MANIPULATION, SELECTION AND THE DESIGN OF TARGETED SOCIAL INSURANCE*

Luca Citino[†] Kilian Russ[‡] Vincenzo Scrutinio[§]

27th March 2022

Abstract

This paper provides a sufficient statistics framework to study the design of optimal targeted social insurance in the presence of manipulation opportunities, through which individuals select into policies not intended for them. We apply our framework to Italian unemployment insurance (UI), which features a discontinuous coverage increase around an age-at-layoff threshold. Using novel bunching techniques, we document pervasive manipulation and a substantial increase in benefit receipts. However, most of this increase is mechanically due to higher coverage, while the implied moral hazard response is modest. We connect these findings to our theory and discuss how they affect welfare conclusions.
JEL classification: E24, J64, J65, J68.

*This paper is a revised version of the paper "Happy Birthday? Manipulation and Selection in Unemployment Insurance". We are extremely thankful to our advisors Stephen Machin, Alan Manning, Jörn-Steffen Pischke and Johannes Spinnewijn for continuous guidance and support. This project also benefited from discussions with Andrea Alati, Andres Barrios Fernandez, Miguel Bandeira, Fabio Bertolotti, Rebecca Diamond, François Gerard, Simon Jäger, Felix König, Camille Landais, Attila Lindner, Clara Martinez-Toledano, Matteo Paradisi, Michele Pellizzari, Frank Pisch, Michel Serafinelli, Enrico Sette, Martina Zanella, Josef Zweimüller, and seminar participants at briq, EUI, INPS and the LSE. This project was carried out while Kilian Russ was visiting the London School of Economics as part of the European Doctoral Programme in Quantitative Economics. Financial support from the London School of Economics and the Bonn Graduate School of Economics is gratefully acknowledged. The realization of this project was possible thanks to the VisitInps initiative. We are very grateful to Massimo Antichi, Elio Bellucci, Mariella Cozzolino, Edoardo Di Porto, Paolo Naticchioni and all the staff of Direzione Centrale Studi e Ricerche for their invaluable support with the data. The opinions expressed in this paper are those of the authors alone and do not necessarily reflect the views of the Bank of Italy or INPS. All remaining errors are ours alone.

[†]Bank of Italy; Via Nazionale 91, Rome, Italy. Corresponding author; Email: luca.citino@bancaditalia.it
[‡]London School of Economics and Political Science, CEP, BGSE; Email: k.russ1@lse.ac.uk
[§]University of Bologna, CEP, IZA; Email: vincenzo.scrutinio@unibo.it

I Introduction

The targeting of public policies on the basis of observable individual characteristics is ubiquitous in OECD countries. Governments tax individuals based on their marital status, provide welfare payments which depend on the number of children in the household, and tie disability insurance to particular medical conditions. Tying public benefits to observable information holds the potential to increase cost-effectiveness while providing assistance and support to those most in need. Strained government budgets and an ever increasing amount of available information are likely to increase both the necessity and possibility of more efficiently targeted interventions in decades to come. The theoretical desirability for targeting based on immutable tags based on efficiency grounds has long been recognized (Akerlof, 1978). In practice, however, policy makers often rely on endogenous tags, which leave room for strategic manipulation and selection into benefit schemes.

How should we design targeted public policies in the presence of manipulation opportunities? In particular, how does manipulation alter the desirability of differentiated policy? Finally, once we know which empirical moments are relevant, how do we estimate them in practice?

This paper breaks new grounds on these questions in the context of optimal social insurance and makes three main contributions. First we propose a simple, yet robust, theoretical framework to study the design of optimal differentiated social insurance in the presence of manipulation. To this end we introduce differentiation and manipulation opportunities into a canonical model for the optimal design of UI (Baily, 1978; Chetty, 2006), which balances insurance and incentive provision. In its simplest form, our sufficient statistic formula reveals three effects through which manipulation alters the desirability of tagging: (i) the extent to which unintended recipients, henceforth manipulators, are selected on moral hazard as measured by the behavioral to mechanical cost ratio, (ii) the extent to which they are selected on consumption smoothing value and (iii) a manipulation externality capturing the extent to which more differentiation induces more manipulation. The latter makes insurance under manipulation more costly and thus calls for less insurance overall. However, selection effects might work to amplify, mitigate or reverse these conclusions depending on their sign and strength. Intuitively, if manipulators value additional benefits more than their social cost, more differentiation – inclusive of manipulation – might be welfare improving. Conversely, if manipulators are adversely selected on moral hazard, manipulation exacerbates the cost of differentiation.

Second, we develop novel bunching techniques to estimate several key parameters of our model. Building on Diamond and Persson (2016), our methodology exploits the local nature of manipulation and combines traditional bunching and regression discontinuity design (RDD) estimates to uncover selection on both observables and unobservables. We illustrate how the latter can reveal selection effects and treatment effect heterogeneity. In particular, our methodology lets us directly estimate the extent to which manipulators are selected on risk and moral hazard, which links to our theoretical results. Estimating selection on moral hazard has proven notoriously difficult in practice, resulting in relatively little empirical work on the topic, with Einav et al. (2013) and Landais et al. (2021) representing two notable exceptions in the context of health and unemployment insurance, respectively. Our methodology requires neither knowledge about manipulators' identity nor reform-induced policy variation over time, making it readily applicable in other settings.

Third, we apply our methodology in the context of Italian unemployment insurance (UI) and connect the empirics to our theory. Exploiting a discontinuous jump from eight to twelve months of UI coverage around an age-at-layoff threshold and rich administrative social security data, we provide clear graphical evidence of manipulation in the form of systematic delays in the exact timing of layoffs. We find that over 15% of all layoffs occurring within six weeks before workers' fiftieth birthday are strategically delayed. Over the subsequent nonemployment spell affected workers collect on average 2,239 additional Euro each, which correspond to a 38,5% increase in total UI benefit receipt. A survival analysis reveals that approximately 80% of this increase in UI benefit receipt is mechanically due to higher coverage, while the remaining 20% is the result of a decrease in job search effort. This implies that the government pays an additional 25 cents for each Euro of mechanical UI transfer to manipulators. Interestingly, we find virtually the same result when studying non-manipulators, i.e. individuals who were laid off just before their fiftieth birthday. This implies that manipulators are not adversely selected on moral hazard and that selection on moral hazard does not alter the design of optimal policy in our setting.

From a positive perspective our findings mitigate concerns about anticipated moral hazard being the prime motive for selection into manipulation. Rather, we document that manipulators are highly selected on long-term nonemployment risk. Even absent manipulation, manipulators would have exhausted eight months of UI benefits with 16.8 percentage points higher probability than non-manipulators. The underlying firm-worker collusion decision to delay the date of layoff thus acts as an effective screening mechanism for long-term nonemployment risk, while preventing selection on moral hazard. In the last part of the paper we investigate this mechanism further by documenting observable worker and firm characteristics that are associated with manipulation. Manipulation is pervasive among permanent contract workers in private sector firms. We find no evidence of manipulation in public sector firms or among temporary contracts. It is relatively more prevalent among female, part-time, white-collar workers and in firms with less than 50 employees. This suggests that lower adjustment costs and proximity between workers and supervisors may facilitate manipulation in our context.

Our work relates to several strands of the literature. The theoretical model introduces the concept of *tagging* (Akerlof, 1978) into the design of optimal social insurance, in the spirit of Baily (1978) and Chetty (2006) on UI benefit levels and Schmieder et al. (2012) and Gerard and Gonzaga (2021) for potential UI duration.[1] In particular, we study the case of endogenous tags which are perfectly observable at zero cost but subject to manipulation. Importantly, we assume the absence of any verification technology that would allow the government to learn about individuals' (un-manipulated) types. This is in contrast to a large literature on tagging in optimal transfer programs and disability insurance which focuses on imperfect tags that are noisy signals about individual types, but verifiable (at some cost) by the planner, see e.g. Stern (1982), Diamond and Sheshinski (1995), Parsons (1996), Kleven and Kopczuk (2011) among others.

Our setup is both empirically relevant and theoretically interesting. Many policies do indeed tag on perfectly observable individual characteristics – such as marital status, number of dependents or, as in our case, age – with often no ability of inferring manipulation at the individual level. Second, our setting gives rise to both selection on risk and moral hazard, which have traditionally been analysed separately.[2] Recent efforts to integrate the two are presented in Landais et al. (2021), Hendren et al. (2020) and Marone and Sabety (2021). All three contributions study the welfare implications of offering some form of choice in (regulated) insurance markets.[3] Importantly, and conceptually different from our setup, these papers focus on policies that do not discriminate between different individuals but rely on self-selection through market prices.[4]

It is worth pointing out that our focus is on how to design optimal differentiated policy under manipulation based on a *given* (endogenous) tag, in our case, age at layoff. Although interesting in its own right, we do not directly speak to the appropriateness of tagging on age *per se*, nor do we empirically evaluate how much differentiation would be optimal. Notably, tagging on age has been discussed in several contexts, including in UI from a life-cycle perspective Michelacci and Ruffo (2015) and in optimal Mirrlessian taxation, see Weinzierl (2011), Best and Kleven (2013), among others.

[1]See Spinnewijn (2020) for a discussion on the importance of more conceptual work on optimal differentiated social insurance as well as for suggestions on how to get started.

[2]While moral hazard is the key concept in most of the work on unemployment insurance design, adverse selection has received a lot of attention in the context of health insurance, in particular, in the US context.

[3]In their work Landais et al. (2021) provide the first assessment of the desirability of a UI mandate in the Swedish context. Adverse selection under a universal mandate has also been studied in the context of health insurance, see Hackmann et al. (2015).

[4]Barnichon and Zylberberg (2021) show that it might be theoretically desirable to offer a menu of contracts to the unemployed screening individuals by how they trade lump-sum severance payments with UI benefits.

The fact that we find positive selection on long-term nonemployment risk also speaks to a literature studying the role of private information and (ex-post) adverse selection in explaining the unravelling of insurance markets, see e.g. Hendren (2017) for unemployment insurance and Cabral (2016) for dental insurance. Our results indicate that individuals hold information about their expected duration of unemployment around the time of layoff. Understanding to what degree this information is held privately is beyond the scope of this paper.

From a methodological perspective, our empirical strategy is most closely related to recent work by Diamond and Persson (2016), who study manipulation of test scores in Swedish high-stakes exams.[5] They propose a bunching estimator to estimate the effect of teacher discretion in grading around important exam thresholds on students future labour market outcomes. They also show how these techniques can be used to study selection on observables. We extend their methodology to investigate selection on unobservables and to uncover treatment effect heterogeneity. We borrow several ideas from standard bunching techniques recently surveyed by Kleven (2016). Conceptually, our empirical insights also relate to the literature on "essential heterogeneity" in instrumental variable settings, in which individuals select into treatment in part based on their anticipated treatment effect, see e.g. Heckman et al. (2006).

On the empirical side, a large body of work studies the disincentives effects of UI exploiting similar policy variation, see e.g. Card et al. (2007), Lalive (2007), Schmieder et al. (2012), Landais (2015), Nekoei and Weber (2017), Johnston and Mas (2018) among others. Contrary to our setting, these papers rely on the *absence* of manipulation to identify the treatment effects of interest, whereas we study the effect of manipulation in a setting where it does occur. Two recent contributions by Doornik et al. (2020) and Khoury (2019) also study manipulation in UI systems around an eligibility and seniority threshold in Brazil and France, respectively. Doornik et al. (2020) provide evidence of strategic collusion between workers and firms who time layoffs to coincide with workers' eligibility for UI in Brazil. Khoury (2019) exploits a discontinuity in benefit levels for workers laid off for economic reasons and estimates an elasticity of employment spell duration with respect to UI benefits of 0.014. While both of these papers suggest that manipulation in social insurance contexts are widespread, neither studies the welfare consequences of manipulation or estimates selection effects as we do.

The remainder of the paper is organized as follows: Section II covers the theoretical analysis for which Section II.A introduces the formal model, Section II.B discusses the main assumptions, Section II.C derives our main results and Section II.D connects our theory to the data; Section III contains our empirical application with Section III.A presenting the institutional setting and data, Section III.B outlining the empirical strategy

[5]For an example of test score manipulation in the US context, see Dee et al. (2019) who study the impact of manipulation of test scores in New York Regents Examinations on students subsequent educational outcomes.

and Section III.C reporting our results and robustness checks; Section IV concludes.

II Theory: Optimal Targeted Social Insurance with Manipulation

This section lays out a model for the design of optimal differentiated social insurance in the presence of manipulation opportunities. We stay deliberately close to our empirical setting to facilitate the connection between the theoretical and empirical part of the paper. Although the model is derived in the context of unemployment insurance duration, our results readily extend to other social insurance settings.

II.A The Model

II.A.1 Setting

We assume there are two groups of individuals, referred to as the "young" and the "old" and denote their exogenous share in the population by G and $1 - G$, respectively. Young and old individuals differ in their utility of consumption, job search costs and their ability to manipulate (more on this below). All individuals are unemployed in $t = 0$, retire at a finite time horizon T and are hand-to-mouth consumers.[6]

The government provides unemployment benefits b, financed through a lump-sum UI tax τ. Young and old individuals enjoy consumption $c_u + b$ when unemployed and covered by UI, c_u when unemployed and not covered, and $c_e = w - \tau$ when employed, where w denotes the exogenous wage rate. The governments sets two separate UI schemes of varying generosity characterized by two different potential benefit durations P_y and P_o, with $P_o \geq P_y$. It targets the longer potential UI benefit duration P_o to the old. When doing so it faces a challenge: young individuals have the ability to manipulate their eligibility status (at some cost) and might endogenously select into the more generous scheme intended for the old.[7] In order to study how a benevolent government should optimally set P_y and P_o in this context, we begin by formally stating individuals' job search problems.

II.A.2 The Old

Preferences and Job Search. Old workers are homogeneous, always eligible for longer potential benefit duration P_o, and face the standard job search problem. They enjoy flow utility $u^o(c)$ at consumption level c and choose job search intensity s_t^o at time t,

[6] The model setup closely follows previous work on optimal potential benefit duration in UI, e.g. Schmieder et al. (2012) and Gerard and Gonzaga (2021).

[7] Since the two policies differ only in terms of their potential benefit duration, with $P_o \geq P_y$, we w.l.o.g. restrict attention to one-sided manipulation.

normalized to the arrival rate of job offers, at utility cost $\phi^o_t(s^o_t)$. Formally, old individuals maximize:

$$V^o(P_o) = \max_{s^o_t} \left\{ \int_0^{P_o} S^o_t u^o(c_u + b) + \int_{P_o}^T S^o_t u^o(c_u) + \int_0^T \left(1 - S^o_t\right) u^o(c_e) - \int_0^T S^o_t \phi^o_t(s^o_t) \right\},$$

where $S^o_t = exp\left(-\int_0^t s^o_{t'} dt'\right)$ denotes the nonemployment survival probability at time t and all integrals are taken w.r.t. dt. We denote the old's implied benefit and nonemployment duration by

$$B^o(P_o) = \int_0^{P_o} S^o_t(P_o)dt \quad \text{and} \quad D^o(P_o) = \int_0^T S^o_t(P_o)dt.$$

II.A.3 The Young

Preferences and Job Search. Young individuals have heterogeneous preferences and are characterized by utility of consumption u^i, job search cost function ϕ^i and fixed cost q^i. Conditional on eligibility for potential benefit duration $\tilde{P}$, young individuals maximize search effort as follows:

$$\tilde{V}^i(\tilde{P}) = \max_{s^i_t} \left\{ \int_0^{\tilde{P}} S^i_t u^i(c_u + b) + \int_{\tilde{P}}^T S^i_t u^i(c_u) + \int_0^T \left(1 - S^i_t\right) u^i(c_e) - \int_0^T S^i_t \phi^i_t(s^i_t) \right\},$$

where $S^i_t = exp\left(-\int_0^t s^i_{t'} dt'\right)$ denotes individuals' nonemployment survival probability at time t and all integrals are w.r.t. dt. Denote an individual's implied benefit and nonemployment duration by:

$$B^i(\tilde{P}) = \int_0^{\tilde{P}} S^i_t(\tilde{P})dt \quad \text{and} \quad D^i(\tilde{P}) = \int_0^T S^i_t(\tilde{P})dt.$$

Manipulation. At time zero, young individuals can engage in manipulation by incurring a fixed cost $q^i \geq 0$ to become eligible for potential benefit duration P_o rather than P_y, with $P_o \geq P_y$. Formally, a young individual i with fixed cost q^i maximizes:

$$\begin{aligned} V^i(P_o, P_y) &= \max_{a^i \in \{0,1\}} \left\{ \left(\tilde{V}^i(P_o) - q^i\right) \cdot \mathbb{1}_{a^i=1} + \tilde{V}^i(P_y) \cdot \mathbb{1}_{a^i=0} \right\} \\ &= \tilde{V}^i(P_y) + \max_{a^i \in \{0,1\}} \left\{ \left(\tilde{V}^i(P_o) - \tilde{V}^i(P_y) - q^i\right) \cdot \mathbb{1}_{a^i=1} \right\}, \end{aligned}$$

where a^i encodes the choice of whether ($a^i = 1$) or not ($a^i = 0$) to manipulate.

Thus, young individual i manipulates if and only if

$$q^i \leq \bar{q}^i(P_o, P_y) \equiv \tilde{V}^i(P_o) - \tilde{V}^i(P_y). \tag{1}$$

Preferences and fixed costs are distributed according to a continuously differentiable pdf $f(u^i, \phi^i, q^i)$. We denote the share of young individuals who manipulate – henceforth manipulators – by $M(P_o, P_y))$ and the benefit and nonemployment durations of manipulators and non-manipulators respectively by:

$$B^m(P_o, P_y) = \mathbb{E}\left[B^i(P_o)|a^i(P_o, P_y)) = 1\right] \quad \text{and} \quad D^m(P_o, P_y)) = \mathbb{E}\left[D^i(P_o)|a^i(P_o, P_y)) = 1\right],$$
$$B^n(P_o, P_y)) = \mathbb{E}\left[B^i(P_y)|a^i(P_o, P_y)) = 0\right] \quad \text{and} \quad D^n(P_o, P_y)) = \mathbb{E}\left[D^i(P_y)|a^i(P_o, P_y)) = 0\right].$$

The average benefit and nonemployment durations for the young are

$$B^y(P_o, P_y)) = M(P_o, P_y)) \cdot B^m(P_o, P_y)) + \left(1 - M(P_o, P_y))\right) \cdot B^n(P_o, P_y))\big)$$
$$D^y(P_o, P_y)) = M(P_o, P_y)) \cdot D^m(P_o, P_y)) + \left(1 - M(P_o, P_y))\right) \cdot D^n(P_o, P_y))\big),$$

and we denote by $V^y(P_o, P_y) = \mathbb{E}\left[V^i(P_o, P_y)\right]$ the average utility of the young and use superscripts to denote conditional expectation operators.

II.A.4 The Planner's Problem

A benevolent social planner sets (P_o, P_y) to maximize ex-ante social welfare taking into account the incentive constraints, including the fact that manipulation might occur. Concretely, the planner's objective is given by:

$$W(P_o, P_y) = (1 - G) \cdot V^o(P_o) + G \cdot V^y(P_o, P_y),$$

subject to the budget constraint:

$$L \cdot \tau = U \cdot b + R,$$

with total labor supply $L = (1 - G)(T - D^o(P_o)) + G(T - D^y(P_y)) + GM(D^m(P_y) - D^m(P_o))$, total unemployment covered by unemployment benefits $U = (1 - G)B^o(P_o) + GB^y(P_y) + GM(B^m(P_o) - B^m(P_y))$ and exogenous government spending R.

II.B Simplifying Assumptions

We assume that the planner's optimization problem is well-behaved warranting a first-order approach. In order to ease the exposition and gain tractability, we impose two additional simplifying assumptions: The first corresponds to a constant elasticity assumption while the second restricts dynamic screening opportunities. The formal derivations in Appendix A make explicit how each assumption is used and how our results generalize. To state our assumptions precisely we introduce two key concepts for the analysis.

