[3]U.S. Department of Labor news releases dated December 30, 2020, and January 5, 2021.

[4]U.S. Department of Labor news release dated March 16, 2021.

[5]Attachment to Unemployment Insurance Program Letter No.15–20, Change 1, U.S. Department of Labor, dated May 9, 2020.

[6]U.S. Department of Labor news release dated August 12, 2020, Lost Wages Supplemental Payment Assistance Guidelines.

[7]Unemployment Insurance Program Letter No. 14–21, U.S. Department of Labor, dated March 15, 2021.

limited amount of hours while on unemployment), hence by every worker with earnings below the threshold determining the access to Partial Unemployment Insurance. Between March and July of 2020, individuals crossing this earnings threshold exhausted all remaining UI benefits and forwent the $600 FPUC benefit. Crucially for identification, this threshold differs across the 21 US states. Table A3 shows the threshold for each state in column 5, as we calculated it based on State-specific UI eligibility rules reported by the Department of Labor (DOL) for the year 2020 in their document titled "The Comparison of State Unemployment Insurance Laws".[8]. In most states, an individual is considered partially unemployed in some week if working less than full-time with earnings less than the weekly benefit amount or less than a percentage of, or less than a multiplier of the weekly benefit amount. Since we do not observe the actual UI benefits in our sample, we focus our analysis on workers who have an earnings history that makes them eligible for maximum UI benefits. In columns 1 and 2 of Table A3 we thus report, from Table 3-5 of the DOL document, the maximum UI weekly benefit amount (WBA) allowed in each state. In three states, the maximum WBA is slightly higher for individuals with dependence. For these three states, we consider the higher maximum WBA as a reference for our calculations. In columns 3 and 4 we report, from Table 3-8 of the DOL document, the maximum amount of labor market earnings allowed to retain eligibility for partial UI benefits and the earnings to be disregarded when this maximum amount is calculated.[9]. In column 6 of Table A3 we report how we have processed the information provided by the DOL to calculate the Partial UI thresholds of column 5. In Michigan, Washington and Wisconsin, eligibility for partial UI benefits is also conditional on workers reducing hours below a certain amount of hours per week. Considering the sample of workers under analysis, in most cases reducing earnings to an amount below the earnings threshold also corresponds to meeting

[8]Available at https://oui.doleta.gov/unemploy/comparison/2020-2029/comparison2020.asp

[9]All states disregard some earnings as an incentive to take part-time or short-term work.

the hour condition. For instance, consider the case of the 32 maximum weekly hour threshold for Wisconsin: given an average hourly wage of $18 and the $500 earnings threshold, a worker would work 27 hours a week, well below the 32-hour condition. Therefore, at the cost of a potential small amount of measurement error, we focus only on the earnings threshold to determine FPUC eligibility also for these three states. During the first month of Covid-19 emergency, Georgia has temporarily increased the earnings amount disregarded for the calculation of the Partial UI threshold. We have considered this temporary change relative to the DOL document in our calculation.

Table A3: State-specific eligibility rules for access to partial UI benefits

		(1)	(2)	(3)	(4)	(5)
State	**Max WBA ($)**	**Max WBA with dependence ($)**	**Definition of Partial UI. Earnings less than:**	**Earnings Disregarded**	**Threshold ($)**	**Calculation**
Alabama	275		WBA	⅓ WBA	367	Max WBA + ⅓*Max WBA
Arizona	240		WBA	$30	270	Max WBA + Earnings Disregarded
California	450		WBA	Greater of $25 or ¼ of wages	600	Max WBA/0.75
Colorado	561	618	WBA	¼ WBA	773	1.25*Max WBA with dependence
Delaware	400		WBA + greater of $10 or ½ WBA	Greater of $10 or ½ WBA	800	Max WBA + 2*0.5*Max WBA
Florida	275		WBA	8 x Federal hourly minimum wage	333	Max WBA + 8*7.25
Georgia	365		WBA	$50	715	Max WBA + Earnings Disregarded + $300
Louisiana	221	284	WBA	Lesser of ½ WBA or $50	334	1.5* Max WBA with dependence
Michigan	362		1.6 x WBA	For each $1 earned, WBA is reduced by 50 cents (benefits and earnings cannot exceed 1.6 WBA)	434	0.6*Max WBA/0.5

State	Max WBA ($)	Max WBA with dependence ($)	Definition of Partial UI. Earnings less than:	Earnings Disregarded	Threshold ($)	Calculation
Mississippi	235		WBA	$40	275	Max WBA + Earnings Disregarded
Missouri	320		WBA + $20 or 20%WBA, whichever is greater	$20 or 20% WBA, whichever is greater	448	Max WBA + 0.2*Max WBA
North Carolina	350			20% WBA		
Oregon	648		WBA	Greater of $120 or ⅓ WBA	864	Max WBA + Max WBA/3
Pennsylvania	561	569	WBA + 40% WBA	Greater of $21 or 30% WBA	967	1.4*Max WBA with dependence + 0.3*Max WBA with dependence
South Carolina	326		WBA	¼ WBA	408	1.25*Max WBA
Tennessee	275		WBA	Greater of $50 or ¼ WBA	344	Max WBA + Max WBA/4
Texas	521		WBA + greater of $5 or ¼ WBA	Greater of $5 or ¼ WBA	782	Max WBA + 2*Max WBA/4
Virginia	378		WBA	$50	428	Max WBA + $50
Washington	790		1.33 WBA + $5	¼ wages over $5	1.414	(1.33*Max WBA + $10)/0.75
West Virginia	424		WBA + $61	$60	545	Max WBA + $60 + $61

State	Max WBA ($)	Max WBA with dependence ($)	Definition of Partial UI. Earnings less than:	Earnings Disregarded	Threshold ($)	Calculation
Wisconsin	370		500	$30 plus 33% of wages in excess of $30	500	No benefits are payable if weekly earnings exceed $500.

D Model Extensions

D.1 WTP Approximation and Bounds

Here we show that the approximation in (8c) holds exactly in the case of kinks and provides a tight lower bound for notches.

First, consider the case of a regression kink design, where the marginal tax rate increases by Δt at m^* and will show that equation (8c) holds. Recall that the definition of labor supply elasticity is $e = \frac{m^o - m^*}{m^*} / \frac{\Delta \tilde{t}}{1-\tilde{t}}$, where $\tilde{t}$ is the implicit tax rate $t + \theta W$. We derive an expression for W by evaluating the elasticity in two risk scenarios with $\theta_L = 0$,θ_H. Assuming that risks are smooth at the threshold, we can use the ratio of the two elasticity expressions to obtain:

$$1 = \frac{m_L^o - m^*}{m_H^o - m^*} \frac{1-t}{1-t-\theta_H W} \tag{11}$$

Next, we can prove the claim by re-arranging this expression and using the definition of $WTP(r) = \frac{rW}{m^*(1-t)}$:

$$WTP(r) = \frac{m_H^o - m_L^o}{m_H^o - m^*} \tag{12}$$

Next consider the case of notches. Here the approximation in equation (8c) provides a lower bound estimate of the true WTP. To see this, recall that $WTP(r) = 1 - \frac{\frac{m_L^o}{m^*}\gamma_L - 1}{\frac{m_H^o}{m^*}\gamma_H - 1}$. The approximation result sets $\gamma_H = \gamma_L = 1$. The difference between such an approximation and the true WTP can be approximated by:

