IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17600

Taxes Today, Benefits Tomorrow

Thomas Le Barbanchon

JANUARY 2025

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17600

Taxes Today, Benefits Tomorrow

Thomas Le Barbanchon
Bocconi University, CEPR, IGIER, IZA and J-PAL

JANUARY 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Taxes Today, Benefits Tomorrow*

This paper tests whether partially unemployed workers value future preserved benefits when they bunch at the kink of the unemployment insurance benefit-withdrawal schedule. I extend the bunching formula of Saez (2010) to a dynamic setting that accounts for the value of future benefits tied to taxation. This yields new tests of tax-benefit linkage based on bunching heterogeneity. I verify in quasi-experiments that UI extension programs that decrease the value of future benefits lead to more bunching and to lower labor supply. Last, a quantification exercise of the dynamic bunching formula provides extra support for a strong tax-benefit linkage.

JEL Classification: J65, H24, H31
Keywords: tax-benefit linkage, bunching, unemployment insurance

Corresponding author:
Thomas Le Barbanchon
Bocconi University
Via Roentgen 1
20136 Milano
Italy
E-mail: lebarbanchon@unibocconi.it

* A previous version of this paper circulated under the title: "Partial Unemployment Insurance" I thank J´erˆome Adda, Massimo Anelli, David Autor, Luc Behaghel, Pierre Cahuc, Raj Chetty, Amy Finkelstein, Giulia Giupponi, Marc Gurgand, Camille Landais, Etienne Lehmann, Ioana Marinescu, Nicola Pavoni, Thomas Piketty, Luigi Pistaferri, Roland Rathelot, Alexandra Roulet, Tom Schmitz and seminar participants at Bocconi, Bristol Univ., CEMFI, Chicago Fed, Copenhagen Univ., Cornell Univ., CREST, Ecole Polytechnique, EUI, IIES, McGill, PSE, QMUL, Sciences Po, UCL, UCSD, UdeM, UNIL, UZH and VATT for helpful remarks. I thank Camille Landais for sharing the data.

Social insurance contributions entitle workers to old-age pensions, and to health and unemployment insurance benefits. Social insurance contributions represent 9.6% of 2022 GDP in OECD countries, and this rate doubled over the past 50 years.[1] In many countries social insurance contributions are levied as payroll taxes on wages and represent more than a quarter of overall taxes. By introducing a wedge between workers' net wages and firms' labor cost, payroll taxes may distort labor market allocation and lead to inefficiently low levels of employment. However, the size of tax distortions depends on the workers' valuation of social insurance and the link they draw between their payroll taxes and the future benefits they become entitle to. If workers value social insurance and understand that benefit eligibility is tied to taxed employment, they may be willing to accept lower net wages internalizing that the overall compensation package includes social insurance. As in the classical example of mandated benefits (Summers, 1989), these conditions dampen distortionary effects of taxation. In this paper, I study whether workers draw the link between taxes paid today and social insurance benefits received later on.

This paper leverages the context of unemployment insurance in the US to advance the literature on several dimensions. First, on the theory side, I extend the bunching formula of Saez (2010) to a dynamic setting that accounts for the value of future benefits tied to taxation. This extended framework allows to design new empirical tests for tax-benefit linkage. Second, I implement my new tests in U.S. data and provide new empirical evidence. I show that UI claimants take into account the future value of benefits when supplying labor. In the most controlled test, I find that policy shocks that decrease the future value of benefits while holding taxation constant lead to higher bunching and to lower labor supply.

I analyze the tax-benefit linkage embedded in the partial unemployment insurance rules in the U.S. (McCall, 1996). Partial UI allows unemployment insurance claimants to receive unemployment benefits while they work in low-earnings jobs – usually part-time or temporary work. Weekly labor earnings below a certain threshold, termed the *disregard*, do not trigger any reduction in current benefits. However, for every dollar earned above the disregard level, current weekly benefits are reduced on a dollar-per-dollar basis. The benefit withdrawal schedule then implies a 100% marginal tax rate on earnings above the disregard level, drawing a large kink in the claimant budget set. Importantly, the reduction in current benefits leads to future entitlement, i.e. withdrawn benefits can be paid later in the claiming spell.[2] I study whether claimants bunch at the disregard level in reaction to the kink in the benefit withdrawal schedule, and whether bunching depends on the expected value of future entitlement.

I compute bunching estimates using UI administrative data from the U.S. Continuous Work and Benefit History (CWBH) project. I find substantial bunching at the kink level. In Idaho

[1]The OECD publishes aggregate statistics on social insurance at the following link: https://data.oecd.org/tax/social-security-contributions.htm

[2]In other words, intertemporal benefit transfers delay the potential benefit exhaustion date.

and Louisiana, the excess mass of workers at the disregard is five times the population density of workers that would earn this amount absent the kink. I also observe that a significant fraction of claimants have earnings above the disregard amount. This observation is consistent with claimants reacting to an *effective* marginal tax rate that is lower than the static 100% benefit-withdrawal rate above the kink.

To show how tax-benefit linkage affects claimants' behavior, I develop a job-search model of partially unemployed workers. In the model, job seekers work in low-earnings jobs while they search for permanent jobs that are ineligible for partial UI. They make their labor supply decisions in the low-earnings labor market based on an *effective dynamic* marginal tax rate that accounts for the present value of benefits preservation.

The dynamic marginal tax rate depends on the claimants' expected probability to find a permanent job and on the horizon over which preserved benefits are rolled over. First, if the claimant expects to rapidly find a permanent job and to exit the UI registers, then she is less likely to profit from the benefit-preservation mechanism and her dynamic marginal tax rate is larger, closer to the static benefit-withdrawal rate. Second, if the claimant is entitled to many weekly payments of unemployment benefits (long potential benefit duration), preserving benefits is less valuable. She is less likely to profit from a delay in the date when her benefit payments exhaust, as she would have found a permanent job before her claim ends. Thus, claimants with longer potential benefit duration have larger effective marginal tax rate.

I then show that excess bunching at the kink of the partial-UI schedule equals the product of the earnings elasticity to the net-of-tax rate and of the change in the effective dynamic marginal tax rate at the kink. This extends the bunching formula of Saez (2010) to contexts where taxes entitle workers to future benefits.

Consistent with the model-based implications of tax-benefit linkage on effective tax rates, I find in my data that claimants with longer potential benefit duration bunch more. This is confirmed when focusing on within-individual variation in potential benefit duration across claims, or when focusing on exogenous variations in potential benefit duration initiated by triggers of emergency UI extension programs. This provides quasi-experimental evidence on the existence of tax-benefit linkages, building on discontinuities in tax schedules à la Saez and on well-identified policy shocks.[3] I also verify in the cross-section that bunching estimates are larger for claimants with a low propensity to remain on the UI registers, such as claimants expecting to be recalled by their previous employer (Katz and Meyer, 1990b).

Lastly, I quantify whether the tax-benefit linkage mechanism may rationalize claimants' observed labor supply. I then assume that claimants have rational expectations about their

[3] Moffitt and Nicholson (1982); Farber et al. (2015) use triggers of emergency UI extension programs, either in the 80s or during the Great Recession, as a source of exogenous variation in the U.S. unemployment insurance generosity.

permanent job finding rate. I estimate a hazard model of exiting the UI registers that depends on workers' socio-demographics (age, gender, education), their claim characteristics (benefit level, potential benefit duration) and their recall expectations. This enables to predict the expected probability of exiting the UI registers for each claimant and to compute their model-based dynamic marginal tax rate. I find that the average effective tax rate amounts to 55%, significantly lower than the 100% static benefit-withdrawal rate. The bunching formula then identifies the earnings elasticity to the net-of-tax rate, whose estimate lies between 0.1 and 0.2. This estimate is in line with the consensus estimates in the literature (see the review of quasi-experimental estimates in Chetty (2012) or Chetty et al. (2011b)). Overall this confirms the explanatory power of perfect tax-benefit linkage under rational expectations.

I provide several discussions of the robustness of my extended bunching formula and of the empirical results. First, while I consider hand-to-mouth risk-neutral workers in the baseline model, I show how relaxing those assumptions affects the bunching formula and the corresponding empirical tests. Risk-averse workers value more transfers of benefits later in the unemployment spell. They seek to smooth consumption. Thus they would work more in low-earnings jobs and bunch less than risk-neutral workers. Quantifying this extra effect, we find that the earnings elasticity estimates would be 25% to 50% higher, but still in the same ballpark as consensus estimates. Second, I discuss how frictions in the low-earnings job market could be introduced in the baseline job search model. Building on the quantification of optimization friction in Gelber et al. (2020), I show that adjusting elasticity estimates to frictions yields robust conclusions of a strong tax-benefit linkage. Third, I discuss how biased beliefs on job finding rates do not affect my bunching-heterogeneity tests, but lead to biased elasticity estimates. I then adjust my elasticity estimates using results from Mueller et al. (2021) and I find that my conclusions are robust to biased beliefs.

My paper contributes to several strands of literature on the effects of taxation for social insurance. I provide new empirical evidence of tax-benefit linkage leveraging bunching estimates and policy shocks in the context of partial unemployment insurance. Close to my paper, Liebman et al. (2009) find that workers are more likely to retire when the marginal effect of labor supply on future social security benefits is low. Their identification uses non-linearities in the value of benefits only, while I combine both sharp discontinuities in the current taxation schedule and quasi-experimental variations in the value of future benefits.[4] More broadly, I contribute to the large literature that tests for tax-benefit linkage using tax incidence on wages (Gruber, 1997b; Anderson and Meyer, 1997a, 2000; Bozio et al., 2020). While this literature finds significant but incomplete pass-through, some of my empirical evidence suggest an almost perfect tax-benefit linkage which could be explained by the

[4] In the context of Liebman et al. (2009), uncertainty smooths discontinuities in the social security benefit rule, which prevents them from adopting a standard regression discontinuity approach. Discontinuities in the schedule of current taxation do not suffer from uncertainty smoothing, so that I can adopt a bunching approach.

high salience of both taxation and benefit entitlement in the partial UI context. Earnings in low-wage jobs are taxed through high-frequency weekly benefits withdrawals, and the horizon of future benefits entitlement is within the next months, which is shorter than the horizon of social security entitlements for example studied in Gruber (1997b). My results also suggest that entitlement effects matter when assessing the dynamic effects of unemployment insurance (Mortensen, 1977; Hamermesh, 1979; Kuhn and Riddell, 2010), of welfare programs with time limits (Grogger and Michalopoulos, 2003), and of social security programs (Coile and Gruber, 2001), especially their earning tests (Friedberg, 1998, 2000; Haider and Loughran, 2008; Gelber et al., 2020).[5]

I contribute to the bunching literature (see Saez (2010) and Chetty et al. (2013), and the review in Kleven (2016)). Kleven (2016) writes that "Extending the bunching approach to dynamic settings is still in its infancy." (p.13) The bunching literature on labor supply adopts a static framework, where workers do not link taxes and future benefits. This assumption is questionable when assessing the earnings elasticity to payroll taxation. I show how the bunching identification strategy and the Saez (2010) formula can be extended to account for these dynamic aspects. The closest papers in the bunching literature are Brown (2013) and Manoli and Weber (2016). They both study bunching at the legal retirement age taking into account that delaying retirement increases either severance payments or future annuities. They develop a static model of lifetime labor supply, and a non-stochastic dynamic model of annual labor supply respectively. The dynamic approach of my paper allows for *stochastic* events affecting benefit payments, which are essential features of social insurance linked to payroll taxation.[6]

My paper also contributes to the literature on the effect of unemployment insurance. I study partial unemployment insurance programs that are widespread in OECD countries. In 2017, 11% of UI claimants in OECD countries work while on claim.[7] In the U.S., McCall (1996), O'Leary (1997) and the early contributions of Holen and Horowitz (1974) and of Kiefer and Neumann (1979) document the behavioral response at the extensive margin, i.e. whether claimants take up low-earnings jobs when partial UI is more generous.[8] O'Leary (1997) also documents behavioral response at the intensive margin, the focus of my paper. I provide the first evidence on significant *intertemporal* response to partial UI rules at the

[5] Gelber et al. (2020) discuss intertemporal aspects of the U.S. Social Security Annual Earnings Test. Reductions in current benefits can lead to increases in future scheduled benefits (i.e. benefit enhancement mechanism). However, benefit enhancement is triggered only when a sufficient amount of current benefits is reduced. Thus there is no difference between the static benefit-reduction rate and the dynamic marginal tax rate at the first kink in the SSAET schedule. Consequently, bunching at the SSAET kink studied in Friedberg (1998, 2000) and Gelber et al. (2020), is not informative about tax-benefit linkage.

[6] le Maire and Schjerning (2013) also consider dynamic aspects in income tax schedule, but they specifically model income shifting by the self-employed.

[7] See OECD data for national shares of partial-UI claimants, such as 33% in Sweden, 22% in Finland and 6% in Portugal (OECD, 2020). See Kyyra (2010) for older figures.

[8] In Europe, Kyyra (2010), Caliendo et al. (2016), Kyyra et al. (2013), Fremigacci and Terracol (2013) and Godoy and Roed (2016) study the effects of partial-UI jobs on regular employment.

intensive margin.[9] Conditional on working in low-earnings jobs, I show how labor earnings react to changes in partial UI generosity.

The paper is organized as follows. In Section I, I describe the U.S. partial unemployment insurance program. In Section II, I introduce the data and I provide visual evidence of bunching patterns. In Section III, I develop a job-search model of claimants working while on claim. In Section IV, I test the tax-benefit linkage using bunching heterogeneity in quasi-experiments. In Section V, I show that perfect tax-benefit linkage quantitatively explains bunching patterns with reasonable earnings elasticity to the net-of-tax rate and under the rational expectation hypothesis. Section VI concludes. Data and Code for replication are available at Le Barbanchon (2024).

I Institutional background

In the U.S., unemployment insurance (UI) claimants who work while on claim, are eligible for *partial* unemployment benefits, provided that they do not earn more than a maximum amount of labor income per week. Partial-UI claimants must still meet the usual UI eligibility requirements (such as actively searching for jobs, see online Appendix A). Partial-UI claimants are allowed to work for any employer, including their past employers; claimants who are temporarily laid off are also eligible for partial UI.[10]

Partial-UI claimants are paid their usual weekly benefit amount (WBA) when their weekly earnings are below some state-specific *disregard* thresholds.[11] When partial-UI claimants earn between the disregard level and the maximum eligibility amount, their current benefits for that week are reduced by their earnings minus the disregard. The static marginal benefit-withdrawal rate is then 100%.

Figure 1 illustrates the partial-UI schedules for the four states and for the time period covered by my dataset: Idaho (ID), Louisiana (LA), New Mexico (NM) and Missouri (MO) in the late 70s and early 80s. I plot the weekly net income (earnings plus UB payments) against the weekly earnings while on claim.[12] I normalize earnings and UB payments by

[9]Following a previous version of my work, Lee et al. (2021) find significant fiscal externalities of the partial UI program when pooling together the intensive- and extensive-margin responses to a small increase in the benefit reduction rate (by 9 percentage point). Their estimates on labor earnings are statistically consistent with the intensive-margin elasticity estimated in this paper.

[10]Also, individuals whose hours have been reduced at their current workplace are eligible for partial UI, as long as they can file a claim based on this reduction in hours worked. Claimants with reduced hours represent only a small share of partial-UI claimants. In the CWBH data used in this paper, I can distinguish between claimants taking up new jobs and claimants with reduced hours (Short Time Compensation) in Louisiana only. From 1982 to 1984, only 15.7% of partial-UI weeks in Louisiana concern claimants with reduced hours.

[11]By definition, the weekly benefit amount (WBA) is the unemployment benefits (UB) payment when claimants do not work, i.e. total unemployment benefits.

[12]The plots and my analysis abstract from income taxes. I expect partially unemployed claimants' labor income to be lower than minimum taxable income thresholds. Similarly, only unemployment benefits over some minimal thresholds are subject to income taxes.

the WBA, as the maximal amount and the disregard are expressed as a fraction of the WBA for three of the four states (see summary in Appendix Table A1). The figure illustrates that the schedule is kinked at the disregard amount. Intuitively, one expects claimants to bunch at the disregard level. In addition, from a pure *static* point of view, there are no incentives to work for wages right above the disregard, as the net income is essentially a plateau above that level.

The graphics also illustrate the notches at the maximal eligibility amount in Louisiana, New Mexico and Missouri (see Munts (1970) for an early discussion on notches in the U.S. partial-UI schedule). Notches generate even stronger disincentives to work than kinks, as claimants lose income when they work above the notch threshold (Kleven and Waseem, 2013). Because of data limitations, I will not analyze the claimants' behavior around notches. The incentives to claim drop discontinuously at the notch value, so that some workers above the notch may leave the UI registers, and hence my data. There are no such data coverage concerns around the kink.[13]

I now turn to the *dynamic* aspects of the partial-UI rules. At the beginning of each claim, the UI administration computes the claimant's weekly benefit amount (WBA) and potential benefit duration (PBD), which both depend on past earnings. The product of the WBA and the PBD is called the (total) benefit entitlement which I denote B_0. The benefit entitlement can be thought of as a kind of UB capital that depreciates over time with UB payments. If claimants are totally unemployed all along their claim, they receive each week their WBA, and their benefits lapse after $PBD = B_0/WBA$ weeks. When claimants are only paid part of their WBA in a given week, the unpaid amount is rolled over to a later week in the claim and the UB capital depreciates at a slower pace. Let me take the example of a claimant entitled to a WBA of 300 dollars over a PBD of 10 weeks. Under total unemployment, the claimant receives 300 dollars every week from week 1 to week 10. If in week 5 the claimant takes up a low-earning job that reduces her UB payment by 150 dollars in that week, she is entitled to a 150 dollar UB payment in week 11.

The partial UI rules clearly draw a link between benefit withdrawal/taxes above the disregard level and future benefits. Intuitively, the benefit preservation mechanism provides extra incentives to earn weekly wages above the disregard level and should lower the amount of bunching at the disregard level. The theoretical model in Section III shows formally how the tax-benefit linkage affects bunching. In words, working while on claim, with earnings above the disregard level, is thus a way to delay the benefit exhaustion date.

In principle, there is one limitation to the possibility to delay exhaustion, as any remaining UB capital is lost one year after the first claiming week, defined as the benefit year. However, in the data, almost all claimants exhaust their benefits or find a regular job before the end

[13] Just above the disregard level, claimants still have strong incentives to remain on the UI rolls. If they leave, their total income drops.

of the benefit year. Consequently, I abstract in the remainder from any horizon effects of the benefit-year rule.

The partial-UI rules remain in place when additional UI programs are triggered because of tough labor market conditions. During the late 70s and early 80s, there were two additional programs in place - the Extended Benefit (EB) program (Tier II) and the Federal Supplemental Compensation (FSC) program (Tier IV) - which both increased the potential benefit duration of claimants (see more details in Appendix A). I leverage the quasi-experimental variation in PBD generated by these policy shocks to empirical test for the tax-benefit linkage in Section IV.

