Partial Unemployment Insurance*

Thomas Le Barbanchon†

October 17, 2016

Abstract

Partial unemployment insurance enables claimants to keep part of their unemployment benefits when they work in low-earnings jobs. The reduction in current benefits leads to an increase in future benefits, so that forward-looking claimants are taxed according to a lower dynamic marginal tax rate than the static benefit-reduction rate. I develop a dynamic model of working claimants and use bunching in U.S. data at kinks of the benefit-reduction schedule to estimate the earnings elasticity to the net-of-tax rate. Consistent with forward-looking behaviors, I document that claimants expecting short spells bunch more. Counterfactual simulations suggest lowering the U.S. benefit reduction rate.

Keywords: unemployment insurance, temporary/part-time work, tax bunching

JEL codes : J65, H24, H31

*A previous version of this paper circulated under the title: "Optimal Partial Unemployment Insurance: Evidence from Bunching in the U.S." I thank Jérôme Adda, David Autor, Luc Behaghel, Pierre Cahuc, Raj Chetty, Marc Gurgand, Francis Kramarz, Etienne Lehmann, Manasa Patnam, Nicola Pavoni, Vincent Pons, Thomas Piketty, Roland Rathelot, Alexandra Roulet, Tom Schmitz, Arne Uhlendorff and seminar participants at Bocconi, Bristol Univ., CEMFI, Chicago Fed, Copenhagen Univ., CREST, Ecole Polytechnique, EUI, IIES, McGill, PSE, QMUL, Sciences Po, UdeM, UNIL, UZH and VATT for helpful remarks. I thank Camille Landais for sharing the data.

†Bocconi University, CEPR and IGIER, lebarbanchon@unibocconi.it

1 Introduction

The design of optimal unemployment insurance (UI) has been addressed by a large academic literature and is regularly at the center of the public debate.[1] Mostly, the literature emphasizes the issue of moral hazard associated with unemployment benefits. By reducing the net gain from employment relative to unemployment, UI reduces incentives for the unemployed to search for a job or to accept low-wage employment. To reduce such disincentives effects, many countries implement partial unemployment insurance rules. Partial UI enables claimants to keep their unemployment benefits (or a percentage of their benefits) while they work in low-earnings jobs – usually part-time or temporary work. In 2012, 12% of UI claimants in OECD countries work while on claim.[2] Despite its wide implementation, partial UI has not been fully incorporated in theoretical or empirical work on unemployment insurance. This research-policy gap is especially apparent in the U.S., where the most recent contribution is the empirical study by McCall (1996), while partial UI concerns almost 20% of claimants in some U.S. states. In this paper, I attempt to fill this gap by studying behavioral responses to the U.S. partial-UI program using administrative weekly UI data.

Partial unemployment insurance can be viewed as a form of in-work benefits. As such, it induces unemployed claimants to work in low-earnings jobs and thus affects the extensive margin of labor supply. This is confirmed by McCall (1996) who shows that part-time work is more prevalent in U.S states where partial UI is more generous. To the best of my knowledge, my paper makes a first contribution on the behavior of claimants at the intensive margin. The benefits of US claimants are reduced when their earnings while on claim are above a certain threshold, termed the *disregard*. I investigate whether conditional on working, UI claimants work less in response to the benefit reduction in the partial-UI schedule. Understanding the intensive margin is important to evaluate the efficiency of partial UI. A reduction of the benefits of partial-UI claimants, on behalf of the UI administration, could induce claimants to reduce their earnings. This resulting reduction in earnings could in turn increase the overall cost of partial UI. Using the kinks in the partial UI schedule and the bunching of claimants at the *disregard* earning levels, I estimate that the earnings elasticity to the net-of-tax rate[3] lies between 0.1 and 0.2.[4] Using this estimate, simulations show that setting the marginal benefit reduction rate above the

[1] Baily (1978) is an early contribution on optimal unemployment insurance and has been followed by Chetty (2006*a*) and Shimer and Werning (2007), among others.

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

[3] This is the elasticity of earnings to one minus the marginal tax rate.

[4] This is consistent with estimates found in previous micro empirical work using annual data (see the review of quasi-experimental estimates in Chetty (2012) or Chetty, Guren, Manoli and Weber (2011)).

disregard at 80% minimizes the benefits paid to partially unemployed claimants, while it improves their welfare compared to the current schedule.

Intertemporal aspects are key to understand the behavioral response of claimants to the partial-UI schedule. In most U.S. states, when claimants earn under the disregard level, the marginal benefit-reduction rate is zero. Then, for every dollar earned above the disregard, current benefits are reduced on a dollar-per-dollar basis: the static marginal benefit-reduction rate is 100%. However, the reduction in benefits because of the partial-UI program is not lost, it can be paid in a later week. The corresponding benefit transfers delay the potential benefit exhaustion date. As a consequence, forward-looking claimants make their labor supply decisions based on a dynamic marginal tax rate, which is lower than the static benefit-reduction rate. To account for these intertemporal effects, I develop a dynamic model of labor supply while on claim, where risk-neutral job-seekers also search for permanent jobs ineligible for partial UI. This model enables me to derive the analytical expression of the dynamic marginal tax rate, which depends not only on the discount factor and on the marginal benefit-reduction rate, but also on the claimant's expected probability to find a permanent job. 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-transfer mechanism and her dynamic marginal tax rate is larger, closer to the static benefit-reduction rate. I therefore estimate a hazard model of exiting the UI registers, and calculate that claimants with rational expectations have on average a dynamic marginal tax rate of approximately 60%. My dynamic model of working while on claim shows that bunching at the kink of the partial-UI schedule identifies the earnings elasticity to the net-of-tax rate. Then I extend the bunching formula of Saez (2010) to contexts where benefits can be transferred to the future.

In my model, I assume that increasing earnings in partial-UI jobs has no effect on the probability of finding permanent jobs (stepping-stone or crowding-out effects). This is reasonable as there is no empirical evidence on these effects in the US (McCall, 1996) and European studies do not reach a consensus.[5] Nevertheless, I discuss how my identification strategy accommodates for these stepping-stone or crowding-out mechanisms. I also discuss how risk-aversion affects my identification strategy.

I compute bunching and elasticity estimates using UI administrative data from four U.S. states: Idaho, Louisiana, New Mexico and Missouri. The data come from the Con-

[5]European studies on partial UI estimate the effects of part-time jobs on the probability to find permanent jobs (Kyyrä, 2010; Caliendo et al., 2012; Kyyrä et al., 2013; Fremigacci and Terracol, 2013; Godoy and Røed, 2014).

tinuous Work and Benefit History (CWBH) project.[6] I find substantial bunching at the disregard level. In Idaho and Louisiana, the excess mass at the disregard is five times the population density that would work at this level 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 the dynamic marginal tax rate, rather than to the static 100% marginal tax rate.[7]

I perform two placebo exercises showing that the bunching identified in the data is not an artifact of labor legislations or social norms unrelated to the partial UI schedule. First, I leverage the fact that disregard amounts vary across states. In particular, I confirm that there is no bunching in Missouri at the disregard levels prevailing in Idaho and Louisiana. Second, I use a policy shock in Louisiana, when the disregard amount was changed for claimants with high previous wages. The location of bunching for this treated group changed from the high previous disregard level to the low new disregard amount.

Bunching heterogeneity is consistent with the fact that claimants react to the benefit-transfer mechanism of the partial-UI program. Claimants with longer potential duration of benefits - i.e. initial number of benefit days if totally unemployed - bunch more in my data. As predicted by the dynamic model, they have lower incentives to use the partial-UI program to delay the benefits exhaustion date. I also verify that bunching estimates are larger for claimants with a low propensity to remain on the UI registers, such as claimants expecting to be recalled to their previous employer (Katz and Meyer, 1990*b*). Additionally, I study the evolution of bunching with unemployment duration, holding the sample of claimants constant to avoid composition effects. The theoretical model predicts that bunching should decrease over the claim.[8] I find a moderate decrease in bunching estimates, close to the exhaustion date.

Finally, I consider the program of the UI administration that chooses the benefit-reduction rate to maximize the welfare of partially-unemployed claimants while maintaining its costs (benefits payments) below a certain exogenous level. I focus on the marginal benefit-reduction rate over the disregard level, as the behavioral effects of this partial-UI parameter are well identified with the bunching strategy above. After deriving the first-order condition of the UI administration, I perform simulations using the elasticities estimated above. Simulations suggest that the UI administration could actually minimize

[6] I thank Camille Landais for sharing the data. See Moffitt (1985) and Landais (2014) for more details about the data.

[7] However this is not a definitive test of the dynamic aspects of my model, as adjustment costs/search frictions for low-wage jobs could also explain why myopic claimants work for earnings above the disregard.

[8] Claimants have a higher probability to receive transferred benefits when they are already at the end of their claim.

its costs by setting the benefit-reduction rate at 80%. Switching from the current schedule with a 100% benefit-reduction rate to this level would also increase the welfare of partially unemployed claimants. When benefit-reduction rates are as high as in the current schedules, job-seekers inefficiently reduce their earnings in low-earnings jobs. Their behavioral response to a decrease in benefit reduction rate is so strong that it overwhelms the initial mechanical increase in benefit payments and leads to a decrease in UI costs.

My paper contributes to the literature on partial UI. It is the first paper that studies behavioral response at the intensive margin and computes the related efficiency costs. In the U.S., McCall (1996) and the early contributions of Holen and Horowitz (1974) and of Kiefer and Neumann (1979) document the behavioral response at the extensive margin.[9] In Europe, Kyyrä (2010), Caliendo et al. (2012), Kyyrä et al. (2013), Fremigacci and Terracol (2013) and Godoy and Røed (2014) study the effects of partial-UI jobs on regular employment.

I contribute to the literature on the estimation of intensive elasticities using bunching triggered by kinks in tax or benefit-reduction schedule. My paper is the first to consider bunching of UI claimants. Saez (2010) and Chetty et al. (2013) study bunching of workers due to kinks in the EITC schedule. Gelber et al. (2013) document bunching among U.S. old-age wage-earners at the kink of the Social Security Annual Earnings Test (SSAET). I also contribute to the bunching literature by adapting the Saez (2010) formula when benefits or taxes can be transferred to the future. In a different context, Brown (2013) studies bunching at the legal retirement age taking into account that delaying retirement increases future annuities.[10]

The paper is organized as follows. In Section 2, I describe the U.S. partial unemployment insurance program. In Section 3, I develop a job-search model of a claimant working while on claim and derive the identification result. In Section 4, I detail the different steps of the estimation procedure and the data. In Section 5, I present my main estimates of bunching and the corresponding earned income elasticities to the net-of-tax rate; I also perform various placebo tests including a difference-in-difference analysis in Louisiana. In Section 6, I document the heterogeneity of bunching and its evolution along the claim. In Section 7, I perform the normative exercise. Section 8 concludes.

[9] Munts (1970) is an early descriptive contribution on partial UI in the U.S.

[10] le Maire and Schjerning (2013) consider dynamic aspects in income tax schedule, but they specifically model income shifting by the self-employed. Gelber et al. (2013) discuss intertemporal aspects of the U.S. Social Security Annual Earnings Test. As in my case, 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 kink in the SSAET schedule.

2 Institutional background

In the U.S., when unemployment insurance (UI) claimants work while on claim, they are eligible for partial unemployment benefits, provided that they do not earn more than a maximum amount of labor income per week.[11] This maximum amount is usually set as a fraction of the weekly benefit amount (WBA), which is the unemployment benefits (UB) payment when claimants do not work, i.e. total unemployment benefits. Partial-UI claimants are paid their weekly benefit amount when their weekly earnings are below the state-specific *disregard* threshold. When partial-UI claimants earn between the disregard and the maximum amount, their current benefits are reduced by their earnings minus the disregard. The static marginal benefit-reduction rate is then 100%. Table 1 displays the parameters of partial-UI rules in the late 70s and early 80s for the four states analyzed in this paper. As partial-UI rules have hardly changed since then, our discussion of the institutions refers to both current rules and rules of the late 70s and early 80s, except when we discuss nominal amounts.[12] The most generous state is Idaho: it has the highest maximal wage threshold and its disregard is also high. The three other states have a maximal amount around the WBA (in practice less than 1.1 times the WBA). Missouri is the least generous state. Its disregard is stated in dollar values and, actually, does not exceed 10% of the WBA. In Louisiana, the disregard was reduced in April 1983. I will use this policy shock as a placebo exercise in a difference-in-difference analysis.

Table 1: 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 current disregard and maximum levels. Taking into account inflation, $10 (resp. $50) in 1978 represent around $37 (resp. $185) in 2016.