$$\begin{aligned}
\Delta_{approx} WTP(r) &= \frac{\partial WTP(r)}{\partial \gamma_L} d\gamma_L + \frac{\partial WTP(r)}{\partial \gamma_H} d\gamma_H && (13) \\
&= -\frac{\frac{m_L^o}{m^*}}{\frac{m_H^o}{m^*}\gamma_H - 1} d\gamma_L + \frac{m_H^o}{m^*} \frac{\frac{m_L^o}{m^*}\gamma_L - 1}{(\frac{m_H^o}{m^*}\gamma_H - 1)^2 d} d\gamma_H && (14) \\
&= \frac{-\frac{m_L^o}{m^*}(\frac{m_H^o}{m^*}\gamma_H - 1) d\gamma_L + \frac{m_H^o}{m^*}(\frac{m_L^o}{m^*}\gamma_L - 1) d\gamma_H}{(\frac{m_H^o}{m^*}\gamma_H - 1)^2} && (15)
\end{aligned}$$

In order to show that the approximation is a lower bound, we want to sign this expression and show that it is negative. First note that the denominator is positive and we can therefore focus on the sign of the numerator to sign the overall expression. We will take the check-and-verify approach:

$$-\frac{m_L^o}{m^*}(\frac{m_H^o}{m^*}\gamma_H - 1)d\gamma_L + \frac{m_H^o}{m^*}(\frac{m_L^o}{m^*}\gamma_L - 1)d\gamma_H < 0 \tag{16}$$

and re-arranging:

$$\frac{m_L^o - m_H^o}{m^*}d\gamma_L + \frac{m_H^o}{m^*}(\frac{m_L^o}{m^*}\gamma_L - 1)(d\gamma_H - d\gamma_L) < 0 \tag{17}$$

Consider the two terms separately. The first term has two components. $m_L^o < m_H^o$ implies that the $\frac{m_L^o - m_H^o}{m^*}$ is negative. Moreover, we can show that $d\gamma_L$ is positive. The approximation sets $\gamma_L = 1$, and hence $d\gamma_L = 1 - \gamma_L$. Using the fact that $\gamma_L < 1$ proves that $d\gamma_L > 0$. The first term is therefore negative.

The second term has three components. The first two components are both positive: $\frac{m_H^o}{m^*} > 0$ because $m > 0$ and $(\frac{m_L^o}{m^*}\gamma_L - 1) > 0$ because $(\frac{m_L^o}{m^*}\gamma_L - 1) = \frac{B(1-t)}{m^*(1-t-\theta W)} \geq 0$. The sign of the final term therefore depends on the final component: $(d\gamma_H - d\gamma_L)$. Using $d\gamma_L = 1 - \gamma_L$ and $d\gamma_H = 1 - \gamma_H$ we can write this term as:

$$d\gamma_H - d\gamma_L = \gamma_L - \gamma_H = \frac{1}{1+e}\left[(\frac{m^*}{m_H^o})^{\frac{1+e}{e}} - (\frac{m^*}{m_L^o})^{\frac{1+e}{e}}\right] < 0 \tag{18}$$

where the last equality uses the definition of γ. We can sign this expression because $m_L^o < m_H^o$ and $e > 0$. Combining this result with the first term means that both terms in (17) are negative and hence:

$$\Delta_{approx}WTP(r) < 0 \tag{19}$$

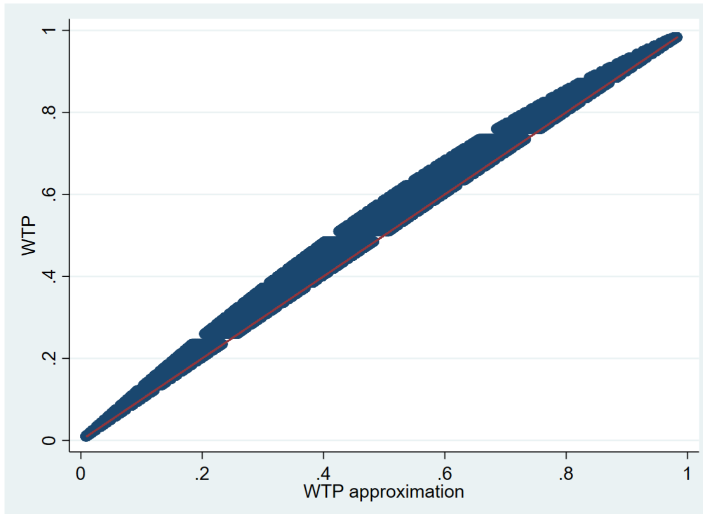
This shows that the approximation is always smaller than the true WTP and hence that the approximation is a lower bound for the WTP.

We now assess how tight this bound is and perform a simulation to compare the approximation to the true WTP for plausible parameter values.[10] In our FPUC case B is \$600 and we run simulations varying the tax rate t between 0 and 0.9, θW between 0 and 0.91, spanning the full range of plausible values. We let m^* vary from \$200 to \$1,000, covering the eligibility thresholds in our sample states. For the labor supply elasticity, there exists a range of estimates from a very large literature on this parameter. The meta-analysis by Chetty (2012) concludes that a plausible estimate is around 0.25. Since there is substantial disagreement about this parameter, we use a wide range between 0.02 and 0.92, which includes most estimates.

Figure A3 shows the results and compares the WTP approximation to the true WTP value. The dots are close to the 45-degree line, meaning that the approximation performs extremely well. As we proved above, the approximation provides a conservative, lower bound, estimate of the WTP and the approximation values are smaller or equal to the structural WTP. In addition, the results show that the lower bound estimate is always fairly close to the true WTP and the approximation thus provides a tight bound. The maximum bias occurs at a WTP of 50 percent of earnings (a fairly high WTP). In this case, the worst approximation estimates WTP to be around 41 percent, even this worst-case scenario thus still provides a very reasonable approximation. On average the bias is 3 percentage points and thus substantially smaller. Using the approximation therefore comes at relatively little cost, but has the major advantage that it allows the researcher to be agnostic about the size of the labor supply elasticity.

[10] Note that equation (15) provides a closed form solution for the bias and could be used to assess the magnitude of the approximation bias. But the equation is hard to interpret and we therefore perform a simulation.

Figure A3: Performance of WTP Approximation



Note: The figure shows the results of the WTP approximation from equation (8c) in simulated data and plots the approximation and the true structural parameter. The 45-degree line presents the line of perfect fit. The fact that the true values lie above the 45-degree line reflects the fact that (8c) provides a lower bound estimate. The simulation uses the following parameter ranges $m^* \in [200, 1000]$,$t \in [0, 0.9]$,$\theta W \in [0, 0.91]$, $e \in [0, 0.92]$.

D.2 Adjustment Frictions

A sizable literature discusses how adjustment frictions affect responses to budget discontinuities and proposes solutions to deal with such frictions (Chetty et al., 2011; Chetty, 2012; Kleven and Waseem, 2013; Einav, Finkelstein, and Schrimpf, 2017). In principle, any of these solutions could be applied to our setting. However, this is not required because our approach can handle frictions in a less parametric way and does not require correction methods that could be sensitive to assumptions (c.f., Einav, Finkelstein, and Schrimpf 2017).

