

# Using Non-Linear Budget Sets to Estimate Extensive Margin Responses: Method and Evidence from the Earnings Test

Alexander M. Gelber, Damon Jones, Daniel W. Sacks, and Jae Song\*

March 2020

## Abstract

We estimate the impact of the Social Security Annual Earnings Test (AET) on older workers' employment. The AET reduces Social Security claimants' current benefits in proportion to their earnings in excess of an exempt amount. Using a Regression Kink Design and Social Security Administration data, we document that the discontinuous change in the benefit reduction rate at the exempt amount causes a corresponding change in the slope of the employment rate, suggesting that the extensive margin of labor supply is more sensitive to this policy than commonly thought. We develop a model and method that allows us to translate the behavioral responses into a lower bound estimate of 0.49 for the extensive margin elasticity, which implies more than a 1 percentage point increase in work in the absence of the AET.

---

\*Gelber: UC San Diego and NBER, amgelber@ucsd.edu; Jones: University of Chicago, Harris School of Public Policy and NBER, damonjones@uchicago.edu; Sacks: Indiana University, Kelley School of Business, dansacks@indiana.edu; Song: Social Security Administration, jae.song@ssa.gov. This research was supported by the U.S. Social Security Administration through grant #RRC08098400-06-00 to the National Bureau of Economic Research as part of the SSA Retirement Research Consortium and by grant #G-2015-14005 from the Alfred P. Sloan Foundation. The paper was completed partly while Gelber was the Alfred P. Sloan Visiting Associate Professor at the Stanford Institute for Economic Policy Research. The findings and conclusions expressed are solely those of the authors and do not represent the views of SSA, any agency of the Federal Government, or the NBER. We also thank the UC Berkeley IRLE, CGIF, and Burch Center for research support. We are extremely grateful to David Pattison for running our early code on the data. We thank David Card, Raj Chetty, Jim Cole, Julie Cullen, Rebecca Diamond, Avi Feller, Jon Gruber, Hilary Hoynes, Pat Kline, Rafael Lalive, Emmanuel Saez, Håkan Selin, Joel Slemrod, Dmitry Taubinsky and seminar participants at Brookings, Case Western Reserve, Cornell, Dartmouth, DePaul, Duke, the Federal Reserve Bank of Chicago, Maryland, NBER, Notre Dame, NYU, SIEPR, SOLE, UC Berkeley, UC Davis, University of Chicago, UIC, UIUC, and UVA for helpful comments and Mark Sheppard for editorial assistance. All errors are our own.

# 1 Introduction

Average U.S. life-expectancy has risen almost twice as fast as the average retirement age during the last few decades<sup>1</sup>, prompting concerns about workers’ ability to finance an ever-growing share of life spent in retirement. In this setting, it is important to consider policies that might indirectly discourage work, such as the Social Security Annual Earnings Test (AET). For Social Security<sup>2</sup> claimants who work, the AET reduces current benefits as a proportion of earnings above an exempt amount, and typically increases future benefits in an actuarially fair way. For example, in 2019, Social Security claimants who are younger than the normal retirement age (NRA) of 65 have their benefits reduced by 50 cents for each dollar earned above \$17,640. The AET was estimated to affect about 520,000 Social Security claimants in 2019 (Congressional Research Service, 2019), and the group affected, claimants younger than the NRA, is only growing over time, as the NRA gradually rises from 65 to 67. In fact, the Senior Citizens’ Freedom to Work Act of 2019 was proposed in Congress in 2019, with the aim of eliminating the AET “penalty” on older workers.<sup>3</sup> Previous studies have demonstrated a moderate impact of the AET on the number of hours worked by claimants (e.g. Burtless and Moffitt, 1985), despite an actuarial adjustment to future benefits that might partially offset the disincentive created by reductions in current benefits. However, less is known about the effect of the AET on the extensive margin, i.e. the decision of whether or not to work at all. Understanding this margin is important because it can have large welfare effects (e.g. Eissa et al. 2008).