The first is a measure of the disincentive or moral hazard effect of UI in the context of extended potential benefit duration (PBD). Note that in the case of PBD, extra statutory coverage may mechanically lead to higher benefit receipts if individuals stay unemployed during the additional months with and without extra coverage. This cost increase for the government is not due to distorted job search incentives but simply reflects nonzero exhaustion risks during the relevant months of nonemployment. Because there is no distortion, such mechanical transfer is not, by itself, welfare relevant. What matters is by how much individuals change their behavior, and thereby increase the cost of UI, for each dollar of such mechanical transfers.

Concretely, we follow Schmieder and von Wachter (2017) and define the behavioral to mechanical cost ratio for individual i when marginally increasing PBD P as:

$$\frac{BC_P^i}{MC_P^i} = \frac{b \cdot \int_0^P \frac{dS_t^i}{dP} dt + \tau \cdot \int_0^T \frac{dS_t^i}{dP} dt}{b \cdot S_P^i}. \tag{2}$$

The above BC/MC ratio has a classical leaking bucket interpretation. It captures by how many additional dollars total UI expenditure goes up for each dollar of mechanical transfer from the government to the unemployed.[8] We illustrate BC/MC ratios graphically in Figure I and refer to it simply as moral hazard throughout.

Second, we define the "marginal" utility of individual i at the point of benefit exhaustion $\tilde{u}_i'$ as

$$\tilde{u}_i' = \frac{1}{b} \int_0^b (u^i)'(c_u + x) dx = \frac{u^i(c_u + b) - u^i(c_u)}{b}. \tag{3}$$

Since we are working with benefit duration extensions, the relevant utility gap is between receiving and not receiving UI benefits during unemployment (which are the numeraire in the right-most term in equation (3)). We conveniently recast this gap into the appropriately weighted marginal utility. Note that neither consumption nor utilities are time dependent in the current setup which makes (3) time-invariant. However, it is straightforward to allow for time dependence in utility and consumption.

[8]An important property of this measure of moral hazard is its comparability across different (groups of) individuals. This is especially important in the context of unemployment duration because individuals might have heterogeneous exhaustion risk and thus face different incentives to respond to PBD extensions. As in previous work, it turns out that it is precisely this re-scaled moral hazard effect that is relevant for optimal policy in our setting.

Equipped with the above concepts we impose the following assumptions. First, we assume a constant, i.e. time-invariant, moral hazard cost for the young. Concretely, we assume:

Assumption 1. *Moral hazard is constant over the UI spell. Formally, for each I subset of the young we have*

$$\frac{BC_P^I}{MC_P^I} = \frac{BC_{P'}^I}{MC_{P'}^I} \quad \textit{for all } P, P'.$$

Second, we assume that exhaustion risks and marginal utilities are uncorrelated.

Assumption 2. *Exhaustion risks and marginal utilities are uncorrelated: Formally, for all I subset of the young we have*

$$\mathbb{C}ov^I \left[S^i_{\tilde{P}}, \tilde{u}'_i \right] = 0 \ \textit{for all} \ \ \tilde{P}.$$

Assumption 1 is akin to a constant elasticity assumption. It requires that the behavioral to mechanical cost ratio remains constant over the UI spell which intuitively assumes a time-invariant responsiveness to UI transfers. Assumption 2 implies that exhaustion risks are uninformative of marginal utilities. On the one hand, high-marginal utility individuals might have stronger incentives to find a job which would violate assumption 2. However, to the extent that unemployed individuals deplete their assets over the UI spell, marginal utilities might in fact increase over the spell which would push the correlation in the opposite direction. Assumption 2 thus requires that such forces exactly offset each other. Both assumptions are assumptions on individual behavior but also implicitly restrict the space of possible selection pattern among the young because they have to hold for each subset of the young.[9] This makes the analysis considerably more tractable but rules out dynamic screening possibilities. For instance, exhaustion risks cannot be used to dynamically screen high marginal utility individuals. We regard our simplified setup as a natural starting point for the analysis and leave its generalization to future work.

II.C Characterizing Optimal Policies

We parameterize policy $(P_o, P_y) = (P + \Delta P, P)$, such that P represents the level of baseline coverage and $\Delta P \geq 0$ reflects the amount of extra coverage. Before turning to the full optimum, we briefly focus on two related (sub-)problems that help building intuition. First we look at the case without manipulation.

Optimum without Manipulation. In Appendix A we show that the optimal policy in the absence of manipulation opportunities is given by:

[9]It suffices if assumptions 1 and 2 hold for all possible sets of manipulators and non-manipulators.

Proposition 1 (Optimum without manipulation)**.** *The optimal policy* (P_o^*, P_y^*) *without manipulation, i.e.* $M \equiv 0$*, satisfies:*

$$\frac{\tilde{u}'_o - \bar{u}'}{\bar{u}'} = \frac{BC^o}{MC^o} \quad and \quad \frac{\tilde{u}'_y - \bar{u}'}{\bar{u}'} = \frac{BC^y}{MC^y},$$

where $\bar{u}' = (1-G)\cdot(T-D^o)\cdot(u^o)'(c_e) + G\cdot(T-D^y)\cdot(u^y)'(c_e)$ *is the average marginal utility of the employed and* $\tilde{u}'_j$ *and* $\frac{BC^j}{MC^j}$ *defined in (3) and (2) for* $j = y, o$*.*

Proposition 1 follows previous results in the literature on optimal UI benefit duration, e.g. Schmieder and von Wachter (2017). As in the classical Baily-Chetty formula, the optimal policy without manipulation equates consumption smoothing benefits with moral hazard costs for the old and the young *separately*.[10]

Introducing Manipulation. To build further intuition, we now study the introduction of manipulation by first imagining a world without the old, i.e. $G = 1$. The extra coverage ΔP now simply represents an alternative contract into which some of the young might self-select. As we show in Appendix A, the (re-scaled) welfare effect of marginally increasing extra coverage ΔP starting from the case where there is none $\Delta P = 0$ is given by:

$$\frac{1}{M\cdot MC^m_{P_y}\cdot\bar{u}}\cdot\left.\frac{dW}{d(\Delta P)}\right|_{\Delta P=0} = \frac{\tilde{u}'_m - \bar{u}'}{\bar{u}'} - \frac{BC^m}{MC^m} \tag{4}$$

What matters for welfare at the margin is the insurance surplus, that is the difference between the consumption smoothing benefits and the moral hazard cost, of manipulators. It is instructive to evaluate this expression at the optimal manipulation-free policy P_y^* from Proposition 1.

Proposition 2 (The marginal welfare effect of manipulation at P_y^*)**.** *The marginal budget-balanced welfare effect of increasing extra coverage at* P_y^* *from Proposition 1 is given by:*

$$\frac{1}{M\cdot MC^m_{P_y}\cdot\bar{u}}\cdot\left.\frac{dW(P_y^*)}{d(\Delta P)}\right|_{\Delta P=0} = \underbrace{\left(\frac{\tilde{u}'_m - \tilde{u}'_y}{\bar{u}'}\right)}_{\substack{\text{selection on consumption}\\ \text{smoothing value}}} - \underbrace{\left(\frac{BC^m}{MC^m} - \frac{BC^y}{MC^y}\right)}_{\substack{\text{selection on moral}\\ \text{hazard cost}}}$$

Proposition 2 shows that the welfare effect of additional coverage depends on the extent to which manipulators are selected on consumption smoothing value and moral hazard cost at the optimally set manipulation-free policy P_y^*. If manipulators have higher insurance surplus than the average young individual, manipulation increases welfare and vice versa. This result mimics that of Hendren et al. (2020) who study the

[10]Note that the current setup imposes a common tax rate for the old and the young and the problem is thus not entirely separable across groups. It is straightforward to allow for different tax schedules across groups.

welfare effect of allowing for choice in a social insurance context.[11]

It turns out that selection effects, like the one in Proposition 2, remain crucial for determining the full optimal policy with manipulation which we turn to next.

Optimum with Manipulation. We now analyze the design of optimal policy in the presence of both groups young and old, i.e. $G \in (0,1)$ *and* with (potential) nonzero manipulation. At the optimum, small budget-neutral changes $d\Delta P$ in extra coverage ΔP which cannot increase welfare. In Appendix A we show that this implies

$$(1-G)\cdot S^o_{P_o}\cdot\left[\frac{\tilde{u}'_o-\bar{u}'}{\bar{u}'}-\frac{BC^o}{MC^o}\right]+G\cdot M\cdot S^m_{P_o}\cdot\left[\frac{\tilde{u}'_m-\bar{u}'}{\bar{u}'}-\frac{BC^m}{MC^m}\right]$$
$$+G\cdot(1-M)\cdot S^n_{P_y}\cdot\epsilon_{1-M,\Delta P}=0, \tag{5}$$

where all variables are defined as above and $\epsilon_{1-M,\Delta P}$ refers to the cost-weighted elasticity of manipulation w.r.t. extra coverage ΔP which we define formally below. Equation (5) generalizes equation (4) by introducing two additional terms (the first and third term). The first term takes into account that the old, who are always entitled to receiving higher coverage P_o, have a direct welfare effect from increases in extra coverage. The third terms captures the fact that marginal manipulators might cause non-marginal changes in the government budget, because we are no longer starting at a point without any additional coverage. Concretely, define the fiscal externality from manipulation, that is the budgetary cost arising from higher benefit receipt and lower tax revenue, of all individuals of type $i=(u_i,\phi_i)$ as

$$FE^i=\Big(B^i(P+\Delta P)-B^i(P)\Big)\cdot b+\Big(D^i(P+\Delta P)-D^i(P)\Big)\cdot\tau \tag{6}$$

and the share of these individuals who end up manipulating because their fixed cost falls below the threshold $\bar{q}^i$ in equation (1), as

$$M^i=\int_0^{\bar{q}^i} f(q|u_i,\phi_i)dq. \tag{7}$$

Equipped with these two quantities we formally define the cost-weighted elasticity of manipulation w.r.t. extra coverage introduced in equation (5) as follows

$$\epsilon_{1-M,\Delta P}=\mathbb{E}^n\left[\frac{FE^i}{MC^n_{P_y}\cdot\Delta P}\cdot\epsilon_{1-M^i,\Delta P}\right]. \tag{8}$$

Thus the elasticity term captures by how much each share M^i, as measured by $1-M^i$,

[11]While Hendren et al. (2020) are interested in price surcharges required for extra coverage, a feature one could also include in our setup, we model manipulation as an entirely private choice without any *direct* financial implications for the government. The manipulation fixed cost q^i is relevant for individual utilities but not for government revenue.

responds to increases in extra coverage weighted by the cost that such changes impose on the government budget.

Turning to the optimal level of baseline coverage, we again have that at the optimum, marginal budget-neutral changes dP in baseline coverage P cannot increase welfare. As shown in Appendix A, by the envelope theorem this implies

$$(1-G)\cdot S^{o}_{P_o}\cdot\left[\frac{\tilde{u}'_o-\bar{u}'}{\bar{u}'}-\frac{BC^o}{MC^o}\right]+G\cdot S^{y}_{P_y}\cdot\left[\frac{\tilde{u}'_y-\bar{u}'}{\bar{u}'}-\frac{BC^y}{MC^y}\right]$$
$$+G\cdot M\cdot\left(S^{m}_{P_o}-S^{m}_{P_y}\right)\cdot\left[\frac{\tilde{u}'_m-\bar{u}'}{\bar{u}'}-\frac{BC^m}{MC^m}\right]+G\cdot(1-M)\cdot S^{n}_{P_y}\cdot\epsilon_{1-M,P}=0, \tag{9}$$

where $\epsilon_{1-M,P}$ is the cost-weighted elasticity of manipulation w.r.t. baseline coverage P, defined analogously as in equation (8) but with respect to baseline coverage P. Intuitively, when deciding how much baseline coverage to provide, the planners weighs the surplus from the old (first term), the young (second term), an adjustment accounting for the fact that a subset of the young are in fact manipulators with now different exhaustion risk (third term) and the effect of baseline coverage on the extent of manipulation (fourth term).

Combining equations (5) and (9) leads to our main proposition regarding the optimal policy under manipulation.

Proposition 3 (Optimum with manipulation)**.** *The optimal policy with manipulation satisfies:*

$$\frac{\tilde{u}'_y-\bar{u}'}{\bar{u}'}-\frac{BC^y}{MC^y}=\underbrace{\epsilon_{1-M,\Delta P}}_{\substack{\text{manipulation externality}\\ \text{of extra coverage}}}-\underbrace{\epsilon_{1-M,P}}_{\substack{\text{manipulation externality}\\ \text{of baseline coverage}}}$$
$$+M\cdot\underbrace{\left(\frac{S^{m}_{P_y}}{S^{y}_{P_y}}\right)}_{\substack{\text{selection on risk}\\ \text{scale factor}}}\cdot\left\{\underbrace{\left(\frac{\tilde{u}'_m-\tilde{u}'_n}{\bar{u}'}\right)}_{\substack{\text{selection on consumption}\\ \text{smoothing value}}}-\underbrace{\left(\frac{BC^m}{MC^m}-\frac{BC^n}{MC^n}\right)}_{\substack{\text{selection on moral}\\ \text{hazard cost}}}\right\} \tag{10}$$

and

$$(1-G)\cdot S^{o}_{P_o}\cdot\left(\frac{\tilde{u}'_o-\bar{u}'}{\bar{u}'}-\frac{BC^o}{MC^o}\right)+G\cdot S^{y}_{P_o}\cdot\left(\frac{\tilde{u}'_y-\bar{u}'}{\bar{u}'}-\frac{BC^y}{MC^y}\right)$$
$$=G\cdot(1-M)\cdot\left(\left(S^{n}_{P_o}-S^{n}_{P_y}\right)\cdot\epsilon_{1-M,\Delta P}-S^{n}_{P_o}\cdot\epsilon_{1-M,P}\right) \tag{11}$$

First note that without manipulation, i.e. $M\equiv 0$, Proposition 3 nests Proposition 1. However, the presence of manipulation induces a wedge in the provision of insurance for both young and old. Equation (10) shows that the wedge for the young is determined by two elasticities, namely that of extra and baseline coverage, and by a selection term, capturing the extent to which manipulators are selected on consumption smoothing value and moral hazard cost. Equation (11) implies that the wedge for the old is the direct counterpart of that for the young together with an effect on the overall level of insurance (RHS). In order to build intuition, it is instructive to consider two special cases.

Fixed, nonzero M. First, consider a scenario in which a fixed subset of young individuals manipulate irrespectively of policy and always obtain higher UI coverage. In this case the share M is nonzero and unresponsive to the design of UI. As a consequence, all elasticity terms in Proposition 3 are zero. It is straightforward to show that equation (10) implies that

$$\frac{\tilde{u}'_n - \bar{u}'}{\bar{u}'} - \frac{BC^n}{MC^n} = 0, \tag{12}$$

which means that consumption smoothing benefits and moral hazard cost for non-manipulators or the 'endogenous young' are equated. Similarly equation (5) shows the same holds true for the 'endogenous old', i.e. the group of the old and manipulators. Note that equation (11) implies that such manipulation induces no distortion in the desired overall level of insurance. Intuitively, this is a case of pure re-labelling, in which the planner regards a subset of the young as old because their manipulation choice is unresponsive to policy.

Homogeneous young. Suppose there is no heterogeneity among the young, except potentially in their manipulation fixed cost. In this case the selection term in equation (10) vanishes and the wedge of the young is governed only by the two elasticities. If one assumes that additional coverage weakly increases the share of manipulators and that additional baseline coverage weakly decreases it, then the wedge of the young is unambiguously negative, calling for overinsurance. Intuitively, it is optimal for the planner to grant the young additional surplus, above and beyond their manipulation-free level, because of their manipulation threat. To the extent that additional coverage mitigates manipulation the planner finds it optimal to provide such insurance to the young. Contrary, by equation (11), the old will be underinsured by more than the wedge for the young representing the fact that shifting insurance surplus is now costly due to the fiscal externality associated with manipulation.

II.D Connecting Theory and Empirics

This section lays out how to connect our theoretical framework to the data. There are several points worth emphasizing. First and foremost, the purpose of our theory is to guide the design of differentiated policy w.r.t. a *given* endogenous tag, not for choosing among several potential tags or assessing their appropriateness more generally. A full implementation of proposition 3 would nevertheless reveal whether or not differentiation w.r.t. to a tag has any potential benefit or if the optimal policy is in fact undifferentiated. Finding out which heterogeneities allow welfare-improving targeting in different policies is a fruitful avenue for future research, although policy makers might ultimately refrain from exploiting some of them, because of e.g. administrative costs or horizontal equity and fairness concerns.[12]

Second, our theory takes the *degree* of initial differentiation, that is, the grouping of individuals, in our setup two groups of young and old, as given. This has important consequences for any empirical implementation in which the classification itself is a policy choice. Our theory does not directly speak to the optimal classification but rather analyses the effect of manipulation for a given grouping of individuals. This implies that any statement about the welfare-relevance of manipulation is always with respect to a reference degree of differentiation, over which there might be empirical ambiguity.

To illustrate this point, consider a scenario in which group membership is defined by a cutoff rule in some cardinal individual characteristic which can be manipulated by individuals at some cost, as will be the case in our empirical application below. If the manipulation cost increases with distance form the threshold, manipulation will tend to be locally concentrated around the threshold. Whether or not manipulation matters for welfare in this setting depends on the definition of what constitutes the relevant groups. For instance, if a large number of individuals is located far away from the threshold and one considers all of these individuals as part of the two groups, one might trivially conclude that manipulation is not globally welfare-relevant, essentially because $M \approx 0$. However, this is precisely the case in which the policy is a two-part policy in a large population and thus not very ambitiously targeted. The importance of manipulation increases mechanically with the degree of differentiation, *ceteris paribus*. The smaller the group of targeted individuals the more relevant manipulation effects become, because it is easier for the share M of manipulators to rise to meaningful levels.

Third and relatedly, given that there is no "correct" classification, our empirical application focuses on developing a methodology to estimate the empirical moments in Proposition 3, rather than to provide a welfare assessment of any one particular policy. We do point out explicitly how to connect our estimates to the theory as well as which other moments might be of interest. Concretely, we illustrate how bunching techniques can be used to reveal the extent of selection on moral hazard even in the absence of policy reforms. Although of equal theoretical interest, we lack the data, variation and methods to estimate the corresponding selection on value counterpart.[13] We do discuss some tentative findings based on our selection on observables analysis in Section III.B.4.

[12]Although not part of the current model, it is straightforward to incorporate other objectives, e.g. welfare weights, in the analysis.

[13]Identifying and estimating the consumption smoothing benefit of UI has proven a considerable challenge in the literature by itself. Our setting features two additional complications: the fact that we are interested in estimating the difference in marginal utilities between two groups of individuals and that this gap is measured at the respective time of benefit exhaustion. We are unaware of any work which estimates marginal utilities of UI exhaustees directly.

III Empirics: Manipulation in Italian Unemployment Insurance

III.A The Italian Unemployment Insurance Scheme

III.A.1 Institutional Setting

We study manipulation in Italy's *Ordinary Unemployment Benefits* (OUB) scheme.[14] The OUB was in effect from the late 1930s until its abolishment and replacement in January 2013.[15] OUB covered all private non-farm and public sector employees who lost their job either due to the termination of their temporary contract, or due to an involuntary termination (a layoff), or a quit for just cause, such as unpaid wages or harassment. Other types of voluntary quits and the self-employed were not eligible for OUB.[16]

To qualify for OUB, workers were also required to have some labor market attachment. Concretely, workers needed to have started their first job spell at least two years before the date of layoff, and to have worked for at least 52 weeks in the previous two years.[17]

Benefit levels were based on the average monthly wage, calculated over the three months preceding the layoff. The replacement rate was declining over the unemployment spell: 60% of the average wage for the first six months; 50% for the following two months and 40% for any remaining period. OUB did not involve any form of experience rating.

PBD under OUB was a sole function of age at layoff and amounted to eight months if the layoff preceded the worker's fiftieth birthday and twelve months if it followed it. This discontinuous change (a notch) in coverage created a strong incentive for workers to delay their date of layoff so that it falls after their fiftieth birthday.