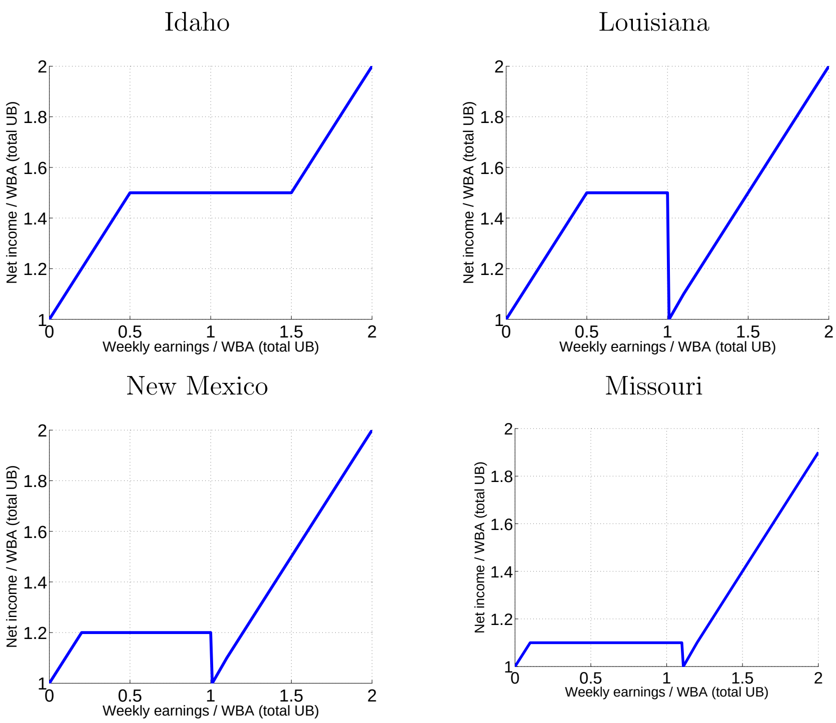
Figure 1 illustrates the partial-UI schedules in Idaho (ID), Lousiana (LA), New Mexico (NM) and Missouri (MO). I plot the weekly net income (earnings plus UB payments)

[11] All components of labor income are considered in the computation. The only exceptions are payments for jury service in New Mexico and wages from "service in the organized militia for training or authorized duty from benefit computation" in Missouri.

[12] See Chapter 3 of the 2013 DOLE booklet "Comparison of the State Unemployment Laws"

against the weekly earnings while on claim. I normalize earnings and UB payments by the WBA, as the maximal amount and the disregard are expressed as a fraction of the WBA for three of the four states. The figure illustrates that the schedule is kinked at the disregard amount. From a static point of view, there are no incentives to work for a wage just above the disregard, as the net income is essentially a plateau above that level. The graphics also illustrate the notch at the maximal amount in Louisiana and New Mexico (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 threshold (Kleven and Waseem, 2013). Because of data limitations, I will not analyze the claimants' behavior around notches. The incentives to claim jump discontinuously at the notch value, so that individuals above the notch should leave the UI registers, and hence my data. As will become clear below, I take advantage of the absence of kinks in Missouri at the disregard level prevailing in Idaho and Louisiana ($0.5 \times WBA$) to perform a placebo exercise.

Figure 1: Partial-UI schedules from 1976 to 1984



Source: U.S. Department of Labor, "Significant Provisions of State Unemployment Insurance Laws." Notes: X-axis corresponds to weekly earnings divided by the weekly benefit amount (UB paid in case of total unemployment). Y-axis corresponds to the net income (earnings + UB payments) divided by the WBA. The different panels show the theoretical schedules. For Louisiana, I plot the schedule before April 1983. For Missouri, I consider a claimant whose WBA is $100.

UI claimants pay income taxes on their labor earnings.[13] Before 1979, there were no federal taxes to be paid on unemployment benefits. Since 1979, unemployment benefits have been taxable for single tax filers with income over \$20,000 and for married filers with income over \$25,000. The thresholds were lowered in 1982 to respectively \$12,000 and \$18,000. The difference between income tax thresholds for labor earnings and for unemployment benefits may affect the relative gains of working while on claim. However, it is very unlikely that the weekly disregard level of partial UI corresponds to another kink in the annual income tax schedule or welfare system. This ensures that my identification strategy below is robust to the existence of other incentives caused by the whole tax and benefit system.

I now turn to the dynamic aspects of the partial-UI schedule. At the beginning of each claim, the UI administration computes a Weekly Benefit Amount (WBA) and a Potential Benefit Duration (PBD), which both depend on past earnings. The PBD typically varies between 10 and 26 weeks (see the Appendix for more details). The product of the WBA and the PBD is called the (total) benefit entitlement which we 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 will 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. Working while on claim is thus a way to delay the benefit exhaustion date. In principle, there is one limitation to the possibility to delay, 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 of the benefit year. Consequently, we abstract in the remainder from any horizon effects of the benefit year rule.

Except for the above earning thresholds, there is no other specific eligibility condition for partial UI. Claimants must only meet the usual UI eligibility requirement (described in the Appendix). 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. Also, individuals whose hours have been reduced at their current workplace are eligible for partial UI, so 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.[14]

[13] In practice, I expect that many UI claimants will be below the minimum income threshold of the income tax schedule.

[14] In the CWBH data, I can only distinguish between claimants taking up new jobs and claimants with reduced hours (Short Time Compensation) in Louisiana since 1982. From 1982 to 1984, only 15.7% of

The partial-UI rules described above 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 increased the PBD of claimants. They are described in more details in the Appendix.

Eligibility to partial UI is based on reported earnings. There is thus scope for claimants manipulating their reports. However the UI administration takes action to limit false statements. The UI administration performs random audits of claimants' declarations.[15] If fraud is detected, it can be severely punished as a Class VI Felony.[16]

3 Theoretical model

In this section, I develop a dynamic model of working while on claim that incorporates the dynamic aspects of the partial-UI program. The model features reasonable assumptions from both the job-search literature and public finance literature on bunching: frictions in the search for permanent jobs, liquidity constraints and quasi-linear utility. 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. The identification of the earned income elasticity then follows a modified bunching formula (Saez, 2010).

3.1 Setup

I consider an infinitely lived individual i, claiming benefits from period 0 on. 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 a 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. I discuss the implications of introducing adjustment costs in the low-wage job market in Section 3.7.

The job-seeker's earnings in the low-wage job in period t are denoted z_t. In line with Saez (2010), I do not make any distinction between wage rates and hours as those

partial-UI weeks in Louisiana concerned claimants with reduced hours.

[15]The UI administration currently cross-checks W-2 and new hires declarations of employers with claimants reported earnings.

[16]Criminal action may result in up to 2 years in prison and fines up to $150,000 for each false statement.

different components are not observed in the data.[17] 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. In the baseline model, I assume that the job-seeker is risk-neutral. I discuss this assumption in Section 3.5. Then it is convenient to parametrize the period utility 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 talent/taste parameter - smoothly distributed in the population - and e is my parameter of interest: it captures the earnings elasticity to the net-of-tax rate. I introduce heterogeneity in e in Section 3.4. As already noted by 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 (see below). However it makes the interpretation of the parameter difficult, as n_i may capture both individual taste for work and ability. In the remainder, I choose to refer to n_i as individual talent.

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. 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. I discuss potential stepping-stone effects or job-search crowding-out effects of low-wage jobs in Section 3.6. I introduce individual heterogeneity in p in Section 3.4. 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 talent in low-wage jobs.

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 is registered and does not work at all (total unemployment), the job-seeker

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

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.[18] 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 may still work for a low-wage job. The partial-UI schedule $T(.)$ is defined as:[19]

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

where z^* is the amount of disregard. 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 above, I will abstract from the second kink as data limitation prevents me to analyze behaviors around the maximal earnings amount.[20] This second kink does not affect my identification strategy that is local and 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

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

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

[19] 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.

[20] 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.

The first constraint of the program is the current budget constraint. I assume that agents cannot lend or borrow, as UI claimants are likely to be low-skilled workers who are credit-constrained. The second constraint captures the endogenous entitlement reduction (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.[21]

3.2 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 $B_t = 0$.[22] 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 the Appendix. For all z_t, such that $T(.)$ is differentiable at z_t, the first order condition is:

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

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 the Appendix. The second term of Equation (4) then simplifies to:

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

[21] 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. Note that I also assume that the value of a permanent job does not depend on current entitlement. The UI rules enable the former claimants who have left the UI registers to work in a new job, to claim past remaining entitlement in the same benefit year if they lose their new job. However, in the data, this happens in very few occasions, so I do not model this possibility.

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

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 talent 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 talent $n^* = z^*$, i.e. the 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 talent $n^* + \delta n(t)$, such that all individuals with talent strictly above $n^* + \delta n(t)$ earn strictly more than the disregard z^*. Such individuals have their current benefits reduced and 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:

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

Equation (8) illustrates that the upper threshold depends on the dynamic marginal tax rate and consequently on its determinants, such as the time period. More fundamentally, it depends on the time to exhaustion, which is endogenous. However, job seekers with a talent close the upper threshold delay their exhaustion date by only one period, so that $\bar{t}$ is known.[23] 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:[24]

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

[23]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.

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

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

3.3 Identification strategy

Suppose that the 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:

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

where the first equality is obtained thanks to Equation (9) and the second equality uses an approximation of the integral of a continuous function. The excess bunching thus identifies the difference in talent 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 talent n^*, a first-order approximation of Equation (8) yields the following expression for the elasticity:[25]

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

The main difference between the static bunching formula of Saez (2010) and the above expression is the definition of the marginal tax rate. In my setting, the dynamic marginal tax rate depends on the discount factor and the probability to exhaust the initial benefit entitlement.[26] I explain in Section 4.1.2 how the dynamic marginal tax rate τ_t can be estimated. Consequently, Equation (11) identifies the earnings elasticity to the net-of-tax rate.

Let me define $\mathcal{B} = \frac{1}{f(z^*)} \int_t \int_{n^*}^{n^*+\delta n(t)} f(n)dndG(t) = \int_t \mathcal{B}_t dG(t)$ 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.

[25] 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 $\tau_t << 1$. Indeed, I estimate below that on average $\tau_t = 0.6$.

[26] Note that one key condition for identification is that claimants are not myopic. Myopic individuals have no incentives to work above the disregard level. Then bunching is not informative about the earnings elasticity to the net-of-tax rate.

3.4 Heterogeneity in bunching

A direct consequence of the bunching formula (11) is that bunching decreases as the claiming duration increases. As the job-seeker approaches the exhaustion date, she is more likely to benefit from the UB transfers and the dynamic marginal tax rate decreases over the spell. This yields the following Proposition:

Proposition 1 (Bunching over the spell) *Bunching decreases over the spell.*

Proof: Bunching is time-dependent through the dynamic marginal tax rate, which decreases over time: $\frac{d\tau_t}{dt} = \log(\beta p) \times (\beta p)^{\bar{t}-t-1} < 0$.

I next consider the predictions of the model when job-seekers differ in their propensity to find a permanent job (heterogeneity in $1-p$). Intuitively, 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 2 (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: $\frac{d\tau_t(p)}{dp} = -(\bar{t}-t) \times (\beta p)^{\bar{t}-t}/p < 0$.

In the Appendix, I further discuss the implications of heterogeneity in the permanent job finding rate. I show that, when there is heterogeneity in p, Proposition 1 is not only valid within a p-strata, but also on average. Dynamic selection over the spell implies that individuals still claiming after a long spell have also a higher probability p to remain claiming. I also introduce heterogeneity in the structural elasticity e. I then show that the model still predicts that bunching decreases over the spell within a p-strata. However, bunching does not necessarily decrease with the permanent job finding rate, as the comparison across p-strata may be confounded by different average structural elasticities of claimants with talent at the lower threshold.

3.5 Risk-aversion

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

$$u_i(c, z; n_i) = \frac{c^{1-\sigma}}{1-\sigma} - \frac{n_i}{1+1/e}\left(\frac{z}{n_i}\right)^{1+1/e} \tag{13}$$

where σ is the coefficient of relative risk-aversion. I show in the Appendix that the FOC then writes:

$$1 - T'(z_t)\left(1 - (\beta p)^{\bar{t}-t-1}\left(\frac{c_t}{c_{\bar{t}-1}}\right)^{\sigma}\right) = \left(\frac{z_t}{n_i}\right)^{1/e}(c_t)^{\sigma}. \qquad (14)$$

Compared to Equation (6), 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 the consumption in the last week of claim. Because consumption drops at the end of the claim, the ratio of current to future consumption is greater than 1 and the dynamic marginal tax rate of risk-averse job-seekers is smaller. 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 elasticity e. In the Appendix, I obtain a first-order approximation of the ratio of the estimates of e with or without risk-aversion: $1 - \sigma\frac{\Delta c}{c_t}\frac{1}{\tau_t}$ where $\Delta c/c_t$ is the relative change in consumption between the current period and the last week of claim before exhaustion. Gruber (1997) 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 around 1 (Chetty, 2006*b*) and τ_t around 0.5 (see below), the ratio of elasticity estimates is 1.12. Anticipating my estimation results - average elasticity around 0.15, the elasticity taking into account risk-aversion would be around 0.18. Given the small difference in elasticity estimates, I abstract from risk-aversion in the remainder of the paper.

3.6 Stepping-stone/crowding-out effects

In this section, I allow the probability of finding a permanent job to depend on earnings in low-wage jobs: $1-p(z)$. I detail in the Appendix the solution of the job-seeker program modified accordingly. I obtain the following FOC:

$$\underbrace{u_c(c_t, z_t; n_i)\left(1 - T'(z_t)\right)}_{(I)} + \underbrace{\beta p(z_t)T'(z_t)U'(B_{t+1}; n_i)}_{(II)} - \underbrace{\beta p'(z_t)\left(W - U(B_{t+1}; n_i)\right)}_{(III)} = -u_z(c_t, z_t; n_i). \qquad (15)$$

Compared to Equation (6), 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 ($p' > 0$).