II Data and bunching patterns

I use individual panel data from the Continuous Wage and Benefit History project.[14] [15] The CWBH project collects weekly claims for a random subsample of UI claimants in the U.S., and the resulting dataset has the unique advantage of including the *weekly* earnings that claimants report to the UI administration and the consecutive UB payments.[16] I can thus characterize whether claimants are partially unemployed. The data cover four U.S. states – Idaho, Louisiana, Missouri and New Mexico – during the late 70s and early 80s.[17] The data include socio-demographics characteristics, pre-unemployment labor history, and all relevant information about the claim. In addition, the data set includes survey information about recall expectations for a subsample of claimants.

Table 1 reports descriptive statistics of partial UI claimants by state. I select unemployment spells where workers are partially unemployed for at least one week. This amounts to 32.1% of the initial sample of spells. The share of men varies between 50% to 70% across states. Claimants are in their early 30s, with around 11 years of education on average. Manufacturing is the most common industry of the pre-unemployment firms except in Louisiana where construction is as important. The weekly benefit amount (WBA) is around \$100 (current dollars) and the average replacement rate is between 40% and 50%. The potential benefit duration (PBD) is greater than 26 weeks (the maximal PBD in Tier 1) as the early 80s is a period of high unemployment, and of UI extensions. The average claiming duration is around four months.

[14] The CWBH data are analyzed in Moffitt (1985), Katz and Meyer (1990b), Katz and Meyer (1990a), Anderson and Meyer (1993), Anderson and Meyer (1997b) and Landais (2015a), among others. To the best of my knowledge, partial UI has never been analyzed in the CWBH data.

[15] The CWBH data can be downloaded from Landais (2015b)

[16] One concern is that claimants manipulate their earning reports to become eligible for partial UI. However the UI administration takes action to limit false statements. The UI administration performs random audits of claimants' declarations. The UI administration currently cross-checks W-2 and new hires declarations of employers with claimants reported earnings. If fraud is detected, it can be severely punished as a Class VI Felony. Criminal action may result in up to 2 years in prison and fines up to \$150,000 for each false statement.

[17] The CWBH project collected data for other states but weekly earnings were missing for these.

Bunching patterns Figure 2 displays the weekly earnings density reported by UI claimants together with the observed partial-UI schedule state by state. In line with the partial-UI rules in Idaho, Louisiana and New Mexico, I normalize the weekly earnings by the weekly benefit amount. The empirical schedules (dashed blue lines), which describe the actual total weekly income (unemployment benefits plus earnings) as a function of weekly earnings, closely follow the theoretical schedules displayed in Figure 1. The upper panels - Idaho and Louisiana - clearly display bunching at the level of the disregard (50% of the weekly benefit amount). In Louisiana (upper right-hand panel), there is also a sharp drop in the density at the weekly benefit amount, when claimants are no longer eligible for partial UI. This may be related to the notch in the schedule, but it can also be due to the fact that individuals have no incentives to stay registered above this "exit" level. In New Mexico, where the disregard level is only 20% of the weekly benefit amount, bunching is less striking (lower left-hand panel). The lower right-hand panel illustrates a placebo test. In Missouri, the level of disregard is $10, so that the schedule is totally flat when earnings amount to $0.5 \times WBA$. There is indeed no bunching at this placebo level. Thus the bunching observed in Idaho or Louisiana at $0.5 \times WBA$ is unlikely to be an artifact of other labor legislations or norms, or of hour constraints according to which claimants take some part-time jobs that provide roughly one fourth of their previous wages.

I quantify the extent of bunching along the lines of Chetty et al. (2011a). I fit a polynomial on the earnings density of partial-UI claimants, taking into account that there may be bunching in a bandwidth around the disregard, and that the bunching mass comes from the earnings distribution above the disregard. I report the details of the estimation procedure in Appendix B. Appendix Figure B1 also shows that the procedure fits well the earnings distribution in each state.[18] Table 2 reports the results of the bunching estimation for each state. I find that in Idaho and Louisiana, the mass bunched at the disregard level is around five times in excess to the mass that would have been at the disregard level in the absence of the kink. Excess bunching is highly statistically significant. In New Mexico, the excess bunching mass amounts to 1.2 times the counterfactual density at the disregard level, but is not statistically significant. In Missouri, there seems to be a missing mass of claimants at the placebo threshold level, confirming the visual placebo test above.

In Appendix E, I conduct a second placebo exercise. In Louisiana, the disregard threshold becomes lower for a subsample of high-WBA workers in 1983. I find that bunching closely follows the new kink location for the treated high-WBA workers, while it remains at its previous and unchanged location for the control low-WBA workers. This further demonstrates that the bunching pattern in the data is related to the incentives embedded in the partial UI rules.

Another important feature of the earnings distribution in Figure 2 is the substantial fraction

[18] The procedure fits a polynomial of degree 7. The bunching bandwidth is between 5 dollars below and 2 dollars above the kink threshold.

of claimants working for earnings above the disregard level. In this paper, I argue that this is partly driven by forward-looking claimants who have incentives to supply labor above the disregard level as they remain entitled for the withdrawn benefits later on.

III Theoretical model

In this section, I develop a model of job seekers working while on claim that incorporates the dynamic aspects of the partial-UI program. The objective is to derive the implications of tax-benefit linkage on claimants' labor supply. I show that claimants make their labor supply decision based on a dynamic marginal tax rate, which is lower than the static marginal benefit-reduction rate, because job seekers value the expected benefit transfers generated by their work while on claim. I derive a modified bunching formula à la Saez (2010) that links bunching, effective marginal tax rate, and earned income elasticity. The modified bunching formula allows me to test for tax-benefit linkage.

III.A Setup

I consider an infinitely lived individual i, claiming benefits from period 0 onwards. Following UI rules, periods are weeks in my model. Until she finds a *permanent* job, the job-seeker may work in a *low-earnings* job, corresponding to short-term or part-time work eligible for partial UI. In the remainder, I also refer to these low-earnings jobs as *low-wage* jobs. I assume that low-wage and permanent jobs are different types of jobs. The market for low-wage jobs is tight, and there are no search frictions. On the contrary, the market for permanent jobs features search frictions.

The job-seeker's earnings in the low-wage job in period t are denoted z_t. In line with Saez (2010) model, I do not make any distinction between wage rates and hours as those different components are not observed in the data.[19] The per period utility $u_i(c_t, z_t)$ of job-seeker i depends on consumption c_t and on labor earnings in low-wage jobs z_t - the latter dependence captures disutility of labor. The individual heterogeneity in preferences is smoothly distributed in the population, so that earnings z_t would also be smoothly distributed in the absence of any kinks in the benefit reduction schedule. This is the key assumption of the bunching identification strategy.[20] In the baseline model, I assume that the job-seeker is risk-neutral. In Appendix C.B, I provide a model extension with risk-averse workers. With risk-neutral workers, it is convenient to parametrize the period utility

[19] Alternatively, one can think of the wage rate as being fixed and that the job-seeker chooses the number of hours worked.

[20] When turning to estimation, we further assume a polynomial shape for the smooth distribution of individual heterogeneity. As discussed in Blomquist et al. (2017), this is an important assumption. We test the robustness of the main estimates when varying the degree of the polynomial.

function as follows:

$$u(c_t, z_t; n_i) = c_t - \frac{n_i}{1 + 1/e} \left(\frac{z_t}{n_i} \right)^{1+1/e} \tag{1}$$

where n_i is an individual ability or preference parameter - smoothly distributed in the population - and e is the parameter capturing the earnings elasticity to the net-of-tax rate. As discussed in Saez (2010), the identification argument also holds with more general utility function as long as the individual heterogeneity is smoothly distributed. Such a parametrized utility function is convenient, as the heterogeneity parameter n_i then equals the earnings level in the absence of any benefit-reduction (as derived below). n_i may capture both individual taste for work and ability, I choose to refer to n_i as individual ability in the remainder.

At each date $t > 0$, the job-seeker may find a permanent job with probability $(1 - p)$. Then she leaves the unemployment registers. Permanent jobs yield the expected intertemporal utility W, which is assumed to be greater than the continuation value of unemployment at any period. Therefore, claimants never decline permanent job offers. Consistent with the view that the markets for low-wage and permanent jobs are separated, I assume that the utility derived from permanent job W is not related to the individual ability in low-wage jobs. In the baseline model, I assume that the probability to find a permanent job does not depend on the amount of earnings in low-wage jobs. This follows the empirical literature that finds small effects of partial-UI jobs on permanent employment (McCall, 1996; Kyyra, 2010; Caliendo et al., 2016; Kyyra et al., 2013; Fremigacci and Terracol, 2013; Godoy and Roed, 2016). For the sake of completeness, I provide in Appendix C.D the model extension to potential stepping-stone effects or job-search crowding-out effects of low-wage jobs. I introduce individual heterogeneity in p in Section III.D.

At the beginning of her claim, the job-seeker has a UB capital - total benefit entitlement - equal to B_0. Weekly benefit payments are deducted from the UB capital, so that B_t, the current entitlement at the beginning of period t, decreases over the spell. At each period that she does not work at all (total unemployment), the job-seeker receives an amount b of unemployment benefits, or the remaining entitlement B_t if her current UB capital is not large enough to pay b. If the job-seeker does not work at all along her unemployment spell, she receives benefits during $\bar{t}^{Utot} = B_0/b$ periods.[21] When she takes up a low-wage job with earnings z_t in a given week, she receives an amount $b - T(z_t)$ of unemployment benefits, where $T(z_t)$ is the reduction in benefits. This reduction in benefits $T(z_t)$ is then "transferred" to a later period within the claim. When benefits are exhausted, the job-seeker leaves the unemployment registers, but she still looks for a permanent job and she

[21] The model parameters b and $\bar{t}^{Utot}$ correspond to the following institutional parameters: Weekly Benefit Amount (WBA) and Potential Benefit Duration (PBD).

may still work for a low-wage job. The partial-UI schedule $T(.)$ is defined as:

$$T(z) = \begin{cases} 0 & \text{if } z < z^* \\ z - z^* & \text{if } z \in (z^*, z^* + \min(b, B_t)) \\ \min(b, B_t) & \text{if } z > z^* + \min(b, B_t) \end{cases} \tag{2}$$

where z^* is the amount of disregard.[22] The partial-UI schedule feature two kinks: the marginal benefit reduction rate jumps from 0% to 100% at the disregard level z^*, and comes back to 0% at the maximum earnings level. Except at the very end of the claim, the remaining UB capital B_t is greater than the weekly benefit amount b and the maximum earnings level equals $z^* + b$. As explained in Section I, my empirical analysis abstracts from the second kink at the maximal earnings amount because of data limitation.[23] This second kink does not affect my identification strategy that is local around the first kink.

Let me define $U(B_t; n_i)$ the value of unemployment of job-seeker i when the UB capital is B_t. At each date, the job-seeker with discount factor β, maximizes the following program:

$$U(B_t; n_i) = \max_{c_t, z_t} u(c_t, z_t; n_i) + \beta \left[pU(B_{t+1}; n_i) + (1-p)W \right] \tag{3}$$

such that

$$\begin{cases} c_t & = z_t + \min(b, B_t) - T(z_t) \\ B_{t+1} & = B_t - \min(b, B_t) + T(z_t) \\ B_{t+1} & \geq 0 \end{cases}$$

The first constraint of the program is the current budget constraint. I assume that workers cannot save or borrow, as UI claimants are likely to be low-skilled workers who are credit-constrained. In Appendix C.C, I provide a model extension where workers are allowed to borrow and save, and discuss how this changes incentives to bunch for risk-averse workers. The second constraint captures the endogenous entitlement reduction over time (or UB capital depreciation). The UB capital is reduced by the UB payment $\min(b, B_t) - T(z_t)$. The last constraint states that job-seekers cannot borrow UB entitlement from the UI administration.[24]

[22] I assume here that current benefit reduction can reach the actual weekly benefit amount b, as in Idaho. In other states, the maximal amount earned by partial-UI claimants is smaller. However this simplification does not affect the identification as the focus is on earnings close to the kink.

[23] I observe earnings reported to the UI administration. When individuals earn more than the maximal amount, there are no incentives to remain on the UI register and report earnings.

[24] For the sake of simplicity, I do not model the fact that any remaining entitlement at the end of the benefit year is lost, as in the data, almost all job-seekers find permanent jobs or exhaust their UB entitlement before that date.

III.B Model solution

I focus on the case where the UB capital is strictly decreasing, and I define $\bar{t} < \infty$ the exhaustion date, i.e. the first date when $B_t = 0$.[25] The exhaustion date is endogenous, as it depends on the solution path of z_t. I describe the model solution for a period when $B_t > b$. This is more relevant to the empirical analysis, as most of my observations are in this case. The case $B_t < b$ is reported in online Appendix C.A. For all z_t, such that $T(.)$ is differentiable at z_t, the first order condition is:

$$\underbrace{u_c(c_t, z_t; n_i)\left(1 - T'(z_t)\right)}_{gain(I)} + \underbrace{\beta p T'(z_t) U'(B_{t+1}; n_i)}_{gain(II)} = -u_z(c_t, z_t; n_i) \tag{4}$$

where u_c is the marginal utility of consumption and u_z the marginal disutility of work. Equation (4) equates the marginal gains of work (on the left-hand side) with the marginal cost of effort (or disutility of work). The marginal gains have two components. The first term on the left-hand side is the current marginal utility of consumption due to one extra dollar of earnings, which is taxed at the marginal benefit-reduction rate $T'(z)$. The second term is the marginal value of an increase in future UB capital due to one extra dollar of earnings. It is scaled by the discount factor β and the survival rate p.

From the envelope theorem - used at every future period -, it is possible to compute the marginal value of UB capital. Computation details are reported in Appendix C.A. The second term of Equation (4) then simplifies to:

$$\beta p T'(z_t) U'(B_{t+1}; n_i) = T'(z_t) \beta^{\bar{t}-t-1} p^{\bar{t}-t-1} u_c(c_{\bar{t}-1}, z_{\bar{t}-1}; n_i). \tag{5}$$

Using Equation (5) and the parametrization of the utility in Equation (1), Equation (4) simplifies to:

$$1 - T'(z_t)\tau_t = \left(\frac{z_t}{n_i}\right)^{1/e} \tag{6}$$

where τ_t is the wedge between the static marginal tax rate $T'(z_t)$ and the dynamic marginal tax rate $\tau_t T'(z_t)$:

$$\tau_t = 1 - \beta^{\bar{t}-t-1} p^{\bar{t}-t-1}. \tag{7}$$

Note that, if there was no benefit reduction at all, all individuals would supply $z_t = n$. The ability n_i of individual i can thus be interpreted as her potential earnings in low-wage jobs. The actual partial-UI schedule features a kink at the disregard level: the marginal benefit-reduction rate jumps from 0% to 100% (see Equation 2). Such a kink implies that some claimants bunch at the disregard amount.

To describe the bunching behavior, I define a first threshold at ability $n^* = z^*$, i.e. the

[25] I show, in Appendix C.A, that such a focus is relevant when studying the behavior of claimants around the disregard level.

disregard level. The FOC implies that all individuals with $n < n^*$ earn $z_t = n$, as $T'(z_t) = 0$ below z^*. I define another threshold of ability $n^* + \delta n$, such that all individuals with ability strictly above $n^* + \delta n$ earn strictly more than the disregard z^*. Such individuals have their current benefits reduced. Their earnings in low-wage jobs are $z_t = n(1-\tau_t)^e$, as $T'(z_t) = 1$. Using the FOC, the upper threshold then verifies:

$$z^* = (n^* + \delta n)\,(1-\tau_t)^e. \tag{8}$$

Equation (8) illustrates that the upper threshold depends on the dynamic marginal tax rate and consequently on its determinants, such as the time period. We highlight this dependence by denoting the ability gap between first and last buncher as $\delta n(t)$. More fundamentally, the dynamic marginal tax rate depends on the time to exhaustion, which is endogenous. As job seekers with an ability just above the upper threshold preserve a small amount of benefits and thus delay their exhaustion date by only one period, their benefit exhaustion date $\bar{t}$ is equal to their potential benefit duration plus one.[26] Finally, all individuals with $n \in (n^*, n^* + \delta n(t))$, earn exactly the disregard amount $z_t = z^*$: they bunch at the kink point of the schedule.

To summarize, the earnings density function $g_t(z)$ at period t verifies:[27]

$$g_t(z) = \begin{cases} f(z) & \text{if } z < z^* \\ \int_{n^*}^{n^*+\delta n(t)} f(n)dn & \text{if } z = z^* \\ f\left(\frac{z}{(1-\tau_t)^e}\right)\frac{1}{(1-\tau_t)^e} & \text{if } z > z^* \end{cases} \tag{9}$$

where $f(n)$ is the ability density of claimants, assumed smoothly distributed.

III.C What bunching identifies

Suppose that the earnings distribution $g_t(z)$ is identified in the data. This yields the bunching mass at the disregard level $g_t(z^*)$ and the left limit of the earnings density at the disregard level $g_t^-(z^*)$. The ratio of these two quantities corresponds to the excess bunching at period t, denoted $\mathcal{B}_t$, which is equal to:

$$\mathcal{B}_t \equiv \frac{g_t(z^*)}{g_t^-(z^*)} = \frac{1}{f(n^*)}\int_{n^*}^{n^*+\delta n(t)} f(n)dn \simeq \delta n(t) \tag{10}$$

[26] From a theoretical point of view, there could be other bunching masses at the earnings levels where the theoretical exhaustion date increases by one period. Because the corresponding changes in the dynamic marginal rate are small, especially at the beginning of the spell, I expect the resulting bunching to be small as well. Indeed, I find none in the data and thus abstract from those further kinks.

[27] $g_t(z)$ is a density with respect to $\lambda + \delta(z^*)$ where λ is the Lebesgue measure and $\delta()$ is the Dirac measure.

where the first equality is obtained thanks to Expression (9) and the second equality uses a first-order approximation of the integral of a continuous function. The excess bunching thus identifies the difference in ability between the first job-seeker bunching from below and the last job-seeker bunching from above: $\delta n(t)$.

Using Equation (10) and the definition of the lower ability threshold n^*, a first-order approximation of Equation (8) yields the following expression for the earnings elasticity:[28]

$$e = \frac{\mathcal{B}_t}{z^* \tau_t}. \tag{11}$$

The main difference between the static bunching formula of Saez (2010) and the above expression is the definition of the marginal tax rate. When taxes above the kink are linked to deferred benefits, the bunching formula holds with *effective* marginal tax rates. In my setting, the effective dynamic marginal tax rate depends on the discount factor and the probability to exhaust the initial benefit entitlement. This shows that provided that the discount factor and the probability to exhaust the initial benefit entitlement are identified, the bunching formula allows to identify the earnings elasticity.