III.A.2 Data

We use confidential administrative data from the Italian Social Security Institute (INPS) on the universe of UI claims in Italy between 2009 and 2012 and combine them with matched employer-employee records covering the universe of working careers

[14] *Indennità di Disoccupazione Ordinaria a Requisiti Normali* in Italian. We are not the first to study the Italian OUB scheme, see e.g. Anastasia et al. (2009), Scrutinio (2018) and Albanese et al. (2020), of which we discuss the last in more detail in Appendix C.

[15] OUB was introduced through *Regio Decreto 14.* in April 1939 and replaced by ASPI on of January 1, 2013.

[16] For convenience, in the rest of the paper we will use the term "layoff" to indicate all job terminations that are eligible for UI.

[17] Two other UI benefit schemes were in place in Italy at the same time of our analysis: Reduced Unemployment Benefits (RUB) and Mobility Indemnity (MI). However, neither one is likely to interfere with our analysis due to different eligibility conditions and less generous benefit coverage. For completeness, we present the two other UI schemes in Appendix B.

in the private sector. Information on UI claims comes from the SIP database,[18] which collects data on all income support measures administered by INPS as a consequence of job separation. For every claim we observe the UI benefit scheme type, its starting date, duration and amount paid. We further observe information related to the job and the firm. This includes details about the type of the contract and a broad occupation category.

The SIP database does not contain the date of re-employment after receiving UI. We therefore retrieve this information from the matched employer-employee database (UNIEMENS) and construct nonemployment durations as the time difference between the layoff date in the SIP and the first re-employment in UNIEMENS.[19] The UNIEMENS database provides additional information on workers' careers in the private sector, including detailed information on wages and the type of contract. We observe individuals in the UNIEMENS database until 2016, which gives us at least four years of observations for all workers. We therefore censor all nonemployment durations at this horizon.

For our main sample we restrict our attention to individuals who lost their job between February 2009 and December 2012, were between 46 and 54 years of age at the time of layoff, and claimed OUB. Unfortunately, our data does not cover the years prior to February 2009 and the introduction of a new UI scheme in January 2013 prevents us from including later years. We further restrict attention to individuals who separate from an employer in the private sector after a permanent contract. The motivation for this is twofold. First, we show in Section III.B.4 that manipulation is confined to permanent contracts in the private sector. Second, the UNIEMENS database does not contain job information for public sector jobs, which means we have no information about the previous work arrangement, nor would we observe re-employment. At this point, one might be worried that we are missing some re-employment events, namely, those into public sector jobs. This in unlikely to affect our results because transitions from private into public sector jobs should be rare for workers at such late stage in their careers. After the exclusion of a few observations with missing key information we are left with 249,581 separation episodes that led to UI claims.

Table I reports summary statistics for our main sample. The average worker receives UI for about 30 weeks (7 months) corresponding to roughly one third of the 90 weeks (21 months) average nonemployment duration. An average of 50% and 39% of workers are still nonemployed after eight and twelve months, respectively, implying substantial exhaustion risk. Our sample of workers is predominately male, on full time contracts, and employed in blue collar jobs. Workers have spent about 27.5 years in the labor market since their first job and almost 6 years in their last firm. In terms

[18] *Sistema Informativo Percettori* in Italian.

[19] We restrict the latter to be later than the former, which excludes a few short-term jobs that are compatible with the continuation of UI benefit receipt.

of geographic distribution, 46% of workers are laid off in the South or the Islands.[20] Workers earned about 70 Euro per day (gross) which is equivalent to $70 \times 26 = 1820$ Euro per month if working full time.[21] The separating firm is relatively old (14 years) and large (28.16 employees), but this is driven by a few very large firms. Indeed, more than 60% of workers come from firms with less than 15 employees while only 18% come from firms with more than 50 employees. Because our main sample contains workers in their late forties and early fifties, one might be concerned that transitions into retirement could play a non-negligible role. However, this is not the case with only about 1,500 or 0.6% of workers in our sample claiming retirement benefits before the end of our observation window (4 years since layoff).[22] We now turn to a description of our objects of interest and identification strategy.

III.B Empirical Strategy

This section sketches our empirical strategy and explains the sources of variation in the data that we use to pin down different parameters of interest. The main idea is to exploit the local nature of manipulation by extrapolating outcomes from regions that are unaffected by it, to learn about what would have happened in a counterfactual world without it. We first assess the range of the manipulation region with standard bunching techniques. We then fit polynomials to the unmanipulated part of the data and interpolate to construct a counterfactual layoff frequency and recover the number (and share) of manipulators. Similarly, we construct counterfactuals of outcomes that are not directly manipulated, such as subsequent benefit receipt or nonemployment survival probabilities, to learn whether these outcomes respond to manipulation. Intuitively, any unusual change in these outcomes near the cutoff together with how many manipulators are causing it, let us recover manipulators' responses. Under plausible assumptions, we also recover the response of non-manipulators, a group of individuals laid off just before their fiftieth birthday. We also illustrate how we can use part of the procedure just described to study selection into manipulation. Our approach is closely related to that of Diamond and Persson (2016).

III.B.1 Quantifying manipulation

Consider a hypothetical manipulated layoff density as in Figure IIa. Absent any manipulation we would expect the frequency of layoffs to be smooth in the neighborhood of the cutoff. Manipulation instead causes a sharp drop in the number or layoffs right before and a spike right after age fifty. We refer to the first region as the "missing" and

[20] This area encompasses the following regions: Abruzzo, Basilicata, Calabria, Molise, Puglia, Sardegna and Sicilia.

[21] This information is consistent with the monthly wage reported in our second data source, the SIP database, which reports an average monthly wage of 1,735 Euro in the three months preceding the layoff.

[22] For these workers we define the nonemployment spell as the period between the end of the previous employment and the date at which they claim their pension.

the later the "excess" region which together make up the "manipulation" region. As in standard bunching techniques, we recover the counterfactual frequency of layoffs by fitting a polynomial to the unmanipulated parts of the data (on the left and right of the cutoff) and interpolate inwards. The difference between the observed frequency and the fitted counterfactual lets us recover missing and excess shares, as well as the number of manipulators in the missing and excess regions. This estimation strategy assumes that manipulation takes the form of a pure re-timing of layoffs that would have occurred anyways and for which we provide supporting evidence in Section III.C.6.

We operationalize this identification strategy following standard bunching techniques, e.g. Saez (2010), Chetty et al. (2011), Kleven and Waseem (2013). First, we group all layoffs into two-week bins based on the workers' age at layoff. Second, we determine the lower bound of the missing region z_L by visual inspection, in our case three bins or six weeks. Last, we iteratively try different upper bounds for the excess region z_U until we balance the missing and excess "mass", that is, the estimated number of manipulators on either side of the threshold. We estimate the number of manipulators by fitting a second order polynomial to the observed layoff frequency, including a full set of dummies for bins in the manipulation region, and retrieving the relevant regression coefficients. In practice, we estimate the following specification:

$$c_j = \alpha + \sum_{p=0}^{P} \beta_p \cdot a_j^p + \sum_{k=z_L}^{z_U} \gamma_k \cdot \mathbb{I}[a_j = k] + \nu_j, \tag{13}$$

where c_j denotes the absolute frequency of layoffs in headcounts in bin j, a_j is the mid-point age in bin j, P denotes the order of the polynomial. The coefficients γ_k recover the differences between the observed data and the counterfactual frequency in the manipulation region $[z_L, z_U]$. Using hat-notation to denote regression coefficients, our estimate for the number of manipulators in the missing and excess region, respectively, is given by:

$$N_{\text{mani}}^{\text{missing}} = \sum_{k \in \text{missing}} |\hat{\gamma}_k| \quad \text{and} \quad N_{\text{mani}}^{\text{excess}} = \sum_{k \in \text{missing}} \hat{\gamma}_k. \tag{14}$$

Note that $\hat{\gamma}_k < 0$ if k belongs to the missing region, while $\hat{\gamma}_k > 0$ if it belongs to the excess region. We repeat the above procedure for different values of z_U until $N_{\text{mani}}^{\text{missing}} \approx N_{\text{mani}}^{\text{excess}}$. In our application we estimate a manipulation region consisting of three bins (six weeks) for the missing and two bins (four weeks) for the excess region.

Because they will be useful in the next steps, let us define estimates for the number of non-manipulators, which is an observable quantity, and the number of individuals in the excess regions who are not manipulators, respectively, as:

$$N_{\text{non-mani}}^{\text{missing}} = \sum_{k \in \text{missing}} c_k \quad \text{and} \quad N_{\text{w/o mani}}^{\text{excess}} = \sum_{k \in \text{excess}} c_k - \hat{\gamma}_k. \tag{15}$$

Note that we deliberately reserve the term "non-manipulator" for individuals in the missing region who at least in principle could have engaged in manipulation but did not. Given the total headcounts, it is straightforward to compute the share of manipulators in the missing and excess region, respectively, as follows:

$$s^{\text{missing}} = \frac{N_{\text{mani}}^{\text{missing}}}{N_{\text{mani}}^{\text{missing}} + N_{\text{non-mani}}^{\text{missing}}} \quad \text{and} \quad s^{\text{excess}} = \frac{N_{\text{mani}}^{\text{excess}}}{N_{\text{mani}}^{\text{excess}} + N_{\text{w/o mani}}^{\text{excess}}}. \tag{16}$$

Analogously, we define the share of manipulators in age bin k by:

$$s_k^{\text{missing}} = \frac{|\hat{\gamma}_k|}{|\hat{\gamma}_k| + c_k} \text{ for } k \in \text{missing} \quad \text{and} \quad s_k^{\text{excess}} = \frac{\hat{\gamma}_k}{c_k} \text{ for } k \in \text{excess}. \tag{17}$$

Equipped with a measure of the size of manipulation, we now turn to studying affected outcomes.

III.B.2 Effects of manipulation

This section outlines our empirical strategy for studying outcome variables that are not directly manipulated but could potentially be affected by manipulation. Figure IIb illustrates the idea for one of our outcomes of interest: nonemployment survival rates. Manipulation provides workers with additional UI coverage from month eight to twelve. Thus, it is likely that nonemployment survival rates respond to the increase in coverage. Consider a hypothetical statistical relationship between nonemployment survival and age at layoff, as in Figure IIb. In order to estimate how manipulators' survival rate responds, we take the difference between two quantities: manipulators' actual survival probability and manipulators' counterfactual survival probability had they not been able to manipulate. As illustrated in Figure IIb, we obtain these quantities by separately studying the missing and excess region. First, we fit a flexible counterfactual on the right-hand side of the threshold and estimate the difference between the observed and predicted survival rates to assess manipulators' actual survival probability. Intuitively, survival rates in the excess region are higher than predicted by the un-manipulated region to the right only due to manipulation. The extent to which observed and predicted nonemployment survival rates differ, together with an estimate of how many manipulators are causing this difference, let us recover manipulators' actual nonemployment survival probability. We use analogous arguments to back out manipulators' counterfactual nonemployment survival probability on the left-hand side of the threshold.

In practice, we start by running the following regression on individual-level data:

$$
\begin{aligned}
y_i = \; & \alpha + \sum_{p=1}^{P} \beta_p^{\leq 50} \cdot a_i^p \cdot \mathbb{I}[a_i \leq 50] + \sum_{p=0}^{P} \beta_p^{>50} \cdot a_i^p \cdot \mathbb{I}[a_i > 50] + \\
& + \sum_{k=z_U}^{z_L} \delta_k \cdot \mathbb{I}[a_i = k] + \xi_i, \qquad (18)
\end{aligned}
$$

where y_i is the outcome of interest, e.g. weeks of UI benefit receipt or probability of still being nonemployed eight months after the layoff, $\beta_p^{\leq 50}$ and $\beta_p^{>50}$ are coefficients of two P-th degree polynomials in age, that are estimated based on information from the left-hand side and right-hand side, respectively. Due to the inclusion of $\mathbb{I}[a_i = k]$ indicator variables, the counterfactual polynomial is estimated as if we were excluding observations from the manipulation region $[z_L, z_U]$. The coefficients δ_k capture the difference in average outcomes between the observed data and the estimated counterfactual in the manipulation region.

Specification (18) allows for a treatment effect of longer PBD on outcomes, i.e. $\beta_0^{>50}$. We refer to $\beta_0^{>50}$ as the "donut" regression discontinuity (RD) coefficient. This coefficient captures the treatment effect of four additional months of PBD for the average individual in the population, as in Barreca et al. (2011) and Scrutinio (2018).[23] We use it to benchmark our results for the response of manipulators (more on this below). Graphically, $\beta_0^{>50}$ recovers the difference between the two grey dots in Figure IIb.

The central idea of our estimation strategy is the re-scaling of the estimated differences ($\hat{\delta}_k$) by the respective share of manipulators. Formally, let Y denote our outcome of interest and $\bar{Y}_l^j$ its average over individuals l in region j. For each bin k in the missing region, we may calculate the difference in average outcomes between manipulators and non-manipulators as:[24]

[23] Alternatively one could derive bounds on the average treatment effect following the method of Gerard et al. (2020). Because manipulation is clearly visible and locally confined in our setting we use a "donut" regression discontinuity design.

[24] Indeed, we can write the coefficient $\hat{\delta}_k$ as:

$$
\hat{\delta}_k = \bar{Y}^{\text{missing}}_{\text{non-mani},k} - \left(s_k \bar{Y}^{\text{missing}}_{\text{mani},k} - (1 - s_k) \bar{Y}^{\text{missing}}_{\text{non-mani},k} \right)
$$

which after some rearrangement leads to our equation 19.

$$\bar{Y}^{\text{missing}}_{\text{non-mani},k} - \bar{Y}^{\text{missing}}_{\text{mani},k} = \frac{\hat{\delta}_k}{s_k^{\text{missing}}}. \tag{19}$$

Note that the average outcome of non-manipulators in bin k is observable and given by

$$\bar{Y}^{\text{missing}}_{\text{non-mani},k} = \frac{\sum_{i=1}^{N} y_i \cdot \mathbb{I}[a_i = k]}{c_k}, \tag{20}$$

which allows us to recover manipulators' counterfactual outcome in bin k as

$$\bar{Y}^{\text{missing}}_{\text{mani},k} = \frac{\sum_{i=1}^{N} y_i \cdot \mathbb{I}[a_i = k]}{c_k} - \frac{\hat{\delta}_k}{s_k^{\text{missing}}} \tag{21}$$

and manipulators average counterfactual outcome over the entire missing region as

$$\bar{Y}^{\text{missing}}_{\text{mani}} = \frac{1}{N^{\text{missing}}_{\text{mani}}} \sum_k |\hat{\gamma}_k| \cdot \bar{Y}^{\text{missing}}_{\text{mani},k}. \tag{22}$$

The logic behind this re-scaling is straightforward: if we found that the absence of 10% of individuals in the missing region, namely the manipulators, resulted in a 100 unit drop starting from a predicted counterfactual of 1000 units, we could infer that the now missing individuals must have had an outcome of $\frac{1000-0.9\times(1000-100)}{0.1} = 1900$ units on average.

Following an analogous argument on the right-hand side of the age cutoff, we first re-scale the regression coefficient for bin k to obtain

$$\bar{Y}^{\text{excess}}_{\text{mani},k} - \bar{Y}^{\text{excess}}_{\text{w/o mani},k} = \frac{\hat{\delta}_k}{s_k^{\text{excess}}}. \tag{23}$$

Notice that the observable average outcome in bin k in the excess region has to satisfy

$$\bar{Y}^{\text{excess}}_{\text{observed},k} = \frac{\sum_{i=1}^{N} y_i \cdot \mathbb{I}[a_i = k]}{c_k} = \frac{\hat{\gamma}_k \cdot \bar{Y}^{\text{excess}}_{\text{mani},k} + (c_k - \hat{\gamma}_k) \cdot \bar{Y}^{\text{excess}}_{\text{w/o mani},k}}{c_k}. \tag{24}$$

Combining the two expressions above and rearranging terms gives us an estimate of manipulators' actual outcome in the form of

$$\bar{Y}^{\text{excess}}_{\text{mani},k} = \frac{\sum_{i=1}^{N} y_i \cdot \mathbb{I}[a_i = k]}{c_k} + (1 - s_k^{\text{excess}}) \cdot \frac{\hat{\delta}_k}{s_k^{\text{excess}}}, \tag{25}$$

for bin k in the excess region. We again calculate manipulators' average actual outcome over the entire excess region by

$$\bar{Y}_{\text{mani}}^{\text{excess}} = \frac{1}{N_{\text{mani}}^{\text{excess}}} \cdot \sum_k \hat{\gamma}_k \cdot \bar{Y}_{\text{mani},k}^{\text{excess}}, \tag{26}$$

which, together with equation (22) lets us define manipulators' response (or treatment effect) as

$$Y_{\text{mani}}^{TE} \equiv \bar{Y}_{\text{mani}}^{\text{excess}} - \bar{Y}_{\text{mani}}^{\text{missing}}. \tag{27}$$

Note that this strategy identifies the average response of a manipulator without recovering by how many weeks each individual manipulator delayed their layoff.

III.B.3 Recovering Responses of Non-manipulators

Having obtained an estimate of manipulators' response, we benchmark these results against the implied response of non-manipulators. As noted above, $\hat{\beta}_0^{>50}$ is an estimate of the effect of four additional months of PBD for an average individual who is moved over the threshold exogenously, i.e. without manipulation. Assuming that manipulators would have shown the same response to additional PBD coverage had they been moved over the threshold exogenously, instead of through manipulation, we can decompose the response for the average individual as follows:

$$s^{\text{missing}} \cdot Y_{\text{mani}}^{TE} + (1 - s^{\text{missing}}) \cdot Y_{\text{non-mani}}^{TE} = \hat{\beta}_0^{>50}. \tag{28}$$

A fraction of s^{missing} of the estimated jump in the polynomial $\hat{\beta}_0^{>50}$ is due to the response of manipulators, the remaining $(1 - s^{\text{missing}})$ has to be due to the response of non-manipulators. Rearranging thus gives us an estimate for non-manipulators' response:

$$Y_{\text{non-mani}}^{TE} = \frac{\hat{\beta}_0^{>50} - s^{\text{missing}} \cdot Y_{\text{mani}}^{TE}}{1 - s^{\text{missing}}}. \tag{29}$$

III.B.4 Selection into manipulation.

The procedure illustrated in Figure IIb also lets us study selection into manipulation by comparing manipulators' counterfactual outcomes to non-manipulators realized outcomes. Figure IIb highlights this comparison and would suggest that even absent manipulation, manipulators would have had a higher nonemployment survival rate than non-manipulators due to the drop in the outcome variable to the left of the cutoff. This is indeed what we show in Section III.C.4. We now turn to our empirical findings and illustrate how they relate to the theoretical results from Section II.

III.C Results

In this section we examine the main findings. We start by presenting graphical evidence of manipulation in the form of strategic delays in the timing of layoffs around the fiftieth birthday threshold. After quantifying the magnitude of manipulation, we estimate the additional increase in UI receipt and nonemployment duration that arises from the change in manipulators' job search behavior. We highlight that most of the increase is mechanically the result of higher coverage due to relatively high long-term nonemployment risk on which manipulators are adversely selected. The implied responsiveness to UI is modest and, in particular, not higher than for non-manipulators. Last, we probe the robustness of our findings and examine observable characteristics on which manipulators are selected.

III.C.1 Evidence of manipulation

To provide graphical evidence of manipulation, Figure III plots the relative frequency of layoffs against workers' age at layoff. Figure IIIb covers the entire age range from 26 to 64 years of age, while Figure IIIa zooms into a narrower, four year window around the age-fifty threshold.[25] Both figures show a clear drop in the frequency of layoffs just before, and a pronounced spike after, the age-fifty threshold.