In this study, we provide new evidence on the effects of the AET on older workers’ decision to work or not. We leverage the fact that the AET has nonlinear effects on worker incentives: the benefit reduction only kicks in above an exempt amount. We show that this generates a “kink,” i.e. a discontinuous slope, in a worker’s average net-of-tax rate (ANTR). The ANTR is defined as the fraction of earnings kept when moving from not working to working, after accounting for tax, transfers, and any benefit reduction due to the AET, and is therefore key in determining extensive margin labor behavior.<sup>4</sup> For workers likely to earn an amount near the AET exempt amount, we can then measure the extent to which a corresponding kink arises in the probability of employment. We relate the kink in the ANTR to the kink in employment using a Regression Kink Design (RKD) (Nielsen *et al.*, 2010; Card *et al.*, 2015) and estimate an extensive margin elasticity of employment for workers who would potentially earn near the AET exempt amount.

To estimate the employment effects of the AET, we use Social Security Administration (SSA) data on a 25 percent random sample of the U.S. population in birth cohorts 1918 to 1923 over the years 1978 to

---

<sup>1</sup>Life expectancy at birth rose from about 70 in 1962 to 78 in 2015 (World Bank Development Indicator, <https://fred.stlouisfed.org/series/SPDYNLE00INUSA>), the average retirement age grew from about 58.5 to 63.5 over the same period (Marketwatch, <https://www.marketwatch.com/story/why-the-average-retirement-age-is-rising-2017-10-09>).

<sup>2</sup>For ease of exposition, we use the term “Social Security benefits” to refer to Social Security Old Age and Survivors Insurance (OASI) only, and not other transfers administered by the Social Security Administration, such as Disability Insurance (DI) or Supplemental Security Income (SSI).

<sup>3</sup>Senior Citizens’ Freedom to Work Act of 2019, H.R. 2663, 116th Congress. (2019)

<sup>4</sup>Formally, for a tax schedule  $T(z)$ , and Social Security benefit function  $B(z)$ , we define the ANTR as  $ANTR \equiv 1 - [(T(z) - B(z)) - (T(0) - B(0))] / z$ , where  $z$  is pre-tax and transfer earnings. For consistency with the previous literature on kink points that has focused on the effect of taxation, we sometimes use “tax” as shorthand for “tax-and-transfer,” as in our use of the term “ANTR” to apply to the AET, while recognizing that the AET reduces Social Security benefits and is not administered through the tax system.

1987. These data cover over 11 million observations of calendar year earnings and other information on 1.4 million individuals in our full sample. Using an RKD, we show that as a function of past earnings—which we use to proxy for potential earnings—the slope of the employment rate among 63 to 64 year-olds (i.e. early retirees) discontinuously decreases around the budget set kink created by the AET at the exempt amount. We conduct several placebo and other robustness checks to verify that we have found a true effect on earnings, as opposed to an underlying nonlinearity in the employment rate.

In our baseline specification, we estimate an employment elasticity of at least 0.49. Simulations relying on these estimates show that eliminating the AET would cause an increase in the annual employment rate at ages 63 to 64 of 1.4 percentage points in the group we study near the kink, reflecting a 2.5 percent increase in employment. Using our extensive margin elasticity from this paper, as well as the AET intensive margin elasticities from Gelber *et al.* (2020), we calculate that earnings in the group we study decrease by 9.8 percent due to the AET. Of the total earnings reductions due to the AET, 27.6 percent are associated with extensive margin exit decisions. Our model and results therefore highlight the potential importance of the AET to older workers’ earnings and employment decisions.