When there is a kink at the disregard level in the partial-UI schedule, the earnings elasticity to the net-of-tax rate is still identified under two specific assumptions. 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 talent only through earnings. Details of the proof are in the Appendix. The second assumption is strong. However, it seems very likely that *marginal* stepping-stone/crowding-out effects are negligible in practice. McCall (1996) in the U.S. and Kyyrä (2010), Caliendo et al. (2012), Kyyrä et al. (2013), Fremigacci and Terracol (2013) and Godoy and Røed (2014) in European countries find small effects of partial-UI jobs on permanent employment.

3.7 Adjustment costs and matching frictions for low-earnings jobs

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. For example, I can assume that there is a fixed cost ϕ to adjust from total unemployment to work in low-wage jobs. Totally unemployed claimants only accept jobs that deliver a net gain exceeding the fixed cost, i.e. jobs around their optimal earnings. Such extensions show that optimization frictions typically smooth bunching. Bunching then no longer identifies the structural elasticity. However it is still informative about the behavioral costs of the partial-UI program.

Frictions also alter the evolution of bunching over the spell. First, they may lead to an increase in bunching during the first claiming weeks. The intuition is as follows. Totally unemployed claimants take the first job that delivers a positive net utility gain. As long as they work in that first job, they will only switch to a new job if it increases their utility, or equivalently if it is closer to their optimal labor supply. This means that Proposition 1 may not hold during the first weeks of claim. Second, frictions may also smooth the decrease in bunching later in the claim (see Proposition 1). Consider, for example, a claimant whose current job is close to her notional earnings, i.e. close to her current optimal earnings if there were no frictions. Notional earnings increase over the spell for claimants working above the disregard level. Claimants facing frictions will increase their earnings if the associated gains exceed the fixed adjustment cost. However, when claimants are close to their notional earnings, switching gains are typically second-order - the derivative of the utility function with respect to the earnings level is zero at the notional earnings. Consequently, inertia is likely to be strong. This is also the case for bunchers, so that debunching is likely to be slow, and the data may not fully support Proposition 1.

4 Estimation and Data

4.1 Estimation

The estimation procedure follows two main steps. First, I estimate the excess bunching $\mathcal{B}$, i.e. the numerator of Equation (12). Second, I estimate the expected dynamic marginal tax rate in the denominator of Equation (12).

4.1.1 Excess bunching estimation

Following the procedure of Chetty, Friedman, Olsen and Pistaferri (2011), 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:

$$C_j \left(1 + \mathbb{1}[j > \overline{R}] \frac{\hat{B}_N}{\sum_{j>\overline{R}} C_j}\right) = \sum_{k=0}^{q} \beta_k (Z_j)^k + \sum_{i=-\underline{R}}^{\overline{R}} \gamma_i \mathbb{1}[Z_j = i] + \epsilon_j \tag{16}$$

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.[27] 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.

[27]Because $\hat{B}_N$ depend 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.

Equation (16) 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:

$$\hat{\mathcal{B}} = \frac{\hat{B}_N}{\sum_{j=-\underline{R}}^{\bar{R}} \hat{C}_j / (\underline{R} + \overline{R} + 1)}. \tag{17}$$

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

4.1.2 Dynamic marginal tax rate

Recall that the dynamic marginal tax rate for a claimant working just above the disregard in period t, is defined as:

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

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 $h_i = h \exp(\beta X_i)$ and predict the survival rate $\hat{p}_i$ for each individual. The details of the hazard model are reported in the Appendix. Note that, by using predicted rates, claimants are assumed to have rational expectations about their compensated unemployment duration.[28]

I obtain an estimate of the denominator of Equation (12) by averaging, over all individuals and weeks, the predicted dynamic marginal tax rate and by multiplying this average by the average of observed disregards. The standard errors of the elasticity estimate are obtained by the delta method.

4.2 Data

I use individual panel data from the Continuous Wage and Benefit History project (CWBH). The project collected weekly claims for a random subsample of UI claimants in the U.S., and the resulting data has the unique advantage of including the weekly earnings that claimants report to the UI administration and the consecutive UB payments. I can thus characterize whether claimants are partially unemployed. The major drawback of the data set, that it covers the late 70s and early 80s, is mitigated by the fact that the partial-UI schedules have been hardly reformed since then. The data cover four U.S. states – Idaho,

[28]Unobserved heterogeneity in the survival rate would bias the elasticity estimate, if it is correlated with unobserved heterogeneity in the elasticity.

Louisiana, Missouri and New Mexico – and include all relevant information about the claim: weekly benefit amount, total entitlement, and highest quarter earnings. Socio-demographics characteristics are also available (gender, age, education, ethnicity, past firm status and industry, past occupation). In addition, the data set includes survey information about recall expectations for a subsample of claimants.

Table 2 reports descriptive statistics of UI claimants by state. The weekly benefit amount is around \$100 (current dollars) and the average replacement rate is between 40% and 50%. The potential benefit duration (PBD) is actually 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 (excluding the waiting week). The last line of Table 2 reports the share of claimed weeks with positive reported earnings. The share of partial unemployment weeks amounts to 17.6% in Idaho where the partial-UI schedule is very generous: claimants can be partially unemployed up to 1.5 times their WBA. In Louisiana and Missouri, respectively 6.1% and 8.1% of claimed weeks concern partially unemployed claimants. The corresponding share in New Mexico is 2.5%, reflecting the fact that the disregard amount is quite low in this state. Appendix Table A2 reports the socio-demographic characteristics of UI claimants by state. Appendix Table A3 compare claimants with at least one week of partial unemployment during the benefit year to claimants on total unemployment. Partial-UI claimants are older, more educated and more frequently whites. Their previous employer is more frequently a firm from the private sector and operates in construction and manufacturing industries. Partial-UI claimants are less frequently in upper occupations, such as professionals, technicians or managers. Their pre-unemployment wage and their entitlement duration are higher.

Table 2: Descriptive Statistics

	Idaho	Louisiana	New Mexico	Missouri
Years	76-84	79-84	80-84	78-84
Inflow (nb)	91,162	95,675	62,030	78,065
Pre-U weekly wage (current dollars)	337	316	265	247
Weekly benefit amount (WBA)	96	131	94	88
Replacement rate	.425	.471	.405	.455
Potential duration (weeks)	28	35.8	34.7	32.7
Actual UB duration	15.3	18.1	15.0	15.6
Share of partial-UI weeks	.176	.067	.025	.085

Source: CWBH. Notes: Means are computed over the sample of claimants, except the share of partial UI computed over all claiming weeks.

5 Main results

5.1 Earned income elasticity to the net-of-tax rate

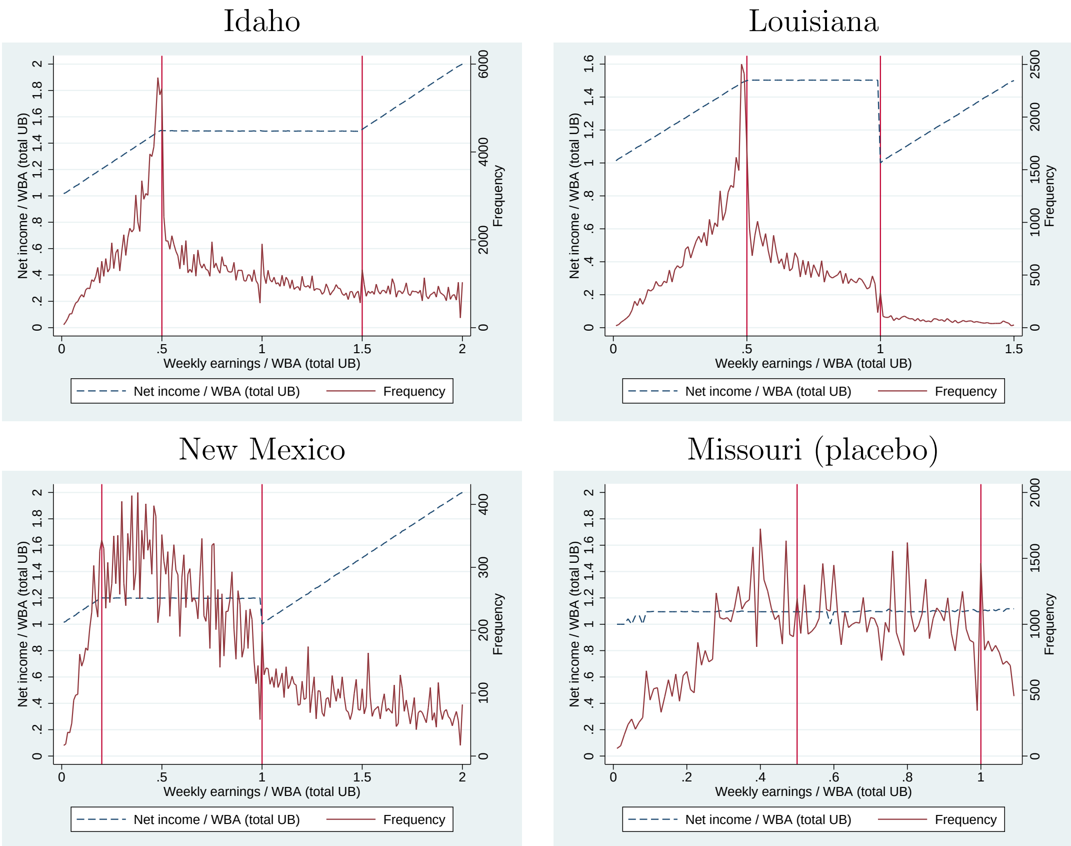
Figure 2 displays the weekly earnings density reported by UI claimants together with the empirical partial-UI schedule for the four different U.S. states. In line with the partial-UI rules in Idaho, Louisiana and New Mexico, I normalize the weekly earnings by the weekly benefit amount (unemployment benefits in case of total unemployment). The empirical schedules, 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 clearly display bunching at the level of the disregard (50% of the weekly benefit amount). On the upper right Louisiana 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 obvious (lower left panel). The lower right panel illustrates a placebo test. In Missouri, the level of disregard is so low at a mere $10 that the schedule is totally flat at the $0.5 \times WBA$ threshold. There is indeed no bunching at this placebo level. Thus the bunching observed in Idaho or Louisiana is unlikely to be an artifact of other labor legislations or norms, or hour constraints according to which claimants take some part-time jobs that provide roughly one fourth of their previous wages (given that the average replacement rate is around 50%).

Another important feature of the earnings distribution in Figure 2 is the substantial fraction of claimants working for earnings above the disregard level. This observation is consistent with claimants reacting to the dynamic marginal tax rate, rather than to the static 100% benefit-reduction rate. Indeed, myopic claimants have no incentive to work above the disregard level. However this observation is not a definitive test of the dynamic aspects of my model, as matching frictions for low-wage jobs could also explain why myopic claimants work for earnings above the disregard.[29]

To conduct the bunching estimation, I consider earnings in absolute levels and I center the earnings density at the disregard level. Note that because the disregard is specified as a fraction of weekly benefit amounts, individuals with different weekly benefit amounts are subject to different disregard levels. The resulting centered earnings densities are plotted in Figure 3. Observations are grouped in bins of 1 dollar width. Figure 3 confirms the

[29] Of course, claimants, who are not aware of the partial UI schedule, would also work above the disregard.

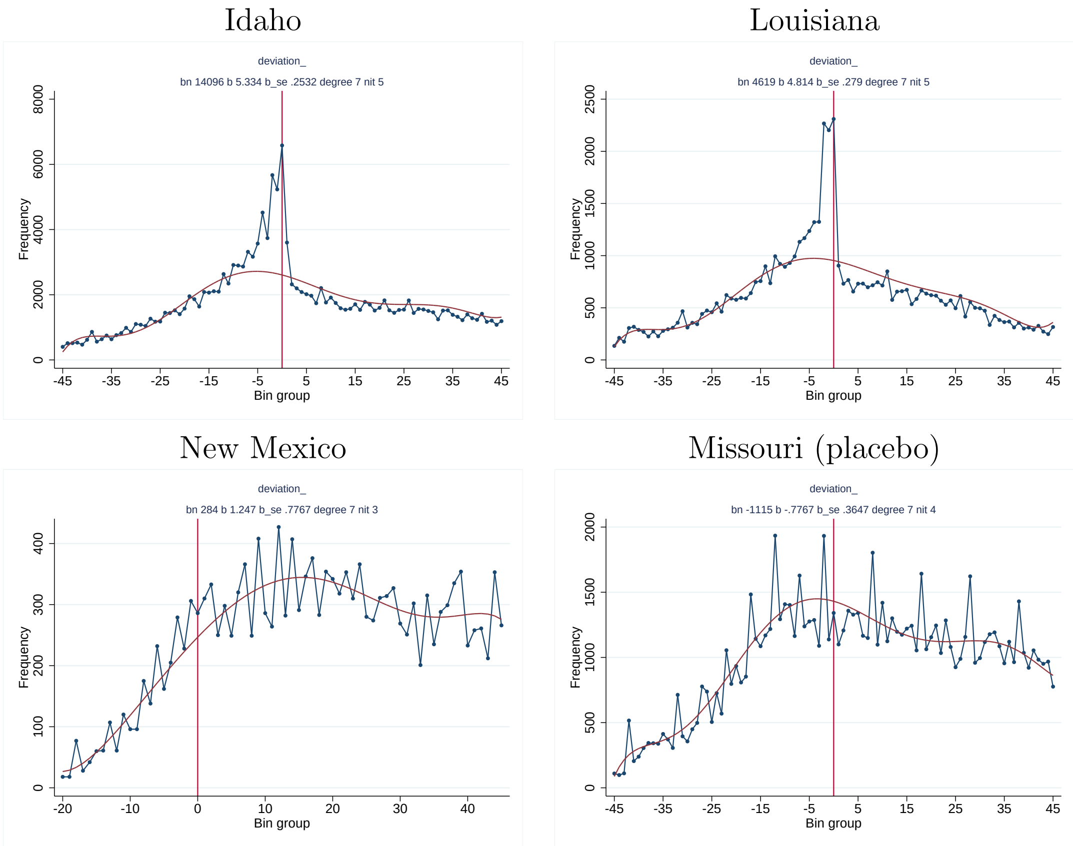
Figure 2: Weekly earnings density and empirical schedule of partial UI.