Following our estimation strategy outlined in Section III.B.1, we find the manipulation region to consist of all age bins from six weeks before (missing region), up to four weeks after the threshold (excess region). Table II reports our estimates for the respective headcounts for the four groups of interest: manipulators in the missing region, non-manipulators in the missing region, manipulators in the excess region and all individuals in the excess region who are not manipulators, as well as share estimates for the missing and excess region. We estimate that a total of 571 layoffs are strategically delayed corresponding to 15.8% of layoffs in the missing region. The counterfactual relationship appears almost perfectly linear and is robust to the choice of the order of the polynomial. The estimated number of manipulators in the excess region, 609, deviates slightly from that in the missing region due to measurement error and corresponds to approximately 20.3% of layoffs in the excess region.

Relating these findings to the theoretical analysis in Section II, we provide clear evidence of the presence of manipulation in our context. It is straightforward to translate the estimated number of manipulators into a share estimate once one decides on the definition of the relevant group. If one, for instances, took six weeks prior to the age threshold as the cutoff for the group definition of the young, the share M

[25]By plotting the layoff frequency over the entire age range in Figure IIIb, we already rule out that manipulation is caused by other mechanisms like (round-) birthday effects. All our estimates for the counterfactual density and counterfactual outcomes are based on the narrower (46-54) window. Section III.C.6 presents additional robustness checks.

would correspond to the above estimate of 15.8% (see our discussion in Section II.D on this point). Unfortunately, we lack sufficient policy variation to credibly estimate the share *elasticities* in Proposition 3. Due to the nature of our manipulation mechanism, namely worker-firm bargaining, one can only speculate about plausible values. It also appears likely that manipulation elasticities are not constant in our setting, e.g. due to non-financial incentives such as warm-glow or reputation concerns playing a role. Importantly, the theory does not require pinning down the exact mechanism as long as one has credible estimates for the share elasticities (or is willing to make additional assumptions).

III.C.2 Effects of manipulation: UI benefit receipt and duration

Manipulation provides workers with four additional months of UI coverage. To study the effect of extra coverage on manipulators' benefit receipt and nonemployment duration we begin by plotting these outcomes against workers' age at layoff in Figure IV. For each outcome we see visible changes around the age threshold indicating that both respond to manipulation. As outlined in Section III.C.2 we combine these changes with the share estimate from the previous section to retrieve manipulators' as well as non-manipulators' responses. We report all estimates with associated 95% confidence intervals in Tables III and IV.[26]

Our results indicate that manipulators would have collected 5814.2 Euro, and spent 27.8 weeks on UI benefits, had they not manipulated (columns 1). Through manipulation these numbers increase to 8053.6 Euro and 41.8 weeks (columns 3), resulting in an additional cost of 2239 Euro per manipulator (columns 5). In order to benchmark these estimates, we compute the same numbers for non-manipulators following the strategy outlined in Section III.B.3. We find that non-manipulators generate a total cost of 1636.9 Euro (colums 6) when receiving additional coverage.

As highlighted in Section II, these numbers alone are not directly welfare relevant, because they reflect both the mechanical transfer as well as possible distortions in job search. The next section provides a decomposition into these two components.[27]

III.C.3 Distinguishing behavioral responses from mechanical effects

The key insight to decomposing behavioral and mechanical cost increases, is to repeat the preceding estimation procedure at different months after layoff to trace out *when* manipulators and non-manipulators respond to additional coverage. We start by plotting nonemployment survival rates against age at layoff at various months after layoff

[26]All confidence intervals in the paper are obtained by simple non-parametric bootstrapping: we operationalize this by resampling layoff events and re-estimating the entire procedure, including the share of manipulators, 5000 times.

[27]It is worth noticing that the cost estimates are relevant for calculating cost-weighted elasticities given in equation (8) because they relate to the fiscal externality defined in equation (6).

in Figure V. Qualitatively, we observe bigger jumps around the thresholds precisely during the months with extra coverage. Similarly to before, we combine these changes with the estimated share of manipulators causing them to trace out monthly survival curves for both manipulators and non-manipulators.

Figure VIa presents our estimated nonemployment survival curves of manipulators under the eight and twelve months PBD schemes. Figure VIb reports the difference between the two curves at any point, with associated bootstrapped 95% confidence intervals. The difference between the two curves reveals the effect of longer PBD along manipulators' survival curve which appears concentrated precisely in the months of extra UI coverage. We replicate the same analysis for non-manipulators and report its findings in Figure VII. The qualitative picture is similar, although confidence bands are much narrower in large part due to the fact that non-manipulators' survival curve under the eight month PBD scheme is observable rather than estimated.

We translate the survival rate responses into BC/MC ratio estimates for manipulators and non-manipulators following equation (2). To do so, we rely on numerical integration and weight responses by statutory benefit rates.[28] We report our BC/MC ratio estimates in Table V. Because there is some disagreement in the literature as to what the appropriate tax rate is this context, columns 1 and 2 provide BC/MC ratios for a no tax $\tau = 0$ and a commonly used UI tax of $\tau = 3\%$, see e.g. Schmieder and von Wachter (2016) and Lawson (2017). As discussed in Section II an estimate of 0.24 for manipulators in column 1 of Table V implies that the government pays an additional 24 cents for each Euro of UI transfer. The estimated BC/MC ratios for manipulators and non-manipulators are strikingly similar suggesting that there is no selection on moral hazard which links directly to equation (10) in Proposition 3.[29] From a positive perspective this finding also mitigates concerns that anticipated moral hazard is a prime motive to engage in manipulation.

III.C.4 Selection on long-term nonemployment risk

The remainder of our empirical analysis provides additional evidence to shed light on the drivers behind manipulation in our context. The previous section ruled out anticipated moral hazard as a key motivation to engage in manipulation. In this section we show that alleviated exhaustion risk is a strong predictor of manipulation.

To do so, Figure VIII combines manipulators' and non-manipulators' eight months

[28]We perform integration using the midpoint rule and impose a non-negativity constraint on the behavioral cost at any point in time. Note that in the first few months the point estimates of the survival rate response is negative for manipulators which would imply that longer PBD increases job finding rates. However, this finding is likely due to noise. As these negative contributions to the overall integral leads us to underestimate BC/MC ratios for manipulators, our estimates are conservative. Results are qualitatively unaltered without imposing the non-negativity constraint.

[29]The reported BC/MC ratios are in the lower range of estimates in the previous literature, see Schmieder and von Wachter (2016) for an overview.

PBD survival curves from Section III.C.3. A clear difference emerges and manipulators exhibit an almost 20 p.p. higher (counterfactual) exhaustion risk under the less generous eight months PBD scheme. This finding provides compelling evidence that anticipated exhaustion risk is a strong motive for manipulation. Note that these estimates also directly relate to the selection of risk scale factor in equation 10 in Proposition 3. The large exhaustion risk is also (partly) responsible for making most of the increase in benefit receipt mechanical, thus lowering the BC/MC ratio, in Section III.C.3.

III.C.5 Characterizing manipulators

This last section of our analysis, provides some suggestive evidence on the underlying manipulation mechanism by documenting observable characteristics that are correlated with manipulation. In Figure IX we start by visually inspecting the age distribution of layoffs for different types of contracts (permanent and temporary) and sectors (private and public). Manipulation is entirely confined to private sector permanent contract workers motivating the choice of our main sample.

Turning to observable worker and firm characteristics for our main sample, Table VI reports a selection on observables analysis.[30] Column 1 and 2 of Table VI report estimated mean characteristics for manipulators and non-manipulators, respectively. Column 3 calculates their difference together with bootstrapped 95% confidence intervals. We find that manipulators are 18 p.p. more likely than non-manipulators to be female, 17 p.p. more likely to be employed in white collar jobs and 7 p.p. less likely to have full-time contracts. Manipulators' wages are 6% lower, although estimates are relatively imprecise. Firm size plays an important role for manipulation: manipulators come from firms that are about 40% smaller. Overall, these findings suggest that adjustment costs, bargaining power and proximity to managers play a role in workers' ability to engage in manipulation. A full investigation into the underlying worker-firm bargaining mechanism is beyond the scope of this paper but we deem it an interesting avenue for future work.

Although more tentative, we view the selection patterns document in this section as evidence consistent with our main conclusion that manipulators are not adversely selected. If anything the findings suggest that manipulators might have higher marginal utilities, e.g. due to part-time work arrangements and lower wages.

III.C.6 Robustness

This section probes the robustness of two identifying assumptions underlying our empirical analysis and its link to the theoretical results from Section II. First, we provide evidence that manipulation is indeed the result of additional UI coverage around the age at layoff threshold. Second, the empirical analysis assumes that the discontinuity

[30]The analysis closely follows Section 6.2 of Diamond and Persson (2016).

in PBD around the age threshold affects layoff decisions in exactly one way, namely, through a delay in an otherwise earlier occurring layoff.

By plotting layoffs across the entire age distribution Figure IIIb already ruled out several alternative explanations such as e.g. round birthday effects. To provide further supporting evidence Figure X plots layoff densities for two Italian UI schemes which replaced the OUB scheme after January 2013 and did not feature any discontinuity in generosity at the age fifty threshold.[31] Reassuringly, we find no evidence of manipulation under any of the these alternative schemes.

The second concern is related to the possible presence of *extensive margin* job separation effects of UI and merits special attention in the light of recent evidence by Albanese et al. (2020) and Jäger et al. (2019). The former documents layoff responses at the eligibility threshold (52 weeks of contributions) in the same Italian OUB scheme we study. Although theoretically possible, we find no empirical evidence of any extensive margin job separation responses in our context through a series of robustness tests presented in detail in Appendix C. Intuitively, the layoff density shown in Figure III, shows no indication of any additional layoffs to the right of the cutoff that are not explained by missing layoffs in the missing region. We discuss this point as well as a series of other robustness tests exhaustively in Appendix C and find no evidence for a violation of our identification assumption.

IV Concluding Remarks

This work lays out a simple, yet robust theoretical framework to guide the design of differentiated social insurance under manipulation. We identify a set of sufficient statistics and illustrate how key moments in the data can be estimated in practice. Our empirical strategy builds on and extends recently proposed bunching techniques which do not require rich policy variation for estimation.

We are optimistic that our empirical methodology might be fruitfully applied in other contexts and, although a full welfare assessment is beyond the scope of this paper, we deem it an interesting area for future research. As pointed out by Spinnewijn (2020) there remains important work to be done in understanding, analysing and justifying frequently used tags in social insurance. We hope that our framework and methodology provide an important first step.

Although the theoretical results hold more generally, our empirical analysis focuses on the case where group membership is defined by a threshold rule in an underlying continuous variable, age-at-layoff. While there are many such cases in practice, another prevalent case is that of discrete variable group membership, e.g. based on gender or

[31] For institutional details regarding both UI schemes see Appendix B.

the number of children. Developing empirical methodologies for such settings is thus of first-order policy interest.

References

Akerlof, G. A. (1978). The economics of "tagging" as applied to the optimal income tax, welfare programs, and manpower planning. *American Economic Review*, 68(1):8–19.

Albanese, A., Picchio, M., and Ghirelli, C. (2020). Timed to say goodbye: Does unemployment benefit eligibility affect worker layoffs? *Labour Economics*, 65:101846.

Anastasia, B., Mancini, M., and Trivellato, U. (2009). Il sostegno al reddito dei disoccupati: note sullo stato dell'arte: tra riformismo strisciante, inerzie dell'impianto categoriale e incerti orizzonti di flexicurity. I Tartufi 32, Veneto Lavoro.

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

Barnichon, R. and Zylberberg, Y. (2021). A menu of insurance contracts for the unemployed. *Review of Economic Studies*, forthcoming.

Barreca, A. I., Guldi, M., Lindo, J. M., and Waddell, G. R. (2011). Saving babies? Revisiting the effect of very low birth weight classification. *Quarterly Journal of Economics*, 126(4):2117–2123.

Best, M. and Kleven, H. J. (2013). Optimal income taxation with career effects of work effort. *mimeo*.

Cabral, M. (2016). Claim timing and ex post adverse selection. *The Review of Economic Studies*, 84(1):1–44.

Card, D., Chetty, R., and Weber, A. (2007). Cash-on-hand and competing models of intertemporal behavior: New evidence from the labor market. *Quarterly Journal of Economics*, 122(4):1511–1560.

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

Chetty, R., Friedman, J. N., Olsen, T., and Pistaferri, L. (2011). Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from danish tax records. *Quarterly Journal of Economics*, 126(2):749–804.

Dee, T. S., Dobbie, W., Jacob, B. A., and Rockoff, J. (2019). The causes and consequences of test score manipulation: Evidence from the new york regents examinations. *American Economic Journal: Applied Economics*, 11(3):382–423.

Diamond, P. and Sheshinski, E. (1995). Economic aspects of optimal disability benefits. *Journal of Public economics*, 57(1):1–23.

Diamond, R. and Persson, P. (2016). The long-term consequences of teacher discretion in grading of high-stakes tests. *NBER working paper no. 22207*.

Doornik, B. F. V., Schoenherr, D., and Skrastins, J. (2020). Strategic formal layoffs: Unemployment insurance and informal labor markets. *Central Bank of Brazil working paper no. 523*.

Einav, L., Finkelstein, A., Ryan, S. P., Schrimpf, P., and Cullen, M. R. (2013). Selection on moral hazard in health insurance. *American Economic Review*, 103(1):178–219.

Feldstein, M. (1976). Temporary layoffs in the theory of unemployment. *Journal of Political Economy*, 84(5):937–957.

Feldstein, M. (1978). The effect of unemployment insurance on temporary layoff unemployment. *American Economic Review*, 68(5):834–846.

Gerard, F. and Gonzaga, G. (2021). Informal labor and the efficiency cost of social programs: Evidence from unemployment insurance in brazil. *American Economic Journal: Economic Policy*, 13(3):167–206.

Gerard, F., Rokkanen, M., and Rothe, C. (2020). Bounds on treatment effects in regression discontinuity designs with a manipulated running variable. *Quantitative Economics*, 11(3):839–870.

Hackmann, M. B., Kolstad, J. T., and Kowalski, A. E. (2015). Adverse selection and an individual mandate: When theory meets practice. *American Economic Review*, 105(3):1030–66.

Heckman, J. J., Urzua, S., and Vytlacil, E. (2006). Understanding Instrumental Variables in Models with Essential Heterogeneity. *The Review of Economics and Statistics*, 88(3):389–432.

Hendren, N. (2017). Knowledge of future job loss and implications for unemployment insurance. *American Economic Review*, 107(7):1778–1823.

Hendren, N., Landais, C., and Spinnewijn, J. (2020). Choice in insurance markets: A pigouvian approach to social insurance design. *NBER working paper no. 27842*.

Jäger, S., Schoefer, B., and Zweimüller, J. (2019). Marginal jobs and job surplus: A test of the efficiency of separations. *NBER working paper no. 25492*.

Johnston, A. C. and Mas, A. (2018). Potential unemployment insurance duration and labor supply: The individual and market-level response to a benefit cut. *Journal of Political Economy*, 126(6):2480–2522.

Khoury, L. (2019). Unemployment benefits and the timing of redundancies. *PSE working paper no. 2019-14*.

Kleven, H. J. (2016). Bunching. *Annual Review of Economics*, 8(1):435–464.

Kleven, H. J. and Kopczuk, W. (2011). Transfer program complexity and the take-up of social benefits. *American Economic Journal: Economic Policy*, 3(1):54–90.

Kleven, H. J. and Waseem, M. (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from pakistan. *Quarterly Journal of Economics*, 128(2):669–723.

Lalive, R. (2007). Unemployment benefits, unemployment duration, and post-unemployment jobs: A regression discontinuity approach. *American Economic Review*, 97(2):108–112.

Landais, C. (2015). Assessing the welfare effects of unemployment benefits using the regression kink design. *American Economic Journal: Economic Policy*, 7(4):243–278.

Landais, C., Nekoei, A., Nilsson, P., Seim, D., and Spinnewijn, J. (2021). Risk-based selection in unemployment insurance: Evidence and implications. *American Economic Review*, 111(4):1315–55.

Lawson, N. (2017). Fiscal externalities and optimal unemployment insurance. *American Economic Journal: Economic Policy*, 9(4):281–312.

Marone, V. R. and Sabety, A. (2021). When should there be vertical choice in health insurance markets? *American Economic Review*, forthcoming.

Michelacci, C. and Ruffo, H. (2015). Optimal life cycle unemployment insurance. *American Economic Review*, 105(2):816–859.

Nekoei, A. and Weber, A. (2017). Does extending unemployment benefits improve job quality? *American Economic Review*, 107(2):527–561.

Parsons, D. O. (1996). Imperfect 'tagging'in social insurance programs. *Journal of Public Economics*, 62(1-2):183–207.

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

Schmieder, J. F. and von Wachter, T. (2016). The effects of unemployment insurance benefits: New evidence and interpretation. *Annual Review of Economics*, 8:547–581.

Schmieder, J. F. and von Wachter, T. (2017). A context-robust measure of the disincentive cost of unemployment insurance. *American Economic Review*, 107(5):343–348.

Schmieder, J. F., von Wachter, T., and Bender, S. (2012). The effects of extended unemployment insurance over the business cycle: Evidence from regression discontinuity estimates over 20 years. *Quarterly Journal of Economics*, 127(2):701–752.

Scrutinio, V. (2018). The medium-term effects of unemployment benefits. *INPS working paper no. 18*.

Spinnewijn, J. (2020). The trade-off between insurance and incentives in differentiated unemployment policies. *Fiscal Studies*, 41(1):101–127.

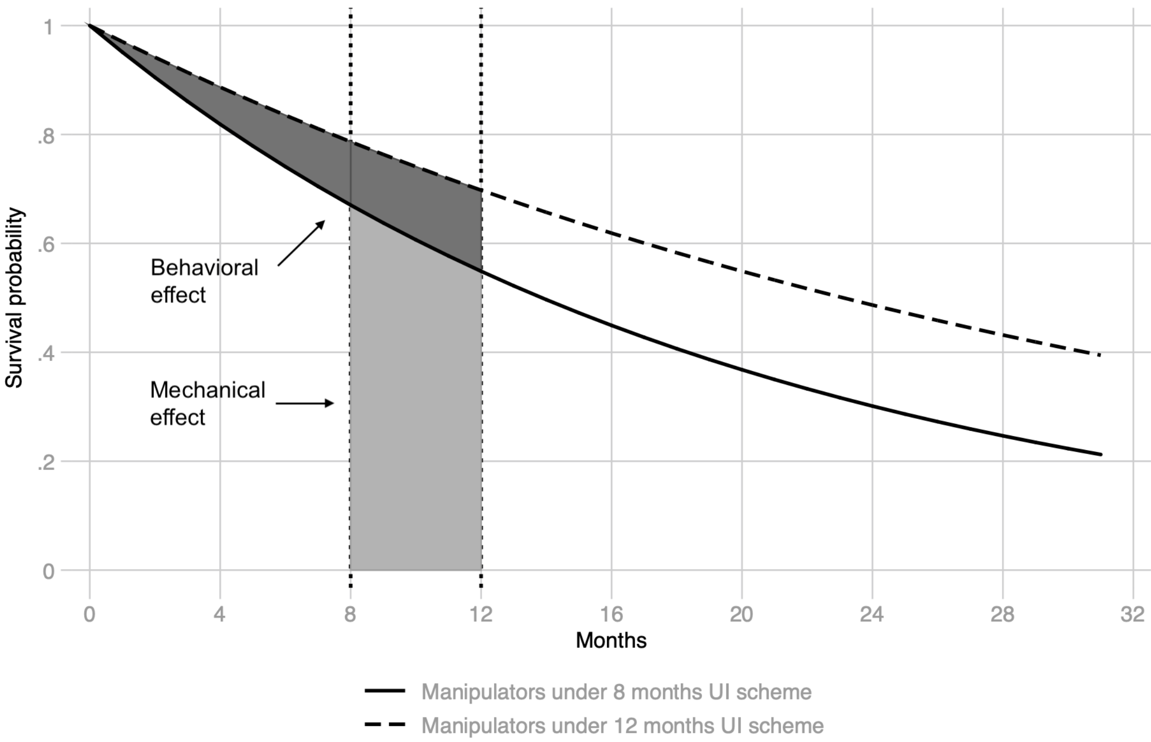
Stern, N. (1982). Optimum taxation with errors in administration. *Journal of Public Economics*, 17(2):181–211.