Our new estimates of the effect of the AET on employment relate to long lines of previous research on the effect of Social Security and other types of pensions on retirement or employment decisions (see the literature review in Blundell *et al.*, 2016). With respect to the AET in particular, much of the previous literature has focused on the policy’s intensive margin effect (*e.g.* Burtless and Moffitt, 1985; Friedberg, 1998; Friedberg, 2000; Song and Manchester, 2007; Gelber *et al.*, 2020), typically finding moderate substitution elasticities at the intensive margin. Given the clear but moderate responses at the intensive margin, it is arguably surprising that nearly all prior studies find little evidence of extensive margin responses. Indeed, much of the earlier empirical literature on the AET concludes that the policy has little meaningful effect on the labor supply of older men.<sup>5</sup> More recent work has examined the effect of the AET on employment decisions using a difference-in-differences framework. Several of these papers find little evidence for an effect on the employment rate (Gruber and Orszag, 2003; Song, 2004; Song and Manchester, 2007; Haider and Loughran, 2008), while Friedberg and Webb (2009) find a significant effect in some specifications in the Current Population Survey.<sup>6</sup> Policymakers appear to have internalized these findings: a recent report prepared for Congress states “research has not found the [AET] to have a large effect on labor force participation.” (Congressional Research Service, 2019, pg. 17).

We interpret our evidence as showing a clear and robust impact of the AET on the annual employment rate, in contrast to the bulk of previous literature and current policy discussions. Although we are not the first to use a large administrative dataset with individual-level microdata (Song, 2004; Song and Manchester,

---

<sup>5</sup>See Viscusi (1979), Burtless and Moffitt (1985), Gustman and Steinmeier (1985, 1991), Vroman (1985), Honig and Reimers (1989), and Leonesio (1990).

<sup>6</sup>Examining the effects of earnings tests in other countries using difference-in-differences designs, Baker and Benjamin (1999) find a significant effect of the Canadian earnings test on weeks worked per year but no significant effect on employment at some point during the year, and Disney and Smith (2002) find inconclusive evidence on the impact of the U.K. earnings test on the employment rate. In contrast to these small or null effects, French (2005) uses method of simulated moments in a lifecycle model of labor supply and retirement; simulations based on these estimates imply that eliminating the AET would cause individuals to retire later on average.

2007; and Haider and Loughran, 2008), we introduce a new identification strategy based on discontinuities in incentives that leads to precise estimates of sizeable elasticities. Our paper is the first that has estimated a significant impact of the AET at the extensive margin through explicitly modeling individuals’ budget sets. It should be noted that because we estimate the effect of the AET on those locating near the exempt amount in the 63 to 64 year-old group—a younger group than those studied in the difference-in-differences literature cited above—our results may not be directly comparable. Nonetheless, our estimates are relevant to the age group for which the AET currently applies, and are therefore the most policy relevant for considering any further changes to the AET.

In the final section of our paper, we develop a model of labor supply and nonlinear budget sets to help clarify the interpretation of the treatment effects estimated using an RKD in this setting. First, we explore under what conditions a kink in the ANTR is expected to translate into a kink in the employment rate, and conclude that a kink arises in the case where some workers face frictions on adjusting their earnings on the intensive margin. Therefore, our empirical discovery of a kink in employment constitutes a joint test of extensive margin responses to the AET *and* the presence of intensive margin frictions, at least among some workers. Next, we show how our elasticity parameter is interpreted in a setting where some workers face intensive margin frictions and some do not, and show that our observed participation elasticity is likely to be a lower bound on the “structural” participation elasticity. The model extends recent methods that use nonlinear budget sets to estimate intensive margin elasticities or learn about intensive margin frictions (Saez, 2010; Chetty et al, 2011; Kleven and Waseem, 2013; Gelber et al., 2020) by showing how extensive margin elasticities may also be recovered in these settings. Furthermore, the model provides a new microfoundation of a fuzzy RKD estimator and, thus, a new application of the method.

The paper proceeds as follows. Section 2 describes the policy environment and basic framework for thinking about labor supply in this setting. Section 3 explains our empirical strategy and Section 4 discusses the data. Section 5 presents empirical evidence on the response to the AET and performs counterfactual simulations. Section 6 presents a model of labor supply that helps to interpret our results, and Section 7 concludes.