Source: CWBH. Notes: X-axis corresponds to weekly earnings divided by the weekly benefit amount (UB paid in case of total unemployment). Earnings density is plotted in red. Corresponding frequencies are on the right Y-axis. The partial-UI schedule corresponds to the blue dashed line. It plots the net income divided by the WBA. Red vertical lines show the kinks and notches of the partial UI schedule, except for Missouri (the placebo state).

bunching patterns observed in Figure 2. It also reveals some periodicity in the earnings distribution in Missouri. Claimants report earnings that are multiples of ten dollars.[30] Each panel also displays the counterfactual density in red, which is estimated along the lines of Chetty, Friedman, Olsen and Pistaferri (2011).[31]

Figure 3: 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.

Table 3 reports the results of the estimation for each state (in columns). 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, had the kink disappeared. Excess bunching is highly statistically significant. In New Mexico, excess bunching is much lower and not statistically significant. In Missouri, there seems to be a missing mass of claimants at the placebo threshold level. As a consequence, the placebo test confirms that in the absence of a kink there is no excess mass at the threshold level. The periodicity in the

[30]The periodicity appears only in Missouri because it has a large mass of individuals at the maximal WBA. Then the disregard is the same for a large fraction of claimants.

[31]The procedure fits a polynomial of degree 7. The bandwidth is such that $-\underline{R} = -5$ and $\bar{R} = 2$.

earnings density in Missouri 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.

To compute the dynamic marginal tax rate, I need to estimate the expected exhaustion probability in the sample. I estimate an exponential model of the hazard rate out of the unemployment registers. I include in this model various characteristics of claimants (gender, age, education and ethnicity), claim characteristics (WBA, PBD, recall expectations) and year fixed effects. Detailed estimation results are reported in the Appendix. The weekly hazard rates vary between 3% and 4% across states. Predicted hazard rates are then used to compute the expected UB exhaustion probability at each date, which takes into account the remaining number of entitled weeks. The average dynamic marginal tax rate is around 54% in Idaho and Louisiana; it is larger in New Mexico, where it amounts to 60%.

Finally, I use the identification relation to compute the earned income elasticities to the net-of-tax rate. I obtain statistically significant elasticities in Idaho and Louisiana, respectively 0.19 and 0.13. The elasticity in Missouri 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.[32] 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.[33] My results 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, Guren, Manoli and Weber (2011)). Note that I cannot be certain that this elasticity is purely driven by behavioral responses from the supply side of the labor market. In the next section, I present another placebo test that indicates that bunching is actually related to the partial-UI schedule.

I can compare my elasticities to estimates obtained specifically with bunching estimators. Most recent papers (Saez, 2010; Chetty et al., 2013; Chetty, Friedman, Olsen and Pistaferri, 2011; le Maire and Schjerning, 2013; Bastani and Selin, 2014) find bunching at the kinks of the annual income tax schedule among the self-employed, but very little bunching among wage-earners (especially among those at the bottom of the wage distribution, as in my case). My sample comprises former wage-earners – indeed, this is an eligibility condition for claiming. While I cannot rule out that some of them are actually

[32]Robustness results are reported in the Supplementary Table A4 in the Appendix

[33]Robustness results are reported in the Supplementary Table A5 in the Appendix.

self-employed while on claim, it is very likely that former wage-earners also work as wage-earners while on claim. This comparison suggests that my elasticity estimate is higher than most results from the bunching literature, with the notable exception of Gelber et al. (2013). Indeed, those authors 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 0.23, which is in the range of my estimates.

Table 3: Bunching, dynamic marginal tax rates and earnings elasticity estimates.

	Idaho	Louisiana	New Mexico	Missouri
Years	76-84	79-83	80-84	78-84
Partial-UI weeks (nb)	230,535	77,602	31,103	91,147
Disregard/kink level (z^*)	\$53	\$64	\$21	\$45
z^* as a fraction of WBA	0.5	0.5	0.2	0.5
				placebo
Excess mass ($\mathcal{B}$)	5.334	4.814	1.247	-.7820
	(.2417)	(.279)	(.8757)	(.3395)
Hazard rate $(1-p)$	.042	.033	.039	
Implicit MTR (τ)	.538	.554	.606	
Earnings elasticity	.187	.134	0.096	
to net-of-tax rate (e)	(.0099)	(.0073)	(.0594)	

Source: CWBH. Notes: Standard errors are in parentheses below estimates.

5.2 Difference-in-difference in Louisiana

Louisiana's rules regarding unemployment insurance changed in April 1983. The change in partial UI affected both the stock of individuals registered in April 1983 and new inflows after that point in time.[34] 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.

I expect that, if bunching is actually related to the partial-UI schedule, the bunching

[34]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.

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.

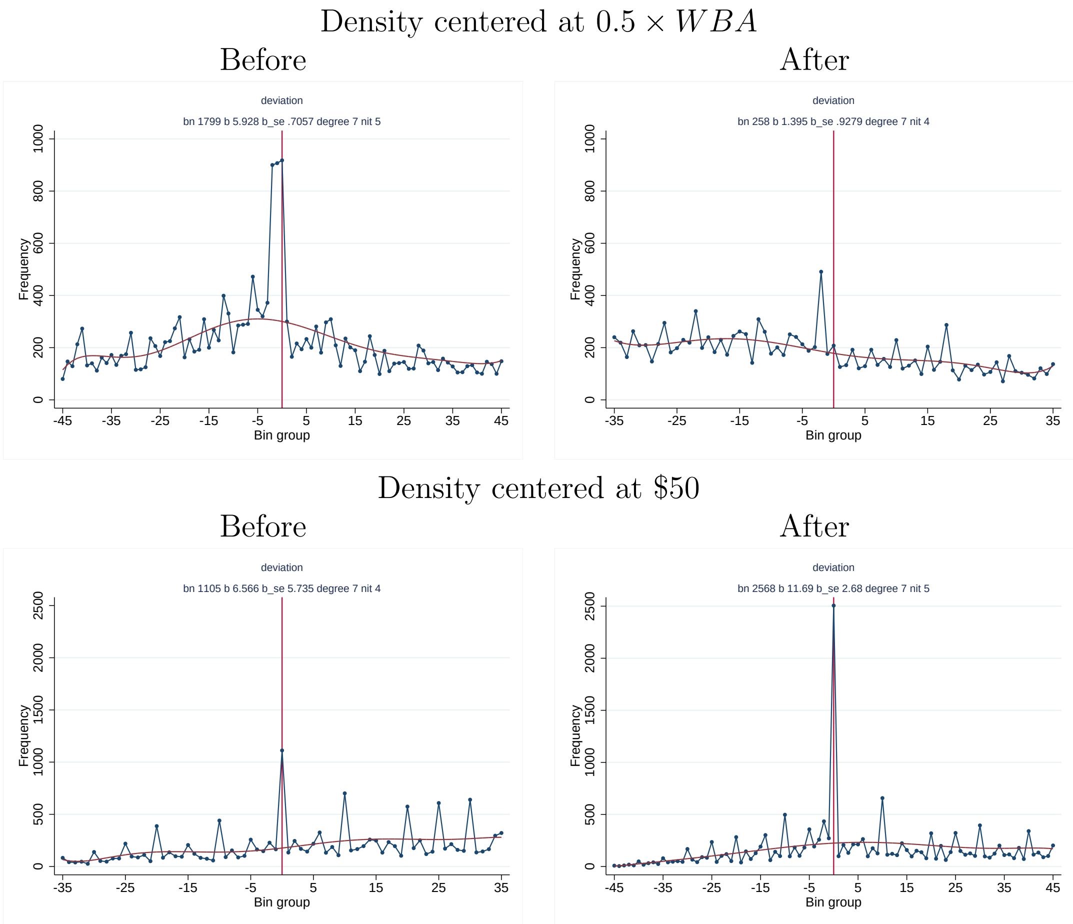
Figure 4 plots the earnings density of partial-UI claimants in the treatment group. In the upper panel, densities are centered at the pre-reform disregard ($0.5 \times WBA$). In the lower panel, they are centered at the post-reform disregard ($50). Starting with the upper panel, bunching is considerably 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.

Appendix Figure A2, 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, Friedman, Olsen and Pistaferri, 2011). 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 at $50 in the control group, which is not the case.

6 Bunching heterogeneity

In this section, I test predictions of the theoretical model regarding heterogeneity in bunching. To maximize statistical power, I jointly analyze partial UI in Idaho and Louisiana (before the reform in April 1983). Indeed, both states share the same disregard level ($0.5 \times WBA$).

Figure 4: 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.

6.1 Potential benefit duration

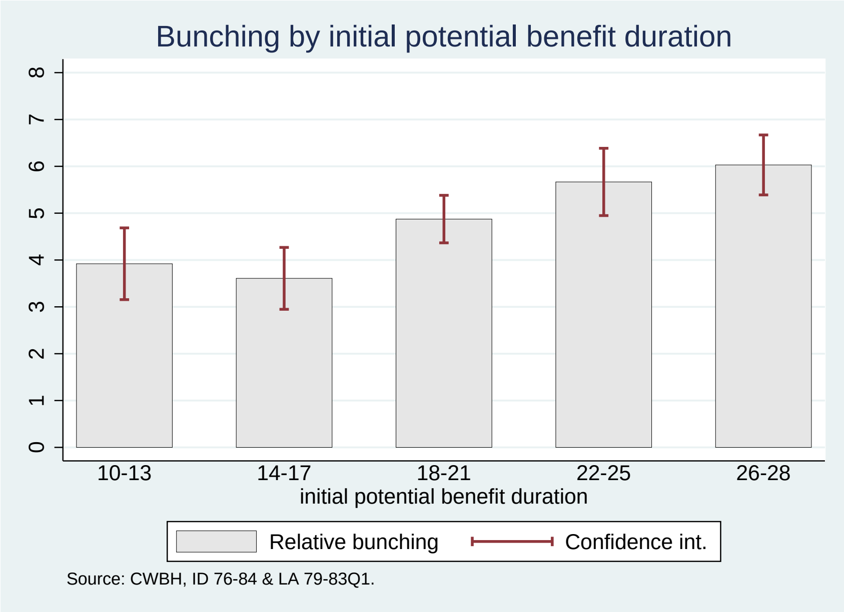
My theoretical model predicts that when total UB entitlement is more generous, bunching increases. The intuition is as follows. Let me consider a baseline claimant (X) with an initial UB entitlement B_0 yielding $\bar{t}^{Utot}$ weeks of unemployment benefits. Her earnings along the claim write $(z_0, z_1, ...)$. Now let me consider an identical claimant (Y) with UB capital enhanced by b. Given that the model depends on the past only through the UB capital, claimant (Y) behaves as claimant (X) with some period lags. Thus claimant (Y) has a higher marginal tax rate during her first claiming weeks. As a consequence, claimant (Y) is more likely to bunch. This prediction is a corollary of Proposition 1. As bunching decreases with time until the exhaustion date, individuals with longer potential benefit duration are more likely to bunch. In Idaho and Louisiana, the potential benefit duration in tier I varies from 10 to 28 weeks, in relation to past work history. Figure 5 shows that bunching is significantly greater when claimants have longer potential benefit durations. 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 Katz and Meyer (1990*a*) for an early contribution or Lalive et al. (2006) for evidence based on regression discontinuity design). Higher survival rates tend to decrease bunching (see Proposition 2), so that Figure 5 underestimates the positive relation between bunching and potential benefit duration.

6.2 Recall expectations

Job-seekers who expect to be recalled to their previous employer have different job-search behaviors than non-expecting claimants. Katz and Meyer (1990*b*) show that their unemployment duration is shorter, i.e. they have a lower probability to remain claimants. According to Proposition 2, claimants expecting recalls would then bunch more than non-expecting claimants. Indeed, expecting claimants have a lower probability to benefit from partial-UI transfers, as they may be recalled by their previous employer even before their calendar exhaustion date (while totally unemployed). In the joint Idaho-Louisiana sample, 65% of partial-UI claimants expect to be recalled to some previous employers.[35] Figure 6 shows that expecting claimants bunch significantly more. The bunching mass at the disregard level is 50% larger. I next propose a more systematic test of Proposition 2.