First, consider the canonical friction case, where only a fraction α of workers can adjust their labor supply. This will reduce the excess mass (η) at the threshold relative to the frictionless benchmark, and η becomes: $\eta = \int_{m^*}^{m^o} d_0 = \alpha(m^o - m^*)d_0$. Now η depends on α and $(m^o - m^*)$, and multiple combinations of α and $(m^o - m^*)$ are consistent with the observed η. Note, however, that the impact of α cancels out in WTP estimates. We can re-write WTP in (7) as the ratio of excess mass in high (η_H)

and low (η_L) risk settings:

$$WTP \approx 1 - \frac{\eta_L}{\eta_H} = 1 - \frac{\alpha(\tilde{m}^o - m^*)d_0}{\alpha(m^o - m^*)d_0} = 1 - \frac{(\tilde{m}^o - m^*)}{(m^o - m^*)} \tag{20}$$

Thus, α affects both the numerator and the denominator proportionally and cancels out. The WTP estimate is thus unaffected by the presence of standard adjustment frictions.

More complex adjustment frictions arise from indivisible shifts, in which workers can add or drop entire shifts but cannot adjust their labor supply by the minute, or when workers negotiate hours with their employer and can only choose from a limited number of shift options. Both of these scenarios are isomorphic in the model and create two distortions that affect the excess mass at the eligibility threshold. First, workers are unable to adjust their labor supply exactly to the threshold earnings m^*, and instead have to reduce their earnings more to become eligible for $\mathcal{B}$. Second, some workers may be deterred from responding to the threshold because the indivisibility friction would force them to take a large earnings cut. Workers are thus less responsive to the threshold than in the frictionless benchmark.

Addressing the first challenge is relatively straightforward. The excess mass, η, now spreads over a wider earnings range. While it may be empirically more difficult to identify the spread out excess mass, such a spread-out mass does not pose any conceptual challenges to our approach.[11] In other words, the first challenge affects the estimation strategy but does not affect the link between the estimates and WTP. The second challenge can be addressed in a similar fashion as the canonical adjustment friction above. Denote the fraction of individuals who do not respond because of the indivisibility friction by $(1 - \alpha)$. If $(1 - \alpha)$ is constant, equation (20) applies again and implies that the WTP estimate is unaffected by this friction. Our framework thus identifies WTP, even if there are indivisibility constraints and hours decisions are not

[11] Canonical bunching methods focus on excess mass right at the threshold and would fail to fully capture more spread-out excess mass.

fully flexible.

D.3 Cobb-Douglas

Consider a case where utility is non-separable in health and cost of effort $U\left((T(m), \frac{m}{z}, a\right)$ $(= U\left(T(m), g(\frac{m}{z}, a)\right)$ and take the Cobb-Douglas case with $g(\frac{m}{z}, a) = m^{\alpha}h^{1-\alpha}$. The FOC becomes:

$$1 - t - \Delta t = (1-r)\alpha\left(\frac{a_0}{m}\right)^{(1-\alpha)} + r\alpha\left(\frac{a_1}{m}\right)^{(1-\alpha)} + \theta[m^{\alpha}a_1^{1-\alpha} -^{\alpha} a_0^{1-\alpha}] \qquad (21)$$

From $u(m^o, a_1) = u(m^o - W(m), a_0)$ we can derive an expression for a_1:

$$m^{\alpha}a_0^{1-\alpha} = W(m) + m^{\alpha}a_1^{1-\alpha}$$

Substituting this in equation (21) and simplifying yields:

$$1 - t - \Delta t - (1+\alpha)\theta W(m) = \alpha\left(\frac{a_0}{m}\right)^{(1-\alpha)}$$

Notice that the implicit tax imposed by the health risk increased by factor α relative to the separable case. This additional cost arises from the health effect on the marginal utility of leisure. A second change is that the marginal cost of a health shock increases the more a worker works ($m \uparrow$). And the value of health (W) now depends on the level of earnings m. This non-linearity in the cost of health shocks makes health risks operate like a non-linear progressive tax system, with increasing cost at higher m.

D.4 Income Effects

The canonical bunching approach uses quasi-linear utilities and thus assumes that there are no income effects. In many contexts where notches are small, the absence of income effects is plausible. Recent work, however, stresses that small notches may not be salient (Chetty, Friedman, and Saez, 2013). Moving to larger notches is thus attractive but leads to the added complication that such notches produce income effects. Structural estimates have previously used utility functions with income effects (Blundell, MaCurdy, and Meghir, 2007). Below we aim to cover a middle ground between the functional form flexibility of structural work and the quasi-experimental approach to identification of the bunching literature. We will show that introducing income effects implies that excess mass does not only appear at m^* but also at lower earnings levels.

Consider a general labor supply function that allows for income effects:

$$\tilde{m^o} = \tilde{z} + e\tilde{w} - \gamma\tilde{y} \tag{22}$$

$\tilde{x}$ indicates log values for x and w is the wage $\gamma\tilde{y}$ captures the income effect. When $\gamma = 0$ this equation collapses to the canonical quasi-linear utility case without income effects.

The introduction of a lump sum benefit $\mathcal{B}$ reduces labor supply if $\gamma < 0$. This effect changes the impact of the non-linear benefit schedule studied above. For a worker with earnings ε above the eligibility notch, introducing $\mathcal{B}$ reduces labor supply to $m^*+\varepsilon-\gamma\mathcal{B}$ which is below m^* if ε is small. The labor supply response thus creates excess mass below m^* and the excess mass at the notch point therefore does not fully capture the labor supply response. Hence, with income effects, excess mass (η) does not appear only at m^* but spreads out across a broader range of earnings. This creates additional identification challenges and we will return to the issue below.

The excess mass η is closely linked to the labor supply response of the marginal buncher. Individuals with pre-period earnings between m^* and the earnings of the marginal buncher $m^* + \Delta m$ make up the excess mass and η is thus given by:

$$\eta = \int_{m^*}^{m^*+\Delta m} d_0 dm$$

$$\Delta m = \eta / d_0 \tag{23}$$

where d_0 is the pre-notch earnings distribution between m^* and $m^* + \Delta m$, and to keep notation simple, we assume that d_0 is constant over this segment.[12]

To compute Δm we need to estimate d_0 and η. If data on the pre-notch distribution is available, we can compute d_0 directly from this data.[13] A second step is to estimate η, the extra mass generated by bunching individuals. η is the difference between the observed post-notch earnings distribution (d_1) and the distribution of non-bunchers (d_0'):

$$d_1 = \eta + d_0', \tag{24}$$

While we observe d_1, d_0' is not directly observed and needs to be estimated. Typically $d_0' \neq d_0$ and the pre-distribution does not provide a valid counterfactual. To see why, consider workers at m^* in the pre-period, they are below the eligibility threshold and thus part of the non-bunchers. Without income effects, their behavior would be unaffected by a lump sum benefit payment $\mathcal{B}$ and the pre-benefit distribution is a valid estimate for the frequency of this group. However, with income effects $\mathcal{B}$ reduces the labor supply of this group to $m^* - \gamma\mathcal{B}$ and no non-buncher is working at m^* after the introduction of $\mathcal{B}$. Now the pre-benefit distribution of earnings d_0 is a bad counterfactual for the distribution of non-buncher after the launch of $\mathcal{B}$ because $d_0'(m^*) = 0 \neq d_0(m^*)$. Using d_0 as counterfactual will bias the results, $d_0'(m^*) = 0$ implies that *all* individuals at $m = m^*$ are bunchers and the spike in density above neighboring cells ($\hat{\eta} = d_1(m^*) - \hat{d}_0(m^*) < \eta$) underestimates the true extend of bunching. Much of the debate about income effects focuses on the difference in compensated

[12]This assumption simplifies notation but is not required and richer baseline distributions can be included in the estimation.