Last, I aggregate bunching over time. Let me define $\mathcal{B} = \frac{1}{f(z^*)} \int_t \int_{n^*}^{n^*+\delta n(t)} f(n) dn dG(t) = \int_t \mathcal{B}_t dG(t)$ where $G(t)$ is the cumulative distribution of time spent claiming. Using Equations (10) and (11), I obtain the aggregate bunching formula:

$$e = \frac{\mathcal{B}}{z^* \int_t \tau_t dG(t)} \tag{12}$$

where $\int_t \tau_t dG(t)$ is the marginal tax rate that new claimants expect.

I show in the online appendix how relaxing several baseline assumptions affects the bunching formula (11). When I allow for stepping-stone or crowding-out effects of low-wage jobs, the bunching formula remains the same. The intuition for this result is that the career concerns above and below the earnings disregard have similar importance. As long as the expected gains (or costs) of working in low-earnings jobs on future career are a smooth continuous function of earnings at the disregard level, career concerns do not contribute to bunching. The detailed proof is in Appendix C.D.

When workers are risk-averse, the bunching formula is still valid but with an augmented effective marginal tax rate (to the right of the disregard level). It includes an extra term capturing the consumption smoothing value of delaying benefit payments to later in the spell. This value classically depends on risk preferences and the consumption drop over the unemployment spell. We discuss how to quantify this extra term in Section V.C, both when workers are hand-to-mouth or when they can save or borrow.

[28] Assuming $\delta n << z^*$, I obtain $e = -\frac{\mathcal{B}_t}{z^* \ln(1-\tau_t)}$. Assuming $\tau_t << 1$, I obtain the formula in the main text. I check below that the estimation results are robust when I do not assume $\tau_t << 1$.

III.D Heterogeneity in bunching

The bunching formula (11) implies that bunching decreases as the probability to find a permanent job and to leave the UI rolls decreases. Intuitively, forward-looking job-seekers with a higher propensity to keep claiming have higher expected returns to partial UI: they are more likely to profit from benefit transfers later in the claim. I then have the following comparative statics result:

Proposition 1 (Bunching across p-strata)**.** *At any given period t, excess bunching decreases with the probability to remain claiming p.*

Proof: The dynamic marginal tax rate decreases with p: $\frac{d\tau_t(p)}{dp} = -(\bar{t} - t) \times (\beta p)^{\bar{t}-t}/p < 0$.

The bunching formula (11) also implies that workers with longer potential benefit duration are more likely to bunch. This yields the following proposition:

Proposition 2 (Bunching effect of potential benefit duration)**.** *Bunching increases with potential benefit duration.*

Proof: Bunching depends on the potential benefit duration $\bar{t}^{Utot}$ through the dynamic marginal tax rate. For bunchers, $\tau_t = 1 - (\beta p)^{\bar{t}^{Utot}-t}$, which increases with $\bar{t}^{Utot}$: $\frac{d\tau_t}{d\bar{t}^{Utot}} = -\log(\beta p) \times (\beta p)^{\bar{t}^{Utot}-t} > 0$.

The bunching heterogeneity, highlighted in the previous propositions, is a direct consequence of the tax-benefit linkage. If workers do not internalize benefit preservation (no tax-benefit linkage), the relevant tax rate in the bunching formula does not depend on survival probability, nor on potential benefit duration. I explore below the corresponding bunching heterogeneity in U.S. data.

Before that, I discuss the robustness of the bunching heterogeneity test. First, Propositions 1 and 2 are robust to allowing for stepping-stone/crowding-out effects (as the bunching formula remains the same). Propositions 1 and 2 are also robust when hand-to-mouth workers are risk-averse, as the value of consumption smoothing through benefit preservation does not depend on permanent job finding probability, nor on PBD. For hand-to-mouth workers, consumption is totally determined by benefits and earnings in low-wage jobs. When risk-averse workers can save or borrow, this is no longer the case, and the value of smoothing consumption over the unemployment spell may indirectly depend on the job finding probability and on PBD. For example, when PBD is longer, UI exhaustees may have lower consumption level, as exhaustion occurs later in the unemployment spell when remaining savings are lower. The consumption smoothing value of delaying benefits may thus be larger after a PBD increase, which goes against Proposition 2. While it is reasonable to assume that partial UI workers have low savings and are credit-constrained, quantifying this counteracting force in the bunching-heterogeneity test should be addressed in future research.

A last important robustness consideration is that Propositions 1 and 2 do not rely on workers having rational expectations. Whatever the process generating the workers' expectations about their job finding rate, Proposition 1 holds. Choosing the relevant measures of job finding expectation is an empirical issue that we discuss below. Regarding Proposition 2, the *sign* of the effect of PBD on bunching does not depend on the expected job finding rate p. Testing the null hypothesis of no relation between PBD and bunching is thus a robust test of the tax-benefit linkage.

IV Bunching heterogeneity in the data

In this section, I implement direct empirical tests of the tax-benefit linkage. I test for heterogeneity in bunching along the lines of Propositions 1 and 2 above. I estimate excess bunching, i.e. the quantity $\mathcal{B}$ in Equation (10), following the procedure in Chetty et al. (2011a). To maximize statistical power, I jointly analyze US states that share the same disregard level ($0.5 \times WBA$): Idaho and Louisiana. I test bunching heterogeneity by potential benefit duration first, and then by expected survival rate (expected time before finding permanent job).

IV.A Potential benefit duration

According to Proposition 2, claimants with longer potential benefit duration (PBD) bunch more. First, I report the correlation between bunching and PBD, within the cross-section. Second, I exploit within-worker variation in PBD, namely across multiple claiming spells. Third, I leverage policy shocks on PBD arising from emergency triggers.

Cross-section design Figure 3 plots bunching estimates by groups of initial potential benefit duration in Tier I (that is before any triggered extensions). The underlying earnings distributions are available in the online appendix (see online Figures G1 and G2). For this design only, I analyze Idaho and Louisiana separately to account for their different PBD distributions. Whatever the state, bunching is significantly greater when claimants have longer potential benefit durations. In Idaho (left-hand panel), excess bunching for claimants with PBD below 17 weeks amounts to around four times the mass of claimants who would have worked at the disregard level, had the kink disappeared. For claimants with 26 weeks of PBD, excess bunching reaches six. I compute the slope of the relation between bunching and PBD for Idaho: it is statistically significant with a p-value of 0.01. Similarly, excess bunching is statistically different across the two PBD groups in Louisiana with a p-value of 0.02. Of course, this comparison may be confounded by other factors correlated with potential benefit duration. For example, it is well-established that longer potential benefit durations cause higher survival rates (see early contributions in Katz and Meyer (1990a) and Lalive et al. (2006), and the review in Schmieder and von Wachter (2016)). Higher survival rates tend to decrease bunching (see Proposition 1), so that Figure 3 likely underestimates the positive relation between bunching and potential benefit duration. Another example of potential confounder - with less obvious direction of bias - is heterogeneous earnings elasticity across workers.

Within-worker design To control for worker heterogeneity in earnings elasticity and other time-invariant factors, I restrict the sample to workers with two claiming spells. I compute the difference in PBD between the first and second spell, and split the sample in three subgroups: workers experiencing a drop in PBD of more than 3 weeks, workers with small changes in PBD (between a 2-week drop or 2-week increase), and workers experiencing an increase in PBD of more than 3 weeks. Figure 4 plots the across-spell change in bunching for these three groups merging data from Idaho and Louisiana. The change in bunching is positively correlated with the change in PBD. Being exposed to a larger PBD changes (switching to the subgroup just above) increases bunching change by 1.5 with bootstrapped standard errors equal to 0.26 (p-value less than 0.01). I further scale the bunching change by actual PBD change across groups: one week increase in PBD leads to a 0.15 increase in excess bunching with standard errors of 0.027. This is also statistically significant at the 1% level.

While the within-worker design controls for permanent unobserved heterogeneity, one may be concerned that the across-claim contrast reflects other variations than PBD change only. The model in the previous section highlights that changes in expected job finding rates may confound the PBD change effect. Workers experiencing a positive PBD shock across claims are likely on an upward labor market trajectory (as PBD increases with pre-unemployment earnings). One expects those workers to have better employment prospects and lower weekly expected survival rate during their second claim. This channel may increase bunching for this group, beyond the direct effect of PBD increase. On the contrary, the causal effect of longer PBD on unemployment duration (just discussed above) pushes job finding rates downwards, and survival rates upwards, leading to lower bunching for the workers group with positive PBD shocks. To inform on the magnitude of those channels, I predict for each worker her expected survival rate (see next section for more details). The duration model uses as predictors both permanent worker heterogeneity, and claim-level characteristics that vary across claims (such as WBA and PBD). In Appendix Figure G3, I show that the change in expected survival rate for each of the three PBD-change groups is rather small: less than 0.004 out of a 0.96 average. While I cannot exclude that other unobserved factors that vary across claims still confound the PBD effect, the absence of strong change in expected survival rates builds confidence in the within-worker design. Overall, the within-worker design provides further evidence consistent with Proposition 2.

Emergency plan design The third strategy leverages state-wide increases in potential benefit duration triggered by emergency plans. These policy shocks are automatically triggered when the state-level unemployment rate reaches certain activation levels. I compute bunching quarter by quarter separately for Idaho and Louisiana together with their respective average potential benefit duration (PBD). I then regress bunching on PBD using extension emergency programs as an instrument for PBD. More precisely, the instrument T_{sq} is the number of triggered extension programs in state s in quarter q. It is equal to one if either EB or FSC is available in the quarter, and to two if both are available. My main regression specification writes:

$$\mathcal{B}_{sq} = \alpha + \beta PBD_{sq} + \delta_s + \gamma_q + \epsilon_{sq} \tag{13}$$

where $\mathcal{B}_{sq}$ is excess bunching in state s quarter q, PBD_{sq} is the average potential benefit duration of claimants in state s and quarter q (assuming they exhaust all benefits in each tier, i.e. standard Tier I and supplemental programs Tier II to IV), δ_s are state fixed effects and γ_q are quarter fixed effects. Controlling for quarter and state fixed effects identifies β as in a difference-in-difference design. As $\mathcal{B}_{sq}$ are estimated quantities, I bootstrap the whole estimation process - both bunching estimation and Regression (13) - to compute the standard errors for the coefficient of interest β. Table 3 reports the effect of potential benefit duration on excess bunching, i.e. the coefficient β. According to Column (2) which corresponds to the specification above (Equation 13), a 10-week increase in PBD yields 0.7 excess bunching (statistically significant at the 5% level). This amounts to around 14% of average bunching reported in Table 2.

One may be concerned that other business cycle state-level factors confound the PBD effect in the previous specification (Equation 13). To address those concerns, I replace the quarter fixed effects of regression (13) by the *state* quarterly unemployment rates (U_{sq}) from BLS (2020), and run the following 2SLS regression:

$$\mathcal{B}_{sq} = \alpha + \beta PBD_{sq} + \delta_s + \nu U_{sq} + \epsilon_{sq} \tag{14}$$

Identification of β then relies on discontinuities in the rule between extended potential benefit duration and local unemployment rate. The point estimate reported in Column (3) is smaller than in Columns (1) and (2), but still maps 10 PBD weeks into a 7% increase in bunching. Note that identification in Column (3) is robust to confounders varying continuously with local unemployment rates, such as expected survival rates. For the sake of completeness, I compute the quarterly state-level average of expected survival rates in unemployment (see next section for more details on the estimation of expected survival rates), and add it as an extra control in Column (4) of Table 3. The results are robust, and the β coefficient is statistically significant at the 10% level.

Overall, this last piece of empirical evidence combines both discontinuities in benefit-withdrawal schedule and policy shocks for credible identification. The results are again consistent with Proposition 2. They show that workers supply labor accounting for the expected value of preserved benefits.

IV.B Survival expectations

I further provide an empirical test of Proposition 1 that leverages survey answers of claimants to direct questions on expectations. In the survey available for a subsample of the CWBH dataset, workers answer whether they expect to be recalled to their previous employers. Katz and Meyer (1990b) find that job-seekers who expect to be recalled have shorter unemployment duration, i.e. they have a lower probability to remain claimants. Figure 5 shows that claimants expecting to be recalled bunch significantly more. The bunching mass at the disregard level is 50% larger. Note that, at this stage, I do not assume that workers have rational expectations on their job finding rate, as I rely on workers' own answers to expectations questions in a survey. In online Appendix Section F, I test for bunching effect of expected survival rates under the assumption of rational expectations. I find converging evidence confirming Proposition 1.

Overall, bunching heterogeneity in the data provides strong support for the tax-benefit linkage. In the next section, I perform a quantitative exercise that relies more extensively on the theoretical model. Namely, I test whether model-based effective tax rates allow to estimate reasonable earnings elasticity to the net-of-tax rate.

V Estimates of effective tax rate and earned income elasticity to the net-of-tax rate

In this section, I estimate the model-based effective marginal tax rate under the assumption of rational expectations. Together with the bunching estimates, this identifies the earning elasticity to the net-of-tax rate (as shown in Equation 11). I compare the elasticity estimates to those found in the literature. This comparison tests whether the data support both the tax-benefit linkage and the rational expectations assumption. It complements the direct bunching-heterogeneity test of the tax-benefit linkage from Section IV.

V.A Dynamic effective marginal tax rate

The theoretical model shows that claimants working just above the disregard in period t react to a dynamic effective marginal tax rate τ_t defined as:

$$\tau_t = 1 - (\beta p)^{\bar{t}^{Utot} - t - 1}.$$

The different components of the effective tax rate are pinned down as follows. I first calibrate the weekly discount factor β to 0.9996, corresponding to an annual discount rate of 4%. Second, I compute for each individual the potential benefit duration under total unemployment: $\bar{t}^{Utot}$. Third, I compute the expected survival rate p taking into account observed individual heterogeneity. More precisely I estimate a proportional hazard model of exiting the UI registers $h_i = h_0 \exp(\beta X_i)$. I include in the proportional hazard model various characteristics of claimants (gender, age, education and ethnicity), claim characteristics (WBA, PBD, recall expectations) and year fixed effects. The model is estimated on the sample of totally unemployed claimants (i.e. without any benefit withdrawal while on claim). Detailed estimation results are reported in Appendix D. The weekly hazard rates vary between 3% and 4% across states (reported in Table 2). I then predict hazard rates at each date, and compute the expected probability of UB exhaustion taking into account the remaining number of entitlement weeks: $(p)^{\bar{t}^{Utot}-t-1}$. By using predicted rates, claimants are assumed to have rational expectations about their covered unemployment duration.

Table 2 reports the estimates of the average dynamic marginal tax rate, over all individuals and weeks. The average estimate is around 55% in Idaho and Louisiana; it is larger in New Mexico, where it amounts to 60%. It is well below the 100% static benefit-withdrawal marginal tax rate above the disregard level.

V.B Earned income elasticity to the net-of-tax rate

I use the identification relation (11) to compute the earned income elasticities to the net-of-tax rate. The standard errors of the elasticity estimates are obtained by the delta method. All components of Equation (11) - bunching, and tax wedge - are already quantified in Table 2. I obtain statistically significant elasticities in Idaho and Louisiana, respectively 0.19 and 0.13. The elasticity in New Mexico has a similar magnitude (0.1), but it is not statistically significant. Elasticity estimates remain between 0.1 and 0.2, when I vary the bunching estimation window and the polynomial degree in the estimation procedure.[29] When I do not use first-order approximation of the logarithm of the net-of-tax rate, elasticity estimates are slightly lower, but still around 0.1.[30]

My elasticity estimates - between 0.1 and 0.2 - are broadly consistent with the estimates of the intensive labor supply elasticity found in previous micro empirical work (see the review of quasi-experimental estimates in Chetty (2012) or Chetty et al. (2011b)). Closer to my context, Gelber et al. (2020) find bunching for both wage-earners and self-employed individuals at the kinks of the Social Security Annual Earnings Test (for workers over the national retirement age). Their estimate of the average earnings elasticity, not taking into account adjustment cost, is 0.19, which is in the range of my estimates.

Overall, the elasticity-comparison test supports tax-benefit linkage. The observed bunching patterns are quantitatively consistent with standard earnings-elasticity claimants supplying labor according to an effective marginal tax rate that fully internalizes the benefit preservation mechanism.

[29]Robustness results are reported in Appendix Table G1.
[30]Robustness results are reported in Appendix Table G2.

V.C Discussion

I discuss whether the elasticity-comparison test is robust to introducing risk-aversion, and frictions in the model and to allowing for biased beliefs in unemployment duration. Last, I discuss saliency issues specific to the partial UI context.

Risk-aversion I discuss how the introduction of risk-aversion in the model affects the elasticity estimates derived from bunching. The detail of the bunching formula with risk-aversion is reported in Appendix Section C.B. Intuitively, risk-averse job-seekers have an extra incentive of transferring benefits to future periods: they insure themselves against future consumption drops. Abstracting from this extra-incentive leads to over-estimate the dynamic marginal tax rate and thus to under-estimate the earnings elasticity e. In Appendix C.B, I obtain a first-order approximation of the ratio of the elasticity estimates with or without risk-aversion: $1 - \sigma \frac{\Delta c}{c_t} \frac{1}{\tau_t}$ where σ is the coefficient of relative risk-aversion and $\Delta c / c_t$ is the relative change in consumption between the current period and the last week of claim before exhaustion. Gruber (1997a) finds that UI claimants experience a 10% drop in consumption when they become unemployed and a further 12% when the replacement rate of UI benefits goes to zero. Then I consider that consumption drops by around 12% between the current period and the last week of claim before exhaustion. Choosing σ between 1 and 2 (Chetty, 2006; Hendren, 2017) and τ_t around 0.5, the ratio of elasticity estimates is between 1.24 and 1.48. The average elasticity in Table 2 is around 0.15, so that the elasticity taking into account risk-aversion would be between 0.186 and 0.222. This is still in the same ballpark of consensus estimates, confirming the evidence on tax-benefit linkage. Note that this robustness test is also valid when I further relax the hand-to-mouth assumption and allow workers to borrow and save (see Appendix C.C).

Optimization frictions Claimants may not be able to find part-time/temporary jobs with earnings that exactly match their desired optimal labor supply. This may be due to firms' constraints in their productive processes and resulting schedules. Alternatively, this may be due to search frictions in the market for low-wage jobs – it takes time for claimants to acquire information about vacancies that fit their labor supply desires. The model can be extended to account for such frictions (e.g. following Chetty et al., 2011a; Kleven and Waseem, 2013; Gelber et al., 2020; Best et al., 2020). For example, I can assume that there is a fixed cost ϕ to adjust from total unemployment to low-earnings work. Totally unemployed claimants only accept jobs that deliver a net gain exceeding the fixed cost, i.e. jobs around their optimal earnings. Intuitively, in such extensions, optimization frictions smooth bunching. Bunching no longer identifies the structural labor supply elasticity, and neglecting frictions in the bunching formula leads to an attenuation bias. This is documented at the kinks of the SSAET by Gelber et al. (2020), GJS thereafter. They estimate that adjusted earnings elasticity amounts to 0.34 when adjustment costs are taking into account. This is almost twice as large as their standard bunching elasticity estimate (0.19). Taking as a benchmark the attenuation bias estimated in GJS, I multiply my unadjusted elasticity estimates by 1.8, yielding a range between 0.18 and 0.36. This adjusted range is still consistent with the elasticity estimates in the various reviews of the literature.