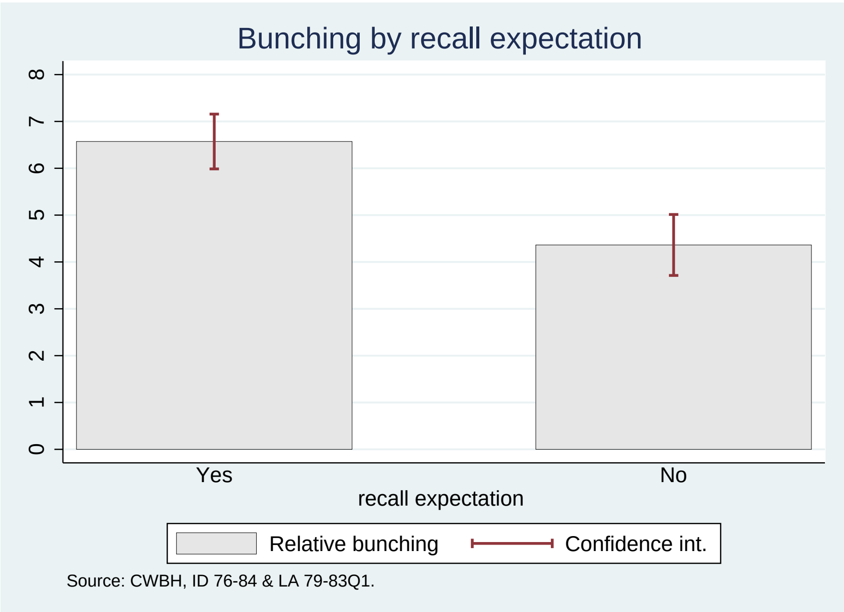
[35] Recall expectations were obtained through a specific survey. As a result, recall expectations are missing for almost 50% of claimants.

Figure 5: *Bunching by initial potential benefit duration*



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

Figure 6: *Bunching by recall expectation*



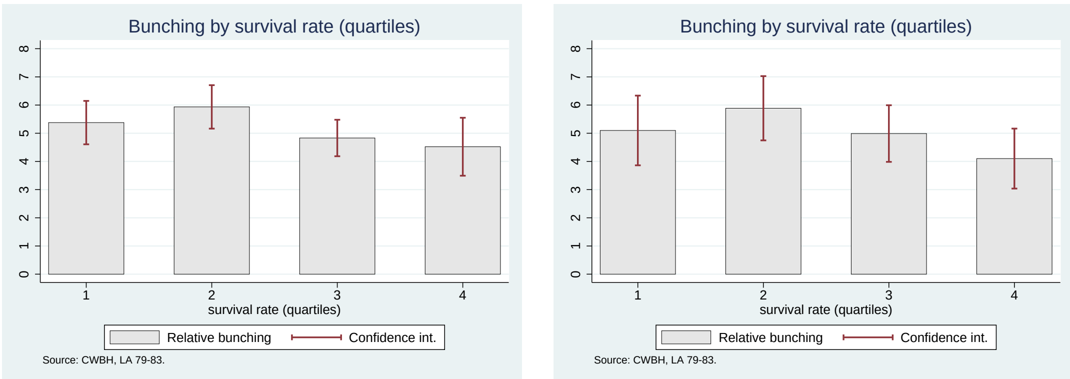
Source: CWBH for Idaho 1976-84 and Louisiana 1979-83, Q1. Notes: Excess mass at the disregard amount (kink) for claimants expecting to be recalled to their previous employer (left bar) and for claimants not expecting any recalls (right bar). Confidence interval at the 95% level in red.

6.3 Survival rate

Proposition 2 states that bunching decreases with the expected survival rate (or increases with the permanent job finding rate). Figure 7 compares bunching across the quartiles of the predicted survival rate distribution. In both states, bunching tends to decrease from the second to the fourth quartile, confirming Proposition 2. Though the differences are not statistically significant (at the 5% level), their magnitude is important. From the second to the fourth quartile, bunching is reduced by one third. The evidence is less clear at the bottom of the survival rate distribution: bunching increases from the first to the second quartile. As above, the comparison across quartiles may be confounded by other factors. Namely, individuals with high survival rate are expected to have on average less remaining benefits (under the assumption that the correlation between the initial potential benefit duration and the survival rate is negligible).

Figure 7: Bunching by survival rate.

Idaho Louisiana



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. Idaho in the left panel, Louisiana in the right panel. Confidence interval at the 95% level in red.

6.4 Bunching over the claim

As described in the theoretical section, two dimensions affect the dynamic marginal tax rate and thus the amount of bunching: weekly survival rates and time to benefit exhaustion. The previous heterogeneity analysis compared bunching along one dimension without holding constant the second dimension. I now take advantage of the panel structure of the data to control for heterogeneity in weekly survival rates. I first select claimants who have worked at least eight weeks while on claim. This cutoff corresponds to the third quartile of the distribution of the number of partial-UI weeks while on claim.[36] I then explore the evolution of bunching over those first eight weeks on partial UI. I further select the sample by excluding claimants expecting to be recalled to their previous employer. Predictions of the evolution of bunching over the claim are less clear for this excluded sample, as they may have a definite recall date.[37] Figure 8 shows that there is a steep increase in bunching between the first and second week of partial UI. This can be explained by learning effects (or alternatively by frictions). Figure 8 also shows that bunching estimates do not evolve after the second week of partial UI. I do not find any decreasing pattern in bunching over the spell. While the sample is held fixed across the different bars of Figure 8, each subgroup pulls together claimants with different horizon until exhaustion (subgroups are defined by the rank of the partial-UI week within the claim). Consequently, changes in horizon vary across subgroups, which may blur decreasing patterns. Moreover, the sample analyzed in Figure 8 may be composed of claimants at the beginning of their claim with a long horizon. This results in small changes in dynamic marginal tax rate over time, which makes it difficult to detect changes in bunching behavior.[38] Indeed, increasing the horizon by one week has a stronger effect on the dynamic marginal tax rate close to the exhaustion date rather than early in the claim. I thus propose another heterogeneity analysis that focuses on claimants later on the spell.

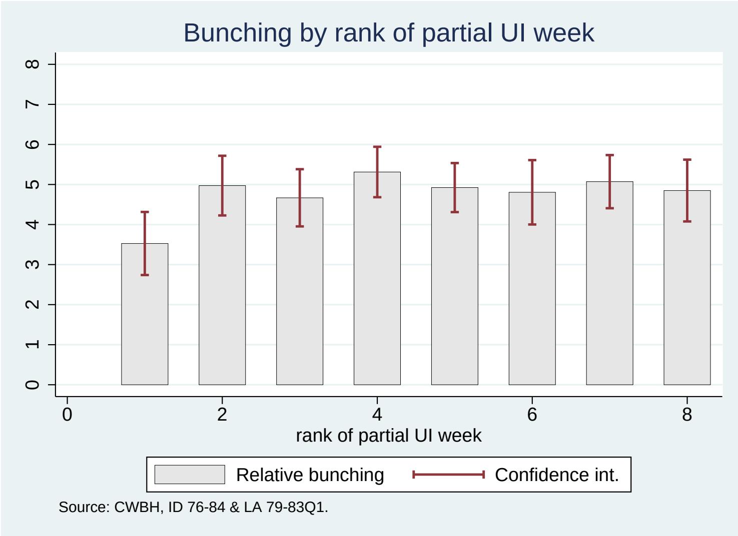
I compute, for each claiming week, the remaining UB entitlement and the corresponding time to benefit exhaustion (under the assumption of total unemployment). The corresponding time to benefit exhaustion can be described as the current potential benefit duration. In the first week of claim, the current potential benefit duration is equal to the initial potential benefit duration (see section 6.1 above). In Figure 9, I select claimants who worked both the month before their exhaustion date and earlier in the claim (when they had at least five months of benefits left in their UB capital). When claimants are far from the exhaustion date, bunching tends to increase while the current potential benefit duration decreases. When there is only one month of benefits left, claimants tend to de-bunch. Confidence intervals (at the 95% level) around bunching estimates are as wide as 2. With bunching estimates around 5, this makes the comparisons not very powerful. Consequently, while the broad picture drawn by the parameter estimates is consistent with the theoretical model amended with learning effects, I cannot formally reject that claimants' bunching behavior is flat over the spell.

[36]See Supplementary Figure A3 in the Appendix.

[37]When claimants have a definite recall date, they have no uncertainty about finding a permanent job and their behavior is different from the predictions of the theoretical model. For this group of claimants, the dynamic marginal tax rate does not evolve over the spell (given that the discount factor is one).

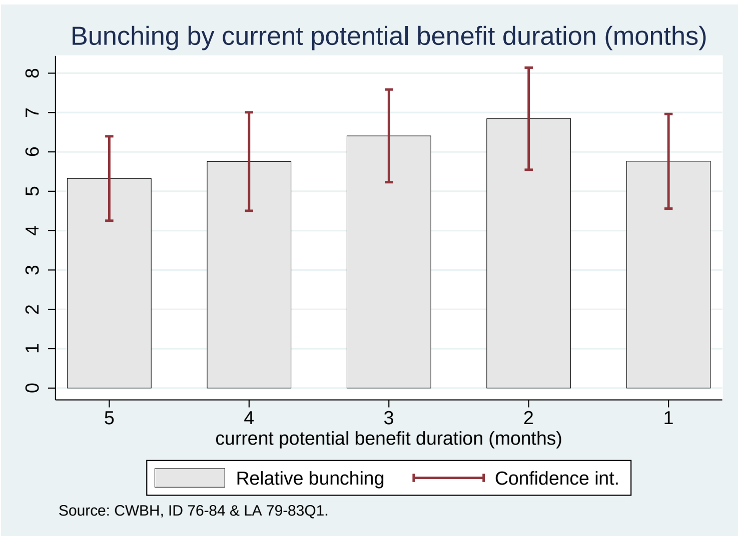
[38]Supplementary Figure A4 in the Appendix show that the change in dynamic marginal tax rate between two consecutive weeks is around -2 percentage points.

Figure 8: *Bunching by rank of the partial-UI week*



Source: CWBH for Idaho 1976-84 and Louisiana 1979-83, Q1. Notes: For every individual, I rank, within her claim, their weeks with positive earnings (partial UI). I select claimants with at least eight weeks with positive earnings. I estimate the excess mass at the disregard amount (kink) for each of those first eight weeks of partial UI. Confidence interval at the 95% level in red.

Figure 9: *Bunching by current potential benefit duration*



Source: CWBH for Idaho 1976-84 and Louisiana 1979-83, Q1. Notes: For every individual and period, I compute her current potential benefit duration. I estimate the excess mass at the disregard amount (kink) at different levels of current PBD. Confidence interval at the 95% level in red.

7 Setting the benefit-reduction rate

This section is devoted to normative considerations. Based on the theoretical model and on the estimate of the earned income elasticity, I propose modifications to the partial-UI schedule with the objective of maximizing claimants' welfare taking into account a UI administration budget constraint.

I focus on modifications to the benefit-reduction rate over the disregard level. The marginal benefit-reduction rate is indeed the partial UI parameter whose behavioral effects are well identified in the previous sections. Modifying the disregard level or the maximum earnings eligible for partial UI would require to predict the earnings elasticity away from the current disregard, while the previous identification strategy is local.

I denote T the benefit-reduction rate above the disregard. In the current schedule, the static marginal benefit-reduction rate is 100% ($T = 1$). The modified partial-UI schedule writes:

$$T(z) = \begin{cases} 0 & \text{if } z < z^*, \\ T \times (z - z^*) & \text{if } z \in (z^*, z^* + \min(b, B_t)), \\ \min(b, B_t) & \text{if } z > z^* + \min(b, B_t), \end{cases} \tag{18}$$

where all notations are defined in the baseline benefit-reduction schedule (Expression 2 in Section 3.1).

I consider the following formal program of the UI administration:[39]

$$SWF = \max_T \int_n U(B_0, n, b, z^*, T) dF(n) \tag{19}$$
$$\text{such that} \int_n C(B_0, n, b, z^*, T) dF(n) \leq \overline{C}.$$

Let me recall that $U(B_0, n, b, z^*, T)$ is the expected utility of a new claimant with talent n, and total initial entitlement B_0 - see Program (3) -, when the UI administration chooses the level of unemployment benefits b and the partial-UI schedule (z^*, T). Let me denote $C(B_0, n, b, z^*, T)$ the expected benefits paid by the UI administration to a new claimant. The UI administration budget constraint states that the aggregate expected benefit payments cannot exceed an exogenous upper threshold $\overline{C}$. In the above program, the UI administration maximizes claimants' welfare modifying the partial-UI schedule, holding the general UI rules constant - benefit level, initial entitlement, and potential

[39] For the sake of simplicity, I assume that the UI administration has the same Pareto weights for all claimants whatever their talent. The analysis below can be easily extended with unequal Pareto weights.

benefit duration. The interaction between these core parameters of the UI schedule and the partial UI rules is left for future research.