## 2 Policy Environment

### 2.1 Annual Earnings Test Rules

Social Security Old-Age and Survivors Insurance (OASI)—hereafter referred to simply as “Social Security”—provides benefits to older Americans and survivors of deceased workers, but delivery of benefits can be affected by whether one is currently working. For those who are simultaneously working and claiming Social Security benefits, the Social Security Annual Earnings Test (AET) reduces current benefits in proportion to earnings above an exempt amount, while typically adjusting future benefits upward in an actuarial fair fashion. For example, consider a 63-year-old earning \$23,640 in 2019, receiving \$1,000 in monthly benefits, and facing a \$17,640 exempt amount and a 50 percent benefit reduction rate (BRR). Her current annual benefits would

be reduced by  $\$3,000 = (\$23,640 - \$17,640) \times 50\%$ , equal to 3 months of benefits. In general, the exempt amount and the benefit reduction rate depend on age. During the years we study, 1978-1987, the exempt amount is much higher for people older than the normal retirement age of 65. People can claim benefits on their own record as early as age 62, the Early Entitlement Age (EEA). For people aged 62 to 65, the exempt amount ranges from \$9,787 to \$11,517, while the exempt amount for those NRA and over is on average around \$3,600 higher. The exempt amount by year and age range is shown in Online Appendix Figure E.1. The AET applied to earnings from ages 62 to 71 between 1978 and 1982, and from ages 62 to 69 between 1983 and 1987. For those under the NRA but above the EEA—the primary group featured in our study—the BRR is 50 percent throughout our study period.

When current Social Security benefits are lost to the AET, future scheduled benefits may be increased in some circumstances. This is referred to as “benefit enhancement.” For beneficiaries below the NRA in particular, the benefit enhancement, known as the “actuarial adjustment,” raises future benefits whenever a claimant earns over the AET exempt amount.<sup>7</sup> Future benefits are raised by 0.55 percent per month of benefits withheld during the years prior to the NRA. Returning to the example above, consider the 63-year-old receiving \$1,000 in monthly benefits due to the AET. Upon reaching the NRA, her monthly benefits would increase by around  $\$16.50 = 0.0055 \times 3 \times \$1,000$ . On average, this adjustment is roughly actuarially fair when considering the timing of claiming Social Security (Diamond and Gruber, 1999).

Even after taking benefit enhancement into account, there are a number of reasons that the AET could create a positive implicit marginal tax rate, and therefore affect earnings and employment decisions. Those whose expected lifespan is shorter than average should expect to collect Social Security benefits for less time than average, implying that the AET is financially punitive. Liquidity-constrained individuals or those who discount faster than average may also reduce work in response to the AET. Finally, many individuals also may not understand AET benefit enhancement or other aspects of Social Security (Honig and Reimers, 1989; Liebman and Luttmer, 2015; Brown *et al.*, 2013). Indeed, the AET is widely viewed as a pure tax. Most popular guides highlight AET benefit reduction but do not note the subsequent adjustment in benefits (Gruber and Orszag, 2003). During the period that we study, the popular guide *Your Income Tax* (J.K. Lasser Institute, 1987), for example, warned readers that if “you are under age 70, Social Security benefits are reduced by earned income,” but did not mention benefit adjustment. This is consistent with survey evidence showing that most older adults understand that the AET reduces current benefits but do not know that later benefits may be affected (Brown and Perron, 2011). We follow most previous work and do not distinguish among the potential reasons for a response to the AET, though Gelber *et al.* (2013) explore some of these potential mechanisms.