The GJS method to account for frictions rests on change in the kink location, and the speed at which workers bunching follows the location change. The slower the bunching location changes, the larger the adjustment costs, the larger the adjusted elasticity estimate relative to the standard elasticity estimate. The Louisiana reform allows to study adjustment speed in the partial UI context. In Louisiana, the disregard threshold becomes lower for a subsample of high-WBA workers in 1983. Appendix Figure E3 plots the monthly evolution of the bunching estimates at the pre-reform disregard for the treated group. The pre-reform bunching pattern disappears within three months after the reform. This is much faster than the two/three year lag in GJS. While the faster speed may be related to the saliency of the location change, it certainly points to lower adjustment costs in our context. The above friction-adjusted range between 0.18 and 0.36 can thus be considered as an upper bound.

Biased beliefs I implement the elasticity-comparison test under the assumption that workers have rational expectations about their job finding rates. Recent surveys eliciting job seeker beliefs on employment prospects document significant over-optimistic bias (e.g. Spinnewijn, 2015; Mueller et al., 2021). When job seekers believe that they will find jobs at a faster rate than they actually do, they will tend to bunch more than predicted by the rational-expectation model. Consequently, when one assumes rational expectation, the structural elasticity estimate is over estimated. Unfortunately, the CWBH data does not include elicited worker beliefs on survival rates. The only available belief is the recall expectation (see Section III.D). I thus rely on the recent estimates of biased beliefs from Mueller et al. (2021). They find that "job-seekers perceive their [job finding] chances to be 20% higher than they are" (p.341). I update the expected job finding rates accordingly: $1 - \tilde{p} = 1.2(1 - p)$, where p is the rational-expectation survival rate (see Section V.A). Appendix Table G3 reports the updated effective marginal tax rates and earnings elasticities. Updated effective marginal tax rates are around 5 percentage points larger than estimates under rational expectation from Table 2. The updated elasticities are .17 and .12 for resp. Idaho and Louisiana. While they are lower than estimates under rational expectation, the difference is quantitatively small. The updated estimates are still in the same ballpark of consensus estimates. Accounting for biased beliefs confirms the conclusion from the elasticity-comparison test.

Saliency The weekly frequency of benefit withdrawal in the U.S. is an important contextual feature that may lead to *salient* taxation. The short horizon - within a year - over which benefits are rolled over is also likely to contribute to the *salience* of the benefits tied to taxation. All in all, the salience of both taxation and benefit entitlement may explain why my elasticity-comparison test points to almost perfect tax-benefit linkage. Quantifying how saliency contributes to the tax-benefit linkage is beyond the scope of my paper and a future promising direction of research (see for example Bozio et al., 2020).

VI Conclusion

This paper provides empirical evidence that workers are willing to pay taxes today in order to become entitled to benefits tomorrow. In the U.S. context, unemployment insurance claimants accept to work more intensively in low-earning jobs and see their current benefits withdrawn, as they value benefit preservation for future payments. The tax-benefit linkage is identified through the consequences of kinks in the withdrawal schedule and their interactions with policy shocks. My approach could be directly applied to study partial-UI schedules in other OECD countries (e.g. Germany features kinks in the partial-UI schedule) or to analyze any social insurance with benefit transfers across periods, such as old-age pensions.

The highlighted tax-benefit linkage has important consequences for the design of optimal benefit withdrawal and benefit preservation rules. It suggests that benefit withdrawal leads to lower distortions than predicted by simpler models where workers do not internalize the preservation rule and the value of future benefit entitlement. Fully integrating the highlighted tax-benefit linkage in a general model of optimal social benefit design would be an interesting topic for future research. Another interesting topic is how partial unemployment insurance interacts with short-time work programs in a dynamic environment. Short-time work programs subsidize workers with temporary low earnings and they are close substitutes to partial UI (see Giupponi et al., 2022, for a comparison between STW and UI in recessions). From a broader perspective, gathering evidence on these dynamic programs would enhance our understanding of intertemporal decisions under uncertainty.

References

Anderson, P. M. and B. D. Meyer (1993): "Unemployment Insurance in the United States: Layoff Incentives and Cross Subsidies," *Journal of Labor Economics*, 11, 70–95.

——— (1997a): "The effects of firm specific taxes and government mandates with an application to the U.S. unemployment insurance program," *Journal of Public Economics*, 65, 119–145.

——— (1997b): "Unemployment Insurance Takeup Rates and the After-Tax Value of Benefits," *The Quarterly Journal of Economics*, 112, 913–37.

——— (2000): "The effects of the unemployment insurance payroll tax on wages, employment, claims and denials," *Journal of Public Economics*, 78, 81–106, proceedings of the Trans Atlantic Public Economics Seminar.

Best, M. C., J. S. Cloyne, E. Ilzetzki, and H. J. Kleven (2020): "Estimating the Elasticity of Intertemporal Substitution Using Mortgage Notches," *The Review of Economic Studies*, 87, 656–690.

Blomquist, S., W. Newey, A. Kumar, and C.-Y. Liang (2017): "On Bunching and Identification of the Taxable Income Elasticity," Working Paper 24136, National Bureau of Economic Research.

BLS (2020): "Local Area Unemployment Statistics, U.S. Bureau of Labor Statistics," Series available at https://www.bls.gov/lau/.

Bozio, A., T. Breda, and J. Grenet (2020): "Does Tax-Benefit Linkage Matter for the Incidence of Social Security Contributions?" Working Papers halshs-02191315, HAL.

Brown, K. M. (2013): "The link between pensions and retirement timing: Lessons from California teachers," *Journal of Public Economics*, 98, 1 – 14.

Caliendo, M., S. Kunn, and A. Uhlendorff (2016): "Earnings exemptions for unemployed workers: The relationship between marginal employment, unemployment duration and job quality," *Labour Economics*, 42, 177 – 193.

Chetty, R. (2006): "A New Method of Estimating Risk Aversion," *American Economic Review*, 96, 1821–1834.

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

Chetty, R., J. Friedman, T. Olsen, and L. Pistaferri (2011a): "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records," *The Quarterly Journal of Economics*, 126, 749–804.

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

Chetty, R., A. Guren, D. Manoli, and A. Weber (2011b): "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, 101, 471–75.

Coile, C. and J. Gruber (2001): "Social Security Incentives for Retirement," in *Themes in the Economics of Aging*, National Bureau of Economic Research, Inc, NBER Chapters, 311–354.

Farber, H. S., J. Rothstein, and R. G. Valletta (2015): "The Effect of Extended Unemployment Insurance Benefits: Evidence from the 2012-2013 Phase-Out," *American Economic Review*, 105, 171–76.

Fremigacci, F. and A. Terracol (2013): "Subsidized temporary jobs: Lock-in and stepping stone effects," *Applied Economics*, 45, 4719–4732.

Friedberg, L. (1998): *The Social Security Earnings Test and Labor Supply of Older Men*, MIT Press, 121–150.

——— (2000): "The Labor Supply Effects of the Social Security Earnings Test," *The Review of Economics and Statistics*, 82, 48–63.

Gelber, A. M., D. Jones, and D. W. Sacks (2020): "Estimating Earnings Adjustment Frictions: Method and Evidence from the Social Security Earnings Test," *American Economic Journal: Applied Economics*, 12, 1–31.

Giupponi, G., C. Landais, and A. Lapeyre (2022): "Should We Insure Workers or Jobs during Recessions?" *Journal of Economic Perspectives*, 36, 29–54.

Godoy, A. and K. Roed (2016): "Unemployment Insurance and Underemployment," *LABOUR*, 30, 158–179.

Grogger, J. and C. Michalopoulos (2003): "Welfare Dynamics under Time Limits," *Journal of Political Economy*, 111, 530–554.

Gruber, J. (1997a): "The Consumption Smoothing Benefits of Unemployment Insurance," *The American Economic Review*, 87, 192–205.

——— (1997b): "The Incidence of Payroll Taxation: Evidence from Chile," *Journal of Labor Economics*, 15, 72–101.

Haider, S. J. and D. S. Loughran (2008): "The Effect of the Social Security Earnings Test on Male Labor Supply: New Evidence from Survey and Administrative Data," *Journal of Human Resources*, 43, 57–87.

Hamermesh, D. S. (1979): "Entitlement Effects, Unemployment Insurance and Employment Decisions," *Economic Inquiry*, 17, 317–332.

Hendren, N. (2017): "Knowledge of Future Job Loss and Implications for Unemployment Insurance," *American Economic Review*, 107, 1778–1823.

Holen, A. and S. Horowitz (1974): "Partial Unemployment Insurance Benefits and the Extent of Partial Unemployment," *The Journal of Human Resources*, 9, 420–422.

Katz, L. and B. Meyer (1990a): "The Impact of the Potential Duration of Unemployment Benefits on the Duration of Unemployment," *Journal of Public Economics*, 41, 45–72.

——— (1990b): "Unemployment Insurance, Recall Expectations, and Unemployment Outcomes," *The Quarterly Journal of Economics*, 105, 973–1002.

Kiefer, N. and G. Neumann (1979): "The effect of alternative partial benefits formulas on beneficiary part-time work behavior," Unemployment Insurance Occasional Paper 79-6, U.S. Department of Labor, Employment and Training Administration, Washington, DC.

Kleven, H. J. (2016): "Bunching," *Annual Review of Economics*, 8, 435–464.

Kleven, H. J. and M. Waseem (2013): "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan," *The Quarterly Journal of Economics*, 128, 669–723.

Kuhn, P. and C. Riddell (2010): "The Long-Term Effects of Unemployment Insurance: Evidence from New Brunswick and Maine, 1940-1991," *Industrial and Labor Relations Review*, 63, 183–204.

Kyyra, T. (2010): "Partial unemployment insurance benefits and the transition rate to regular work," *European Economic Review*, 54, 911–930.

Kyyra, T., P. Parrotta, and M. Rosholm (2013): "The effect of receiving supplementary UI benefits on unemployment duration," *Labour Economics*, 21, 122–133.

Lalive, R., J. V. Ours, and J. Zweimuller (2006): "How Changes in Financial Incentives Affect the Duration of Unemployment," *Review of Economic Studies*, 73, 1009–1038.

Landais, C. (2015a): "Assessing the Welfare Effects of Unemployment Benefits Using the Regression Kink Design," *American Economic Journal: Economic Policy*, 7, 243–78.

——— (2015b): "Replication data for: Assessing the Welfare Effects of Unemployment Benefits Using the Regression Kink Design." American Economic Association, Inter-university Consortium for Political and Social 2019-10-13, https://doi.org/10.3886/E114581V1.

Le Barbanchon, T. (2024): "Data and Code for Replication: Taxes Tioday, Benfits Tomorrow." American Economic Association, Inter-university Consortium for Political and Social, project openicpsr-208173.

le Maire, D. and B. Schjerning (2013): "Tax bunching, income shifting and self-employment," *Journal of Public Economics*, 107, 1–18.

Lee, D. S., P. Leung, C. J. O'Leary, Z. Pei, and S. Quach (2021): "Are Sufficient Statistics Necessary? Nonparametric Measurement of Deadweight Loss from Unemployment Insurance," *Journal of Labor Economics*, 39, S455–S506.

Liebman, J. B., E. F. Luttmer, and D. G. Seif (2009): "Labor supply responses to marginal Social Security benefits: Evidence from discontinuities," *Journal of Public Economics*, 93, 1208 – 1223.

Manoli, D. and A. Weber (2016): "Nonparametric Evidence on the Effects of Financial Incentives on Retirement Decisions," *American Economic Journal: Economic Policy*, 8, 160–82.

McCall, B. (1996): "Unemployment Insurance Rules, Joblessness, and Part-Time Work," *Econometrica*, 64, 647–682.

Moffitt, R. (1985): "Unemployment insurance and the distribution of unemployment spells," *Journal of Econometrics*, 28, 85–101.

Moffitt, R. and W. Nicholson (1982): "The Effect of Unemployment Insurance on Unemployment: The Case of Federal Supplemental Benefits," *The Review of Economics and Statistics*, 64, 1–11.

Mortensen, D. T. (1977): "Unemployment Insurance and Job Search Decisions," *Industrial and Labor Relations Review*, 30, 505–517.

Mueller, A. I., J. Spinnewijn, and G. Topa (2021): "Job Seekers' Perceptions and Employment Prospects: Heterogeneity, Duration Dependence, and Bias," *American Economic Review*, 111, 324–63.

Munts, R. (1970): "Partial Benefit Schedules in Unemployment Insurance: Their Effect on Work Incentive," *The Journal of Human Resources*, 5, 160–176.

OECD (2020): "Labour market programmes: expenditure and participants," *OECD Employment and Labour Market Statistics (database)*.

O'Leary, C. J. (1997): "An Evaluation of the Washington State Unemployment Insurance Earnings Deduction Experiment," Report prepared for ui program analysis, Washington State Employment Security Department.

Saez, E. (2010): "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy*, 2, 180–212.

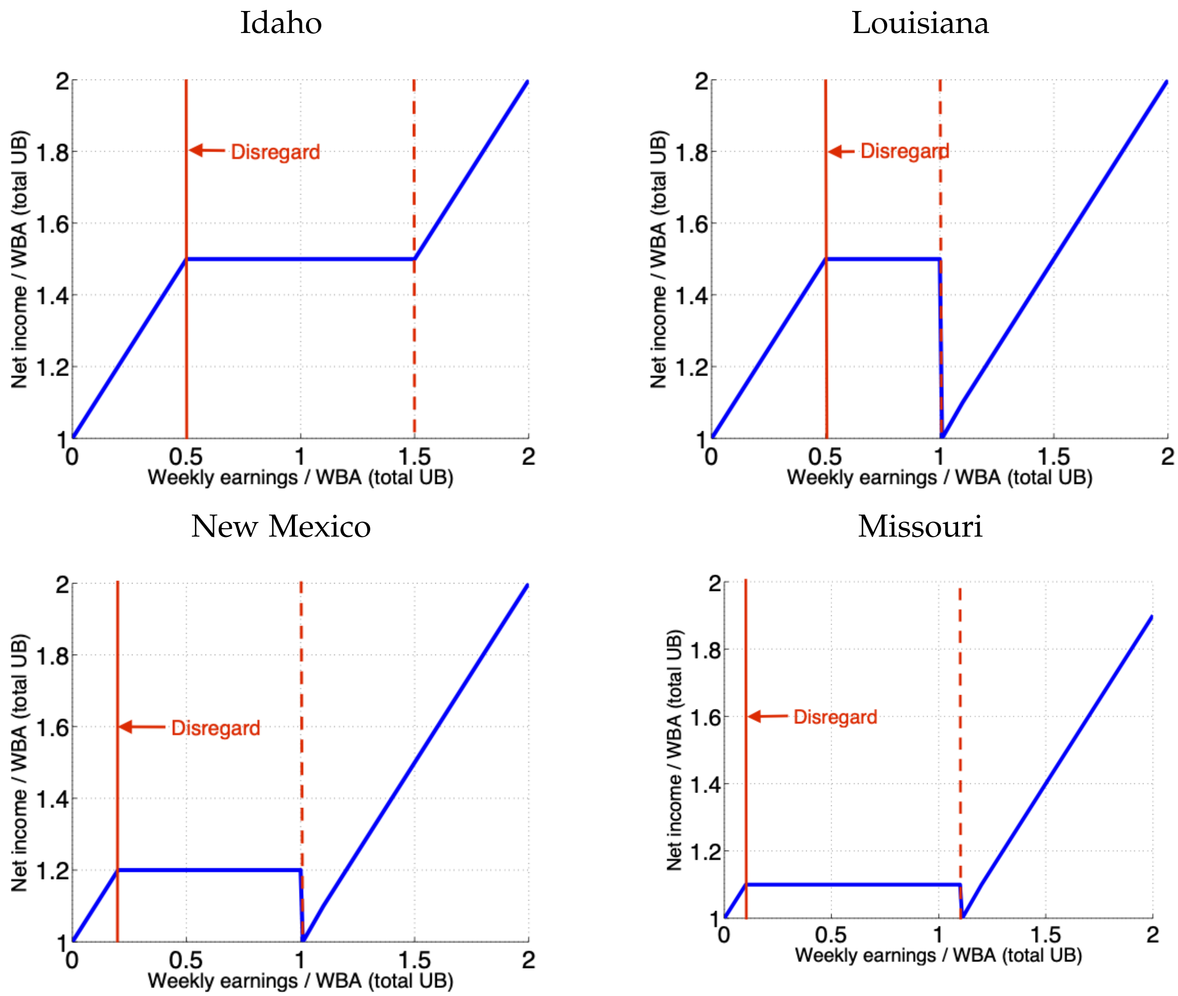
Schmieder, J. F. and T. von Wachter (2016): "The Effects of Unemployment Insurance Benefits: New Evidence and Interpretation," *Annual Review of Economics*, 8, 547–581.

Spinnewijn, J. (2015): "Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs," *Journal of the European Economic Association*, 13, 130–167.

Summers, L. H. (1989): "Some Simple Economics of Mandated Benefits," *American Economic Review*, 79, 177–183.