The expected cost for the administration of a UI claim equals:

$$C(B_0, n, b, z^*, T) = \begin{cases} \Sigma_{t=0}^{\bar{t}-1}(\beta p)^t b & \text{when } n < n^* \\ \Sigma_{t=0}^{\underline{t}-1}(\beta p)^t b + \Sigma_{t=\underline{t}}^{\bar{t}-1}(\beta p)^t(\min(b, B_t) - T.(z_t - z^*)) & \text{when } n > n^* \end{cases} \tag{20}$$

where I denote $\underline{t}$ the debunching date, i.e. the first date when claimants work strictly above the disregard level.[40] Job-seekers with a talent below the threshold $n^* = z^*$ earn less than the disregard level all along their claim. Job-seekers with a talent just over n^* bunch at the disregard level at the beginning of their claim. Then, at a given week $\underline{t}$, their talent becomes greater than the upper talent threshold $n^* + \delta n(\underline{t})$ (see Proposition 1). Starting from this date, they earn strictly above the disregard and their benefits are reduced until they lapse.[41]

To solve Program (19), I consider the effect of a small increase in the benefit-reduction rate $dT > 0$. This manipulation does not affect benefits paid to claimants with talent below n^*. Let me consider an individual with a greater talent n such that $\underline{t} < \bar{t}$. For the sake of the exposition, I consider a claimant, whose debunching and exhaustion dates do not change in response to the small change in benefit-reduction rate.[42] I also assume that it is only in the last period before exhaustion that the remaining UB capital is lower than the weekly benefit amount. First, the manipulation generates a mechanical decrease in benefits paid:

$$d\mathcal{M} = \Sigma_{t=\underline{t}}^{\bar{t}-1}(\beta p)^t(z_t - z^*)dT - (\beta p)^{\bar{t}-1}\left(\Sigma_{t=\underline{t}}^{\bar{t}-1}(z_t - z^*)dT\right), \tag{21}$$

$$= \Sigma_{t=\underline{t}}^{\bar{t}-1}\left((\beta p)^t - (\beta p)^{\bar{t}-1}\right)(z_t - z^*)dT \tag{22}$$

$$= \Sigma_{t=\underline{t}}^{\bar{t}-1}(\beta p)^t \tau_t (z_t - z^*)dT \tag{23}$$

where the second term in Equation 21 corresponds to the increase in the remaining UB capital in period $\bar{t}-1$ due to the mechanical decreases in benefits paid all along the claim. Second, the decrease in benefits paid creates a social welfare loss. According to the usual arguments (envelope theorem and quasi-linear utility), the social welfare loss is only due

[40] The debunching date is zero if claimants do not even bunch in the first week of claim. It is equal to the potential benefit duration if claimants bunch all along their claim.

[41] I abstract from claimants with very high talents who typically exit the partial-UI program before benefits lapse. By definition, the costs associated with these claimants are low

[42] Only claimants who are at the margin of debunching would change their debunching date, e.g. claimants such that there exists t verifying $n = n^* + \delta n(t)$. The same applies to the exhaustion date. This is admittedly a small fraction of the population and we abstract from it in the main text.

to the decrease in net income. The corresponding social welfare loss - in terms of UI administration funds - equals:

$$d\mathcal{W} = -1/\lambda . d\mathcal{M} \tag{24}$$

where λ is the Lagrange multiplier associated to the UI administration budget constraint. Third, the manipulation triggers a decrease in labor earnings due to behavioral response of the claimant and a corresponding decrease in remaining UB capital that is paid to the claimant in the period just before exhaustion. The corresponding decrease in UI administration revenue is:[43]

$$\begin{aligned} d\mathcal{R} &= \Sigma_{t=\underline{t}}^{\bar{t}-1}(\beta p)^t T dz_t - (\beta p)^{\bar{t}-1}\left(\Sigma_{t=\underline{t}}^{\bar{t}-1} T dz_t\right) && (25)\\ &= \Sigma_{t=\underline{t}}^{\bar{t}-1}(\beta p)^t \tau_t T z_t \frac{d\ln z_t}{d\ln(1-T\tau_t)} d\ln(1-T\tau_t) && (26)\\ &= -\Sigma_{t=\underline{t}}^{\bar{t}-1}(\beta p)^t \tau_t z_t e \frac{T\tau_t}{1-T\tau_t} dT. && (27) \end{aligned}$$

The overall effect of the manipulation for a claimant with talent n equals:

$$\Sigma_{t=\underline{t}}^{\bar{t}-1}(\beta p)^t \tau_t (z_t - z^*)\left(1 - 1/\lambda - a_t e \frac{T\tau_t}{1-T\tau_t}\right) dT \tag{28}$$

where a_t is defined as $a_t = \frac{z_t}{z_t - z^*}$. The above formula abstracts from changes in $(\underline{t}, \bar{t})$, which are second order. To obtain the overall effect of the manipulation, I aggregate Expression (28) over all claimants who earn above the disregard level during at least one week. I denote $\mathcal{Q}$ the corresponding set of claimants. For the sake of simplicity, I assume, as in the baseline model, that there is no heterogeneity in the elasticity e, nor in the permanent-job finding probability $(1-p)$. The overall effect then equals:

$$\begin{aligned} dSWF = (1-1/\lambda)\left(\int_{n\in\mathcal{Q}} \Sigma_{t=\underline{t}(n)}^{\bar{t}(n)-1}(\beta p)^t \tau_t \left(z_t(n) - z^*\right) dF(n)\right) dT && (29)\\ - e\left(\int_{n\in\mathcal{Q}} \Sigma_{t=\underline{t}(n)}^{\bar{t}(n)-1}(\beta p)^t \tau_t z_t(n) \frac{T\tau_t}{1-T\tau_t} dF(n)\right) dT \end{aligned}$$

The welfare-maximizing schedule is obtained when Equation (29) is set to zero. In general, there does not exist an analytical expression of the benefit-reduction rate T^* solution to Equation (29). To draw normative considerations, I then perform numerical simulations. I consider a representative claimant earning above the disregard and I search for

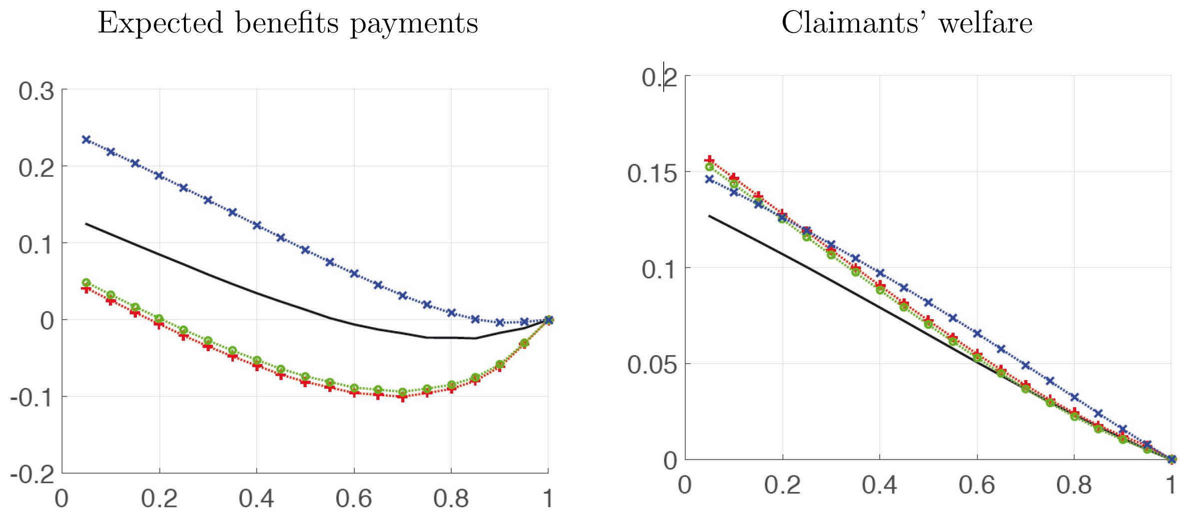
[43] I adjust the FOC of the baseline model to account for T: $1 - T\tau_t = \left(\frac{z_t}{n}\right)^{1/e}$. I obtain that the behavioral response is: $\frac{d\ln z_t}{d(1-T\tau_t)} = e$.

the cost-minimizing benefit-reduction rate. This corresponds to the welfare-maximizing rate when the Lagrange multiplier λ is infinitely large. In the baseline simulation, I set the different elements of Equation (20) at their average values estimated in Idaho: $p = 0.96$ and $e = 0.2$. I obtained them setting the discount factor equal to 0.9996. I consider a claimant entitled to 26 weeks of benefits with weekly benefit amount equal to \$100. Her talent is normalized to 100. In the current schedule - $z^* = \$50$ and $T = 1$ -, the claimant earns \$73 in the first week of her claim. Her earnings increase over the claim, up to \$99 in the exhaustion week.

The left panel in Figure 10 shows the expected cost born by the UI administration for different benefit-reduction rates. The solid black line corresponds to the baseline simulation. The expected cost is minimum when the benefit-reduction rate is 80%. It is 2% lower than in the current schedule with a 100% marginal benefit reduction rate. As explained above, a decrease in the benefit reduction rate leads to a mechanical increase in benefits paid - when job-seekers do not change their labor supply ($d\mathcal{M}$). However, the behavioral response of job-seekers counteracts the mechanical increase ($d\mathcal{R}$). When the initial level of benefit reduction rate is high, the simulations suggest that the behavioral response actually outweighs the mechanical increase. As also explained above, the decrease in the benefit-reduction rate leads to an increase in the welfare of partially unemployed job-seekers ($d\mathcal{W}$), shown by the solid black line in the right panel in Figure 10. This holds whatever the initial benefit reduction rate. However the increase in welfare becomes costly to the UI administration when the benefit-reduction rate is lower than 80%, and whether the UI administration is willing to do so then depends on the social marginal welfare weight $1/\lambda$.

When I decrease the elasticity e to 0.1 compared to the baseline simulation - holding other parameters constant -, the cost-minimizing benefit-reduction rate increases to 90% (see the blue line with crosses in the left panel of Figure 10). The behavioral response of claimants counteracts less the initial mechanical increase in benefit payments following a reduction in benefit-reduction rate. On the contrary, when I increase the permanent job finding rate by 50%, such that $p = 0.94$, or when I decrease the discount factor to 0.98, I obtain a cost-minimizing benefit-reduction rate around 70% (see the red line with '+' markers and the green line with circles in the left panel of Figure 10). These changes increase the wedge τ_t, leading to a decrease in the cost-minimizing benefit-reduction rate. Overall, this sensitivity analysis suggests that the cost-minimizing benefit-reduction rate belongs to the interval $[70\%; 90\%]$.

Figure 10: *Expected benefits payments and claimants' welfare by static benefit-reduction rates*



Note: X-axis corresponds to different static benefit-reduction rates. Y-axis corresponds to the percentage variation in the expected total amount of benefits paid to the claimant over her claim (left panel) and to the expected utility of the claimant over the benefit year, relative the corresponding level in the current schedule with a 100% marginal benefit-reduction rate. The black solid line corresponds to the baseline simulation. The red line with '+' markers corresponds to $p = 0.94$, the green line with circles to $\beta = 0.98$ and the blue line with crosses to $e = 0.1$.

8 Conclusion

I study claimants' behavioral response to the rules of partial unemployment insurance in the U.S. I observe that claimants bunch at the kinks of the partial-UI schedule, and interpret this bunching mass with a dynamic model of working while on claim. The model highlights that forward-looking claimants react to the dynamic marginal tax rate of the partial-UI schedule, which is lower than the static benefit-reduction rate. The wedge between the two rates is primarily due to benefit transfers. Claimants internalize that the reduction in current benefits leads to an increase in future benefits. I extend the approach of Saez (2010) to account for this mechanism and estimate an earnings elasticity to the net-of-tax rate between 0.1 and 0.2. Simulations of the model, calibrated with these estimates, suggest that decreasing the static benefit-reduction rate from 100% to 80% would decrease overall benefits paid to partial claimants by the UI administration and still improve the claimants' welfare.

Empirical evidence on bunching heterogeneity is broadly consistent with the predictions of the dynamic model. The theoretical model predicts (i) that claimants with higher survival probability should bunch less and (ii) that claimants should bunch less and less over time within their claim. While I find empirical support for the first prediction, the decrease in bunching over the spell is not as steep as the dynamic model predicts. Moreover, I find an increase in bunching at the beginning of the spell. This could be the result of adjustment costs/frictions in the search for low-earnings jobs or the consequences of claimants progressively learning the partial-UI schedule. Understanding the roles of these two mechanisms is a promising direction for future work.

The normative exercise provides insights on the efficiency of the partial-UI schedule and guidelines to policymakers. It focuses on modifications of the static benefit-reduction rate, a parameter of great interest. In France, the benefit-reduction rate was extensively discussed during the 2014 UI reform. Policymakers could also be interested in the effect of other parameters of the schedule, such as the disregard level. My work could be extended to study such modifications. It could also be extended, in a more theoretical direction, to derive the optimal flexible partial-UI schedule.

Finally, 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. For example in the U.S., a reduction in current pension benefits increases future benefits (SSAET). Gathering evidence on such schemes would help us understand more broadly how individuals make intertemporal decisions under uncertainty.

References

Baily, M. (1978), 'Some aspects of optimal unemployment insurance', *Journal of Public Economics* **10**(3), 379–402.

Bastani, S. and Selin, H. (2014), 'Bunching and non-bunching at kink points of the Swedish tax schedule', *Journal of Public Economics* **109**, 36–49.

Brown, K. M. (2013), 'The link between pensions and retirement timing: Lessons from california teachers', *Journal of Public Economics* **98**(0), 1 – 14.

Caliendo, M., Künn, S. and Uhlendorff, A. (2012), Marginal Employment, Unemployment Duration and Job Match Quality, IZA Discussion Papers 6499, Institute for the Study of Labor (IZA).

Chetty, R. (2006*a*), 'A general formula for the optimal level of social insurance', *Journal of Public Economics* **90**(10-11), 1879–1901.

Chetty, R. (2006*b*), 'A New Method of Estimating Risk Aversion', *American Economic Review* **96**(5), 1821–1834.

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

Chetty, R., Friedman, J., Olsen, T. and Pistaferri, L. (2011), 'Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records', *The Quarterly Journal of Economics* **126**(2), 749–804.

Chetty, R., Friedman, J. and Saez, E. (2013), 'Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings', *American Economic Review* **103**(7), 2683–2721.