Topel, R. H. (1983). On layoffs and unemployment insurance. *American Economic Review*, 73(4):541–559.

Weinzierl, M. (2011). The surprising power of age-dependent taxes. *The Review of Economic Studies*, 78(4):1490–1518.

Figures

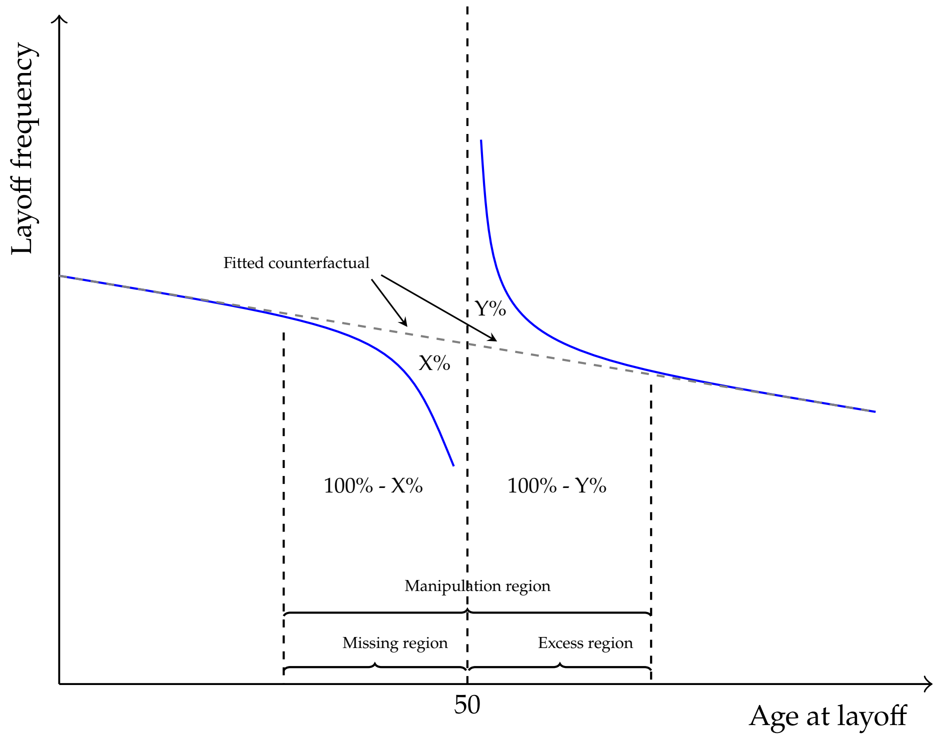
Figure I. The moral hazard cost of extended UI coverage



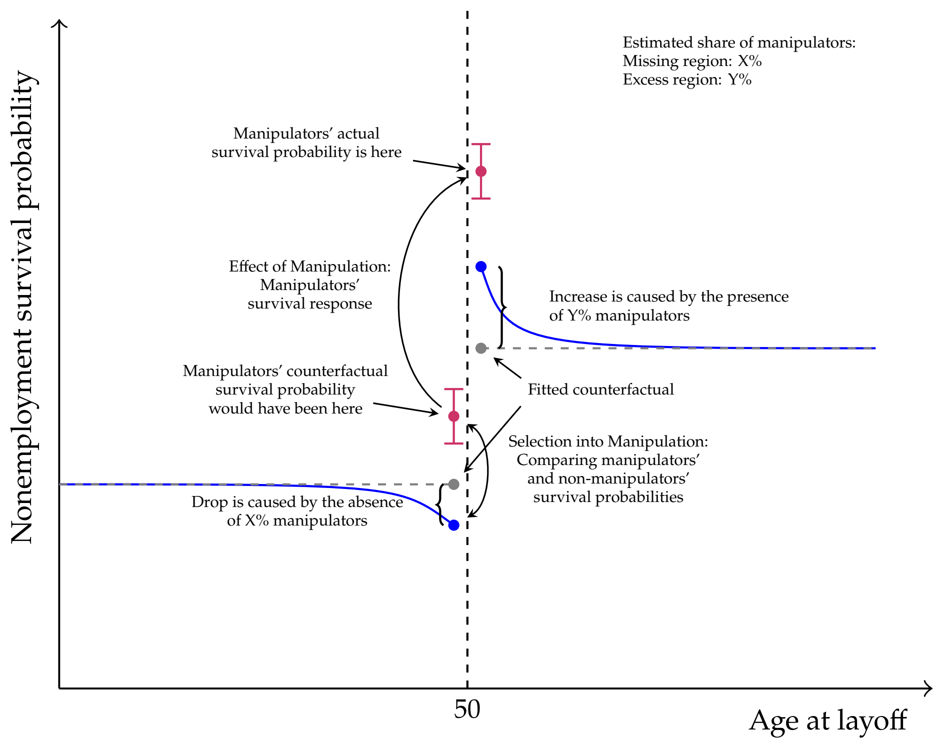
Note: The figure displays two hypothetical nonemployment survival curves for manipulators, namely, under eight months of PBD (solid line) and twelve months of PBD (dashed line). The dashed line is above the solid line assuming that higher PBD lowers the exit hazard rate from nonemployment. The curves are simulated as negative exponentials with a constant hazard rate of 5% and 3%, respectively. The total increase in UI benefit receipt due to higher coverage (shaded areas) consists of two components: (1) a mechanical part (light grey area) which captures additional UI benefit payments that would occur even absent any behavioral change; (2) a behavioral component (dark grey area) which is due to a shift in the survival curve. The BC/MC ratio defined in equation 2 is given by the ratio of (2) and (1).

Figure II. Illustration of identification strategy

(a) Quantifying manipulation



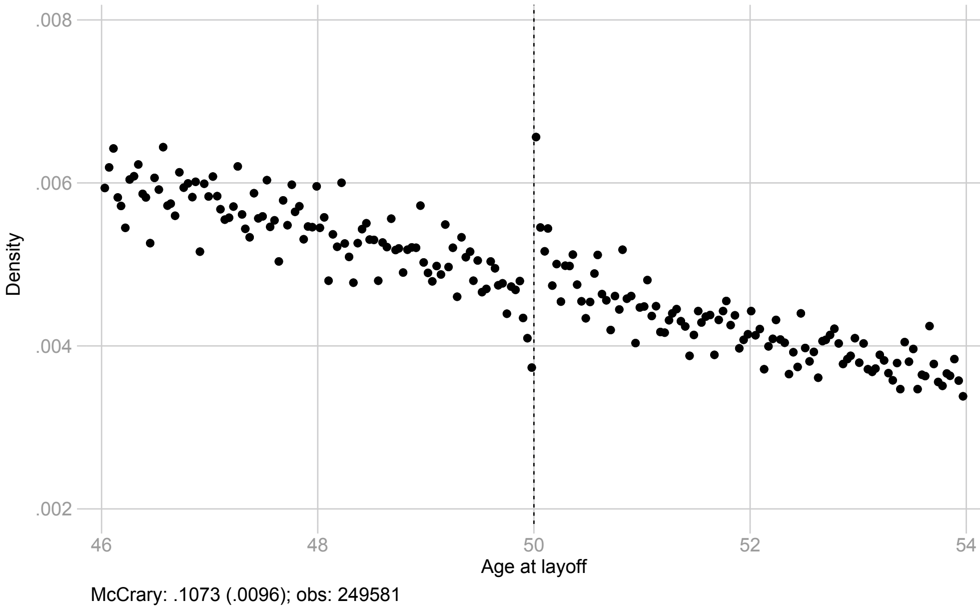
(b) Effect of and selection into manipulation



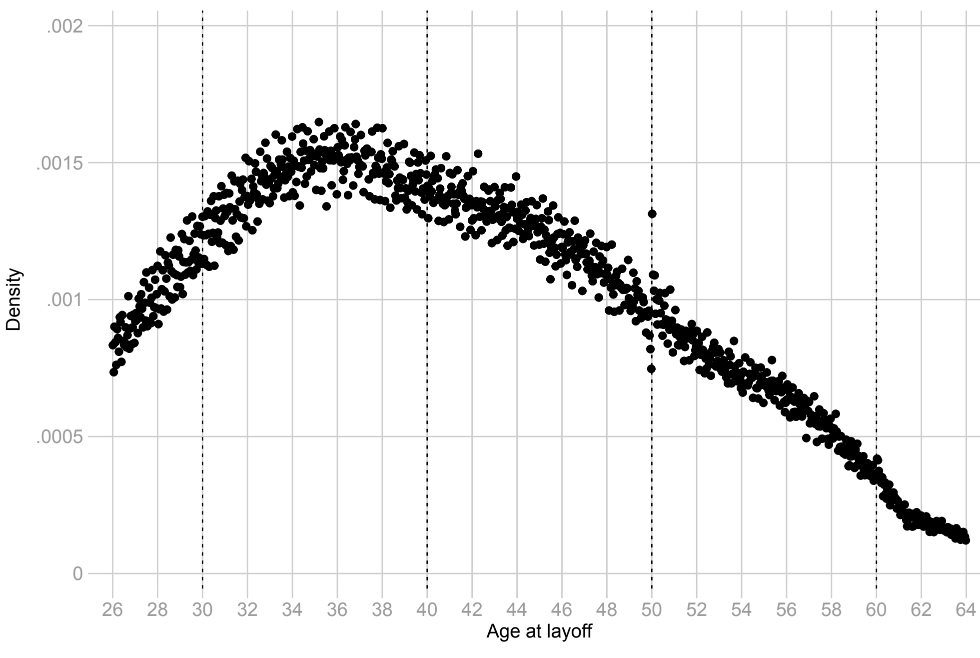
Note: The figure visualizes our identification strategy. Panel (a) illustrates how we estimate the number and respective share of manipulators in both the missing and excess region. Panel (b) constructs manipulators' survival response and illustrates the relevant comparison when studying selection into manipulation. Section III.B lays out how we estimate the fitted counterfactuals in practice.

Figure III. Layoff frequency for permanent contract private sector workers

(a) Age-at-layoff between 46 and 54 years



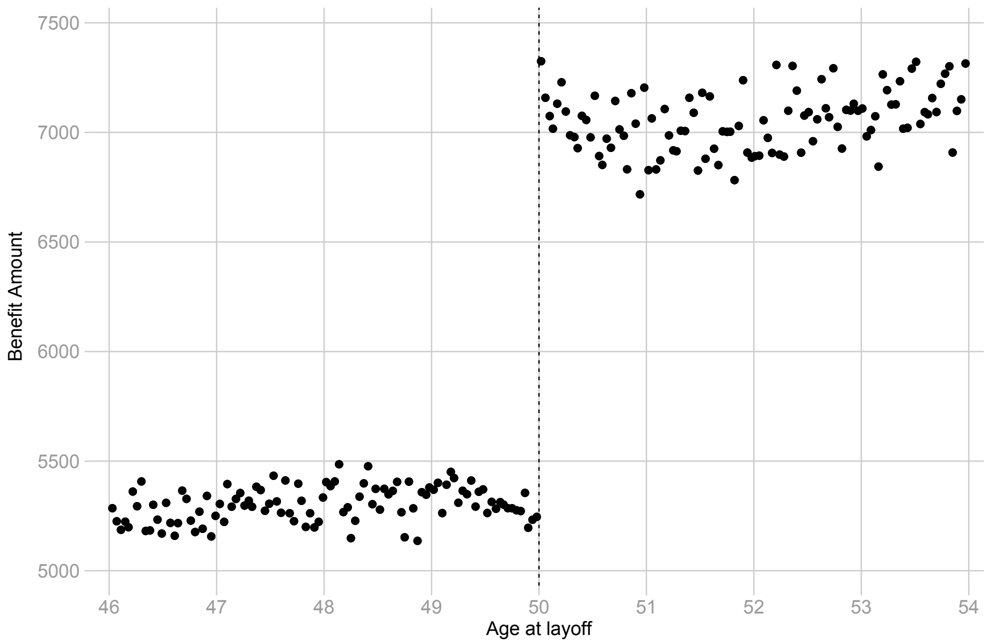
(b) Age-at-layoff between 26 and 64 years



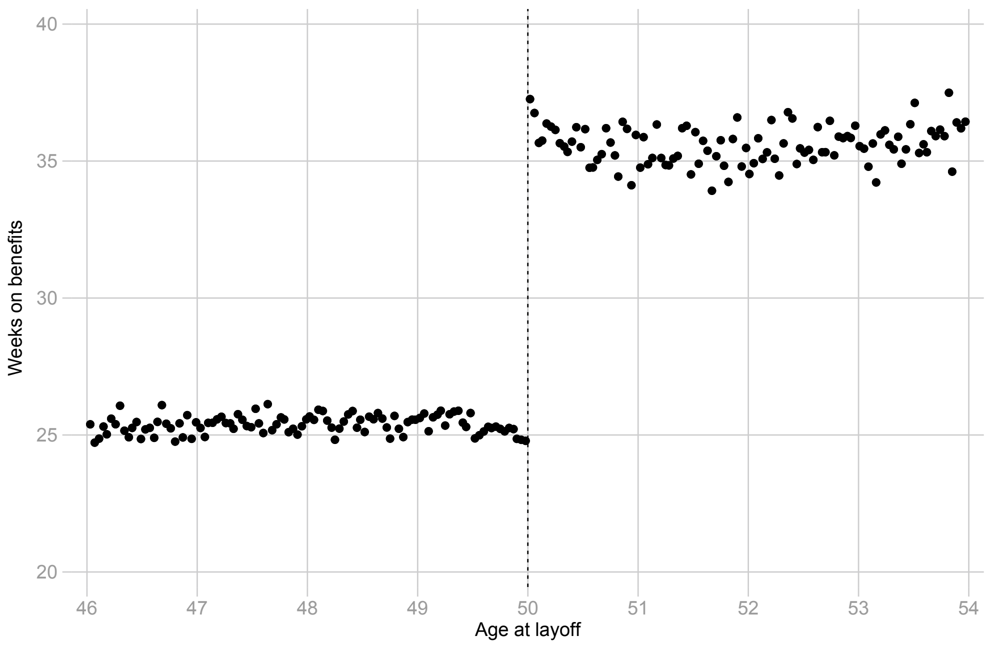
Note: The figure shows the density of layoffs in the private sector, for individuals working on a permanent contract and claiming regular UI (OUB). The data cover the period from Feb 2009 to Dec 2012. Panel (a) plots the density for the age range from 46 to 54 years, while Panel (b) does so for the entire age range from 26 to 64 years of age. In both panels each dot represents a two-week bin. The underlying data in Panel (a) consists of 249,581 layoffs.

Figure IV. Benefit receipt and duration

(a) average UI receipt (in Euro)



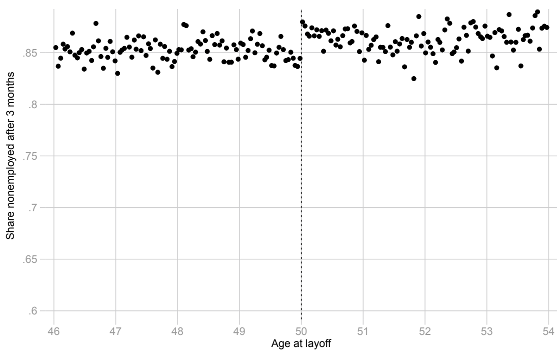
(b) average UI benefit duration (in weeks)



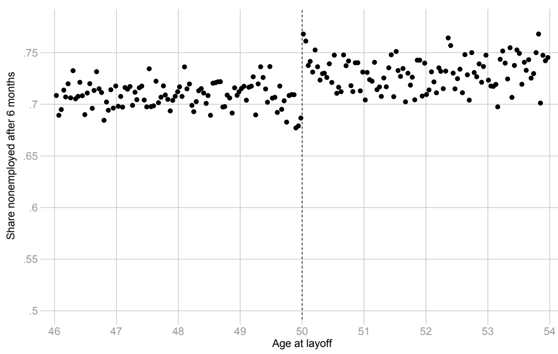
Note: The figure displays the average UI receipt in Euro (panel (a)) and average UI benefit duration in weeks (panel (b)) by age-at-layoff. In both panels each dot represents a two week bin. The sample includes all individuals working on a permanent contract and claiming regular UI (OUB). The data cover the period from Feb 2009 to Dec 2012. The underlying data consists of 249,581 layoffs.

Figure V. Nonemployment survival probabilities

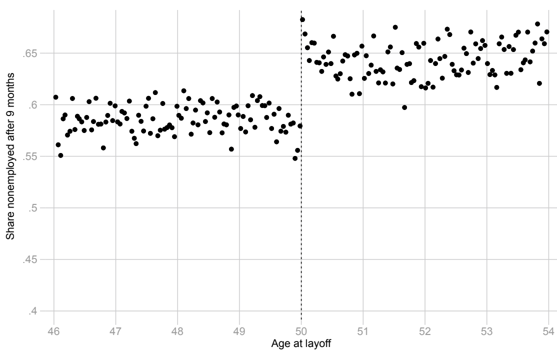
(a) Probability of still not being in employment at 3 months



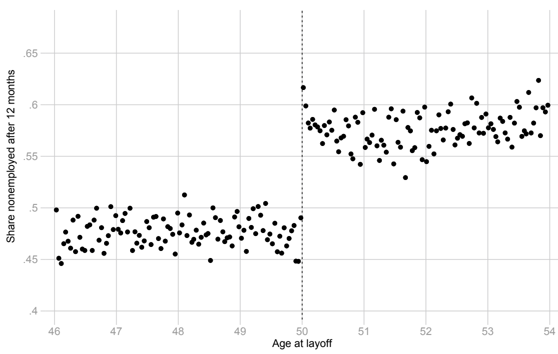
(b) Probability of still not being in employment at 6 months



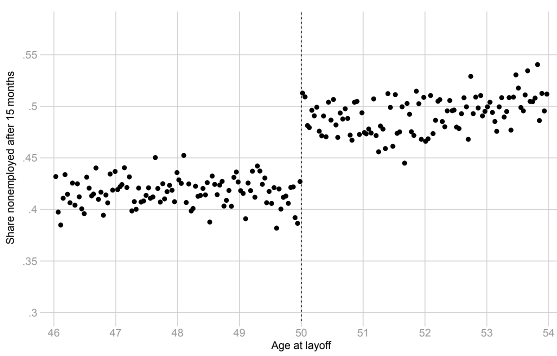
(c) Probability of still not being in employment at 9 months



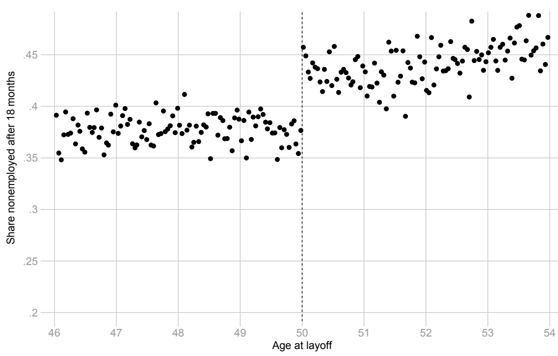
(d) Probability of still not being in employment at 12 months



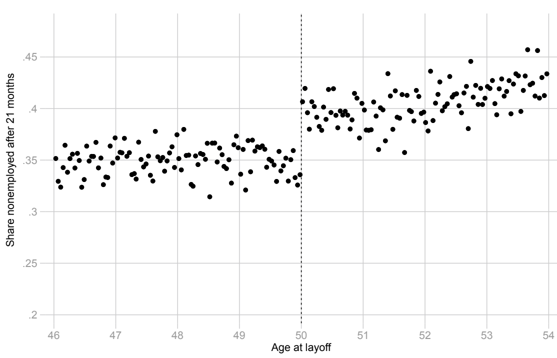
(e) Probability of still not being in employment at 15 months



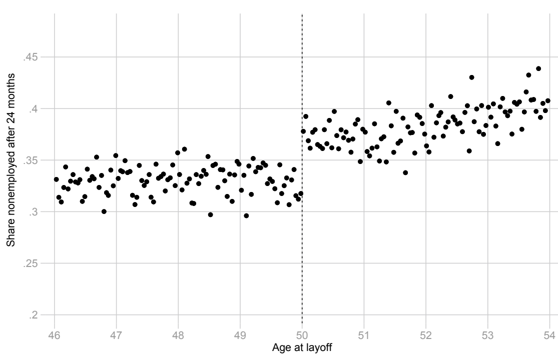
(f) Probability of still not being in employment at 18 months