[13]Without data on the pre-period, d_0 can still be estimated with "untreated" earnings ranges away from the notch point. This requires estimating d_0 in such untreated earnings ranges and then extrapolating to earnings levels in the treatment range. The researchers will need to make an assumption about which earnings ranges are untreated, and this requirement of an ad-hoc assumption has been controversial (Blomquist et al., 2021). The presence of income effects worsens the problem. Bunching is more spread out with income effects and less sharp at the threshold, making it harder to define untreated earnings bins.

and uncompensated labor supply elasticities. It is important to note that the impact is more severe in the context of bunching estimates. Here, income effects not only affect the interpretation of the elasticity as (un)compensated but additionally bias the labor supply response estimate itself.

Valid estimates can be obtained with a difference in difference analysis. A first advantage of the difference-in-differences approach is that it can detect any deviations from the pre-notch distribution, not just spikes in one specific location. As we saw above, this is important with income effects. Additionally, the difference in difference approach can overcome the identification challenge created by $d_0' \neq d_0$. When leisure is a normal good ($\gamma < 0$), the introduction of benefits reduces labor supply among the non-bunchers. Note, that while the local distribution of m is changed, the total mass of non-bunchers below m^* is unaffected by the notch:

$$\int_0^{m^*} d_0' = \int_0^{m^*} d_0 \equiv \pi$$

Using this result in (24), we can show that the notch generates total excess mass:

$$\int_0^{m^*} \eta = \int_0^{m^*} d_1 - \int_0^{m^*} d_0$$

which is the difference in the total density below the notch before and after the notch reform. $\int_0^{m^*} \eta$ can be estimated in a difference in difference regression that compares the density below m^* before and after the introduction of the notch. In difference in differences notation:

$$Pr(I = m)_{t,m} = \phi \cdot \mathbb{1}[t > t^*] + \pi \cdot \mathbb{1}[m < m^*] + \bar{\eta} \cdot \mathbb{1}[t > t^*] \cdot \mathbb{1}[m < m^*] + \varepsilon_{t,m}$$

where t^* is the time of the reform, π is captured by the coefficient on the dummy $\mathbb{1}[m < m^*]$. The coefficient $\bar{\eta}$ captures the average rise in density below m^*. Substituting this estimate into (23) yields the labor supply response of interest Δm.

The setting also yields an identification check in the spirit of a parallel trend check. This test is based on the distribution of the excess mass relative to the notch point. If the notch generates the excess mass, the excess mass should peak near the notch and decline as we move away from the notch. To test this, we estimate a specification similar to a dynamic DiD, and let the η coefficient vary across earnings ranges:

$$Pr(I = m)_{t,m} = \phi \cdot \mathbb{1}[t > t^*] + \pi \cdot \mathbb{1}[m < m^*] + \eta_m \cdot \mathbb{1}[t > t^*] \cdot \mathbb{1}[m < m^*] + \varepsilon_{t,m}$$

Plotting η_m provides a visual check on the assumption that the notch generates excess mass. The excess mass should peak at m^*, and its mirror image, missing mass, should peak above m^*. Finally, for m further from m^*, the effects should diminish.

Similar "difference in bunching" approaches have been used in the literature (Brown, 2013; Best et al., 2015), typically as a check on the identification assumption of canonical bunching estimators. In the set-up above we explicitly leverage the additional degrees of freedom to broaden the applicability of bunching methods to preferences with income effects.

D.4.1 Compensated Elasticity

The observed uncompensated labor supply elasticity reflects both an income and a substitution effect. The canonical bunching approach assumes that the latter is zero and that compensated and uncompensated elasticities coincide. In the more general case, we need to know the income effect γ to quantify the compensated elasticity from observed uncompensated elasticities. With this additional unknown parameter we require one additional moment condition. This section will show that the dispersion of excess mass away from m^* can be used as an extra moment condition. Without income effects all excess mass would arise at m^*, while the excess mass is more spread out over larger earnings ranges the bigger the income effect.

To derive a solution for γ, we take advantage of the location of the bunching. Note that all bunchers below m^* are at an interior solution, we call them "interior bunchers". At the earnings level m^*, there are several individuals who are at a corner solution and one individual for whom m^* is an interior solution; call this person the marginal buncher from the left. Before the notch the earnings of this person were $d_0 = m^* + p$. And using those two labor supply decisions in (22), we can show that:

$$m^* + p - \tilde{z} - e\tilde{w} + \gamma\tilde{y} = m^* - \tilde{z} - e\tilde{w} + \gamma(\tilde{y} + \mathcal{B})$$

$$\gamma = p/\mathcal{B}$$

We can thus solve for γ by deriving p. Notice that everyone with $d_0 \leq m^* + p$ is an interior buncher and the total mass of interior bunchers is thus:

$$I = \int_{m^*}^{m^*+p} d_0$$

The excess mass below the notch point (I) thus pins down p, e.g. with d_0 constant $p = I/d_0$. And using p, we can solve for $\gamma = \frac{I}{d_0\mathcal{B}}$. If all excess mass arises at the notch point then $I = 0$ and consequently $\gamma = 0$ and the analysis collapses to the quasi-linear case. This approach can thus be used to check the validity of canonical bunching estimates. But more powerfully, it can be used to identify labor supply responses from large and salient notches in budget constraints.

E Robustness Checks

E.1 Raw Earnings Density before and after Covid

The figure below shows the distribution of weekly earnings around the budget notches before and after the start of Covid. Since the thresholds are at different income levels in different states, we stack the densities from all states and show the distribution relative to the threshold level denoted by 0. All states are given equal weight in this figure. We show the density for a window from $600 below to $1,200 above the threshold earnings. Negative x-axis values indicate earnings below the threshold and positive values earnings above the threshold. In line with the main results, we show results for the pre- and post-period (defined by C_t). Both densities integrate to 1, so we ignore any mass outside the window. The timing of Covid and the start of FPUC coincide almost perfectly, so we use those two terms interchangeably.