Chetty, R., Guren, A., Manoli, D. and Weber, A. (2011), 'Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins', *American Economic Review* **101**(3), 471–75.

Fremigacci, F. and Terracol, A. (2013), 'Subsidized temporary jobs: Lock-in and stepping stone effects', *Applied Economics* **45**(33), 4719–4732.

Gelber, A. M., Jones, D. and Sacks, D. W. (2013), Earnings Adjustment Frictions: Evidence from the Social Security Earnings Test, NBER Working Papers 19491, National Bureau of Economic Research, Inc.

Godoy, A. and Røed, K. (2014), Unemployment Insurance and Underemployment, IZA Discussion Papers 7913, Institute for the Study of Labor (IZA).

Gruber, J. (1997), 'The consumption smoothing benefits of unemployment insurance', *The American Economic Review* **87**(1), 192–205.

Holen, A. and Horowitz, S. (1974), 'Partial Unemployment Insurance Benefits and the Extent of Partial Unemployment', *The Journal of Human Resources* **9**(3), 420–422.

Katz, L. and Meyer, B. (1990*a*), 'The Impact of the Potential Duration of Unemployment Benefits on the Duration of Unemployment', *Journal of Public Economics* **41**(1), 45–72.

Katz, L. and Meyer, B. (1990*b*), 'Unemployment Insurance, Recall Expectations, and Unemployment Outcomes', *The Quarterly Journal of Economics* **105**(4), 973–1002.

Kiefer, N. and Neumann, G. (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. and Waseem, M. (2013), 'Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan', *The Quarterly Journal of Economics* **128**(2), 669–723.

Kyyrä, T. (2010), 'Partial unemployment insurance benefits and the transition rate to regular work', *European Economic Review* **54**(7), 911–930.

Kyyrä, T., Parrotta, P. and Rosholm, M. (2013), 'The effect of receiving supplementary UI benefits on unemployment duration', *Labour Economics* **21**, 122–133.

Lalive, R., Ours, J. V. and Zweimuller, J. (2006), 'How Changes in Financial Incentives Affect the Duration of Unemployment', *Review of Economic Studies* **73**(4), 1009–1038.

Landais, C. (2014), 'Assessing the Welfare Effects of Unemployment Benefits Using the Regression Kink Design', *American Economic Journal: Economic Policy* . forthcoming.

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

McCall, B. (1996), 'Unemployment Insurance Rules, Joblessness, and Part-Time Work', *Econometrica* **64**(3), 647–682.

Moffitt, R. (1985), 'Unemployment insurance and the distribution of unemployment spells', *Journal of Econometrics* **28**(1), 85–101.

Munts, R. (1970), 'Partial Benefit Schedules in Unemployment Insurance: Their Effect on Work Incentive', *The Journal of Human Resources* **5**(2), 160–176.

Saez, E. (2010), 'Do Taxpayers Bunch at Kink Points?', *American Economic Journal: Economic Policy* **2**(3), 180–212.

Shimer, R. and Werning, I. (2007), 'Reservation Wages and Unemployment Insurance', *The Quarterly Journal of Economics* **122**(3), 1145–1185.

Online Appendix

A Institutional background

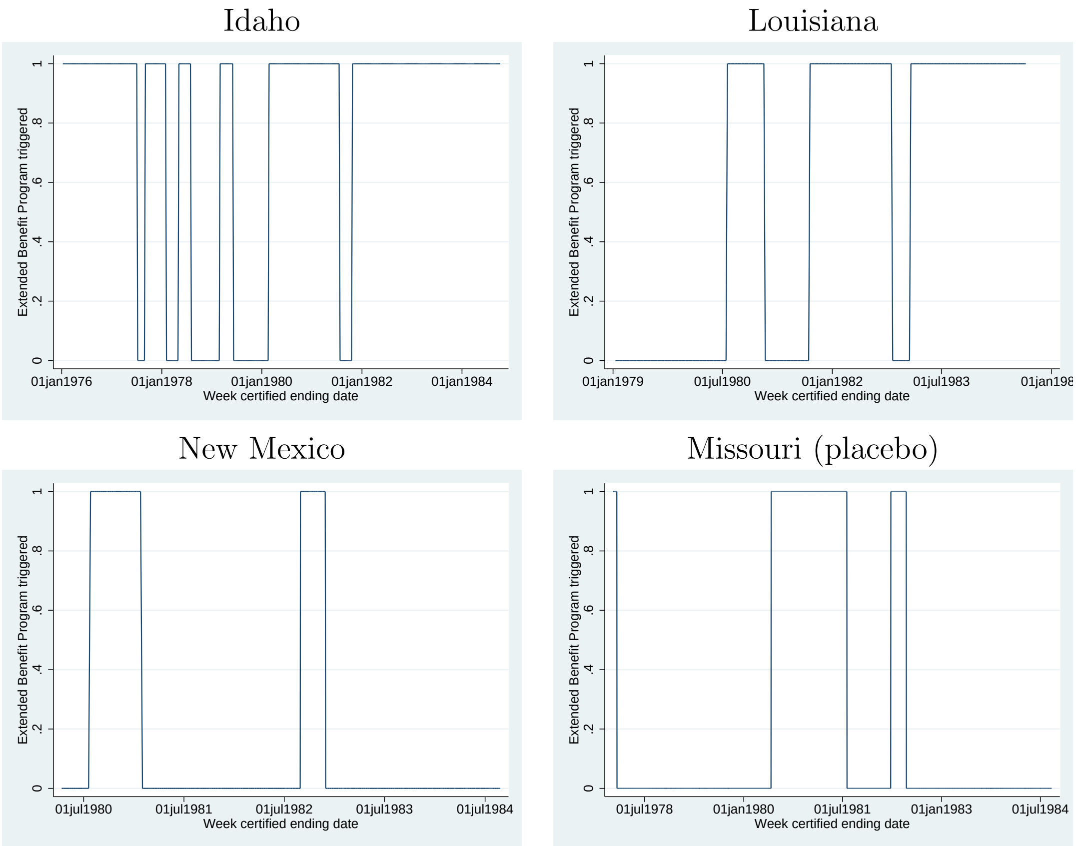
Between the late 70s and early 80s, the unemployment insurance (UI) rules, in Idaho, Louisiana, New Mexico and Missouri, are as follows. First, UI claimants must meet a monetary eligibility requirement. They must have accumulated a sufficient amount of earnings during a one-year base period before job separation. Second, UI claimants must meet nonmonetary 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 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. In Figure A1, I plot the EB periods in each state. 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.

Figure A1: Extended Benefit Program.



Source: Trigger reports.

B Theoretical model

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

B.1 Model Solution

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

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

such that

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

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: $B_{t+1} \leq B_t$. I first discuss the existence of stationary solutions.

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:

$$u_c(c, z; n_i) + u_z(c, z; n_i) = 0. \tag{30}$$

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

$$U_i = \frac{u(c, z; n_i) + \beta(1 - p)W}{1 - \beta p}. \tag{31}$$

The two previous equations (30) and (31) show that the stationary value of unemployment U_i and the corresponding earnings z do not depend on the level of UB capital B.

Recall that, when $B > b$, the typical partial-UI schedule is such that there exists a unique $\overline{z}$ such that for any $z \geq \overline{z}$, $T(z) = b$ and the marginal tax rate is 100% just below $\overline{z}$. Let me consider the marginal individual whose talent is consistent with supplying $\overline{z}$, 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 talent with $B > b$. Of course there may exist some very talented individuals whose stationary z is well above $\overline{z}$. To rule out the existence of such individuals, it is sufficient to assume that there is a fixed flow cost to claim. Such a cost decreases the relative gain of stationary claiming.

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 talent 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:

$$u_c(c_t, z_t; n_i)\left(1 - T'(z_t)\right) + \beta p(z_t) T'(z_t) U'_{t+1}(B_{t+1}; n_i) = -u_z(c_t, z_t; n_i). \qquad (32)$$

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 $t = \overline{t} - 1$. Consequently, the third term of the marginal gain of labor earnings can be written as:

$$\beta p T'(z_t) U'(B_{t+1}; n_i) = T'(z_t) \beta^{\bar{t}-t-1} p^{\bar{t}-t-1} u_c(c_{\bar{t}-1}, z_{\bar{t}-1}; n_i) \tag{33}$$

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

Using Equation (46) and the utility definition, the FOC in Equation (32) can be simplified. The rest of the derivation is in the main text.

B.2 Heterogeneity

Heterogeneity in the permanent job finding rate. In this extension, job-seekers are characterized by their talent n_i and their probability to remain claimants p_i, distributed according to a joint distribution $f_t(n, p)$ at period t. The excess bunching is then a weighted average of excess bunching in each p-strata: $\mathcal{B}_t \simeq \int_p f_t(p|n = z^*) \delta n(\tau(t, p)) dp$. Consequently, the identification formula 11 is only slightly modified:

$$e = \frac{\mathcal{B}_t}{z^* \mathbb{E}_t[\tau(t, p)|n = z^*]}. \tag{34}$$

The relevant dynamic marginal tax rate is then averaged over the population still claiming at period t. When job-seekers differ by their permanent job finding rate, there is dynamic selection over the spell. The fraction of job-seekers with high p increases over the spell. This reinforces the decrease in the marginal tax rate across periods. As a consequence, Proposition 1 is not only valid within a p-strata, but also on average.

Heterogeneity in the structural elasticity. I can also extend the model by allowing for different individual elasticities e_i. At time t, talents and elasticities are distributed according to a joint distribution $f_t(n, e)$. The excess bunching at time t then equals: $\mathcal{B}_t \simeq \int_e f_t(e|n = z^*) \delta n(\tau_t, e) de \simeq z^* \mathbb{E}_t[e \tau_t | n = z^*]$. I obtain the following identification formula:

$$\mathbb{E}_t[e|n = z^*] = \frac{\mathcal{B}_t}{z^* \tau_t}. \tag{35}$$

Equation (35) shows that bunching at kinks identifies a specific local average of structural elasticities. Note that Proposition 1 is still relevant in this context, as $\mathbb{E}_t[e|n = z^*]$ does not depend on t.

Heterogeneity in the permanent job finding rate and in the structural elasticity I define $f_t(n, p, e)$ the joint distribution of talents, permanent job finding probabilities and elasticities, at period t. The excess bunching within a p-strata is:

$$\begin{aligned}\mathcal{B}_t(p) &= \int_e \delta n(\tau(t,p),e) f_t(e|p, n = z^*) de \\ &= E_t[\delta n(\tau(t,p),e)|n = z^*, p] \\ &= z^* \tau(t,p) E_t[e|n = z^*, p].\end{aligned}$$

When there are heterogeneity in all three dimensions, Proposition 1 still holds within a p-strata, because $E_t[e|n = z^*, p] = E_{t+1}[e|n = z^*, p]$. However Proposition 2 is not verified. Heterogeneity across p can be confounded by heterogeneity in elasticities (e), as there may exist p_1 and p_2 such that $E_t[e|n = z^*, p_1] \neq E_t[e|n = z^*, p_2]$.

B.3 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:

$$u(c, z; n_i) = \frac{c^{1-\sigma}}{1-\sigma} - \frac{n_i}{1+1/e} \left(\frac{z}{n_i}\right)^{1+1/e} \tag{36}$$

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 B.1. We focus on the case when $B_t > b$ as in the main text. Solving the job-seekers' program yields the following FOC:

$$u_c(c_t, z_t; n_i)\,(1 - T'(z_t)) + T'(z_t)(\beta p)^{\bar{t}-t-1} u_c(c_{\bar{t}-1}, z_{\bar{t}-1}; n_i) = -u_z(c_t, z_t; n_i). \tag{37}$$

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

$$1 - T'(z_t)\left(1 - (\beta p)^{\bar{t}-t-1}\left(\frac{c_t}{c_{\bar{t}-1}}\right)^{\sigma}\right) = \left(\frac{z_t}{n}\right)^{1/e} (c_t)^{\sigma}. \tag{38}$$

Compared to Equation (6), 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 talent n^* satisfies the following FOC:

$$1 = \left(\frac{z^*}{n^*}\right)^{1/e} (b + z^*)^{\sigma}. \tag{39}$$

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

$$(\beta p)^{\bar{t}-t-1} (c_{\bar{t}-1})^{-\sigma} = \left(\frac{z^*}{n^* + \delta n(t)} \right)^{1/e} . \tag{40}$$

As $c_{\bar{t}-1}$ is a function of the talent $n^* + \delta n(t)$ and of the other parameters of the model (β, p, σ, B_t, b, z^* and e), Equations (39) and (40) 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 $B_{\bar{t}-1} < b$:

$$1 = \left(\frac{z_{\bar{t}-1}}{n^* + \delta n(t)} \right)^{1/e} (c_{\bar{t}-1})^{\sigma} . \tag{41}$$

The budget contraint 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 (41) 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 (40), we obtain that excess bunching identifies e from Equations (39) and (40).

We now quantify the order of magnitude of the bias in the estimate of e when risk-aversion is neglected. We re-write Equations (39) and (40):

$$n^* = z^* (c_t)^{e\sigma} , \tag{42}$$

$$n^* + \delta n(t) = z^* (1 - \tau_t)^{-e} (c_{\bar{t}-1})^{e\sigma} . \tag{43}$$

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:

$$e = \frac{\delta n(t)}{z^* \left(\tau_t + \sigma \frac{\Delta c}{c_t} \right)} . \tag{44}$$

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.