(g) Probability of still not being in employment at 21 months



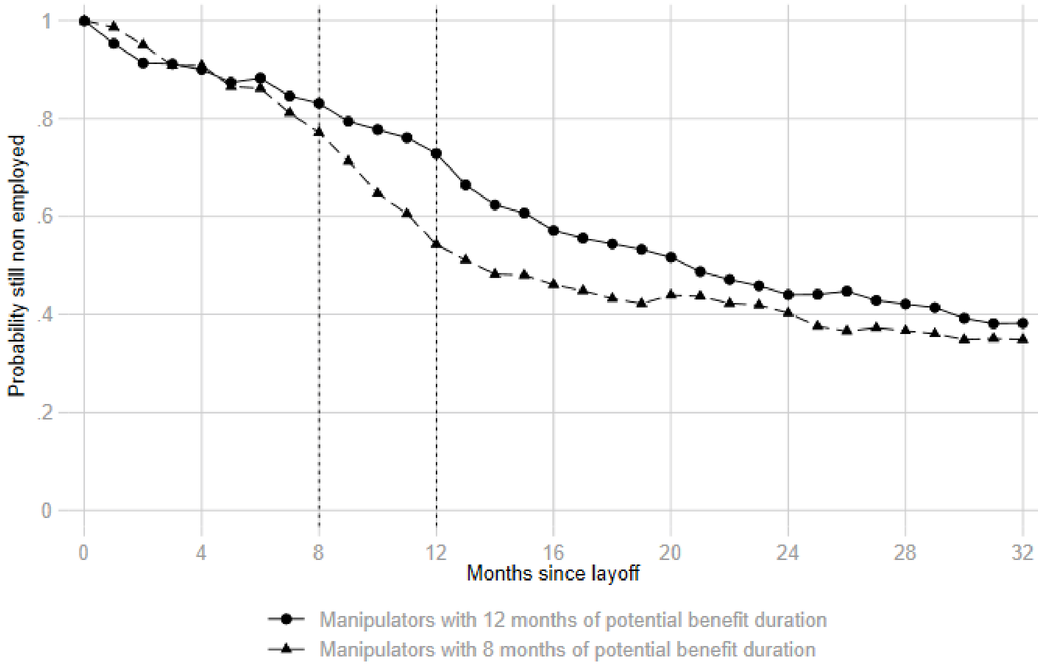
(h) Probability of still not being in employment at 24 months



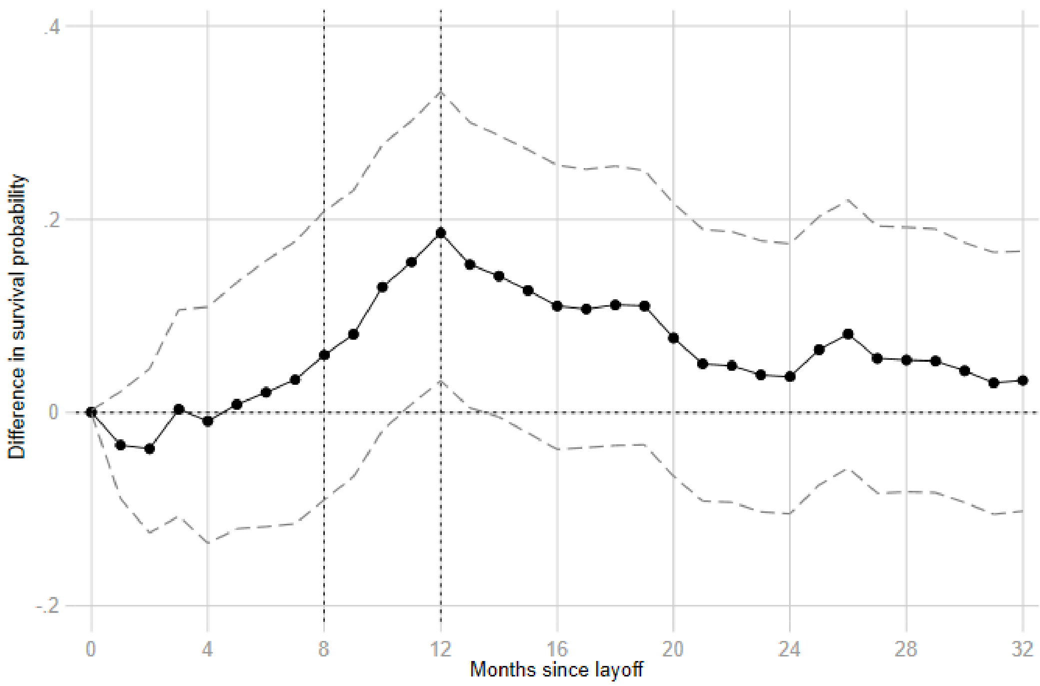
Note: The figures show the share of laid off workers, who are still not in employment after 3, 6, ..., 24 months. In all panels each dot represents a two week bin. The sample includes all individuals working on a permanent contract and claiming regular UI (OUB). The data cover the period from Feb 2009 to Dec 2012. The underlying data consists of 249,581 layoffs.

Figure VI. Manipulators with 8 and 12 months of potential benefit duration

(a) Nonemployment survival rates



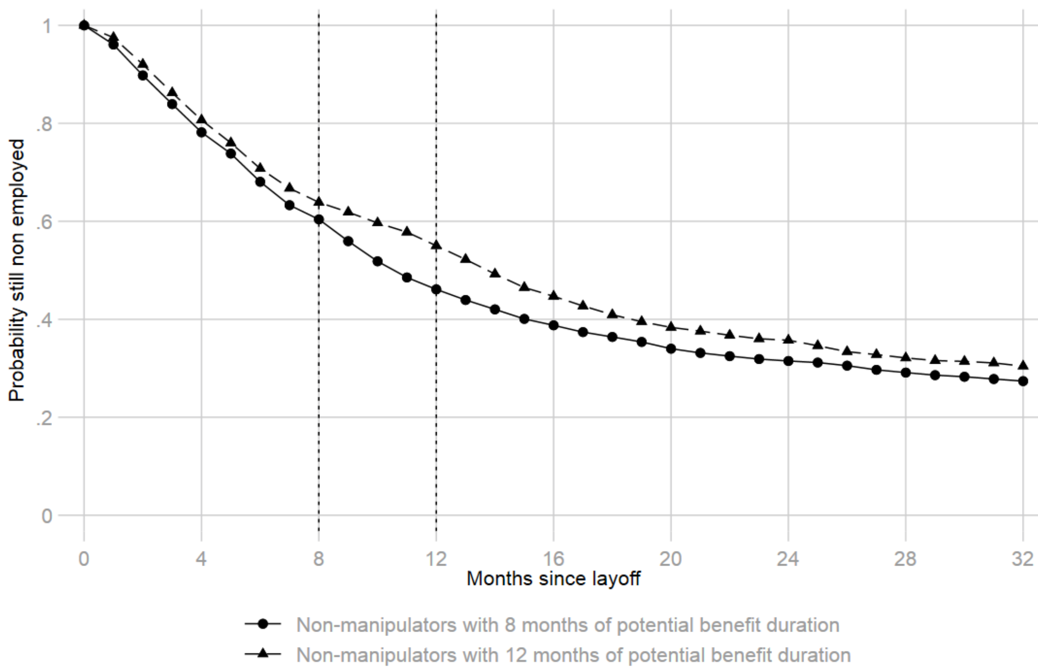
(b) Difference in survival rates



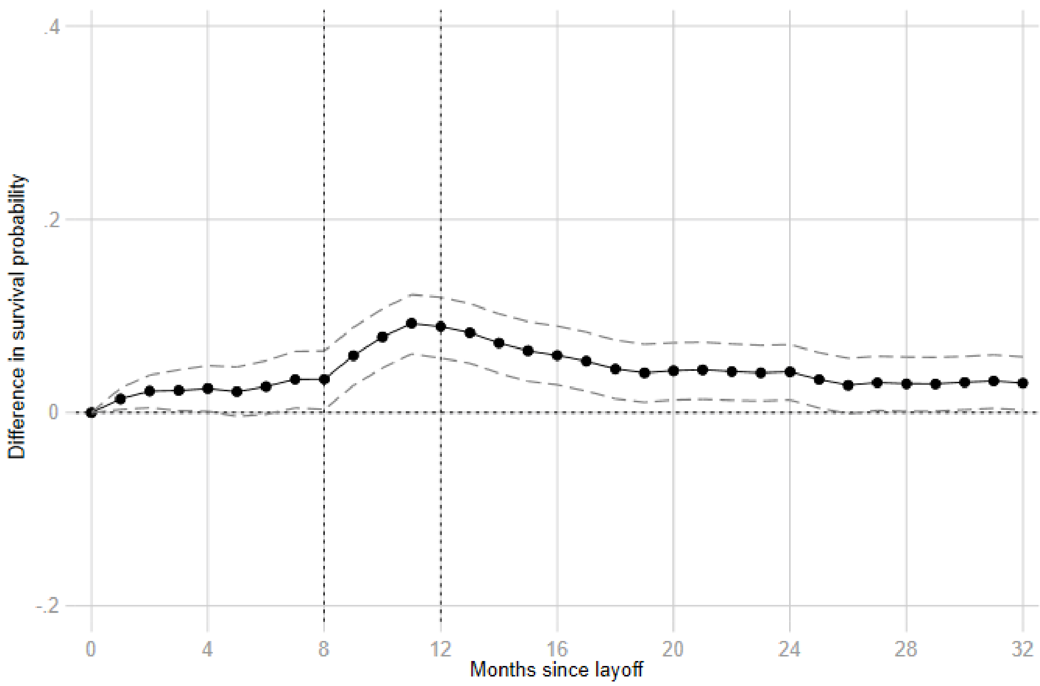
Note: Panel (a) plots point estimates of manipulators' actual and counterfactual nonemployment survival for the first 32 months after layoff. Our estimation strategy is outlined in Section III.B. Panel (b) shows the difference between the two survival curves and contains bootstrapped 95% confidence intervals testing against zero difference.

Figure VII. Manipulators with 8 and 12 months of potential benefit duration

(a) Nonemployment survival rates



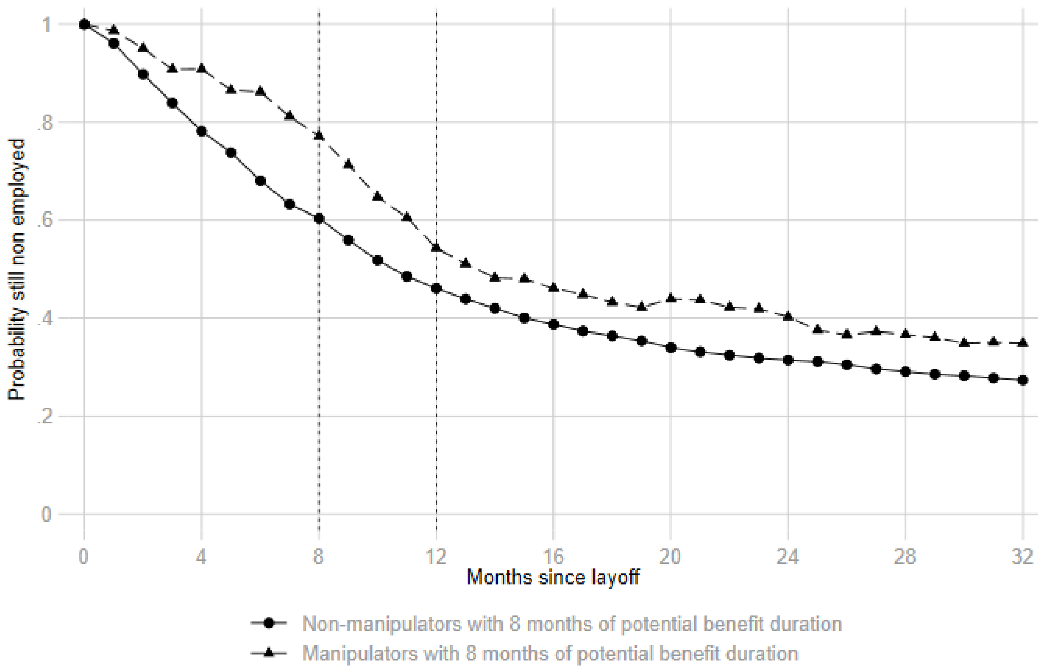
(b) Difference in survival rates



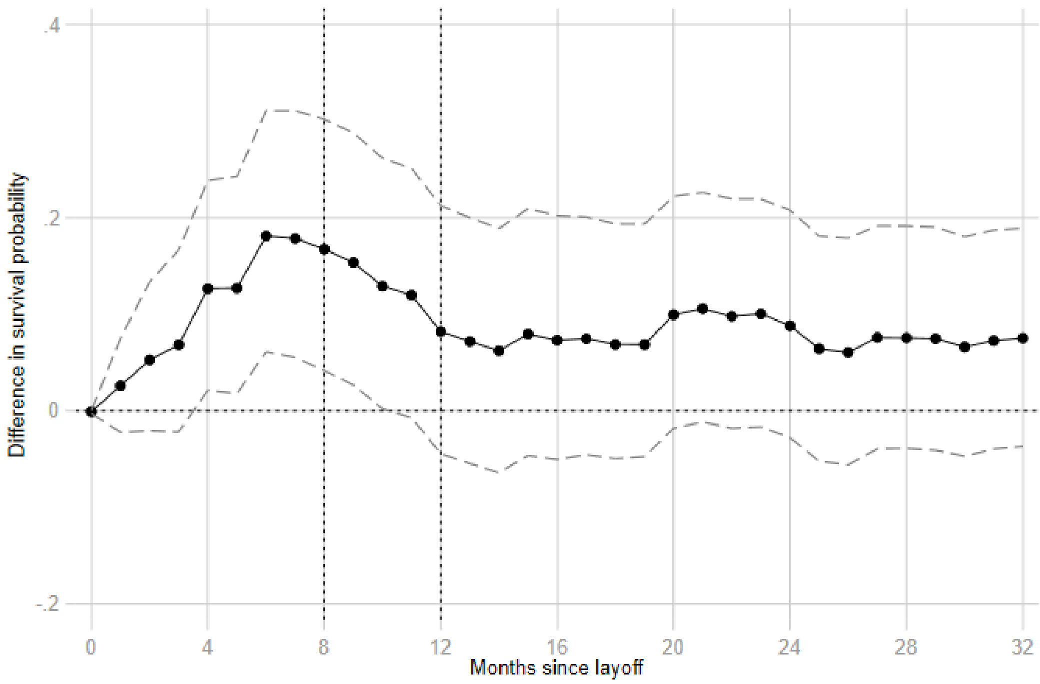
Note: Panel (a) plots point estimates of non-manipulators' actual and counterfactual nonemployment survival for the first 32 months after layoff. Our estimation strategy is outlined in Section III.B. Panel (b) shows the difference between the two survival curves and contains bootstrapped 95% confidence intervals testing against zero difference.

Figure VIII. Manipulators and non-manipulators with 8 months of potential benefit duration

(a) Nonemployment survival rates



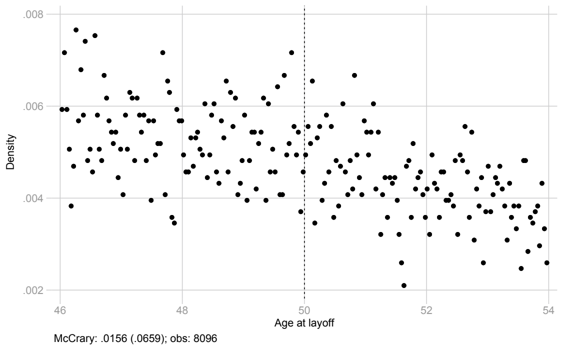
(b) Difference in survival rates



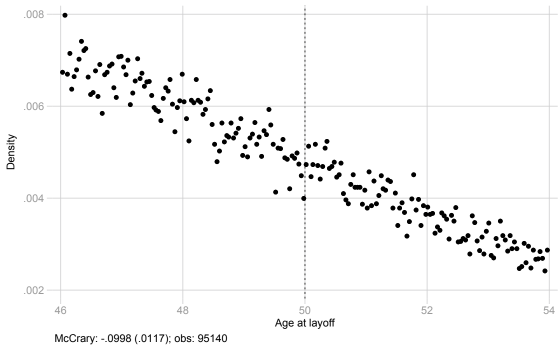
Note: Panel (a) plots point estimates of manipulators' and non-manipulators' nonemployment survival over the first 32 months after layoff under eight months of PBD. The estimation of the former is outlined in Section III.B. The latter represents the observed mean survival rate in the missing region. Panel (b) shows the difference between the two survival curves and contains bootstrapped 95% confidence intervals testing against zero difference.

Figure IX. Density of Layoff by Private and Public sector and by Contract Type

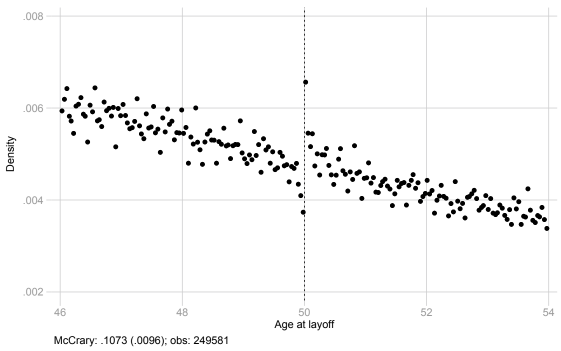
(a) Public Sector: Permanent



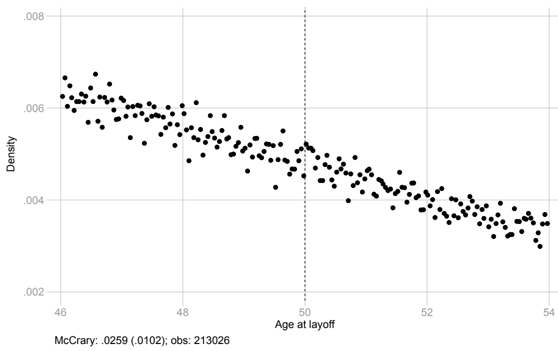
(b) Public Sector: Temporary



(c) Private Sector: Permanent



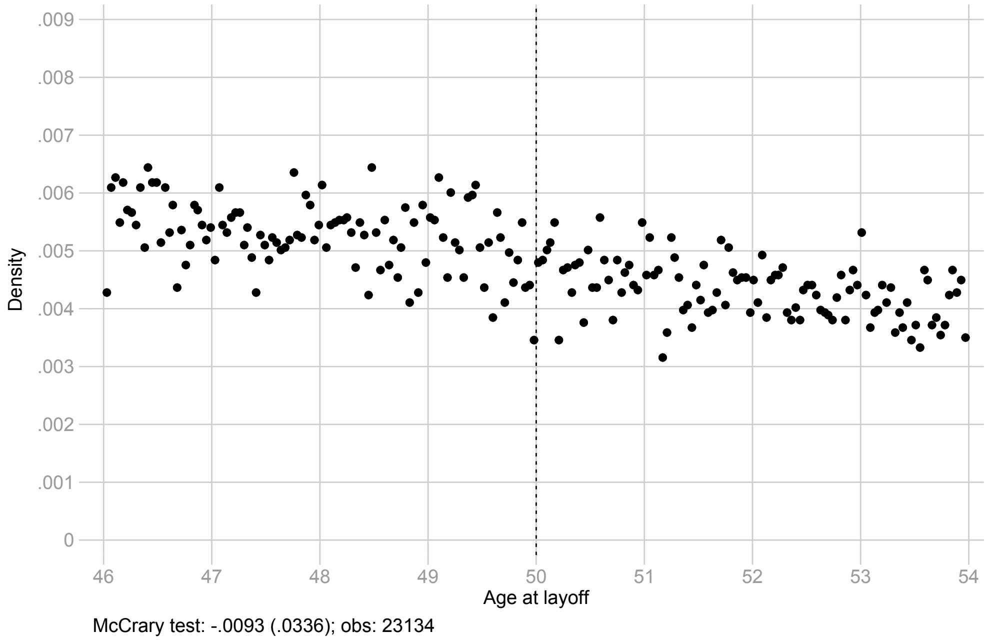
(d) Private Sector: Temporary



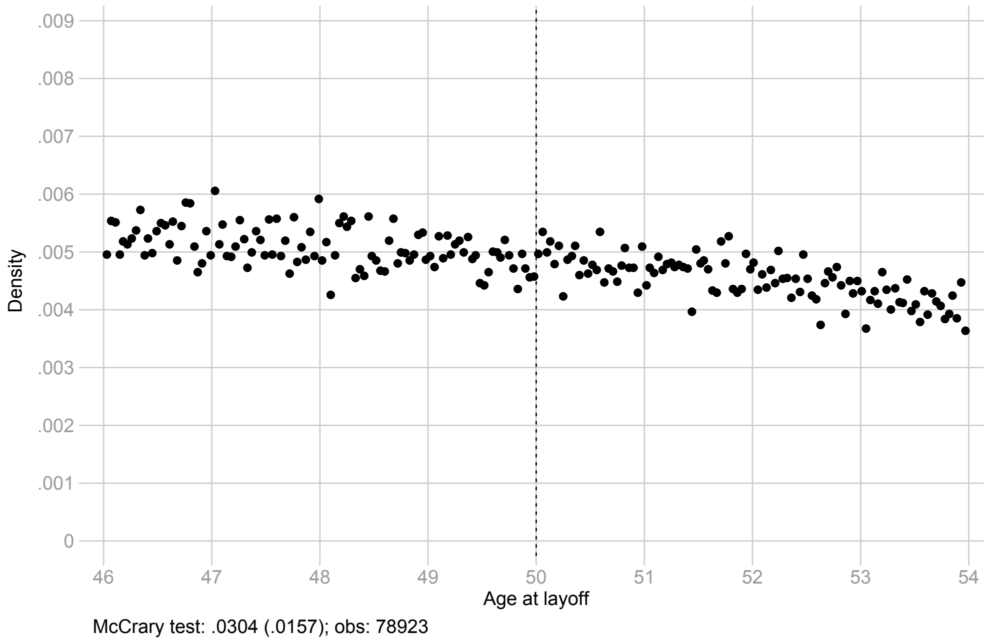
Note: The figure shows the density of layoffs by contract type. The data cover the period from Feb 2009 to Dec 2012. In all panels each dot represents a two-week bin. Individuals are classified as "public sector" workers if they cannot be matched to an employment spell in the private sector database (UNIEMENS).