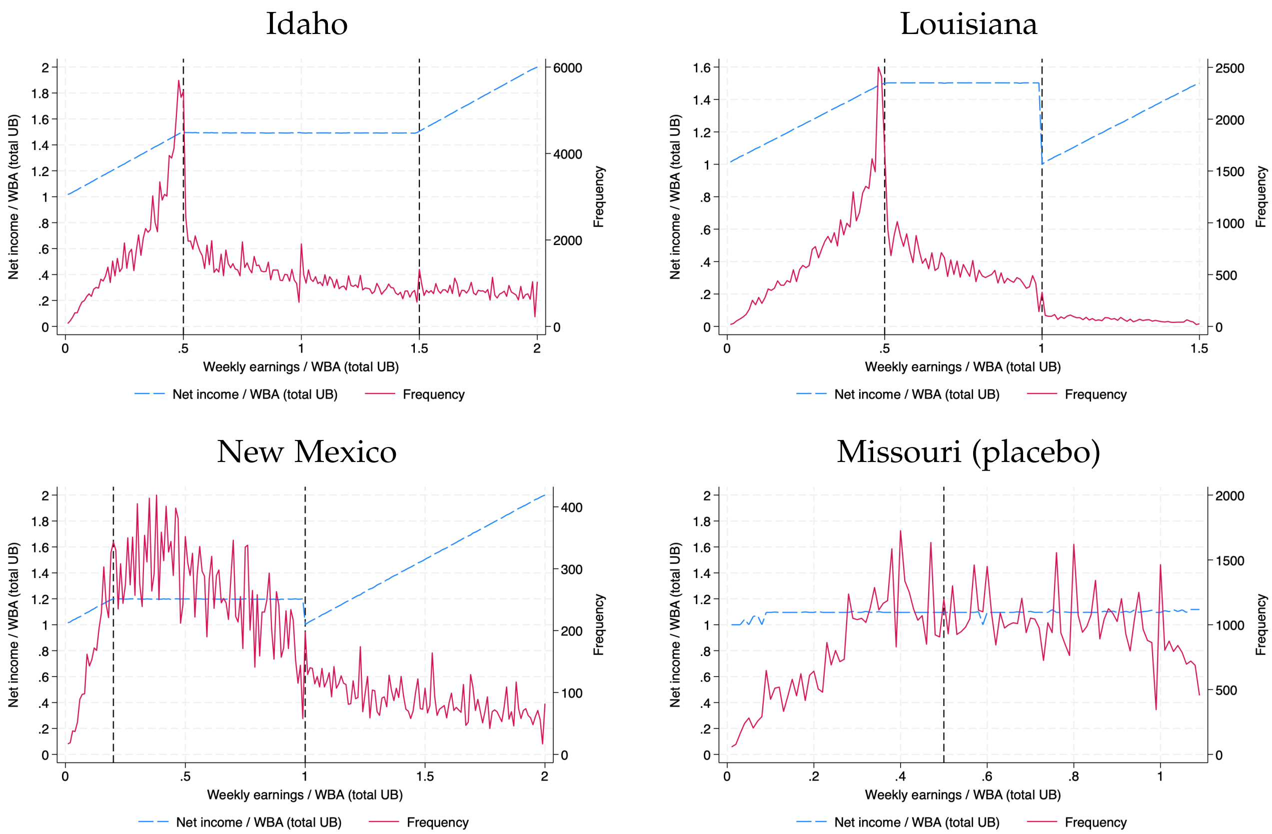
Figures

Figure 1: Partial-UI schedules from 1976 to 1984



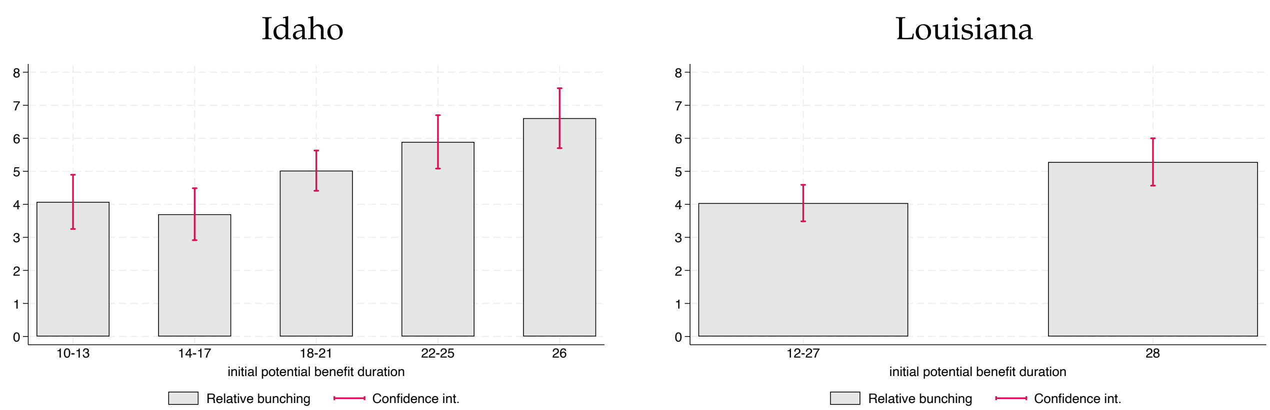
Source: U.S. Department of Labor, "Significant Provisions of State Unemployment Insurance Laws." Notes: This figure plots the theoretical schedules of partial unemployment insurance for the four U.S. states in the CWBH dataset. Each panel yields the net income (weekly labor earnings + unemployment benefit payments) of partially unemployed claimants as a function of their weekly labor earnings. Both net income and earnings are divided by the weekly benefit amount (UB paid in case of total unemployment). The solid red vertical lines correspond to the disregard level. The dashed red vertical lines correspond to the maximum amount of earnings to be eligible for partial UI benefits. For Louisiana, I plot the schedule before April 1983. For Missouri, I consider a claimant whose WBA is $100.

Figure 2: Weekly earnings density and empirical schedule of partial UI.



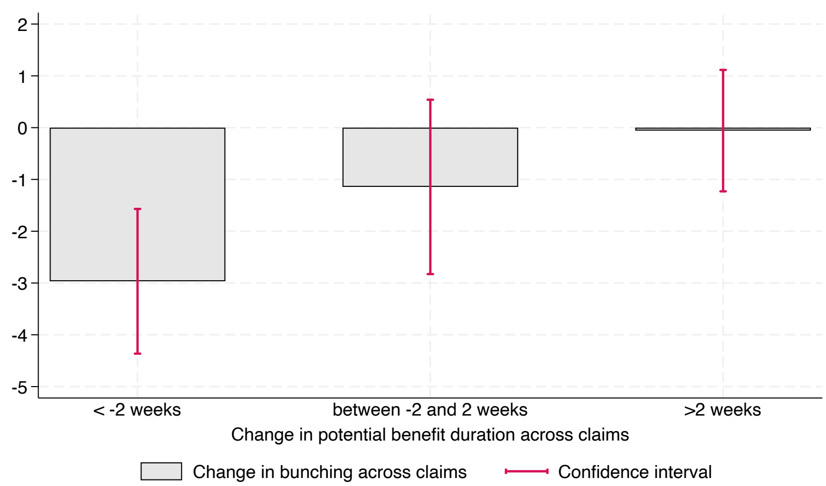
Source: CWBH. Notes: This figure plots in solid red line the distributions of weekly labor earnings of partially unemployed claimants divided by the weekly benefit amount (UB paid in case of total unemployment). Corresponding frequencies are on the right-hand Y-axis. Each panel corresponds to one of the four U.S. states in the CWBH dataset. The figure also plots in blue long dashed line the empirical schedule of partial unemployment insurance: net income (left-hand Y-axis) as a function of labor earnings. The red solid vertical line (resp. dash) shows the kink (resp. the notch) of the partial UI schedule, except for Missouri (the placebo state). For Louisiana, I use data before April 1983.

Figure 3: *Bunching by initial potential benefit duration*



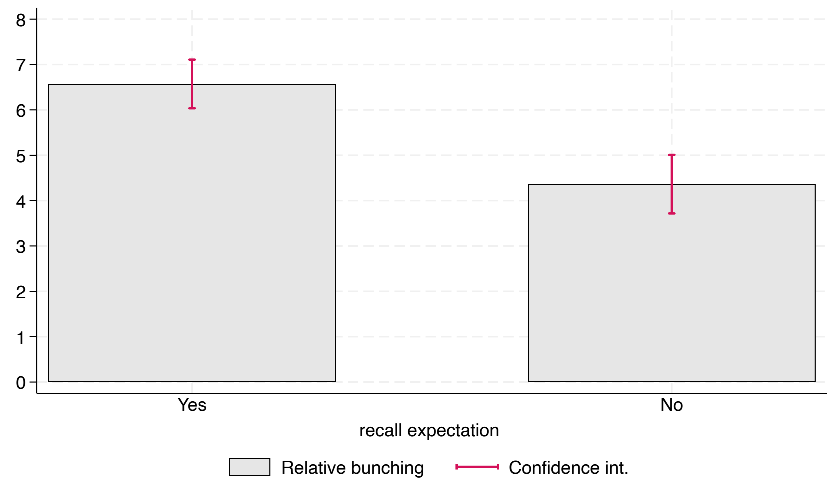
Source: CWBH for Idaho 1976-84 and Louisiana 1979-83Q1. Notes: The figures plot excess mass at the disregard amount (bunching at the kink) by potential benefit duration at the beginning of the claim (in weeks). Confidence interval at the 95% level in red.

Figure 4: *Change in bunching by change in potential benefit duration across claiming spells: within-worker design*



Source: CWBH for Idaho 1976-84 and Louisiana 1979-1984. Notes: The figure plots the across-spell change in excess mass at the disregard amount (bunching at the kink) for claimants experiencing a large negative shock in potential benefit duration across spells (left-hand bar), a small change in PBD (center bar) or a large positive shock in PBD (right-hand bar). Confidence interval at the 95% level in red.

Figure 5: *Bunching by recall expectation*



Source: CWBH for Idaho 1976-84 and Louisiana 1979-83Q1. Notes: The figure plots the excess mass at the disregard amount (bunching at the kink) for claimants expecting to be recalled to their previous employer (left-hand bar) and for claimants not expecting any recalls (right-hand bar). Confidence interval at the 95% level in red.

Tables

Table 1: Descriptive Statistics

	Idaho	Louisiana	New Mexico	Missouri
Male	.635	.704	.608	.504
Age	31.6	34.9	34.1	36
Education (years)	11.8	11.3	11.7	11
White	.947	.625	.461	.885
Pre-U Industries				
Construction	.146	.269	.125	.0667
Manufacturing	.387	.264	.27	.589
Trade	.218	.114	.231	.121
Services	.123	.144	.196	.16
Pre-U Occupations				
Prof., tech. and managers	.0589	.0678	.105	.051
Clerical and sales	.149	.132	.197	.153
Structural work	.234	.264	.18	.0741
Pre-U firm in private sector	.859	.69	.702	.946
Pre-U weekly wage (current $)	332	327	277	222
Weekly Benefit Amount (current $)	103	137	106	89.5
Replacement rate	.449	.468	.435	.484
Potential Benefit duration (weeks)	22.8	28.8	38.7	33.5
Actual Benefit duration (weeks)	17.5	19.7	16.8	16.1
Inflow (no)	40,792	26,363	12,163	26,907
Sample Years	76-84	79-84	80-84	78-84

Source: CWBH. Notes: Means are computed over the sample of partial UI claimants. Inflow reports the number of new claims with at least one week of partial unemployment.

Agriculture, Mining, Transportation, Finance, Insurance and Real Estate industries are not reported for the sake of space. Occupations correspond to the standard DOT (Dictionary of Occupation Titles). I only report the most common occupations and exclude service, agricultural, processing, machine trades and benchwork occupations from the table.

Wages and benefits amounts are in current dollars. Replacement rate is the ratio of weekly benefit amount over pre-unemployment wages. Potential Benefit duration is the maximum number of weekly benefit payments the worker is entitled to under total unemployment. It accounts for extended benefits programs, triggered in high-unemployment periods. Actual benefit duration is the number of weeks within the spell with actual benefit payments.

Table 2: Bunching, dynamic marginal tax rates and earnings elasticity estimates.

	Idaho	Louisiana	New Mexico	Missouri *placebo*
Excess bunching mass ($\mathcal{B}$)	5.33	4.81	1.25	-.777
	(.285)	(.255)	(.804)	(.357)
Hazard rate $(1-p)$	.042	.033	.039	
Effective Marginal Tax Rate (τ)	.538	.576	.606	
Earnings elasticity to net-of-tax rate (e)	.187	.129	0.096	
	(.0010)	(.0068)	(.062)	
Partial-UI weeks (no)	230,535	69,024	31,103	99,451
Sample Years	76-84	79-83	80-84	78-84

Source: CWBH. Notes: This table reports the estimates for excess bunching mass at the disregard level in the distribution of weekly earnings of partial UI claimants, for four US states (in columns). The earnings disregard level is at 50% of weekly benefit amounts (WBA) for Idaho and Louisiana (up to April 1983) and 20% for New Mexico. For Missouri, I consider a placebo earnings disregard at 50% WBA: this amounts to 45 current dollars on average, while the actual disregard level in Missouri is $10. For non-placebo states, the table also reports the average predicted hazard rate out of compensated unemployment $(1-p)$, which is an important building block of the dynamic effective marginal tax rate (τ) (see Equation 7). Quantities in the first three rows allow to estimate the earnings elasticity to the net-of-tax rate in the last row (see bunching relation in Equation 12). Standard errors are in parentheses below estimates. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Bunching and potential benefit duration: IV strategy with UI extension programs (EB and FSC).

	Bunching ($\mathcal{B}$)			
Potential Benefit Duration	.073	.075	.037	.042
	(.013)	(.033)	(.019)	(.020)
State FE	Y	Y	Y	Y
Quarter FE		Y		
State quarterly unemployment rate			Y	Y
State quarterly expected job finding rate				Y
# observations				
state X quarter	46	46	46	46
weekly claims	322,450	322,450	322,450	322,450

Source: CWBH, Idaho and Louisiana 1979-1984. Notes: This table reports the coefficient of potential benefit duration in a regression of excess bunching. I instrument potential benefit duration by the number of benefit extension programs available in state s in quarter t (2SLS estimation). The data of over 300,000 weekly earnings are collapsed at the state X quarter level into 23 observations for each state. Bootstrapped standard errors, accounting for the first-step estimation of excess bunching, are in parentheses below estimates. In Column (1), I control for state fixed effects. In Column (2), I add quarter dummmies from 1979Q1 to 1984Q3. Column (2) is my preferred design (short-period difference-in-difference). In Column (3), I control for state fixed effects, and the quarterly state unemployment rate. I also add dummies for Q1 to Q4 to control for seasonality. IN Column (4), I add the quarterly state estimate of expected job finding rates (see Section V.A). The mean outcome is around 5 (see bunching estimates in Table 2).

Online Appendix

Taxes Today, Benefits Tomorrow

Thomas Le Barbanchon (Bocconi University)

The online appendix is divided in seven sections from A to G. Section A provides a detailed description of the Unemployment Insurance rules in the U.S., specifically for claimants who work while on claim. Section B describes the estimation process for excess bunching. Section C provides the full solution of the main theoretical model, and some model extensions (risk-averse workers, borrowing/saving margins, and stepping-stone/crowding-out effects). Section D estimates the hazard model out of unemployment, used to compute rational expectations of survival. Section E performs a difference-in-difference exercise, as a placebo test for the bunching source. Section F presents supplementary test for heterogeneity in bunching across groups with different job finding predictions. Section G displays supplementary Figures and Tables.

A Institutional background

Between the late 70s and early 80s, the unemployment insurance (UI) rules, in Idaho, Louisiana, New Mexico and Missouri, are as follows. First, UI claimants must meet a monetary eligibility requirement. They must have accumulated a sufficient amount of earnings during a one-year base period before job separation. Second, UI claimants must meet non-monetary eligibility requirements. They must not have quit their previous job, they must not have been fired for misconduct. They must search and be available for work.

When claimants meet the above requirements, states compute their weekly benefit amount (WBA). This would be their weekly unemployment benefit payment when they earn less than the partial UI disregard. The WBA is a fraction (between 1/20 and 1/26) of the high quarter wages (HQW), defined as the wages earned in the quarter of the base period (BP) with the highest earnings. The BP is the first four calendar quarters of the five completed quarters before job separation. The WBA is subject to a maximum and minimum benefit level. As maximum levels are quite low, a large fraction of claimants have their WBA capped. For example, in the first quarter of 1980, the maximum amount was $121 in Idaho. The above rule implies a decreasing gross replacement rate between 50% and 40%. States also compute a potential benefit duration (PBD). This is usually a fraction (between 2/5 and 3/5) of base period wages (BPW), subject to a minimum and maximum number of weeks. The maximum PBD is 26 weeks, except in Louisiana before 1983 where it is 28 weeks. The total entitlement is defined as the product of the WBA and of the PBD. It represents the total amount of unemployment benefits that the claimant can be paid over the benefit year (BY), i.e. the continuous one-year period starting at the first claim. Note that, after the end of the BY, no unemployment benefits can be paid from the corresponding claim, but the claimant can be eligible for a new claim. States observe a waiting period of one week at the beginning of the claim, during which no unemployment benefits are paid.

During periods of high unemployment, the potential duration of unemployment benefits is extended, either by the Federal-state extension benefit (EB) program, or the federal supplemental compensation (FSC) program. Those programs are triggered, when federal or state unemployment are over certain levels. The EB program extended the initial entitlement period by 50% up to a total of 39 weeks when the state unemployment rate reached a certain trigger. The FSC program, in action from September 1982 to March 1985 in all four states considered, extended the entitlement period of individuals who had exhausted their regular and EB entitlement, by a rate ranging from 50% to 65% up to a maximum of weeks depending on the FSC phase and the U.S. state (see Grossman 1989 for more details on the FSC).

There was one major change in UI rules in Louisiana in April 1983. The partial-UI disregards have been capped at $50. In addition, the maximal potential duration of usual benefits was reduced from 28 weeks to 26 weeks. Last, Appendix Table A1 summarizes parameters of the partial UI rules detailed in the main text.

Table A1: Partial-UI rules from 1976 to 1984

	Disregard	Maximum earnings
Idaho	$0.5 \times WBA$	$1.5 \times WBA$
Louisiana bef. Apr. 1983	$0.5 \times WBA$	WBA
Louisiana aft. Apr. 1983	$\min(0.5 \times WBA, \$50)$	WBA
New Mexico	$0.2 \times WBA$	WBA
Missouri	\$10	WBA+\$10

Source: U.S. Department of Labor, "Significant Provisions of State Unemployment Insurance Laws.".

Note: the table reports disregard and maximum levels. Taking into account inflation, \$10 (resp. \$50) in 1978 represent around \$37 (resp. \$185) in 2016.

B Excess bunching estimation

In this appendix, I detail how I estimate excess bunching. I follow the procedure of Chetty et al. (2011a). I fit a polynomial on the earnings density of partial-UI claimants, taking into account that there is bunching in a bandwidth around the disregard, and that the bunching mass comes from the earnings distribution above the disregard.

First, the earnings distribution is centered around the disregard amount. Let C_j be the count of individuals earning between j and $j+1$ dollars above the disregard level (when they earn below the disregard, j is negative), and let Z_j be the dollar amount earned by claimants in bin j ($Z_j = j$), centered around the disregard level. I estimate the following equation:

$$C_j \left(1 + \mathbb{1}[j > \overline{R}] \frac{\hat{B}_N}{\sum_{j>\overline{R}} C_j}\right) = \sum_{k=0}^{q} \beta_k (Z_j)^k + \sum_{i=-\underline{R}}^{\overline{R}} \gamma_i \mathbb{1}[Z_j = i] + \epsilon_j \tag{15}$$

where $\hat{B}_N = \sum_{i=-\underline{R}}^{\overline{R}} \hat{\gamma}_i$ is the excess mass taken off the earnings distribution above the disregard.[31] The order of the polynomial q and the width of the bunching window $(-\underline{R}, \overline{R})$ are not estimated, but set after visual inspection. Robustness checks of the estimation results with respect to those two parameters are presented below.