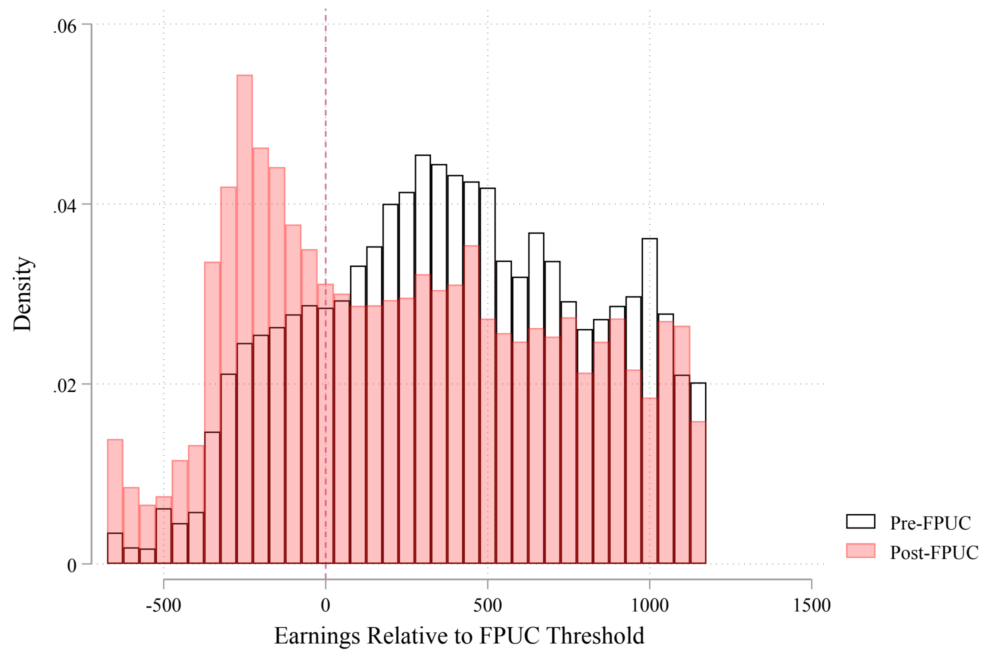
Figure A4 shows substantial excess mass below the earnings threshold in the post-FPUC period. There is no excess mass before FPUC was introduced. In line with the regression results, we see that excess mass is strongest near but below the earnings threshold and similarly the missing mass is most pronounced near but above the threshold level. The difference between the pre and post densities aligns closely with the regression coefficients shown in the main text.

E.2 Border Design

In this section we narrow our sample to counties along state boundaries, and thus with similar characteristics but facing different UI eligibility rules. The border counties are shown in Figure A5. These border communities generally have integrated labor markets and thus share many of the same demand shocks. In such a setting, empirical identification relies on comparing equally-paid workers across state borders with different incentives: such workers are likely to face similar demand shocks, however, one might be eligible for UI while the other might not, simply because of differences in the pre-determined exogenous eligibility thresholds. Our data is comprised of observations from 21 states, between which there are 24 state borders. In the border sample we exclude borders where we don't have data from border counties on both sides, which leaves us with the 17 unique state borders highlighted in Figure A5.

In the first step, we repeat the baseline analysis on the sample of border counties

Figure A4: Earnings Distribution Around FPUC Threshold



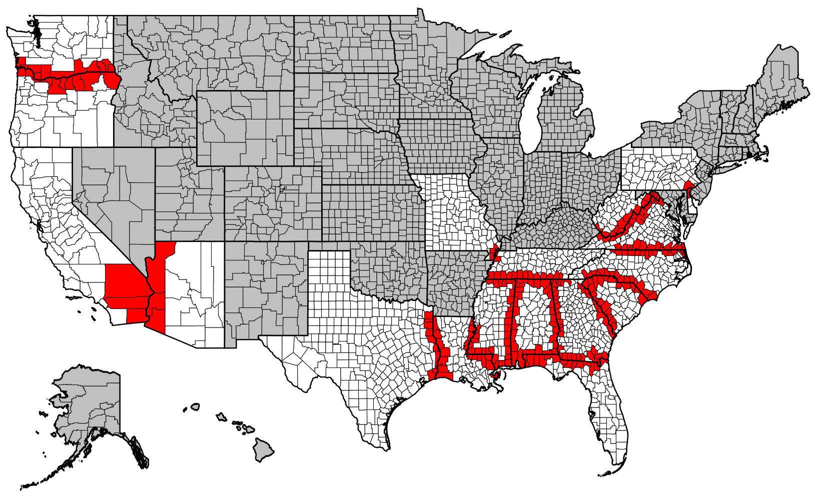
Note: The figure shows the earnings distribution before and after the beginning of FPUC. Pre-FPUC is the period defined by $C_t = 0$ and post-FPUC implies $C_t = 1$. The figure stacks distributions from the sample states for a window from $600 below to $1,200 above the FPUC eligibility threshold. To give each state threshold equal weight in the figure, we weight observations by the density at the absolute earnings level in the pre-FPUC period. Observations in the tails of the absolute distribution thus get a higher weight. The two densities integrate out to 1.

and find very similar effects to the baseline (Column 1 of Table A4). Next, we exploit the idea that neighboring counties experience similar demand shocks and allow all fixed effects to be specific to each border stretch. In practice, this implies that each border stretch is its own DiD experiment and we stack the 17 border DiDs into a single regression. The results are again close to our baseline estimates (Column 2).

E.3 Labor Supply Elasticity

We estimate labor supply elasticities at the industry level to use as a control variable in the main analysis. These estimates regress log hours worked on hourly wage changes at the worker level. A concern with the OLS regression is that wage changes are endogenous and capture promotions or changes in worker productivity. We therefore

Figure A5: Border Counties in Sample



Note: The figure shows counties along the state borders that are included in our border sample. There are in total 17 borders for which we have data from counties on both sides of the border.

Table A4: Excess Mass around UI Eligibility Threshold - Border Counties Sample

	(1)	(2)
Excess Mass (ptp)	0.965	0.914
	(0.146)	(0.146)
Interact income x time FE with		border IDs
Observations	20,596	20,596

Note: The Table reports results from equation (10). The border sample is restricted to counties at state borders shown in Figure A5. Source: Homebase.

use firm-wide wage updates as instrument. This, for example, captures that many firms update all worker wages at once at regular intervals. The firm-wide wage change is built by taking the average wage change among all workers at the firm , excluding the focal worker (leave one out). The measure has several advantages. An important feature is that it is based on hourly wages, rather than weekly earnings. The wage measure thus avoids a mechanical relation of wages and work hours. Second, our firm-wide wage change is unaffected by compositional changes since it takes the average of worker level wage changes. Employee turnover thus doesn't affect our measure of firm wage changes. The analysis is restricted to the pre-covid period.

The first stage is extremely strong with a first stage F-statistic of 10,000. This shows that wage changes among co-workers are a strong predictor of an update to my own wage. One concern might be seasonally patterns. If most firms change wages in January and January is a low hour month, the elasticity estimates would be biased downward. We find that controlling for month fixed effects doesn't change the results, suggesting that seasonality doesn't bias the results.

Finally, we estimate the 2SLS regression and allow for different elasticity coefficients by industry. Specifically, we estimate several 2SLS regressions, one each per two digit NAICS code. To preserve power, we aggregate ten industries with fewer than 1,000 observations into an "other" category. The elasticity of the median industry is 0.86.

E.4 Controls for demand shocks and school closures

In an additional robustness test, we add controls for demand shocks to our baseline specifications. Specifications that control for local employment, revenues of small businesses, business closures, school closures, combinations of these or all of these yield results close to the baseline estimate (Table A5).[14] This provides further evidence that the estimation strategy is not confounded by changes in the state of the local economy.