Figure X. Placebo checks: MiniASpI and NASpI and density of recipients at 50 years of age

(a) MiniASpI



(b) NASpI



Note: The figure shows the density of layoffs for workers laid off in the private sector and receiving MiniASpI (Mar 2013 to Apr 2015) or NASpI (from Jan 2016). In both panels each dot represents a two-week bin. The sample has been restricted to workers coming from permanent contracts in the private sector.

Tables

Table I. Summary Statistics

	Mean	SD	Min	Max
Nonemployment outcomes				
UI Benefit receipt duration (in weeks)	29.853	15.923	0.14	52.00
Nonemployment duration (in weeks)	89.995	79.092	0.00	208.00
Nonemployment survival prob. 8 months	0.502	0.500	0.00	1.00
Nonemployment survival prob. 12 months	0.388	0.487	0.00	1.00
Individual characteristics				
Female (share)	0.311	0.463	0.00	1.00
Experience (in years)	27.656	8.552	2.00	40.00
White collar (share)	0.208	0.406	0.00	1.00
North (share)	0.367	0.482	0.00	1.00
Center (share)	0.174	0.379	0.00	1.00
South and islands (share)	0.459	0.498	0.00	1.00
Previous job characteristics				
Full time (share)	0.807	0.395	0.00	1.00
Tenure (in years)	5.931	6.113	0.08	30.00
Daily income (in Euro)	69.900	70.300	0.04	13,981.01
Firm age (in years)	14.367	12.115	0.00	109.83
Firm size	28.158	259.010	1.00	14,103.00
Firm size below 15 (share)	0.606	0.489	0.00	1.00
Firm size between 15 and 49 (share)	0.213	0.409	0.00	1.00
Firm size above 49 (share)	0.181	0.385	0.00	1.00

Note: The table reports summary statistics of our main sample consisting of all OUB claims from Feb 2009 to Dec 2012 from individuals who are employed in permanent private sector work arrangements and are between 46-54 years of age at the time of layoff. The sample contains a total of 249,581 nonemployment spells from 210,041 individual workers. Nonemployment duration is censored at four years and defined as the time distance between the date of layoff and the date of the first re-employment event that leads to UI benefit termination. Experience is equal to the number of years since the first social security contribution. Tenure is defined as the total number of years (not necessarily uninterrupted) spent with the last employer. The geographical South and Islands dummy encompasses employment in one of the following regions: Abruzzo, Basilicata, Calabria, Molise, Puglia, Sardegna and Sicilia.

Table II. Headcount and share estimates

(1) Headcount manipulators missing region	(2) Headcount non-manipulators missing region	(3) Headcount manipulators excess region	(4) Headcount all other ind. excess region	(5) Share estimate missing	(6) Share estimate excess
571.2	3038.0	608.6	2390.4	0.158	0.203
(458.5, 680.0)	(2931.0, 3150.0)	(496.0, 718.5)	(2379.4, 2401.3)	(0.127, 0.188)	(0.172, 0.231)

Note: The table reports estimates of the total number of individuals in four groups: (1) manipulators in the missing region, (2) non-manipulators in the missing region, (3) manipulators in the excess region and (4) all other individuals in the excess region. Column (5) and (6) contain estimates for the share of manipulators in the missing and excess region, respectively. We formally define all quantities in Section III.B. All results are based on our main sample consisting of 249,581 observations. Bootstrapped 95% confidence intervals are in parentheses.

Table III. UI Benefit receipt estimates (in Euro)

(1) Benefit receipt manipulators missing region	(2) Benefit receipt non-manipulators missing region	(3) Benefit receipt manipulators excess region	(4) Benefit receipt all other ind. excess region	(5) Benefit receipt response manipulators	(6) Benefit receipt response non-manipulators
5814.2	5223.5	8053.6	7044.2	2239.4	1636.9
(5178.5, 6459.2)	(5125.0, 5325.7)	(7326.9, 8836.5)	(6974.5, 7112.4)	(1276.7, 3261.6)	(1410.9, 1849.6)

Note: The table reports estimates of the mean UI benefit receipt (in Euro) of individuals in four groups: (1) manipulators in the missing region, (2) non-manipulators in the missing region, (3) manipulators in the excess region and (4) all other individuals in the excess region. Column (5) and (6) contain estimates of the UI benefit receipt response of manipulators and non-manipulators, respectively. We formally define all quantities in Section III.B. All results are based on our main sample consisting of 249,581 observations. Bootstrapped 95% confidence intervals are in parenthesis.

Table IV. Benefit duration estimates (in weeks)

(1) Benefit duration manipulators missing region	(2) Benefit duration non-manipulators missing region	(3) Benefit duration manipulators excess region	(4) Benefit duration all other ind. excess region	(5) Benefit duration response manipulators	(6) Benefit duration response non-manipulators
27.8 (25.2, 30.6)	24.8 (24.4, 25.2)	41.8 (38.3, 45.6)	35.8 (35.5, 36.2)	13.9 (9.4, 18.7)	9.9 (8.9, 10.9)

Note: The table reports estimates of the mean benefit duration (in weeks) of individuals in four groups: (1) manipulators in the missing region, (2) non-manipulators in the missing region, (3) manipulators in the excess region and (4) all other individuals in the excess region. Column (5) and (6) contain estimates of the benefit duration response of manipulators and non-manipulators, respectively. We formally define all quantities in Section III.B. All results are based on our main sample consisting of 249,581 observations. Bootstrapped 95% confidence intervals are in parenthesis.

Table V. BC/MC Ratio Estimates

	(1) without taxes ($\tau = 0\%$)	(2) with taxes ($\tau = 3\%$)
(a) Manipulators	0.24	0.32
	(0.02, 0.89)	(0.03, 1.13)
(b) Non-manipulators	0.26	0.32
	(0.12, 0.41)	(0.15, 0.50)

Note: The table reports BC/MC ratio estimates for (a) manipulators and (b) non-manipulators. BC/MC ratios are defined in equation (2). Bootstrapped 95% confidence intervals in parentheses.

Table VI. Selection on Observables

	(1) Manipulators	(2) Non-Manipulators	(3) Difference (1)-(2)
Female (share)	0.450	0.270	0.180
			(0.100, 0.281)
White Collar (share)	0.351	0.180	0.170
			(0.101, 0.239)
Southern Region (share)	0.483	0.471	0.012
			(-0.072, 0.098)
Full Time (share)	0.754	0.822	-0.067
			(-0.134, -0.000)
Tenure (in years)	6.577	5.718	0.859
			(-0.142, 1.853)
Daily Wage (in logs)	4.115	4.176	-0.0610
			(-0.142, 0.023)
Firm Age (in years)	14.546	14.335	0.211
			(-1.945, 2.320)
Firm Size (in logs)	1.862	2.258	-0.395
			(-0.640, -0.155)

Note: The table reports differences in observable characteristics between manipulators and non-manipulators in our main sample. Column 1 and 2 report estimated means of observable characteristics for manipulators and non-manipulators, respectively. Column 3 reports their difference and associated 95% bootstrapped confidence intervals in parentheses.

Appendices

For online publication

A Proofs

This appendix lays out the formal derivation of Proposition 3, which implies Proposition 1 for $M \equiv 0$. We further illustrate how to derive equation (4) as well as Proposition 2.

The government problem parameterized with baseline coverage P and extra coverage ΔP, such that $P_y = P$ and $P_o = P + \Delta P$, reads:

$$\max_{P,\Delta P,\tau} W = (1-G)\cdot V^o(P+\Delta P) + G\cdot \mathbb{E}^y\left[\tilde{V}^i(P)\right] + G\cdot M\cdot \mathbb{E}^m\left[\tilde{V}^i(P+\Delta P) - \tilde{V}^i(P) - q^i\right]$$

subject to the budget constraint:

$$\begin{aligned}&\tau\cdot[(1-G)\cdot(T-D^o(P+\Delta P)) + G\cdot(T-D^y(P)) + G\cdot M\cdot(D^m(P)-D^m(P+\Delta P))]\\ &= b\cdot[(1-G)\cdot B^o(P+\Delta P) + G\cdot B^y(P) + G\cdot M\cdot(B^m(P+\Delta P)-B^m(P))] + R.\end{aligned}$$

When considering small changes in extra coverage ΔP we may, by the envelope theorem, ignore all direct welfare effects of changes in job search intensities or manipulation choices.[32] Thus, small budget-neutral changes in ΔP have a welfare effect of:

$$\begin{aligned}\frac{dW}{d\Delta P} &= (1-G)\cdot\frac{dV^o(P_o)}{d\Delta P} + G\cdot M\cdot\mathbb{E}^m\left[\frac{d\tilde{V}^i(P_o)}{d\Delta P}\right] - \bar{u}'\cdot L\cdot\frac{d\tau}{d\Delta P} && (30)\\ &= (1-G)\cdot S^o_{P_o}\cdot(u^o(c_u+b)-u^o(c_u))\\ &+ G\cdot M\cdot\mathbb{E}^m\left[S^i_{P_o}\cdot\left(u^i(c_u+b)-u^i(c_u)\right)\right]\\ &- \bar{u}'\cdot\left[(1-G)\cdot\left(BC^o_{P_o}+MC^o_{P_o}\right) + G\cdot M\cdot\left(BC^m_{P_o}+MC^m_{P_o}\right)\right]\\ &+ \bar{u}'\cdot G\cdot b\cdot(1-M)\cdot\epsilon_{1-M,\Delta P} && (31)\\ &= (1-G)\cdot\left(MC^o_{P_o}\cdot\tilde{u}'_o - \left(BC^o_{P_o}+MC^o_{P_o}\right)\cdot\bar{u}'\right)\\ &+ G\cdot M\cdot\left(\mathbb{E}^m\left[MC^i_{P_o}\cdot\tilde{u}'_i\right] - \left(BC^m_{P_o}+MC^m_{P_o}\right)\cdot\bar{u}'\right)\\ &+ G\cdot(1-M)\cdot MC^n_{P_y}\cdot\bar{u}'\cdot\epsilon_{1-M,\Delta P}, && (32)\end{aligned}$$

where we used the implicit differentiation of the government budget constraint $\tau\cdot L =$

[32]These changes matter only to the extent that they operate through the government budget constraint.

$b \cdot B + R$ and Leibniz rule to obtain:

$$L \cdot \frac{d\tau}{d\Delta P} = b \cdot \frac{dB}{d\Delta P} - \tau \cdot \frac{dL}{d\Delta P} \tag{33}$$

$$\begin{aligned} &= (1-G) \cdot \left(b \cdot \frac{dB^o}{d\Delta P} + \tau \cdot \frac{dD^o}{d\Delta P} \right) \\ &+ G \cdot \frac{d}{d\Delta P} \int_i \underbrace{\left(b \cdot \left(B^i(P_o) - B^i(P_y) \right) + \tau \cdot \left(D^i(P_o) - D^i(P_y) \right) \right)}_{= FE^i \text{ by equation (6)}} \cdot \mathbb{I}_{q^i \leq \bar{q}^i} \, df(u^i, \psi^i, q^i) \end{aligned} \tag{34}$$

$$\begin{aligned} &= (1-G) \cdot \left(BC^o_{P_o} + MC^o_{P_o} \right) + G \cdot \int_i \frac{dFE^i}{d\Delta P} \cdot \mathbb{I}_{q^i \leq \bar{q}^i} \, df(u^i, \psi^i, q^i) \\ &+ G \cdot \int_{u^i, \phi^i} FE^i \cdot \frac{d}{d\Delta P} \underbrace{\int_0^{\bar{q}^i} f(q|u_i, \phi_i) \, dq}_{= M^i \text{ by equation (7)}} \, df(u^i, \phi^i) \end{aligned} \tag{35}$$

$$\begin{aligned} &= (1-G) \cdot \left(BC^o_{P_o} + MC^o_{P_o} \right) + G \cdot M \cdot \left(BC^m_{P_o} + MC^m_{P_o} \right) \\ &- G \cdot MC^n_{P_y} \cdot \int_{u^i, \phi^i} \frac{(1-M^i) \cdot FE^i}{MC^n_{P_y} \cdot \Delta P} \cdot \epsilon_{1-M^i, \Delta P} \, df(u^i, \phi_i) \end{aligned} \tag{36}$$

$$\begin{aligned} &= (1-G) \cdot \left(BC^o_{P_o} + MC^o_{P_o} \right) + G \cdot M \cdot \left(BC^m_{P_o} + MC^m_{P_o} \right) \\ &- G \cdot (1-M) \cdot MC^n_{P_y} \cdot \epsilon_{1-M, \Delta P}, \end{aligned} \tag{37}$$

by the definition in equation (8). Exploiting assumptions 1 and 2, we rewrite (32) as:

$$\begin{aligned} \frac{1}{\bar{u} \cdot b} \cdot \frac{dW}{d\Delta P} &= (1-G) \cdot S^o_{P_o} \left(\frac{\tilde{u}'_o - \bar{u}'}{\bar{u}'} - \frac{BC^o}{MC^o} \right) \\ &+ G \cdot M \cdot S^m_{P_o} \left(\frac{\tilde{u}'_m - \bar{u}'}{\bar{u}'} - \frac{BC^m}{MC^m} \right) \\ &+ G \cdot (1-M) \cdot S^n_{P_y} \cdot \epsilon_{1-M, \Delta P}, \end{aligned} \tag{38}$$

which proves equation (5) in the main text.

Similarly, small budget-neutral changes in baseline coverage P have a welfare effect of:

$$
\begin{aligned}
\frac{dW}{dP} =\ & (1-G)\cdot\frac{dV^o(P_o)}{dP} + G\cdot\mathbb{E}^y\left[\frac{d\tilde{V}^i(P_o)}{dP}\right] \\
+\ & G\cdot M\cdot\mathbb{E}^m\left[\frac{d\tilde{V}^i(P_o)}{dP} - \frac{d\tilde{V}^i(P_y)}{dP}\right] - \bar{u}'\cdot L\cdot\frac{d\tau}{dP} && (39)\\
=\ & (1-G)\cdot\left(MC^o_{P_o}\cdot\tilde{u}'_o - \left(BC^o_{P_o} + MC^o_{P_o}\right)\cdot\bar{u}'\right) \\
+\ & G\cdot\left(MC^y_{P_y}\cdot\tilde{u}'_y - \left(BC^y_{P_y} + MC^y_{P_y}\right)\cdot\bar{u}'\right) \\
+\ & G\cdot M\cdot\left(\mathbb{E}^m\left[MC^i_{P_o}\cdot\tilde{u}'_i\right] - \left(BC^m_{P_o} + MC^m_{P_o}\right)\cdot\bar{u}'\right) \\
-\ & G\cdot M\cdot\left(\mathbb{E}^m\left[MC^i_{P_y}\cdot\tilde{u}'_i\right] - \left(BC^m_{P_y} + MC^m_{P_y}\right)\cdot\bar{u}'\right) \\
+\ & G\cdot(1-M)\cdot MC^n_{P_y}\cdot\bar{u}'\cdot\epsilon_{1-M,P}, && (40)
\end{aligned}
$$

where, again, we used the implicit differentiation of the government budget constraint $\tau\cdot L = b\cdot B + R$ and Leipniz rule to obtain:

$$
\begin{aligned}
L\cdot\frac{d\tau}{dP} =\ & b\cdot\frac{dB}{dP} - \tau\cdot\frac{dL}{dP} && (41)\\
=\ & (1-G)\cdot\left(b\cdot\frac{dB^o}{dP} + \tau\cdot\frac{dD^o}{dP}\right) + G\cdot\left(b\cdot\frac{dB^y}{dP} + \tau\cdot\frac{dD^y}{dP}\right) \\
+\ & G\cdot\frac{d}{dP}\int_i FE^i\cdot\mathbb{I}_{q^i\le\bar{q}^i}\, df(u^i,\psi^i,q^i) && (42)\\
=\ & (1-G)\cdot\left(BC^o_{P_o} + MC^o_{P_o}\right) + G\cdot\left(BC^y_{P_y} + MC^y_{P_y}\right) \\
+\ & G\cdot\int_i\frac{dFE^i}{dP}\cdot\mathbb{I}_{q^i\le\bar{q}^i}\, df(u^i,\psi^i,q^i) + G\cdot\int_{u^i,\phi^i} FE^i\cdot\frac{dM^i}{dP}\, df(u^i,\phi^i) && (43)\\
=\ & (1-G)\cdot\left(BC^o_{P_o} + MC^o_{P_o}\right) + G\cdot\left(BC^y_{P_y} + MC^y_{P_y}\right) \\
+\ & G\cdot M\cdot\left[\left(BC^m_{P_o} + MC^m_{P_o}\right) - \left(BC^m_{P_y} + MC^m_{P_y}\right)\right] - G\cdot(1-M)\cdot MC^n_{P_y}\cdot\epsilon_{1-M,P}, && (44)
\end{aligned}
$$

and define

$$
\epsilon_{1-M,P} := \mathbb{E}^n\left[\frac{FE^i}{MC^n_{P_y}\cdot P}\cdot\epsilon_{1-M^i,P}\right]. \tag{45}
$$

Under assumptions 1 and 2, we may rewrite (40) to:

$$
\begin{aligned}
\frac{1}{\bar{u}\cdot b}\cdot\frac{dW}{dP} =\ & (1-G)\cdot S^o_{P_o}\left(\frac{\tilde{u}'_o-\bar{u}'}{\bar{u}'} - \frac{BC^o}{MC^o}\right) + G\cdot S^y_{P_y}\left(\frac{\tilde{u}'_y-\bar{u}'}{\bar{u}'} - \frac{BC^y}{MC^y}\right) \\
+\ & G\cdot M\cdot\left(S^m_{P_o} - S^m_{P_y}\right)\cdot\left(\frac{\tilde{u}'_m-\bar{u}'}{\bar{u}'} - \frac{BC^m}{MC^m}\right) + G\cdot(1-M)\cdot S^n_{P_y}\cdot\epsilon_{1-M,P}, && (46)
\end{aligned}
$$

which proves equation (9) in the main text.