Equation (15) defines the counterfactual distribution (with no benefit reduction): $\hat{C}_j = \sum_{k=0}^{q} \hat{\beta}_k (Z_j)^k$. Then the estimator of excess bunching equals:

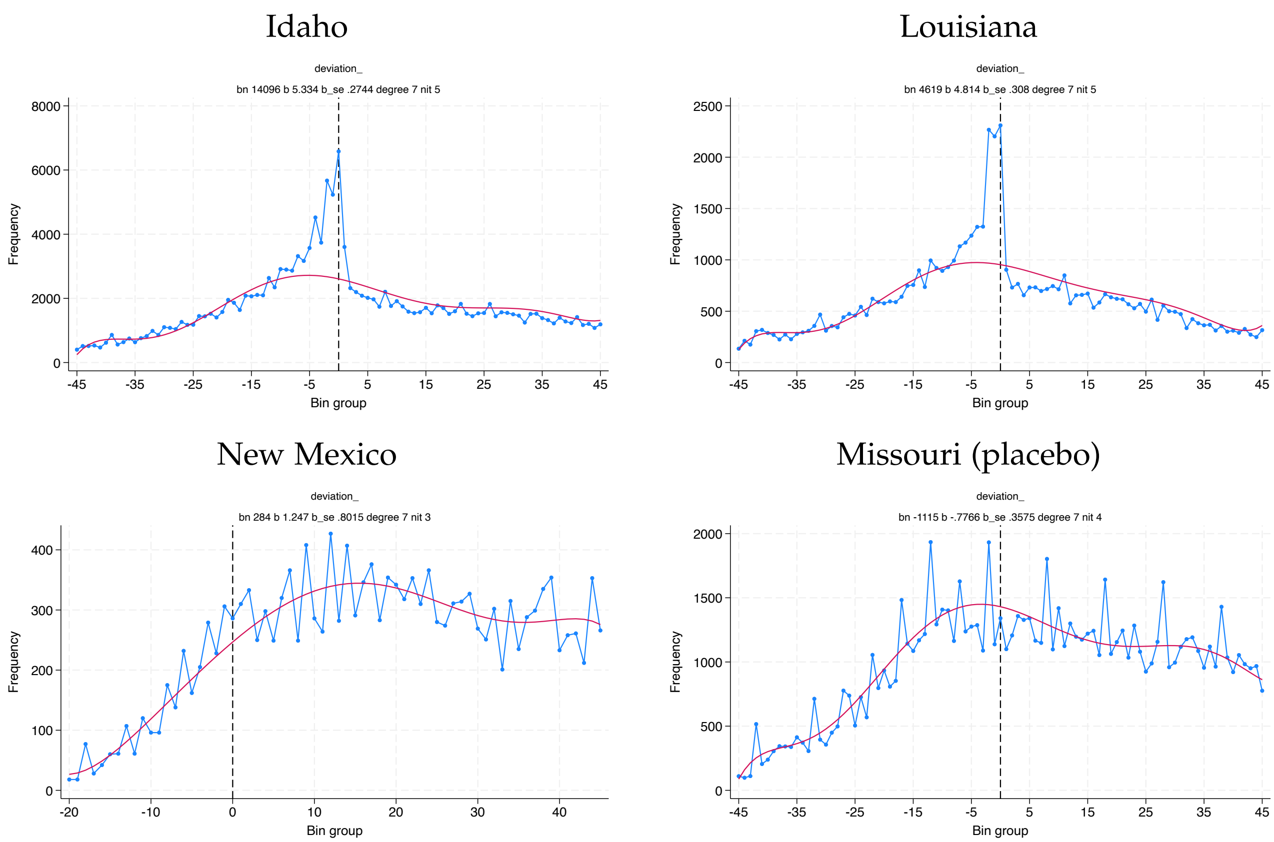
$$\hat{\mathcal{B}} = \frac{\hat{B}_N}{\sum_{j=-\underline{R}}^{\overline{R}} \hat{C}_j / (\underline{R} + \overline{R} + 1)}. \tag{16}$$

The recursive estimation is bootstrapped to obtain standard errors. The bootstrap procedure draws new error terms (ϵ_j) among the estimated distribution.

Appendix Figure B1 illustrates the estimation procedure for each state. It plots the partial claimants' earnings density in bins of one dollar centered around the disregard level, together with the counterfactual density estimated along the lines of Chetty et al. (2011a). In practice, the procedure fits a polynomial of degree 7. The bandwidth is such that $-\underline{R} = -5$ and $\bar{R} = 2$. Appendix Figure B1 confirms that the counterfactual density compares well to the actual data. Appendix Figure B1 reveals some periodicity in the earnings distribution in Missouri. Claimants report earnings that are multiples of ten dollars. This may bias bunching estimates, especially if there are heaps in the window where bunching is expected. I verify that the bunching estimate does not change if I modify the earnings density by smoothing the heaping points.

[31] Because $\hat{B}_N$ depends on $\hat{\gamma}_i$, I follow an iterative procedure to estimate the equation. At each step, $\hat{B}_N$ is computed with past estimates of $\hat{\gamma}$, and the procedure stops when a fixed point is obtained.

Figure B1: Centered weekly earnings density of partial-UI claimants.



Source: CWBH. Notes: Earnings are in dollars centered at the disregard level. Empirical earnings density in blue. Counterfactual density in red.

C Theoretical model

In this appendix, I derive in detail the solution of the claimants' program. Then I introduce risk-aversion, borrowing, and stepping-stone/crowding-out effects in the theoretical model and discuss identification in those cases.

C.A Model Solution

I derive in detail the solution of the claimants' program:

$$U(B_t; n_i) = \max_{c_t, z_t} u(c_t, z_t; n_i) + \beta \left[pU(B_{t+1}; n_i) + (1-p)W \right]$$

$$\text{such that}$$

$$\begin{cases} c_t &= z_t + \min(b, B_t) - T(z_t) \\ B_{t+1} &= B_t - \min(b, B_t) + T(z_t) \\ B_{t+1} &\geq 0. \end{cases}$$

By definition of the partial-UI schedule, I have that $T(z_t) \leq b$ when $B_t > b$ and $T(z_t) \leq B_t$ when $B_t < b$. As a consequence, the capital stock B_t depreciates or stays constant over time: $B_{t+1} \leq B_t$. The main model insights belong to the case of UB capital decreasing over time. Before solving the model under this case, I explain why stationary solutions with constant UB capital can be ruled out. My reasoning is by contradiction. I characterize the stationary solution and explain why deviations increase workers' welfare.

A stationary solution U_i for individual i with $B > b$ (resp. $b > B$) satisfies $T(z) = b$ (resp. $T(z) = B$). Then the program simplifies as:

$$U_i = \max_{c,z} u(c, z; n_i) + \beta \left[pU_i + (1-p)W \right] \text{ such that } c = z.$$

The first order condition writes:

$$u_c(c, z; n_i) + u_z(c, z; n_i) = 0. \tag{17}$$

This determines the level of consumption, while the Bellman equation determines the value of claiming U_i:

$$U_i = \frac{u(c, z; n_i) + \beta(1-p)W}{1 - \beta p}. \tag{18}$$

The two previous equations (17) and (18) show that the stationary value of unemployment U_i and the corresponding earnings z do *not* depend on the level of UB capital B. This is an important property of stationary solutions.

Recall that, when $B > b$, the typical partial-UI schedule is such that there exists a unique maximum earnings $\overline{z}$ such that for any $z \geq \overline{z}$, $T(z) = b$ and the marginal tax rate is 100% just below $\overline{z}$. Workers with stationary consumption have earnings above $\overline{z}$. Let me consider the marginal individual who would supply $\overline{z}$, she would benefit from deviating from the stationary path during one period by decreasing her labor supply by δz. Actually, her flow income is not affected, while she enjoys more leisure. A consequence of this manipulation is that her UB capital is depreciated. However her future utility is not affected as the value of stationary unemployment does not depend on UB capital. Then, this deviation necessarily increases her welfare and a stationary equilibrium does not exist for this worker.

The previous reasoning also applies when $B \in (0, b)$. Recall that, for any $B \in (0, b)$, the typical partial-UI schedule is such that there exists $\overline{z}(B) = B + z^*$ an exit point from partial UI. Let me consider as above the marginal claimant supplying $\overline{z}(B)$. The similar reasoning as above applies: the marginal claimant finds it beneficial to deviate from the stationary path and consume her UB capital. The previous argument does not apply to individuals with $z > \overline{z}(B)$. In the remainder, I implicitly restrict the analysis to individuals with preferences inconsistent with stationarity. An alternative solution could be to introduce a fixed flow cost to claim. This would make the group of job-seekers with ability consistent with stationary claiming arbitrarily small.

While claiming, UB capital is thus strictly decreasing over the spell. I define $\overline{t} < \infty$ the finite exhaustion date (first date when $B_t = 0$). The program becomes stationary only when job-seekers run out of benefits. I denote U_i the value of unemployment of job-seeker i when benefits are exhausted.

Let me now solve the program. When $B_t > b$, it simplifies as:

$$U(B_t; n_i) = \max_{z_t} u(z_t + b - T(z_t), z_t; n_i) + \beta \left[p.U(B_t - b + T(z_t); n_i) + (1-p)W \right].$$

When $B_t \in (0, b)$, it is given by:

$$U(B_t; n_i) = \max_{z_t} u(z_t + B_t - T(z_t), z_t; n_i) + \beta \left[p.U(T(z_t); n_i) + (1-p)W \right].$$

Both sub-programs share the same first order condition:

$$u_c(c_t, z_t; n_i) \left(1 - T'(z_t)\right) + \beta p T'(z_t) U'(B_{t+1}; n_i) = -u_z(c_t, z_t; n_i). \tag{19}$$

Using the envelope theorem, I show that the marginal value of UB capital satisfies the following recursive equation:

$$U'(B_t; n_i) = \begin{cases} \beta p U'(B_{t+1}; n_i) & \text{when } b < B_t, \\ u_c(c_t, z_t; n_i) & \text{when } 0 < B_t < b. \end{cases}$$

For simplicity I assume that the individual only claims one period when $0 < B_t < b$. This can be rationalized by introducing a fixed flow cost of claiming. Then this period verifies $t = \bar{t} - 1$. Consequently, the third term of the marginal gain of labor earnings can be written as:

$$\beta p T'(z_t) U'(B_{t+1}; n_i) = T'(z_t) \beta^{\bar{t}-t-1} p^{\bar{t}-t-1} u_c(c_{\bar{t}-1}, z_{\bar{t}-1}; n_i) \tag{20}$$

where $p^{\bar{t}-t-1}$ is the probability to exhaust benefits conditional on claiming at date t.

Using Equation (39) and the utility definition, the FOC in Equation (19) can be simplified. The rest of the derivation is in the main text.

C.B Risk-aversion

In this section, I consider the behavior of risk-averse job-seekers. I assume that the per-period utility of a claimant writes:

$$u(c, z; n_i) = \frac{c^{1-\sigma}}{1-\sigma} - \frac{n_i}{1+1/e} \left(\frac{z}{n_i} \right)^{1+1/e} \tag{21}$$

where σ is the coefficient of relative risk-aversion. The derivation of the solution path follows the same lines as in the main text and Appendix C.A. We focus on the case when $B_t > b$ as in the main text. Solving the job-seekers' program yields the following FOC:

$$u_c(c_t, z_t; n_i) \left(1 - T'(z_t)\right) + T'(z_t) (\beta p)^{\bar{t}-t-1} u_c(c_{\bar{t}-1}, z_{\bar{t}-1}; n_i) = -u_z(c_t, z_t; n_i). \tag{22}$$

Using the definition of utility of risk-averse job-seekers, the FOC simplifies as:

$$1 - T'(z_t) \left(1 - (\beta p)^{\bar{t}-t-1} \left(\frac{c_t}{c_{\bar{t}-1}} \right)^{\sigma} \right) = \left(\frac{z_t}{n} \right)^{1/e} (c_t)^{\sigma}. \tag{23}$$

Compared to Equation (34), the wedge between the static benefit-reduction rate and the dynamic marginal tax rate depends on the coefficient of risk-aversion and on the ratio of current consumption to consumption in the last week of claim.

While the behavior of risk-averse job-seekers is more complex, bunching still identifies the parameter e *if the coefficient of risk-aversion is separately identified*. The intuition follows. First, consider the individual who bunches from below. Her ability n^* satisfies the following FOC:

$$1 = \left(\frac{z^*}{n^*} \right)^{1/e} (b + z^*)^{\sigma}. \tag{24}$$

Second, consider the individual who bunches from above. Her ability $n^* + \delta n(t)$ satisfies the following FOC:

$$(\beta p)^{\bar{t}-t-1} (c_{\bar{t}-1})^{-\sigma} = \left(\frac{z^*}{n^* + \delta n(t)} \right)^{1/e}. \tag{25}$$

As $c_{\bar{t}-1}$ is a function of the ability $n^* + \delta n(t)$ and of the other parameters of the model (β, p, σ, B_t, b, z^* and e), Equations (24) and (25) identify e when excess bunching is observed in the data. More precisely, we obtain the consumption in the last week of claim $c_{\bar{t}-1}$ using the FOC of the program when $B_{\bar{t}-1} < b$:

$$1 = \left(\frac{z_{\bar{t}-1}}{n^* + \delta n(t)}\right)^{1/e} \left(c_{\bar{t}-1}\right)^{\sigma}. \tag{26}$$

The budget constraint in the last week of claim writes: $c_{\bar{t}-1} = B_{\bar{t}-1} + z_{\bar{t}-1}$. Assuming that the remaining UB capital in the last week of claim is negligible, we have $c_{\bar{t}-1} = z_{\bar{t}-1}$. Then Equation (26) shows that $c_{\bar{t}-1}$ only depends on $n^* + \delta n(t)$, e and σ. Replacing the implicit expression of $c_{\bar{t}-1}$ in Equation (25), we obtain that excess bunching identifies e from Equations (24) and (25).

We now quantify the order of magnitude of the bias in the estimate of e when risk-aversion is neglected. We re-write Equations (24) and (25):

$$n^* = z^* \left(c_t\right)^{e\sigma}, \tag{27}$$

$$n^* + \delta n(t) = z^*(1 - \tau_t)^{-e} \left(c_{\bar{t}-1}\right)^{e\sigma}. \tag{28}$$

Taking the difference between these two equations and using first-order approximations, we obtain:

$$\frac{\delta n(t)}{z^*} = e\left(\tau_t + \sigma\frac{\Delta c}{c_t}\right)$$

where $\Delta c = c_{\bar{t}-1} - c_t$. Rearranging terms, we obtain the following identification formula:

$$e = \frac{\delta n(t)}{z^*\left(\tau_t + \sigma\frac{\Delta c}{c_t}\right)}. \tag{29}$$

Taking the ratio of the above expression and Equation (11), we obtain the ratio of elasticity estimates with or without risk-aversion: $1 - \sigma\frac{\Delta c}{c_t}\frac{1}{\tau_t}$. In the main text, I quantify the order of magnitude of this ratio.

C.C Borrowing-saving

In this section, I no longer assume that job-seekers are hand-to-mouth. They have access to safe assets with return rate r. This introduces a new state variable in the job-seeker's program: asset capital A_t. Asset A can be transferred across state, so that the value of permanent jobs also depends on A_t. The job-seeker program is modified as follows:

$$U(B_t, A_t; n_i) = \max_{c_t, z_t, A_{t+1}} u(c_t, z_t; n_i) + \beta \left[pU(B_{t+1}, A_{t+1}; n_i) + (1-p)W(A_{t+1}) \right]$$

such that

$$\begin{cases} c_t + A_{t+1} & = z_t + \min(b, B_t) - T(z_t) + (1+r)A_t \\ B_{t+1} & = B_t - \min(b, B_t) + T(z_t) \\ B_{t+1} & \geq 0 \\ A_{t+1} & \geq \underline{A} \end{cases}$$

where the modified budget constraint allows for saving and borrowing through A. The last constraint prevents the job-seeker from borrowing infinite amounts.
The solution is characterized by two FOCs (when constraint inequalities do not bind). For the first FOC, I differentiate wrt z_t, and for the second FOC wrt A_{t+1}:

$$u_c(c_t, z_t; n_i) \left(1 - T'(z_t)\right) + \beta p T'(z_t) U_B(B_{t+1}, A_{t+1}; n_i) = -u_z(c_t, z_t; n_i) \quad (30)$$

$$-u_c(c_t, z_t; n_i) + \beta \left[pU_A(B_{t+1}, A_{t+1}; n_i) + (1-p)W_A(A_{t+1}) \right] = 0 \quad (31)$$

where U_B and U_A denote the differentials of the intertemporal value U wrt to its first and second state variable. Using the envelope theorem when $B_t > b$, the FOCs write:

$$U_B(B_t, A_t; n_i) = \beta p U_B(B_{t+1}, A_{t+1}; n_i), \quad (32)$$

$$U_A(B_t, A_t; n_i) = (1+r) u_c(c_t, z_t; n_i). \quad (33)$$

When $B_t < b$, I have the same terminal condition as in the baseline model: $U_B = u_c(c_{\bar{t}}, z_{\bar{t}}; n_i)$. For the sake of simplicity, I do not write down the full program when the worker works in permanent jobs. I assume that the marginal value of wealth in this state relates to the marginal utility of consumption $W_A(A_{t+1}) = (1+r)\bar{u}_c$.

Before solving further, I note that under risk neutrality, the borrowing-saving margin is irrelevant for bunching. The marginal utility of consumption is equal to one and it does not depend on consumption level. The FOC (30) simplifies as before and this pins down the path of low-wage earnings z_t independently of the wealth sequence A_t:

$$1 - T'(z_t)\tau_t = \left(\frac{z_t}{n_i}\right)^{1/e} \quad (34)$$

The more interesting case is that of risk-aversion. I assume that job-seekers are risk-averse as in Section C.B with per-period utility:

$$u(c,z;n_i)=\frac{c^{1-\sigma}}{1-\sigma}-\frac{n_i}{1+1/e}\left(\frac{z}{n_i}\right)^{1+1/e}$$

As in Section C.B, solving the job-seekers' program forward yields the following FOC wrt z_t:

$$u_c(c_t,z_t;n_i)\left(1-T'(z_t)\right)+T'(z_t)(\beta p)^{\bar{t}-t-1}u_c(c_{\bar{t}-1},z_{\bar{t}-1};n_i)=-u_z(c_t,z_t;n_i). \tag{35}$$

The previous expression can be re-written as:

$$1-T'(z_t)\left(1-(\beta p)^{\bar{t}-t-1}\frac{u_c(c_{\bar{t}-1},z_{\bar{t}-1};n_i)}{u_c(c_t,z_t;n_i)}\right)=\left(\frac{z_t}{n}\right)^{1/e}\frac{1}{u_c(c_t,z_t;n_i)}. \tag{36}$$

We note as in Section C.B that the tax wedge depends on the ratio of marginal utility across periods. When intertemporal transfers are allowed on financial markets, the ratio of marginal utility is pinned down by the Euler equation, expressed in the FOC wrt A_t in Equation (31). Substituting relation (32) and the expression of W_A, Equation (31) writes:

$$u_c(c_t,z_t;n_i)=\beta\left[p(1+r)u_c(c_{t+1},z_{t+1};n_i)+(1-p)(1+r)\bar{u}_c\right]$$

Iterating the previous expression at future dates allows to express current marginal utility as an expression of marginal utility at date $\bar{t}-1$:

$$u_c(c_t,z_t;n_i)=\left(\beta p(1+r)\right)^{\bar{t}-1-t}u_c(c_{\bar{t}-1},z_{\bar{t}-1};n_i)+(1-p)\sum_{k=1}^{\bar{t}-1-t}\left(\beta(1+r)\right)^k p^{k-1}\bar{u}_c$$

I further assume that discount factor and interest rates are such that: $\beta(1+r)=1$. I use the analytical formula for the sum of first terms of a geometric sequence and obtain a compact expression of the current marginal utility:

$$u_c(c_t,z_t;n_i)=p^{\bar{t}-1-t}u_c(c_{\bar{t}-1},z_{\bar{t}-1};n_i)+\left(1-p^{\bar{t}-1-t}\right)\bar{u}_c$$

It allows to express the ratio of marginal utility as:

$$\frac{u_c(c_{\bar{t}-1},z_{\bar{t}-1};n_i)}{u_c(c_t,z_t;n_i)}=\frac{1-\frac{\bar{u}_c}{u_c(c_t,z_t;n_i)}}{p^{\bar{t}-1-t}}+\frac{\bar{u}_c}{u_c(c_t,z_t;n_i)} \tag{37}$$

From the above expression, it is not clear how the ratio compares to one. Intuitively, the ability to borrow and save smooths marginal utility across periods. This would bring the ratio of marginal utility closer to one and bunching behavior of risk-averse workers would be closer to that of risk-neutral workers than to that of hand-to-mouth risk-averse workers.