E.5 Alternative Measure of Covid-19 Exposure

In this section, we estimate the labor supply response to the increase in workplace risk using an alternative measure of Covid-19 exposure θ_{tci} and use the local Covid-19 fatality rate in the county, measured as in deaths per 100,000 people.

In Figure A6, the excess/missing mass in grey represents the behavioral response to FPUC in counties with zero recorded new deaths and the black area represents the

[14]Employment and Small Businesses daily data are obtained from Chetty et al. (2020a), while the share of in-class instruction is obtained from Parolin and Lee (2021a)

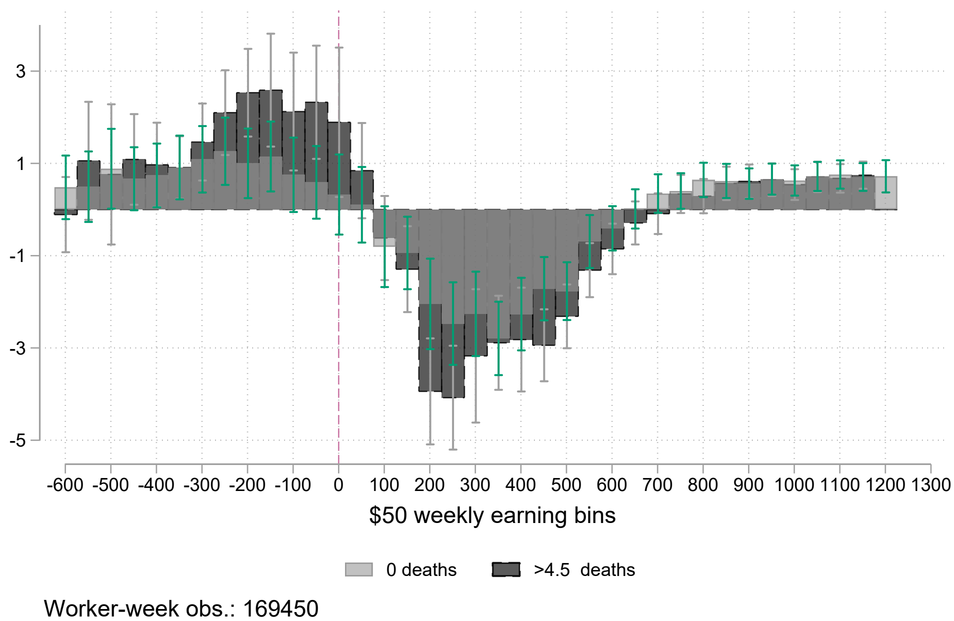
Table A5: Robustness to Labor Demand Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Additional Excess Mass			
Workplace Risk	0.260	0.254	0.256	0.255	0.254	0.261	0.254
(std. dev.)	(0.0527)	(0.0523)	(0.0524)	(0.0523)	(0.0523)	(0.0527)	(0.0524)
Controls		# Employees	Small Business Revenues	Change in # merchants	Revenues X Merchants	Share of in-class instruction	All

Note: Columns (2) through (7) supplement the main specification of Panel B of Table 2 (also presented in column (1)) by controlling for demand shock proxies, interacted with a dummy for the Covid-19 period and a continuous earnings variable. Column (2) controls for the number of active employees from Paychex, Intuit, Earnin and Kronos, varying at state-week-industry level. Column (3) controls for the percent change in net revenue for small businesses from Womply, varying at state-week-industry level. Column (4) controls for the percent change in the number of small businesses open from Womply, varying at state-week-industry level. Column (5) interacts the percent change in net revenue with the percent change in the number of small businesses from Womply. Employment, revenue and merchants data are downloaded from Opportunity Insights Economic Tracker. Column (6) controls for the share of in-class instruction from Parolin and Lee (2021), varying at county-month level. The share of in-class instruction is defined as the complement of the share of all schools in an area with at least 50% year-over-year decline in visitors, consistent with the Parolin and Lee definition. Column (7) controls for all demand shock proxies together. Sources: Chetty et al. (2020a); Chetty et al. (2020b); Parolin and Lee (2021a); Parolin and Lee (2021b).

"magnified response" in very high-risk settings (more than 4.5 weekly new deaths per 100,000 people). The excess mass in these high-risk settings is visibly larger, consistent with the results presented in Figure 6.

Figure A6: Excess and Missing Mass around the Partial UI Notch – Fatality Rate in County



Note: The figure shows $\eta_{k,\theta}$ coefficients from equation (10) for the highest and lowest levels of Covid-19 risk (θ_{tci}). Different from Figure 6 θ_{tci} is measured in deaths per 100,000 in the week in the local area. The gray bars represent the response in area-weeks with no new deaths and black bars in areas with more than 4.5 deaths per 100,000. The other details are the same as in Figure 6.

E.6 Sample Selection and Extensive Margin

In this section, we explore alternative samples and show that the results are robust to alternative choices. The baseline sample studies weeks with positive earnings among more attached workers who are employed by the same establishment before and after the onset of the pandemic. In the baseline sample, we additionally ensure that the earnings distributions before and after the onset of the pandemic are based on the same number of observations by using a window of 15 pre- and 15 post-weeks. For workers with a missing week, we keep an equivalent shorter symmetric window (e.g. a worker with 13 pre- and 15 post-weeks, we keep 13 weeks on both sides of the pandemic onset). This symmetry restriction guarantees that changes in excess and missing mass are not driven by exit effects but by workers moving up or down in the earnings distribution.

In this Appendix, we show WTP estimates for alternative samples, relaxing each of the restrictions above. Column (1) of Table A6 uses the baseline sample and replicates the baseline estimate from the main analysis. Column (2) relaxes the symmetry restriction and allows workers to have more work weeks before or after the pandemic. This leaves the total number of workers unchanged but extends the number of observations (i.e., worker-weeks). In Column 3, we add workers with less workforce attachment and include workers whom we observe exclusively before or after the onset of the pandemic. This doubles the number of workers and worker-week observations relative to our baseline specification. Panel A shows that the results remain very close to the baseline estimates.

We next consider extensive margin responses. For workers who leave the Homebase data, we cannot tell whether they stopped working or started a new job outside the Homebase sample. In the baseline analysis, we thus exclude exits and focus on the intensive margin of hour adjustment. In Panel B row 1 we add a zero-earnings week at the end of the work spell for workers who leave the data. The estimate is similar and in line with potential measurement error, slightly attenuated with a WTP of 23%.[15] In row 2, we take a more conservative approach and only include temporary exits in the analysis. It seems more likely that temporarily absent workers remained with the Homebase employer and the absence from the data are true zero earnings weeks. Using this in the analysis, we again find similar results to the baseline. In line with reduced measurement error, the effects are slightly larger than before at 26%. Finally, row 3 of Panel B combines the two extensive margin approaches and yields a WTP estimate of 24%. The results thus remain in the same ballpark for alternative sample choices.