To prove equation (10) in Proposition 3, we substitute equation (5) into (9), which gives:

$$
\begin{aligned}
& S^y_{P_y} \cdot \left(\frac{\tilde{u}'_y - \bar{u}'}{\bar{u}'} - \frac{BC^y}{MC^y} \right) - M \cdot S^m_{P_y} \cdot \left(\frac{\tilde{u}'_m - \bar{u}'}{\bar{u}'} - \frac{BC^m}{MC^m} \right) \\
+ & \; (1-M) \cdot S^n_{P_y} \cdot (\epsilon_{1-M,P} - \epsilon_{1-M.\Delta P}) \; = \; 0.
\end{aligned} \tag{47}
$$

Noting that,

$$
\begin{aligned}
S^y_{P_y} \cdot \left(\frac{\tilde{u}'_y - \bar{u}'}{\bar{u}'} - \frac{BC^y}{MC^y} \right) = & \; M \cdot S^m_{P_y} \cdot \left(\frac{\tilde{u}'_m - \bar{u}'}{\bar{u}'} - \frac{BC^m}{MC^m} \right) \\
& + \; (1-M) \cdot S^n_{P_y} \cdot \left(\frac{\tilde{u}'_n - \bar{u}'}{\bar{u}'} - \frac{BC^n}{MC^n} \right),
\end{aligned} \tag{48}
$$

we rewrite (47) to obtain:

$$
\frac{\tilde{u}'_n - \bar{u}'}{\bar{u}'} - \frac{BC^n}{MC^n} \; = \; \epsilon_{1-M,\Delta P} - \epsilon_{1-M,P}. \tag{49}
$$

For expositional ease we define

$$
s \; = \; \frac{S^m_{P_y} - S^n_{P_y}}{S^n_{P_y}}, \tag{50}
$$

and introduce shorthand notation for the social surplus from insurance for group $j \in \{n, m, y, o\}$:

$$
SSP^j \; = \; \left(\frac{\tilde{u}'_j - \bar{u}'}{\bar{u}'} - \frac{BC^j}{MC^j} \right). \tag{51}
$$

Finally, we rewrite (48) as follows:

$$
\begin{aligned}
S^y_{P_y} \cdot SSP^y = & \; S^n_{P_y} \cdot SSP^n + M \cdot \left[S^m_{P_y} \cdot SSP^m - S^n_{P_y} \cdot SSP^n \right] \\
= & \; S^n_{P_y} \cdot SSP^n + M \cdot \left[S^m_{P_y} \cdot SSP^m - S^n_{P_y} \cdot SSP^n + S^n_{P_y} \cdot SSP^m - S^n_{P_y} \cdot SSP^m \right] \\
= & \; S^n_{P_y} \cdot SSP^n + M \cdot \left[S^n_{P_y} \cdot (SSP^m - SSP^n) + \left(S^m_{P_y} - S^n_{P_y} \right) \cdot SSP^m \right] \\
= & \; S^n_{P_y} \cdot (SSP^n + M \cdot (SSP^m - SSP^n) + s \cdot M \cdot SSP^m) \\
= & \; S^n_{P_y} \cdot ((1 + s \cdot M) \cdot SSP^n + (1 + s) \cdot M \cdot (SSP^m - SSP^n)),
\end{aligned} \tag{52}
$$

which, since $S^y_{P_y} = S^n_{P_y} \cdot (1 + s \cdot M)$ and $\frac{1+s}{1+s\cdot M} = \frac{S^m_{P_y}}{S^y_{P_y}}$, implies

$$SSP^y \;=\; SSP^n + M \cdot \frac{S^m_{P_y}}{S^y_{P_y}} \cdot (SSP^m - SSP^n)\,. \tag{53}$$

Substituting (49) and (51) completes the proof of equation (10) in Proposition 3.

Last, we derive equation (11) in Proposition 3 by rewriting (5) as:

$$\begin{aligned}(1-G)\cdot S^o_{P_o}\cdot SSP^o \;&=\; -G\cdot\left[M\cdot S^m_{P_o}\cdot SSP^m + (1-M)\cdot S^n_{P_y}\cdot \epsilon_{1-M,\Delta P}\right]\\ &=\; -G\cdot\left[S^y_{P_o}\cdot SSP^y - (1-M)\cdot S^n_{P_o}\cdot SSP^n + (1-M)\cdot S^n_{P_y}\cdot \epsilon_{1-M,\Delta P}\right],\end{aligned}$$

which with equation (49) implies:

$$\begin{aligned}(1-G)\cdot S^o_{P_o}\cdot SSP^o + G\cdot S^y_{P_o}\cdot SSP^y \;&=\; G\cdot(1-M)\cdot\left[S^n_{P_o}\cdot SSP^n - S^n_{P_y}\cdot \epsilon_{1-M,\Delta P}\right]\\ &=\; G\cdot(1-M)\cdot\left[\left(S^n_{P_o} - S^n_{P_y}\right)\cdot \epsilon_{1-M,\Delta P} - S^n_{P_o}\cdot \epsilon_{1-M,P}\right],\end{aligned}$$

and concludes the proof.

B Additional Institutional Details

This section provides additional information about the Italian unemployment insurance schemes in place from 2009. Our main sample covers the period from February 2009 until December 2012. There were two alternative UI schemes in place simultaneously to the main OUB scheme which we study in our analysis.

II.A Alternative UI schemes in Italy from 2009 to 2012

During the years from 2009 to 2012 two other UI schemes were in place: the Reduced Unemployment Benefits (RUB) and the Mobility Indemnity (MI).[33]

The RUB scheme targeted similar workers as OUB albeit different contribution requirements. While still requiring the first contribution to social security to have happened at least two years before, the RUB scheme only required 13 weeks (78 days) of contributions over the past year (instead of 52 weeks within the last two years as in OUB). The milder eligibility requirements went hand in hand with less generous benefits. Potential benefit duration was proportional to the days worked in the previous year (up to 180 days), while the replacement rate granted 35% of the average wage earned in the previous year for the first 120 days and 40% for the following 60 days. Because RUB is significantly less generous it is unlikely to interfere with our analysis of the OUB.[34]

The MI scheme (active until 2017) and was targeted to workers fired during mass layoffs or business re-organizations. It provided long and generous income support with active labor market reintegration and retraining programs. During the period under study the potential duration of this scheme depended on the worker's age at layoff and geography, with a maximum PBD of 48 months in the south and of 36 months in northern regions. UI benefits amounted to 80% of the salary for the first 12 months (with a cap annually set by law) and 64% during the following months. MI benefits represented a particularly attractive alternative for individuals involved in mass layoffs and could be responsible for an under-representation of these types of workers in our sample. What is more relevant for our analysis however is that selection into MI is largely beyond the control of the worker. Indeed, eligible firms needed to be undergoing significant economic restructuring and have a minimum size, while workers needed to meet additional tenure requirements.

II.B UI schemes in Italy after 2012

The Italian welfare system underwent significant reform after 2012 all aiming at reducing the fragmentation of benefit schemes. In January 2013, both the OUB and the

[33] *Indennità di Disoccupazione Ordinaria a Requisiti Ridotti* and *Indennità di Mobilità* in Italian, respectively.
[34] For additional information, please refer to Anastasia et al. (2009).

RUB were replaced respectively by the ASpI and MiniASpI.[35]

The ASpI mimicked many aspects of the OUB both in terms of requirements and structure. Eligibility requirements of the ASpI followed those of the OUB scheme. Potential benefit duration was also identical initially, however, it was reformed several times in 2014 and 2015 which makes it difficult to include the ASpI in our analysis. Benefit levels differed with a replacement rate of 75% for the first six month, 60% for month seven to twelve and 45% thereafter (all as fractions of the average wage in the preceding two years before layoff).

The MiniASpI was aimed at workers who did not meet the requirement for the ASpI, but had accumulated at least thirteen weeks of work in the last year. Potential benefit duration was equal to half of the weeks worked over that time period. Benefit receipt was proportional to past wages: workers received 75% of the average wage received during the two previous years.

Since April 2015, both measures are replaced by a single UI scheme which provides homogeneous coverage to workers from all types of layoffs. The new scheme, the NASpI, is based on the structure of the MiniASpI. To qualify, workers need at least 78 days of contributions in the year before layoff. Potential benefit duration is equal to half of the weeks worked over the previous four years. Benefit levels are proportional to past wages following a declining profile starting at 75% replacement rate with a 3 p.p. reduction for every month after the first four. Importantly for our analysis, there is no longer a discontinuity is potential benefit duration thus removing incentives for workers to delay their layoff.

[35] *Assicurazione Sociale per l'Impiego* in Italian.

C Additional Robustness Tests

This section provides additional evidence in support of the identifying assumptions. Concretely, our analysis assumes that the discontinuity in PBD around the age threshold affects layoff decisions only through the delay of otherwise earlier occurring layoffs. The main threat to this assumption is the possibility of *extensive margin responses*, i.e. increases in the rate of job separations due to the incentives generated by the UI system. This is worrisome for two reasons. First we would be mis-measuring the upper bound of the manipulation region (z_U). Second, if the extra layoffs are systematically different, we would be altering the composition of layoffs in the manipulation region for reasons other than manipulation, introducing bias.

Extensive margin responses to UI have been studied both theoretically, see e.g. early work by Feldstein (1976), Feldstein (1978) and Topel (1983), as well as in recent empirical work e.g. Albanese et al. (2020) and Jäger et al. (2019).

Albanese et al. (2020) find alleviated job separation rates as a response to the same Italian OUB scheme that we study but exploit the eligibility discontinuity of 52 contribution weeks within the last two years after which a worker qualifies for *any* UI. Although closely related there are several reasons why we might not find job separation effects in our context. Their variation is from zero to some PBD, whereas we study a PBD extension from a nonzero level. Because we are exploiting intensive rather than extensive margin incentives, extensive margin responses are likely significantly smaller. This is especially true because all workers in our sample are eligible for UI and have thus already "survived" the eligibility threshold Albanese et al. (2020) exploit.

The work by Jäger et al. (2019) documents job separation effects of a large PBD reform in Austria which raised PBD from one to four years. They exploit this large variation to form a test for the efficiency of job separations by studying differences in separation rates of surviving job cohorts that were differentially treated by the reform. Again, there are several reasons to caution against extrapolating from their setting to ours. First, the sheer size of the PBD extension in Austria was unusually large. Second, it was targeted at relatively old workers who, as Jäger et al. (2019) document, used it (in part) as a gateway into early retirement. Last, their setting is likely to produce larger extensive margin responses because the Austrian UI scheme covers voluntary quits and not just layoffs as in Italy.

Although there exists recent important evidence on the extensive margin job separation effects of UI programs we see reason to believe that such effects are significantly smaller or entirely absent in our context. Of course, the presence of job separation effects is ultimately an empirical question. In the following we provide three tests all of which support the absence of extensive margin responses in our setting.

III.A Testing for Shifts in the Layoff Density

The first test is based on the shape of the layoff density. Concretely, we investigate whether there is a persistent increase in layoffs after the age fifty threshold. One might expect a persistent increase in the density if, for instance, firms that experience negative productivity shocks, dis-proportionally lay off workers above fifty due to the extended UI coverage. We operationalize this approach by estimating versions of a classical regression discontinuity design and estimate the following specification once for the entire sample and by excluding (an extended version of) the manipulation region:

$$d_j = \alpha + \lambda \cdot a_j + \gamma \cdot \mathbb{I}[a_j \geq 50] + \delta \cdot \mathbb{I}[a_j \geq 50] \cdot a_j + \nu_j, \qquad (54)$$

where d_j denote the density of layoffs in two-week age bin j, a_j denotes the mid-point age and ν_j is an error term. The coefficient of interest γ is indicative of any discontinuity in the density at the age fifty threshold. While we expect a positive γ coefficient when estimating specification (54) capturing the presence of manipulation, once we (successfully) exclude the manipulation region, γ should be close to zero in the absence of extensive margin responses. This is precisely what we find with results of all three regressions presented in Table A1. Column 1 presents estimates from the full sample where we do find a positive and significant γ coefficient of 0.027, consistent with the visual evidence in Figure III. More importantly, once we exclude the manipulation region in column 2, the estimated γ becomes indistinguishable from zero lending support to our identifying assumption. Column 3 repeats the previous analysis but with a modified definition of the manipulation region. Concretely, we extend the manipulation region to nine age bins prior to age fifty and four age bins after the threshold. The choice of this extended region is motivated by a simple quantitative heuristic. For the missing (excess) region we include the longest sequence of age bins from the threshold that are associated with negative (positive) regression coefficients in a simple OLS regression that allows for a separate effect of each age bin on the layoff frequency.[36] Reassuringly, the estimated γ coefficient in Table A1 remains quantitatively small and insignificant.

III.B Testing for the presence of extra excess mass

In this section we provide a second test based on the empirical layoff density. This time we investigate the possibility that extensive margin responses are concentrated right after the threshold. Rather than leading to a persistent increase in the density, which we tested for in the preceding section, we are concerned with the presence of additional layoffs just after the threshold that are not due to re-timing. Indeed such additional layoffs might occur if there are jobs that "mature" into negative surplus and

[36] In order to reduce the influence of very small coefficients, we ignore the sign of a coefficient if its absolute value is smaller or equal to 1/1000 of the average density across all bins. This is roughly equal to a deviation of three workers from the predicted counterfactual.

such separate precisely when the worker crosses the eligibility threshold for higher UI coverage. We probe this concern with the following analysis. First, in the absence of such additional layoffs missing and excess "mass", or numbers of manipulators, should balance exactly. If there are more excess manipulators one might be worried that these are the result of an extensive margin response thus violating our identifying assumption. We thus test the extent to which missing and excess mass balance around the threshold. To do so, we rely on the same definition of an extended manipulation region as in Section III.A. Concretely, we estimate the following specification

$$c_j = \alpha + \beta \cdot a_j + \sum_{k=A}^{50^-} \tilde{\gamma_k} \cdot \mathbb{I}[a_j = k] + \sum_{k=50^+}^{B} \tilde{\delta_k} \cdot \mathbb{I}[a_j = k] + \zeta_j, \tag{55}$$

where c_j corresponds to the number of layoffs in age bin j and a_j refers to the mid-point age in bin j. The set of $\tilde{\gamma_k}$ and $\tilde{\delta_k}$ coefficients capture the estimated number of manipulators in the respective bin k in the missing and excess region, respectively. The lower and upper bounds $A < z_L$ and $B > z_U$ are set to eighteen weeks (nine bins) and four eight weeks (four bins) as in the previous section. We calculate the difference between the sum of all $\tilde{\gamma}$ coefficients and the sum of all $\tilde{\delta}$ coefficients and re-scale it by the latter. The estimated 1.3% represents the share of the estimated manipulators in the excess region which is not explained by manipulators in the missing regions. Reassuringly, this number is very small lending further support to our main identification assumption.

III.C Testing for discontinuities in observable characteristics

Last we turn to a set of robustness tests based on observable characteristics around the age threshold. Intuitively, observable characteristics around the age cutoff should also differ due to manipulation. Similar to the density test in Section III.A we investigate if individuals differ based on their observable characteristics outside of the manipulation region. Concretely and for comparison, we run two regression models. The first is a standard regression discontinuity specification run on the full sample:

$$x_i = \alpha + \sum_{p=1}^{P} \lambda_p^{\leq 50} \cdot a_i^p \cdot \mathbb{I}[a_i < 50] + \sum_{p=0}^{P} \lambda_p^{>50} \cdot a_i^p \cdot \mathbb{I}[a_i \geq 50] + \xi_i, \tag{56}$$

where x_i denotes individual i's characteristic, a_i denotes age and P refers to the degree of the polynomial, in our case 2. In this standard RD specification the coefficient $\lambda_0^{>50}$ captures the jump at the threshold and is thus the coefficient of interest. The second model adds indicator variables for each age bin in the manipulation region and is specified as follows:

$$\begin{aligned} x_i = \kappa + \sum_{p=1}^{P} \theta_p^{\leq 50} \cdot a_i^p \cdot \mathbb{I}[a_i < 50] + \sum_{p=0}^{P} \theta_p^{>50} \cdot a_i^p \cdot \mathbb{I}[a_i \geq 50] \\ + \sum_{k=z_U}^{z_L} \delta_k \cdot \mathbb{I}[a_i = k] + \nu_i, \end{aligned} \tag{57}$$

where we use the main definition of the manipulation region, namely six weeks prior and four weeks after the age cutoff.

Each row of Table A2 reports the estimated $\lambda_0^{>50}$ coefficients from specification (56) and $\theta_0^{>50}$ coefficients from specification (57) for a given observable characteristics. Consistent with our main identifying assumption we find no significant estimates of $\theta_0^{>50}$ coefficients despite several of the estimates for $\lambda_0^{>50}$ being significant. Together these results show that once manipulation is taken into account, observable characteristics appear similar on either side of the age threshold, again consistent with the absence of extensive margin job separation effects.

D Additional Tables

Table A1. Test for Discontinuity in Layoff Density

	(1) Whole sample	(2) Without manipulation region	(3) Without manipulation region (alternative definition)
Age	-0.0366	-0.0335	-0.0319
	(0.0027)	(0.0023)	(0.0026)
$\mathbb{I}[age \geq 50] \times$ Age	-0.0000	0.00029	0.0002
	(0.0042)	(0.0032)	(0.0033)
$\mathbb{I}[age \geq 50]$	0.0270	0.0100	0.0015
	(0.0105)	(0.0075)	(0.0079)
Mean	0.48	0.48	0.48
R^2	0.866	0.898	0.904
N	208	203	195

Note: The table reports a parametric test to detect any discontinuity in the density of layoff around the 50 years of age threshold. Column (1) includes all age bins. Column (2) excludes the manipulation region which encompasses the three bins before the cutoff and the two bins after the cutoff. Column (3) excludes an extended manipulation defined in Section III.A. Robust standard errors are reported in parentheses.

Table A2. Test for Discontinuity in Observables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Simple RD model			"Donut" RD model			Baseline
Variable	$\lambda_0^{>50}$	s.e.	T-stat	$\theta_0^{>50}$	s.e.	T-stat	mean
Female	0.011	0.005	**2.43**	0.000	0.005	-0.03	0.31
Experience	0.177	0.095	1.85	0.093	0.107	0.87	27.34
White Collar	0.017	0.005	**3.71**	0.005	0.005	0.86	0.20
Southern Region	-0.003	0.006	-0.56	-0.005	0.007	-0.74	0.47
Full Time	0.001	0.005	0.26	0.005	0.005	1.09	0.81
Tenure (in years)	-0.040	0.063	-0.63	-0.095	0.078	-1.22	5.85
Daily Wage (in logs)	0.000	0.006	0.03	0.005	0.007	0.69	4.17
Firm Age (in years)	-0.116	0.130	-0.89	-0.122	0.137	-0.89	14.269
Firm Size (in logs)	-0.038	0.014	**-2.72**	-0.015	0.016	-0.94	2.02

Note: The table reports results for the robustness test outlined in Section III.C. Columns 1 to 3 report estimates of $\lambda_0^{>50}$ with associated standard error and t-stat from the RD specification (56). Columns 4 through 6 present the corresponding results for $\theta_0^{>50}$ from the "donut" RD model of specification (57). Each row represents a separate observable characteristic. T-stats are highlighted in bold if coefficients are significantly different from zero at the 5% level. Column 7 reports baseline averages for individuals fired between 49 and 50 years of age. The analysis is based on 249,581 spells of individuals laid off from a permanent contract from Feb 2009 to Dec 2012. Standard Errors clustered at the local labour market level.