The bias quantification exercise performed at the end of Section C.B is still valid when borrowing-saving is allowed. It relies on the same FOC (wrt z_t) that is left unchanged when introducing borrowing and saving. The bias quantification essentially relies on supplementary data on the curvature of the utility function and on consumption drops over the spell. Only the interpretation of the consumption drop changes when borrowing and saving are allowed. In the borrowing-saving case, it follows Formula (37).

However, without consumption data, it seems more difficult to exactly identify the earnings elasticity e using the bunching formula derived from the FOC (30) taken for marginal bunchers. The identification proof used for hand-to-mouth risk-averse workers does not hold any longer (even when one assumes identification of the utility curvature). Compared to the hand-to-mouth case, earnings levels are not sufficient to identify consumption, which the proof leverages in Equation (24) and (26) of Section C.B. In the borrowing-saving case, savings/borrowings introduce an unobserved wedge between consumption and earnings. Moreover, the Euler equation is unlikely to solve the identification problem as the marginal utility while on claim depends on the marginal utility in permanent jobs, which is neither observed.

C.D Stepping-stone/crowding-out effects

In this section, I account for stepping-stone and/or crowding-out effects of low-earnings jobs. I assume that the probability to find a permanent job depends on the earnings level in the current low-wage job: $1 - p(z_t)$. The claimants' program then becomes:

$$U(B_t; n_i) = \max_{c_t, z_t} u(c_t, z_t; n_i) + \beta \left[p(z_t) U(B_{t+1}; n_i) + (1 - p(z_t)) W \right]$$

such that

$$\begin{cases} c_t &= z_t + \min(b, B_t) - T(z_t) \\ B_{t+1} &= B_t - \min(b, B_t) + T(z_t) \\ B_{t+1} &\geq 0. \end{cases}$$

I can show, as in the previous appendix C.A, that there exists a solution where the UB capital is strictly decreasing up to a finite exhaustion date $\bar{t} < \infty$. Considering the case when $B_t > b$, we obtain the following FOC:

$$\underbrace{u_c(c_t, z_t; n_i) \left(1 - T'(z_t)\right)}_{(I)} + \underbrace{\beta p(z_t) T'(z_t) U'(B_{t+1}; n_i)}_{(II)} - \underbrace{\beta p'(z_t) \left(W - U(B_{t+1}; n_i)\right)}_{(III)} = -u_z(c_t, z_t; n_i). \tag{38}$$

Compared to the FOC in the baseline model (Equation 34), a third term (III) appears on the left-hand side. When working while on claim increases the future probability to find a permanent job - stepping-stone effect ($p' < 0$) -, the job-seeker is induced to work more. She has the opposite reaction when working while on claim crowds out job search for permanent jobs - crowding-out effect ($p' > 0$).

The FOC makes clear that it is only the *marginal* stepping-stone/crowding-out effect that matters: $p'(z)$. I expect *marginal* effects to be smaller than *average* stepping-stone/crowding-out effects that compare outcomes of partial UI claimants to those of total UI claimants who do not work while on claim. The empirical literature - McCall (1996) in the U.S. and Kyyra (2010), Caliendo et al. (2016), Kyyra et al. (2013), Fremigacci and Terracol (2013) and Godoy and Roed (2016) in European countries - provides estimates of *average* effects that are small. It seems then reasonable to neglect stepping-stone/crowding-out effects when studying the intensive margin of labor supply for low-earnings jobs. For the sake of completeness, this appendix further discusses assumptions that are sufficient to obtain the bunching formula when one does not neglect *marginal* stepping-stone/crowding-out effects.

Following the same reasoning as in the previous appendix C.A, I obtain a simplified expression of the second term in Equation (38):

$$\beta p(z_t) T'(z_t) U'_{t+1}(B_{t+1}; n_i) = T'(z_t) \beta^{\bar{t}-t-1} \left(\Pi_{i=t}^{\bar{t}-2} p(z_i) \right) u_c(c_{\bar{t}-1}, z_{\bar{t}-1}; n_i) \tag{39}$$

where $\Pi_{i=t}^{\bar{t}-2} p(z_i)$ is the probability to exhaust benefits conditional on claiming at date t. Using Equation (39) and the utility definition, the FOC in Equation (38) simplifies to:

$$1 - T'(z_t)\tau_t(z_t) - \beta p'(z_t) \left(W - U(B_{t+1}; n_i) \right) = \left(\frac{z_t}{n_i} \right)^{1/e} \tag{40}$$

where the wedge τ_t now depends explicitly on z_t:

$$\tau_t(z_t) = 1 - \beta^{\bar{t}-t-1} \Pi_{j=t}^{\bar{t}-2} p(z_j). \tag{41}$$

We now state two assumptions that are sufficient to obtain identification of the earnings elasticity to the net-of-tax rate. First, the marginal effect of earnings on the permanent job finding probability $p'(z)$ is continuous. Second, the net gain of permanent jobs ($W - U(B_{t+1}; n_i)$) depends continuously on earnings z_t and depends on individual ability only through earnings. These assumptions imply that there exists a continuous function π_t such that $\pi_t(z_t) = \beta p'(z_t) \left(U(B_{t+1}; n_i) - W \right)$.

Consequently, the FOCs can be written as:

$$1+\pi_t(z_t) = \left(\frac{z_t}{n_i}\right)^{1/e} \quad \text{when } z_t < z^*, \tag{42}$$

$$1-\tau_t(z_t)+\pi_t(z_t) = \left(\frac{z_t}{n_i}\right)^{1/e} \quad \text{when } z_t > z^*. \tag{43}$$

This leads me to define a lower threshold n_t^* and an upper threshold $n_t^* + \delta n_t$, such that:

$$n_t^* = \frac{z^*}{(1+\pi_t^-(z^*))^e}, \tag{44}$$

$$n_t^* + \delta n_t = \frac{z^*}{\left(1-\tau_t(z^*)+\pi_t^+(z^*)\right)^e}. \tag{45}$$

where π^+ and π^- are respectively the upper and lower limits of π. Because π is assumed continuous, the marginal gains induced by stepping-stone/crowding-out effects cancel out of the identifying relation (as long as $\pi_t(z^*) << 1$). Then the elasticity verifies the same identification relation: $e = \frac{\mathcal{B}_t}{z^*\tau_t(z^*)}$.

D Hazard model

In this appendix, I report results of the estimation of the hazard model used to compute the probability to remain claiming the following week (p). I follow the baseline assumptions of the theoretical model and neglect any duration dependence (p does not depend on t). I estimate the following exponential hazard model where covariates enter proportionally. For individual i, the hazard model is: $h_i = h_0.\exp(\beta X_i)$. The hazard model is estimated on a subsample of claimants, according to the local nature of the bunching estimate. I am interested in the hazard rate of claimants, close to bunching. I thus restrict the estimation to claimants whose benefits are not reduced because of partial UI.

It is well-established that hazard rates out of UI registers feature spikes at benefit exhaustion date. I verified that I obtain such patterns in the data from the Continuous Work and Benefit History (CWBH) project, as Katz and Meyer (1990b) do. As I want to capture the probability to remain claiming for individuals who are still entitled to unemployment benefits, observations are censored before exhaustion spikes. I use the theoretical exhaustion date in Tier 1 when claimants are totally unemployed along the whole claim ($\bar{t}^{Utot}$), in order to censor observations.

My objective is to estimate claimants' expectation about their hazard rates. Rational forward-looking claimants would use all available information to form their expectations. Consequently, covariates X capturing individual heterogeneity include: gender, age (and its square), years of initial education (and its square), ethnicity, calendar year of first week of claim, potential benefit duration (in Tier 1), weekly benefit amount and recall expectation. For each covariate, a specific dummy is included to account for missing values. Table D1 reports the coefficient estimates of the hazard model for each state (in columns).

Table D1: Results of hazard model estimation

	Idaho	Louisiana	New Mexico	Missouri
Male	.078***	.288***	.080***	.175***
	(.021)	(.015)	(.015)	(.015)
Age	-.021***	-.006**	-.031***	-.006*
	(.004)	(.003)	(.004)	(.003)
Age (square)	.0001**	-.00005	.0003***	-.00002
	(.00005)	(.00003)	(.00004)	(.00004)
Education (years)	-.111***	-.061***	.012	-.036
	(.021)	(.009)	(.011)	(.062)
Education (square)	.006***	.004***	.0007	.004
	(.0009)	(.0004)	(.0005)	(.003)
Black	-.042	-.216***	-.172***	-.583***
	(.103)	(.013)	(.049)	(.020)
Hispanic	.296***	.097*	-.223***	-.207
	(.044)	(.053)	(.014)	(.153)
American Indian	-.164*	-.084	-.231***	-.211
	(.093)	(.122)	(.024)	(.378)
Asian	.127	-.139*	-.264***	.230
	(.113)	(.075)	(.091)	(.218)
Potential benefit duration	.053***	.031***	.059***	.044***
	(.002)	(.002)	(.007)	(.002)
Weekly benefit amount	-.001***	-.003***	-.001***	-.005***
	(.0003)	(.0001)	(.0002)	(.0004)
No recall expectation	-.484***	-.249***	-.398***	-.600***
	(.025)	(.016)	(.013)	(.016)
Years fixed effects	Yes	Yes	Yes	Yes
No. spells	25274	55519	37937	41663
Log-likelihood	-32412.47	-75213.41	-53243.84	-56269.63

Source: CWBH. Notes: The reference is a white female with recall expectation whose claim starts in the first year of the sample. The estimation includes a constant and missing categories for ethnicity and recall expectation that we do not report in the table. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

E Difference-in-difference

In April 1983, Louisiana changed UI rules. The change in partial UI affected both the stock of individuals registered as unemployed in April 1983 and new inflows after that point in time.[32] The disregard level was reduced from $0.5 \times WBA$ to \$50 for all claimants whose WBA is more than \$100. This is the treatment group. For all claimants with a WBA below \$100, the disregard was not reduced and remained equal to $0.5 \times WBA$. This is the control group. I select claims around the policy shocks, from April 1982 to March 1984. The sample covers a full year before the policy change and another full year after the new rules were implemented.

Placebo test I expect that, if bunching is actually related to the partial-UI schedule, the bunching location would switch from the old to the new threshold in the treatment group, and remain the same in the control group. If bunching is due to norms or policies unrelated to the partial-UI program, bunching (in the treatment group) should not be altered by the policy change.

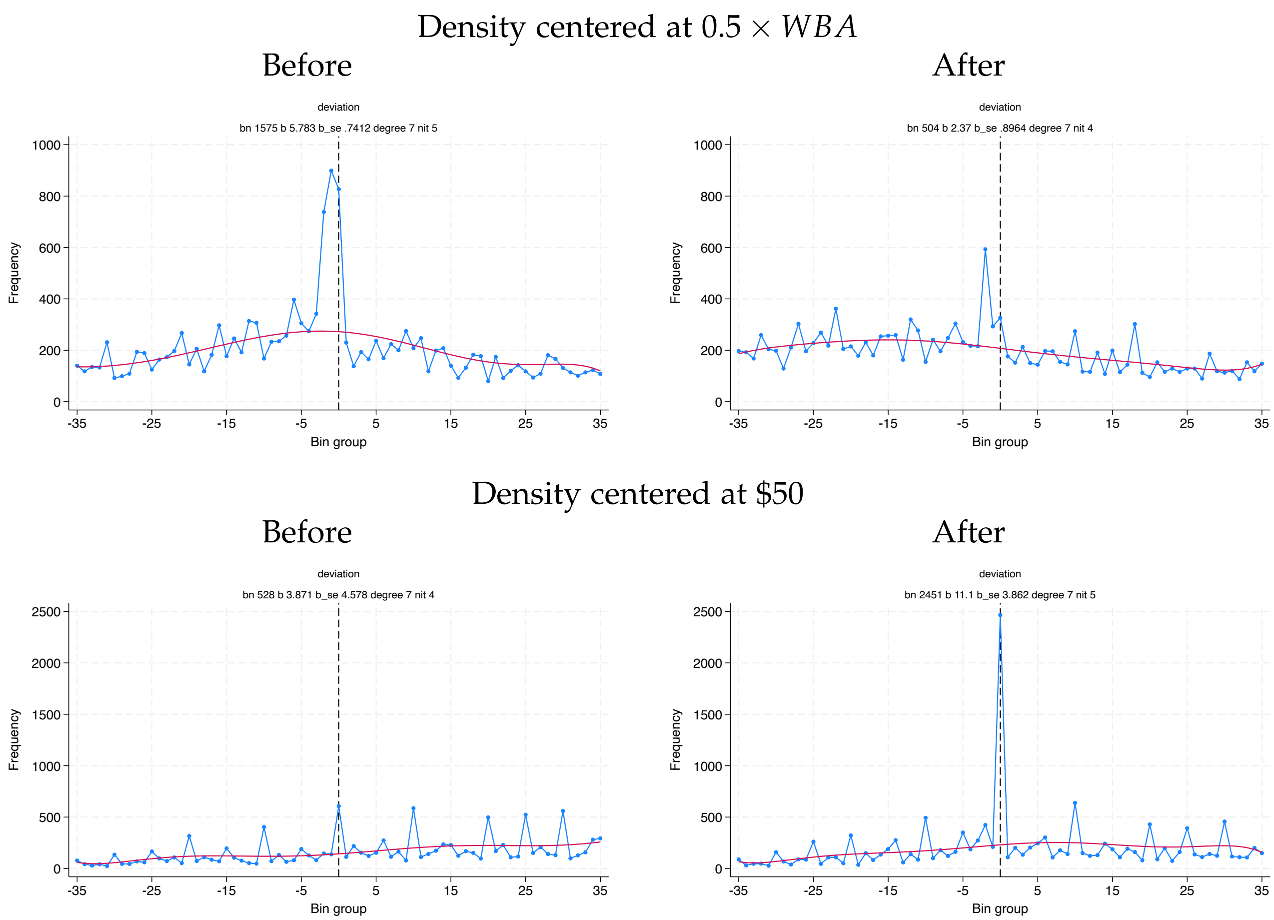
Figure E1 plots the earnings density of partial-UI claimants in the treatment group. In the upper panel, densities are centered at the pre-reform disregard ($0.5 \times WBA$). In the lower panel, they are centered at the post-reform disregard (\$50). Starting with the upper panel, bunching is considerably reduced from before the reform (left graph) to after the reform (right graph). Bunching estimate at the pre-reform disregard level is no longer statistically significant after the reform. The lower panel shows that claimants actually switch to the post-reform disregard after the reform. The mass of bunchers at \$50 doubles after April 1983. Note that there were actually some claimants at the \$50 threshold before the reform. This may be explained by norms unrelated to the partial-UI program. The important point here is that bunching increases after the reform. Note also that bunching is sharper when disregards are rounded amounts.

Figure E2, in which I repeat the same exercise for the control group, does not display any fundamental changes in the bunching pattern after the reform. Claimants in the control group continue to bunch at their relevant disregard amount ($0.5 \times WBA$). They do not switch to the post-reform disregard of the treatment group (\$50). The absence of bunching after the reform in the control group also suggests that bunching incentives mediated by the demand side of the labor market are weak in Louisiana. Suppose that firms actually internalize the partial-UI program and post wages at the disregard level. Because they cannot direct their search to claimants with certain disregard levels, it is likely that they would post the most common disregard (Chetty et al., 2011a). In Louisiana, the mode of the disregard distribution is \$50 (the treatment group is twice as large as the control group). If the bunching incentives were mainly mediated by firms, I would expect to see bunching

[32] There was also a reduction in the maximum number of entitlement weeks from 28 to 26 weeks. This could have affected the amount of bunching, but not its location.

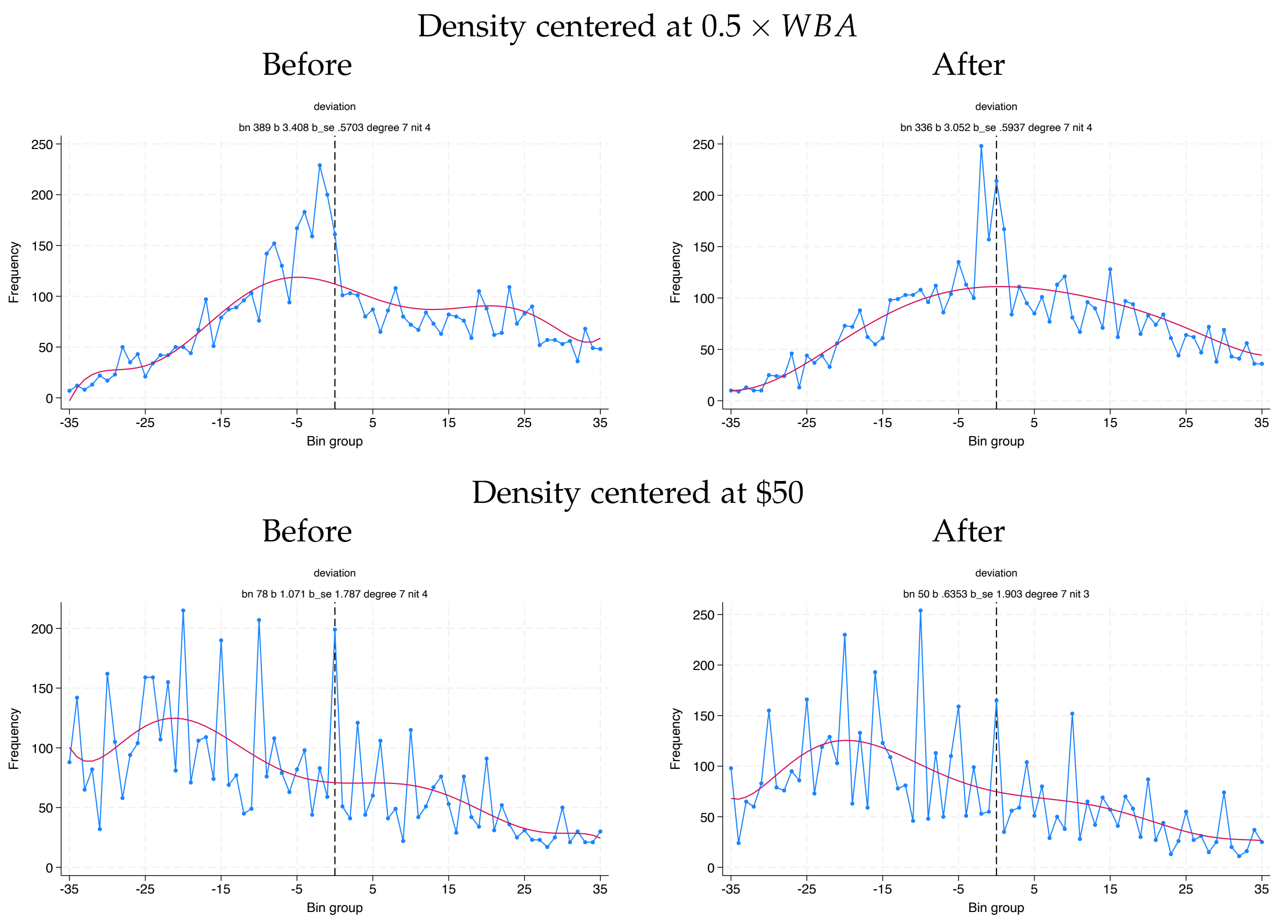
at $50 in the control group, which is not the case.