Table A6: Robustness to sample selection and extensive margin

	(1) Baseline	(2) Asymmetric Sample Window	(3) Less Attached Workers
A - Intensive Margin	0.303	0.295	0.285
Worker weeks	169,450	228,591	315,566
Workers	9,063	9,063	21,418
B - Extensive Margin:			
1) zero earning for last week	0.226	0.216	0.226
Worker weeks	177,108	236,249	331,805
Workers	9,063	9,063	21,418
2) zero earnings for temporary absences	0.258	0.235	0.241
Worker weeks	182,350	241,749	333,648
Workers	9,063	9,063	21,418
1) + 2) zero for inner and last week	0.237	0.215	0.232
Worker weeks	186,183	245,324	345,003
Workers	9,063	9,063	21,418

[15] An additional reason why the results are smaller than the baseline is that at corner labor supply solutions, the WTP approach may yield a lower bound of the true WTP.

F Transmittable vs. Non-Transmittable Health Risks

This section analyzes the difference between the WTP for a transmittable illness (with externalities) and a non-transmittable one. Denote the utility weight of other household members by Ω, the number of other household members by n and the intra-household secondary fatality rate by s. The relation of WTP for a transmittable (WTP_T) and a non-transmittable disease (WTP_{nT}) is: $WTP_T = (1 + \Omega \cdot s \cdot n) \cdot WTP_{nT}$. For the back of the envelope calculation, note that the intra-household secondary fatality rate is $s = 0.002$ and assume that the worker cares as much about others' utility as her own ($\Omega = 1$).[16] For household size, consider a four-person household, i.e. a household at the 90th percentile of the US size distribution (three other household members: $n = 3$). In this case, $WTP_T = 1.006 \cdot WTP_{nT}$ and WTP_{nT} is thus only 0.6% smaller than our baseline estimate. In other words, our baseline estimate of 30.3% would be reduced to roughly 30.3%/1.006=30.1% of weekly earnings for a non-transmittable disease. Quantitatively, the concern for one's own health is thus the main component of the WTP estimate, with a quantitatively small additional contribution from pro-social concerns.[17]

G The Value of Enjoyable Jobs

A strength of the WTP approach is that it can be used widely for different types of amenities. To illustrate this, we present a second case study that estimates the value of enjoyable work. Enjoyment of work scores are widely collected in labor market surveys and provide information on the perceived quality of work. Yet, it is hard to interpret categorical enjoyment scores without a money metric for these scores. Work enjoyment

[16] s is obtained multiplying the 30% intra-household transmission rate (Lewis et al., 2020) with the 0.68% infection fatality rate, that is the fatality rate conditional to being infected (Meyerowitz-Katz and Merone, 2020)

[17] Pro-social concerns will play a more important role for diseases with more aggressive transmission rates and play a minor role in this setting because s is small.

can be affected by various factors and captures an aggregate (net) value of amenities in a given job. This value of "good" and "bad" jobs has been central in labor economics (for an overview, see Lavetti (2023a)).

The empirical strategy analyzes bunching around the U.S. early retirement age threshold. Workers accumulate social security entitlements for each quarter worked and once they reach age 62 the marginal value of additional quarters changes, creating a kink in the lifetime budget constraint.[18] At age 62 individuals also become eligible to claim retirement benefits, potentially alleviating liquidity constraints. We restrict the sample to individuals with sufficient savings to delay retirement and exclude people with less than a year's income in savings to mitigate the impact of the liquidity channel. We then study how bunching at the 62 age threshold differs for workers in high and low-enjoyment jobs. An important limitation relative to the workplace safety application is that we lack panel data and now use cross-sectional data comparing individuals with different job enjoyment. To interpret this heterogeneity, we must ensure that both groups of workers would behave similarly if they held similarly enjoyable jobs.

The analysis uses data from the US Health and Retirement Survey (HRS) between 1992 to 2018. Figure A7 plots retirement rates per quarter and shows that there is substantial bunching at the age 62 threshold.[19] The figure plots retirement rates separately for workers in enjoyable and less enjoyable jobs.[20] During ordinary quarters, around 2% of satisfied workers retire, while this rate spikes to around 10% in the quarter they turn 62 (Panel A of Table A7 shows excess mass estimates). Using these

[18]Each additional month worked beyond 62 increases the retirement benefit by 0.4%-0.6% until individuals reach the full-benefit retirement age (65-66 depending on birth year). The rewards for working extra months beyond this age are 0. 5% - 0. 8%. It is important to note that only convex kinks generate bunching. While in principle, it is ambiguous whether the kink in the lifetime budget constraint at age 62 is convex (the answer depends on the replacement rate, life expectancy, and the discount rate), in our context, an overwhelming majority of individuals do face a convex kink, and we, therefore, treat the kink as convex.

[19]We restrict the sample people who were in the workforce before turning 60. The figure shows the share of this restricted sample retiring each quarter.

[20]Enjoyment is measured in the previous year. If data is missing (12% of cases) we use the next closest year we have data. High enjoyment are people who strongly agree with the statement, "I really enjoy going to work." Workers who disagree or strongly disagree are coded as not enjoying their work.

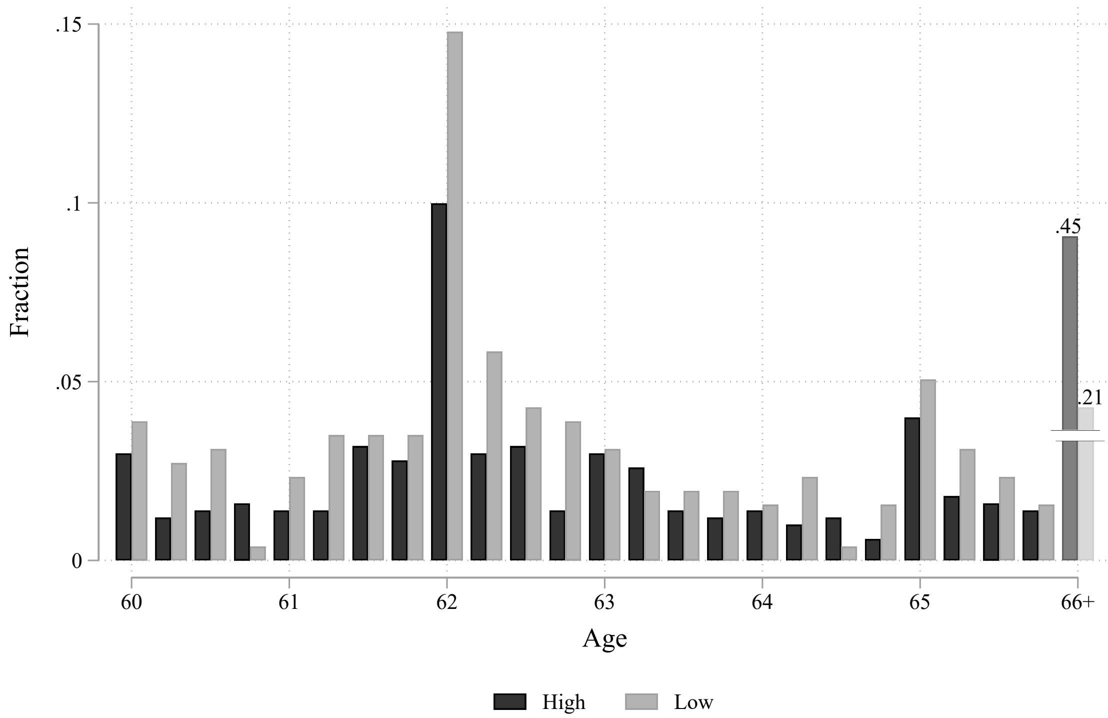
estimates in the traditional bunching framework implies that workers who enjoy their job retire (0.1-0.02)/0.02=4 quarters early because of the kink. The opportunity cost of reducing work time is smaller for individuals who enjoy their work less, and as predicted in our model, they indeed respond more to the kink. Their excess mass jumps to 15%, implying that workers in less enjoyable jobs retire (0.14-0.02)/0.02= 6 quarters earlier and reduce their retirement age by 2 quarters more than workers who *are* enjoying their jobs (see Panel B). Using these results in the WTP formula from above, we find that an enjoyable job is worth an extra (6-4)/4 /4 = 12.5% of annual income (see Panel C). On average, a worker would thus accept a 12.5% wage cut to move from a less enjoyable job to a more enjoyable one.[21]