B.4 Stepping-stone/crowding-out effects

To account for stepping-stone and crowding-out effects, 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, 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:

$$\underbrace{u_c(c_t,z_t;n_i)\left(1-T'(z_t)\right)}_{(I)} + \underbrace{\beta p(z_t)T'(z_t)U'(B_{t+1};n_i)}_{(II)} - \underbrace{\beta p'(z_t)\left(W-U(B_{t+1};n_i)\right)}_{(III)} = -u_z(c_t,z_t;n_i). \quad (45)$$

Compared to the FOC in the baseline model, a third term appear on the left-hand side. When $p' < 0$, it represents the marginal gain induced by the stepping-stone effect of low-wage jobs. When $p' > 0$, it corresponds to the marginal cost induced by the crowding-out effect.

Following the same reasoning as in the previous Appendix, I obtain a simplified expression of the second term in Equation (45):

$$\beta p(z_t)T'(z_t)U'_{t+1}(B_{t+1};n_i) = T'(z_t)\beta^{\bar{t}-t-1}\left(\Pi_{i=t}^{\bar{t}-2}p(z_i)\right)u_c(c_{\bar{t}-1},z_{\bar{t}-1};n_i) \quad (46)$$

where $\Pi_{i=t}^{\bar{t}-2}p(z_i)$ is the probability to exhaust benefits conditional on claiming at date t.

Using Equation (46) and the utility definition, the FOC in Equation (45) simplifies to:

$$1-T'(z_t)\tau_t(z_t) - \beta p'(z_t)\left(W-U(B_{t+1};n_i)\right) = \left(\frac{z_t}{n_i}\right)^{1/e} \quad (47)$$

where the wedge τ_t now depends explicitly on z_t:

$$\tau_t(z_t) = 1-\beta^{\bar{t}-t-1}\Pi_{j=t}^{\bar{t}-2}p(z_j). \quad (48)$$

The assumptions specified in Section 3.6 of the main text 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:

$$1 + \pi_t(z_t) = \left(\frac{z_t}{n_i}\right)^{1/e} \text{ when } z_t < z^*, \tag{49}$$

$$1 - \tau_t(z_t) + \pi_t(z_t) = \left(\frac{z_t}{n_i}\right)^{1/e} \text{ when } z_t > z^*. \tag{50}$$

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

$$n_t^* = \frac{z^*}{(1 + \pi_t^-(z^*))^e}, \tag{51}$$

$$n_t^* + \delta n_t = \frac{z^*}{\left(1 - \tau_t(z^*) + \pi_t^+(z^*)\right)^e}. \tag{52}$$

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: $e = \frac{\mathcal{B}_t}{z^* \tau_t(z^*)}$.

C 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: $h_i = h.\exp(\beta X_i)$. 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 (1990*b*) 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 capture 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 A1 reports the coefficient estimates of the hazard model for each state (in columns).

Table A1: 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)
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)
Constant	-2.846***	-3.327***	-3.929***	-3.279***
	(.143)	(.081)	(.198)	(.358)
Years fixed effects	Yes	Yes	Yes	Yes
Nb. 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.

D Supplementary Tables and Figures

Table A2: Descriptive Statistics

	Idaho	Louisiana	New Mexico	Missouri
Years	76-84	79-84	80-84	78-84
Inflow (nb)	91,162	95,675	62,030	78,065
Male	.673	.708	.670	.606
Age	31.1	34.5	33.6	34.7
Education (years)	11.8	11.3	11.6	11.2
White	.946	.622	.423	83.8
Private firm	.889	.952	.912	.939
Industries				
Construction	.148	.308	.224	.157
Manufacturing	.330	.178	.131	.397
Trade	.218	.135	.199	.142
Services	.139	.168	.199	.202
Occupations				
Prof., tech. and managers	.076	.095	.121	.072
Clerical and sales	.145	.153	.192	.162
Structural work	.221	.312	.278	.147

Source: CWBH. Notes : Agriculture, Mining, Transportation, Finance, Insurance and Real Estate industries are not reported for the sake of space (around 20% of the sample). Occupations corresponds 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.

Table A3: Selection into partial UI (Louisiana 1980-82)

	At least one week of partial UI	No partial UI	Diff- test p-value
Male	0.70	0.71	0.38
Age	34.90	33.98	0.00
Education (years)	11.39	11.26	0.00
White	0.62	0.60	0.00
Private firm	0.97	0.95	0.00
Industries			
Construction	0.32	0.31	0.03
Manufacturing	0.23	0.19	0.00
Trade	0.10	0.14	0.00
Services	0.14	0.17	0.00
Occupations			
Prof., tech. and managers	0.07	0.09	0.00
Clerical and sales	0.12	0.15	0.00
Structural work	0.31	0.30	0.03
Pre-U weekly wage (current dollars)	352.18	315.34	0.00
Weekly benefit amount (WBA)	144.47	130.58	0.00
Replacement rate	0.46	0.47	0.00
Potential duration (weeks)	38.99	37.28	0.00
Actual UB duration	24.06	20.18	0.00
Inflow (nb)	13,174	37,139	

Source: CWBH. This table compares claimants with at least one week of partial unemployment during the benefit year in Column 1 to claimants on total unemployment in Column 2. Column 3 reports the p-value of the test of equality between the two first columns. Notes: Agriculture, Mining, Transportation and the FIRE industries (Finance, Insurance and Real Estate) are not reported for the sake of space (around 20% of the sample). Occupations corresponds 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.

Table A4: Robustness of earnings elasticities to the net-of-tax rate varying estimation parameters

	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]
Poly. Deg.	7	7	7	7	7	7	9	5	3
Idaho	0.187	0.287	0.264	0.134	0.188	0.184	0.167	0.216	0.276
	(0.010)	(0.021)	(0.014)	(0.010)	(0.010)	(0.011)	(0.008)	(0.012)	(0.019)
Louisiana	0.134	0.215	0.168	0.108	0.142	0.129	0.129	0.145	0.187
	(0.007)	(0.017)	(0.009)	(0.006)	(0.007)	(0.007)	(0.008)	(0.008)	(0.011)
New Mexico	0.096	.	0.197	0.100	0.054	0.071	0.053	0.057	0.039
	(0.065)	.	(0.214)	(0.044)	(0.048)	(0.068)	(0.063)	(0.054)	(0.049)

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 3) 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 9, I decrease the degree of the polynomial fitting the density. Standard errors are in parentheses. 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 A5: Earnings elasticity estimates without first-order approximation of the marginal tax rate.

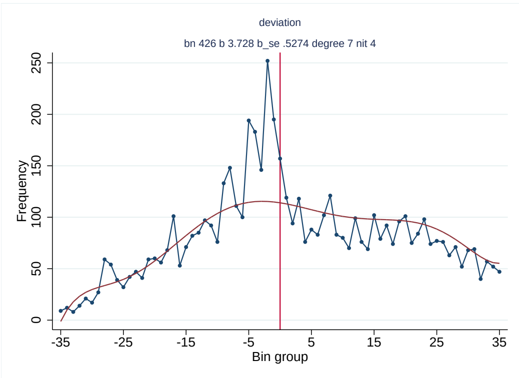
	Idaho	Louisiana	New Mexico
Earnings elasticity to net-of-tax rate (e)	.130	.092	.063
	(.0065)	(.0055)	(.0457)

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

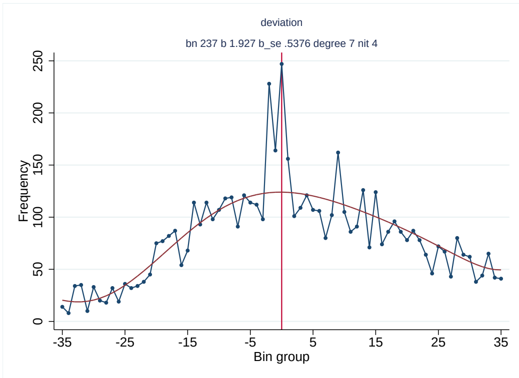
Figure A2: Centered weekly earnings density of partial-UI claimants in the control group.

Density centered at $0.5 \times WBA$

Before

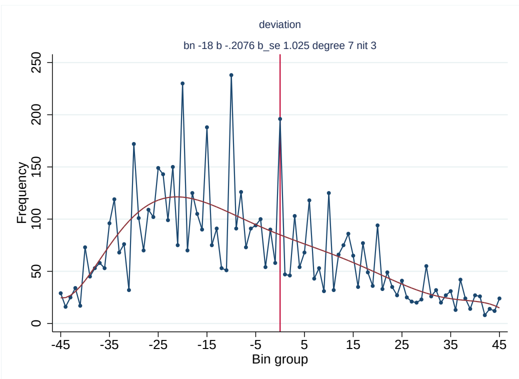


After

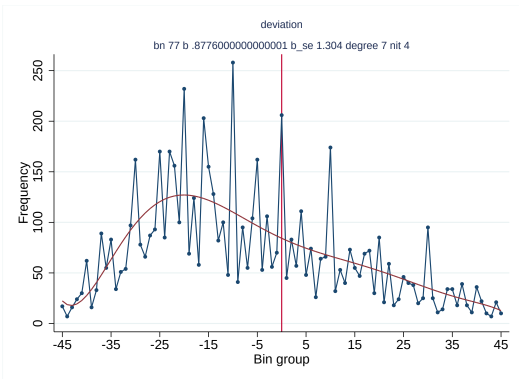


Density centered at $50

Before

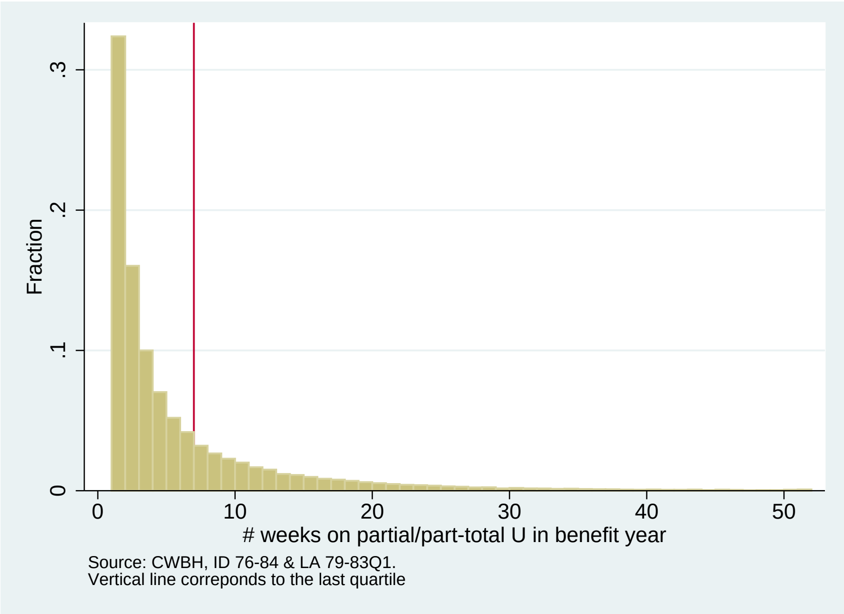


After



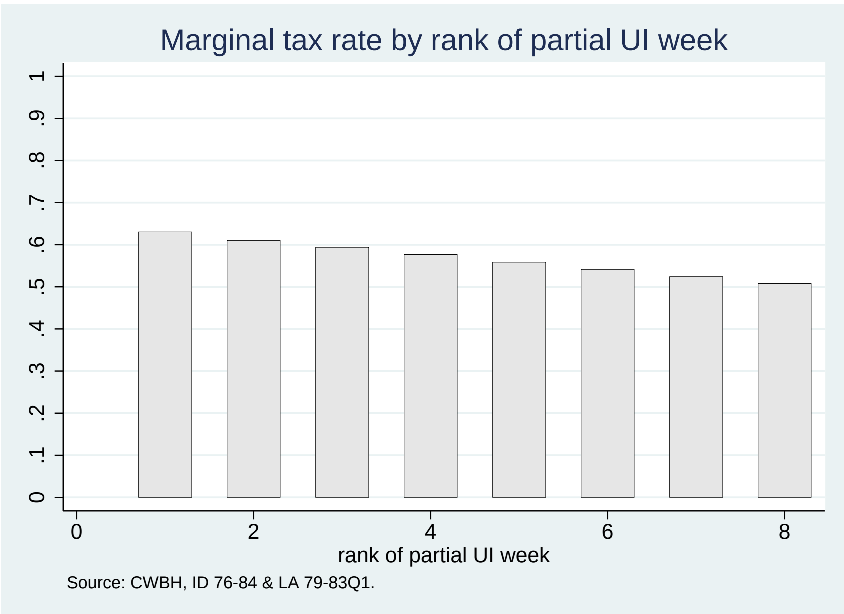
Source: CWBH. Notes: Earnings are in dollars. Empirical earnings density in blue. Counterfactual density in red.

Figure A3: *Distribution of the number of partial-UI weeks over the claim*



Source: CWBH ID 76-84 & LA 79-83Q1. Notes: For each individual claim, I compute the total number of weeks with positive earnings (partial UI).

Figure A4: *Dynamic marginal tax rate by rank of the partial-UI week*



Source: CWBH for Idaho 1976-84 and Lousiana 1979-83, Q1. Notes: For every individual, I rank, within her claim, their weeks with positive earnings (partial UI). I select claimants with at least eight weeks with positive earnings. I estimate the average dynamic marginal tax rate for each of those first eight weeks of partial UI.