Figure E1: Centered weekly earnings density of partial-UI claimants in the treatment group.



Source: CWBH. Notes: Earnings are in dollars. Empirical earnings density in blue. Counterfactual density in red.

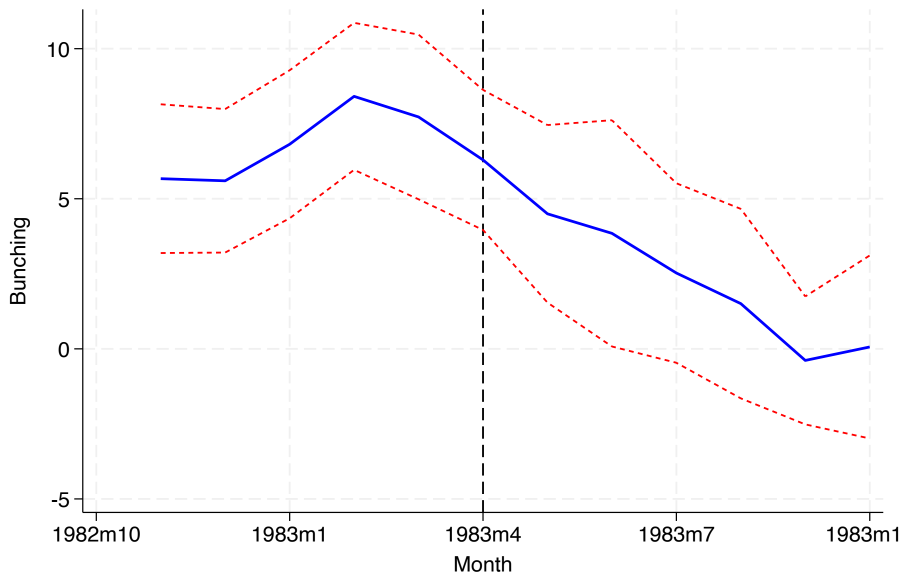
Figure E2: Centered weekly earnings density of partial-UI claimants in the control group.



Source: CWBH. Notes: Earnings are in dollars. Empirical earnings density in blue. Counterfactual density in red.

Speed of adjustment Figure E3 plots the monthly evolution of the bunching estimates at the pre-reform disregard for the treated group. The pre-reform bunching pattern disappears within a few months after the reform. This suggests that adjustment costs in the market for low-wage jobs are small.

Figure E3: Bunching before and after the April 1983 reform in Louisiana

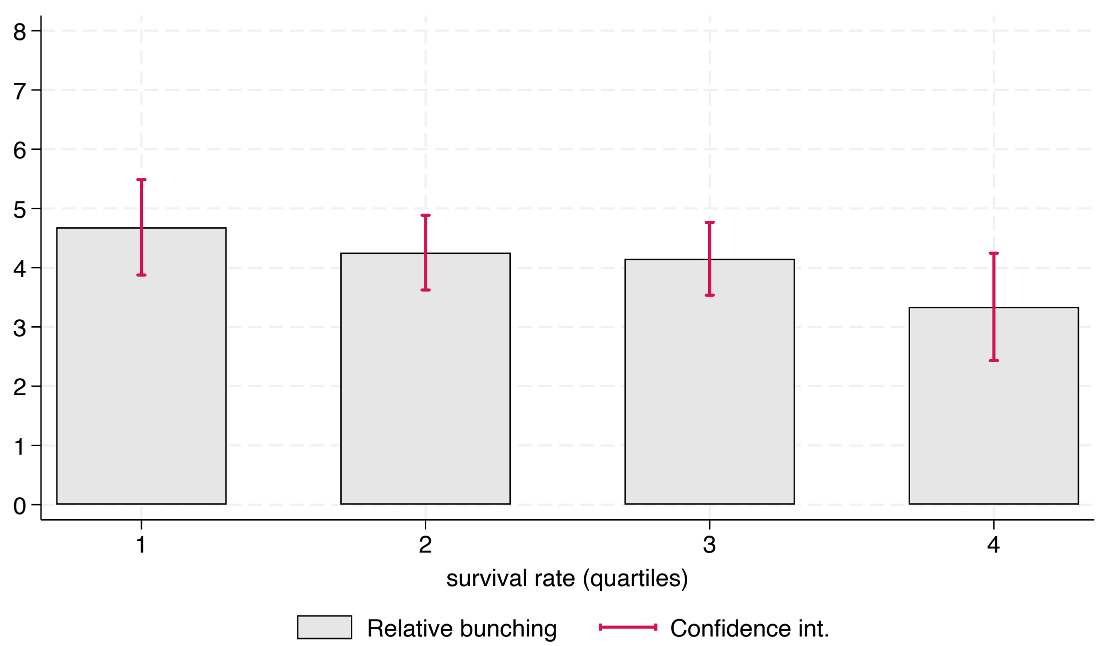


Source: CWBH, Louisiana October 1982 to September 1983. Notes: the figure plots the monthly bunching estimates at the earnings disregard (kink) before the reform (50% of WBA). The sample is restricted to workers with weekly benefit amount (WBA) greater than 100$, for whom the reform changes the disregard level. 95% confidence interval in dashed lines.

F Supplementary bunching heterogeneity test

Appendix Figure F1 provides further evidence that supports Proposition 1. Under the assumption of rational expectations for claimants' job finding rates, I predict for each worker her expected survival rate (see Section V.A for more details). I then compare bunching across the quartiles of the predicted survival rate distribution. Overall, bunching tends to decrease from the first to the fourth quartile, confirming Proposition 1. The relation between bunching and predicted survival rate is significantly negative: the test of zero slope is rejected with p-value 0.03. The empirical evidence is consistent with Proposition 1.
Note that biased beliefs in job finding rates do not compromise the heterogeneity test in Appendix Figure F1 to the extent that biased beliefs preserve the rank of workers from rational-expectations estimates. However, it may lead to an attenuation bias when estimating the bunching elasticity to the expected survival rate, as workers with high realized job finding rates tend to be over-pessimistic and workers with low realized job finding rates over-optimistic (see Mueller et al., 2021).

Figure F1: Bunching by predicted survival rate.

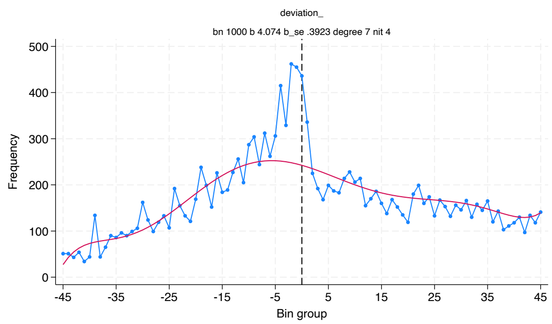


Source: CWBH for Idaho 1976-84 and Louisiana 1979-83, Q1. Notes: Excess mass at the disregard amount (kink) by quartile of predicted survival rates. Confidence interval at the 95% level in red.

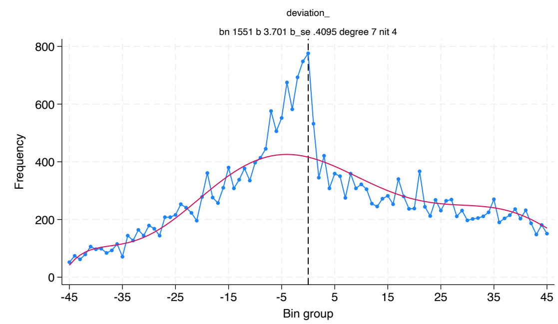
G Supplementary Figures and Tables

Figure G1: Bunching by initial potential benefit duration in Idaho: Centered weekly earnings density of partial-UI claimants.

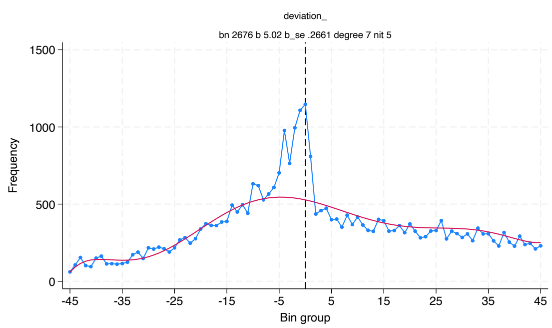
Panel A: 10-13 PBD weeks



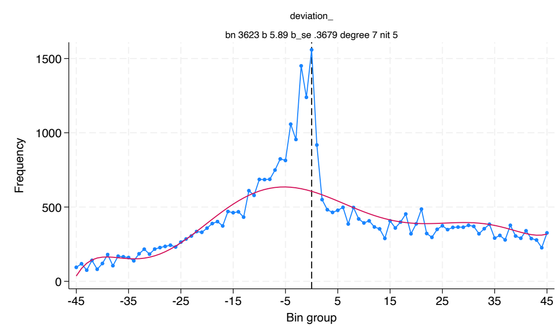
Panel B: 14-17 PBD weeks



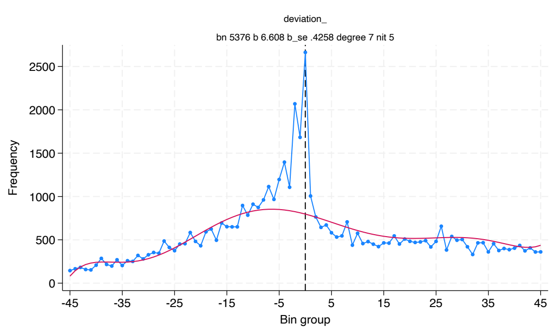
Panel C: 18-21 PBD weeks



Panel D: 22-25 PBD weeks



Panel E: 26 PBD weeks

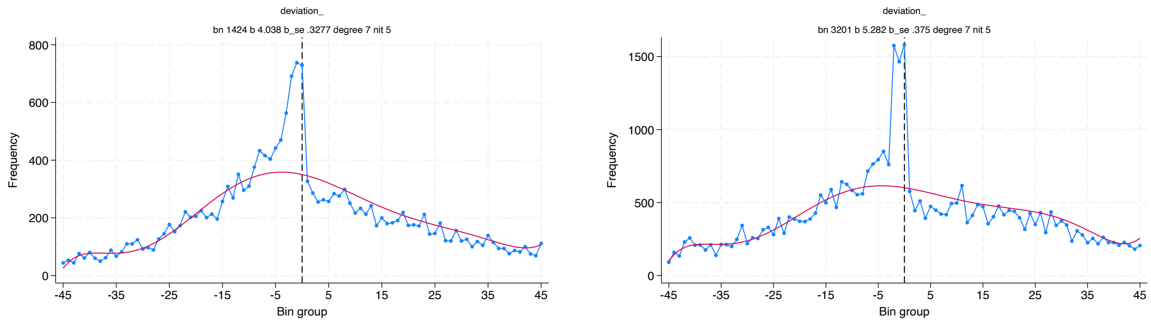


Source: CWBH. Notes: This figure plots centered weekly earnings density of partial-UI claimants, underlying the bunching estimates in Figure 3. Earnings are in dollars centered at the disregard level. Empirical earnings density in blue. Counterfactual density in red.

Figure G2: Bunching by initial potential benefit duration in Louisiana: Centered weekly earnings density of partial-UI claimants.

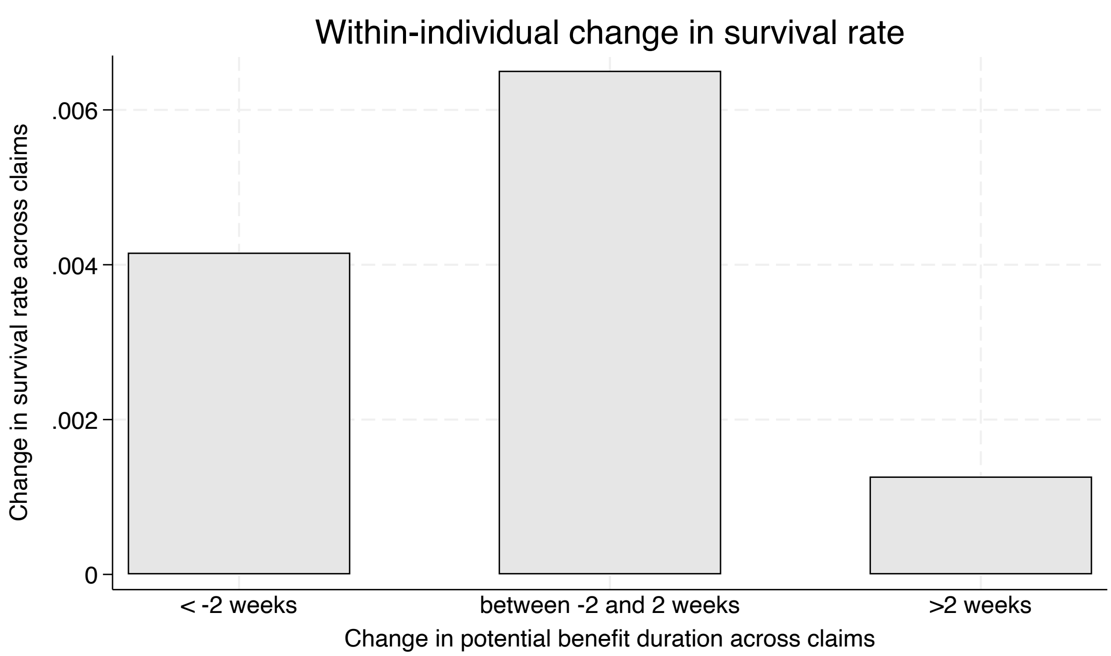
Panel A: 12-27 PBD weeks

Panel B: 28 PBD weeks



Source: CWBH. Notes: This figure plots centered weekly earnings density of partial-UI claimants, underlying the bunching estimates in Figure 3. Earnings are in dollars centered at the disregard level. Empirical earnings density in blue. Counterfactual density in red.

Figure G3: Change in expected weekly survival rate by change in potential benefit duration across claiming spells: within-worker design



Source: CWBH for Idaho 1976-84 and Louisiana 1979-1984. Notes: The figure plots the across-spell change in expected survival for claimants experiencing a large negative shock in potential benefit duration across spells (left-hand bar), a small change in PBD (center bar) or a large positive shock in PBD (right-hand bar). The average weekly survival rate is around 0.96.

Table G1: Robustness of earnings elasticities to the net-of-tax rate varying estimation parameters

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Baseline	Lower bound			Upper bound		Polynomial degree					
Bandwidth	[-5,2]	[-15,2]	[-10,2]	[-3,2]	[-5,1]	[-5,3]	[-5,2]	[-5,2]	[-5,2]	[-5,2]	[-5,2]	[-5,2]
Poly. Deg.	7	7	7	7	7	7	9	8	6	5	4	3
Idaho												
Elasticity	.187	.287	.264	0.134	0.188	0.184	0.167	0.175	0.196	0.218	0.228	0.275
s.e.	.00907	.0194	.0144	0.010	0.010	0.010	0.010	0.010	0.011	0.011	0.014	0.019
Louisiana												
Elasticity	.129	.207	.161	0.104	0.136	0.124	0.124	0.128	0.136	0.144	0.148	0.182
s.e.	.0071	.018	.011	0.007	0.007	0.008	0.008	0.009	0.008	0.009	0.008	0.010
New Mexico												
Elasticity	.0962	.	.197	0.100	0.054	0.071	0.053	0.102	0.099	0.081	0.094	0.054
s.e.	.0646	.	.17	0.040	0.050	0.065	0.080	0.070	0.066	0.053	0.047	0.056

Source: CWBH. Notes: This Table reports estimates of the earnings elasticity to the net-of-tax rate varying the estimation parameters. Column 1 recalls the results of the baseline estimation (in Table 2) for the three U.S. states: ID, LA and NM. In Columns 2 to 4, I increase the lower bound of the bunching window. In Columns 5 and 6, I increase the upper bound of the bunching window. In Columns 7 to 12, I decrease the degree of the polynomial fitting the density. Because the disregard level is around \$20 in NM, it does not make sense to consider a lower bound at -15, and the estimation results are not reported.

Table G2: Earnings elasticity estimates without first-order approximation of the marginal tax rate.

	Idaho	Louisiana	New Mexico
Elasticity	.13	.0867	.0626
s.e.	.00694	.00459	.0404
Obs.	230535	69024	31103

Source: CWBH. Notes : This Table reports estimates of earnings elasticity to the net-of-tax rate, computed with the exact identifying formula $e = -\mathcal{B}/z^* / \ln(1-\tau_t)$.

Table G3: Bunching, dynamic marginal tax rates and earnings elasticity estimates with biased beliefs correction.

	Idaho	Louisiana	New Mexico	Missouri *placebo*
Excess bunching mass ($\mathcal{B}$)	5.33	4.81	1.25	-.777
	(.315)	(.299)	(.811)	(.351)
Hazard rate ($1-p$)	.0505	.04	.0472	.
Effective Marginal Tax Rate (τ)	.588	.631	.661	.
Earnings elasticity	.171	.118	.0882	.
to net-of-tax rate (e)	(.0101)	(.00734)	(.0574)	.
Partial-UI weeks (no)	230,535	69,024	31,103	99,451
Sample Years	76-84	79-83	80-84	78-84

Source: CWBH. Notes: This table replicates Table 2 of the main text, updating the expected hazard rates to account for biased beliefs. Standard errors are in parentheses below estimates.