We probe the importance of confounders that could give rise to spurious differences in retirement patterns. We interact all controls with age to allow for different retirement patterns across the demographic groups. The first strategy introduces industry fixed effects to absorb industry-wide retirement practices and exploits variation in enjoyment *within* a given industry. The results remain similar to the baseline (see Table A7, Column 2). Similarly, adding proxies for health has little impact on the results, suggesting that our threshold design is orthogonal to variation in health (column 3). Next, we add several proxies for human capital. Adding occupation-fixed effects again has little impact on the results (column 4). Proxies for education and location also have little impact on the results (columns 5 and 6).[22]

[21]We are not aware of a directly comparable estimate. Work that estimates the importance of non-wage amenities for inequality found that amenities explain between 15% and 26% of inequality (Lavetti (2023a); Taber and Vejlin (2020); Sorkin (2018)).

[22]A further potential concern is inflated bunching at the eligibility threshold from delayed retirement reporting. Individuals have little incentive to report a retirement age before the age of 62 to the Social Security Administration since it would not lead to additional benefits, potentially resulting in a spike in reports at age 62. Instead of admin data reports, we use survey reports on the age people stop working, which is less likely to suffer from target-date reporting problems.

Figure A7: Retirement Age by Work Enjoyment



Note: The figure shows the share of people retiring at any given age among people who had not retired by age 60, separately for those with high and low work enjoyment before retiring. The last bar shows the share of people who had not retired by age 66.

Online Appendix References

Best, Michael Carlos, Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn, and Mazhar Waseem. 2015. "Production versus Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan." *Journal of Political Economy* 123 (6):1311–1355.

Blundell, Richard, Thomas MaCurdy, and Costas Meghir. 2007. "Chapter 69 Labor Supply Models: Unobserved Heterogeneity, Nonparticipation and Dynamics." In *Handbook of Econometrics*, vol. 6 part A, edited by James J. Heckman and Edward E. Leamer. Elsevier, 4667–4775.

Table A7: Effect of Work Enjoyment on Retirement

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: High Work Enjoyment					
Excess Mass at Age 62	0.079	0.078	0.079	0.078	0.079	0.079
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
	Panel B: Low Work Enjoyment					
Excess Mass at Age 62	0.119	0.118	0.119	0.119	0.119	0.120
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
	Panel C: WTP for Enjoyable Work					
% of income	0.125	0.127	0.125	0.130	0.126	0.127
FE, interacted w/ age		industry	health	occupation	location	education
Observations	18192	18096	18192	18120	18179	18168

Note: The table shows how work enjoyment affects the likelihood of retiring at age 62 (first three months), which is when workers become eligible for social security benefits. The baseline specification (1) shows results from a pooled OLS regression on a balanced quarterly panel of 758 individuals aged 60-65. Only people who had not retired by 60 are included. We exclude people with less wealth than annual income before age 60. The dependent variable is an indicator for the age of retirement. Panel (A) shows the additional mass retiring at the threshold age among people with high enjoyment of work, and panel (B) among those with low enjoyment. Columns (2) to (6) show the results when controlling for industry, general health level, occupation, region-division of residence, and years of education. All specifications include a linear age control and a dummy for retiring at 66 or later.

Brown, Kristine M. 2013. "The Link between Pensions and Retirement Timing: Lessons from California Teachers." *Journal of Public Economics* 98:1–14.

Chetty, Raj, John N. Friedman, Nathaniel Hendren, Michael Stepner, and The Opportunity Insights Team. 2020a. "Data for: The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data." URL `https://tracktherecovery.org/`.

———. 2020b. "The Economic Impacts of COVID-19: Evidence from a New Public

Database Built Using Private Sector Data." NBER Working Paper 27431.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, and Michael Westberry. 2021. "Integrated Public Use Microdata Series, Current Population Survey: Version 9.0 [dataset]." URL `https://doi.org/10.18128/D030.V9.0`. Minneapolis, MN.

Lewis, Nathaniel M., Victoria T. Chu, Dongni Ye, Erin E. Conners, Radhika Gharpure, Rebecca L. Laws, Hannah E. Reses, Brandi D. Freeman, Mark Fajans, Elizabeth M. Rabold, Patrick Dawson, Sean Buono, Sherry Yin, Daniel Owusu, Ashutosh Wadhwa, Mary Pomeroy, Anna Yousaf, Eric Pevzner, Henry Njuguna, Katherine A. Battey, Cuc H. Tran, Victoria L. Fields, Phillip Salvatore, Michelle O'Hegarty, Jeni Vuong, Rebecca Chancey, Christopher Gregory, Michelle Banks, Jared R. Rispens, Elizabeth Dietrich, Perrine Marcenac, Almea M. Matanock, Lindsey Duca, Allison Binder, Garrett Fox, Sandra Lester, Lisa Mills, Susan I. Gerber, John Watson, Amy Schumacher, Lucia Pawloski, Natalie J. Thornburg, Aron J. Hall, Tair Kiphibane, Sarah Willardson, Kim Christensen, Lindsey Page, Sanjib Bhattacharyya, Trivikram Dasu, Ann Christiansen, Ian W. Pray, Ryan P. Westergaard, Angela C. Dunn, Jacqueline E. Tate, Scott A. Nabity, and Hannah L. Kirking. 2020. "Household Transmission of Severe Acute Respiratory Syndrome Coronavirus-2 in the United States." *Clinical Infectious Diseases* 73 (7):e1805–e1813.

Meyerowitz-Katz, Gideon and Lea Merone. 2020. "A Systematic Review and Meta-analysis of Published Research Data on COVID-19 Infection Fatality Rates." *International Journal of Infectious Diseases* 101:138–148.

Parolin, Zachary and Emma Lee. 2021a. "U.S. School Closure & Distance Learning Database." Data retrieved from OSF.

Parolin, Zachary and Emma K. Lee. 2021b. "Large Socio-economic, Geographic and

Demographic Disparities Exist in Exposure to School Closures." *Nature Human Behaviour* 5:522–528.

U.S. Census Bureau. 2021. "Quarterly Workforce Indicators (1990–2021) [computer file]." URL `https://ledextract.ces.census.gov`. Washington, DC: U.S. Census Bureau, Longitudinal-Employer Household Dynamics Program [distributor].