Though the AET rules are more complicated in the case of a married couple, the key parameters and outcomes for the purposes of our study—individual earnings and the 50 percent BRR—remain unchanged. For married couples in our study period, the couple’s total benefit is reduced at the rate of 50 percent

---

<sup>7</sup>Social Security Administration (2012); Gruber and Orszag (2003).

for every dollar of each spouse’s individual earnings above the respective spouse’s exempt amount. How this occurs depends on how benefits are claimed. If each spouse is a primary beneficiary, then the AET reduces each spouse’s separate benefit by the BRR multiplied by that spouse’s individual current earnings in excess of the exempt amount. If one spouse is a primary beneficiary and the other is a secondary or dual-entitled beneficiary, the couple’s total benefits are reduced by the BRR multiplied by the primary beneficiary’s current individual earnings in excess of the exempt amount, and further reduced by the BRR multiplied by the secondary or dual beneficiary’s current individual earnings in excess of the exempt amount. In either case, the relevant amount for applying the AET is each individual’s current annual earnings, which we observe in our data.

## 2.2 Modeling the Employment Response to the AET

In this section, we outline a basic static model of employment that captures the key features of the AET and underlies our empirical strategy. We discuss how the AET impacts an individual’s decision of whether or not to have a positive amount of earnings, which we refer to as the “employment” decision. Importantly, we focus on the reduction in current benefits due to the AET and not the accompanying increase in future benefits. What matters for our model is that individuals view the net effect of these adjustments as a decrease in benefits. This is consistent with the empirical finding in prior studies that some individuals bunch at the exempt amount. We therefore follow prior literature in modeling the AET as effectively creating a positive marginal tax rate above the exempt amount (*e.g.* Friedberg, 1998; Friedberg, 2000). This introduces a kink in the budget set, which in turn creates a kink in the incentives to work or not.<sup>8</sup> Although we presently abstract from the adjustment to future benefits that is also triggered by the AET, we explain In Online Appendix B.4 how relaxing this assumption affects the interpretation of our results.

We index two potential states of the world by  $j \in \{0, 1\}$ . Individuals choose a level of gross earnings,  $z$ , and receive Social Security benefits,  $B_j(z)$ , which may depend on current earnings. In state 0 individuals have no AET, and therefore benefits are constant with respect to earnings, *i.e.*  $B_0(z) = b$ . In state 1, the AET is introduced, with an exempt amount of  $z^{AET}$ . Benefits are reduced by a BRR of  $db$  for every dollar earned above  $z^{AET}$ . We therefore have:

$$B_1(z) = \begin{cases} b & \text{if } z < z^{AET} \\ b - db \cdot (z - z^{AET}) & \text{if } z^{AET} \leq z < z^{AET} + b/db \\ 0 & \text{if } z^{AET} + b/db \leq z \end{cases} \quad (1)$$

Thus, there is a convex kink in net income,  $z + B_1(z)$ , at  $z^{AET}$ , where the marginal dollar kept goes from 1 to  $1 - db$ . Note that while the exempt amount  $z^{AET}$  is the same for all workers of a given age, the location of a second, non-convex kink at benefit exhaustion varies with an individual’s level of benefits  $b$  and therefore any responses to this kink will tend to be smoothed out in the data. Because of this, and because we do

<sup>8</sup>To clarify, we use “kink” in two senses—both to describe a discontinuity in the effective marginal tax rate in the budget set, and to describe a discontinuity in the slope of an outcome variable (in our case the employment rate).

not observe the relevant benefit amount needed to locate the second, non-convex kink, we focus on the first convex kink. The agent also faces a tax (and transfer) schedule  $T(z)$ , which does not vary by state and, for simplicity, we assume to be locally linear:  $T(z) = \tau_0 z$ .

As is done in much of the previous literature on employment responses to kinks (*e.g.* Hausman, 1981; Saez, 2010; Kleven and Waseem, 2013), we assume individuals maximize a utility function that is increasing in consumption and decreasing in earnings—reflecting the disutility of labor supply. As in Saez (2010) and much subsequent literature, we model the determination of earnings, rather than hours worked, as earnings (but not hours worked) are observed in many administrative datasets. The tradeoff between consumption and expending effort to increase earnings is governed by an ability parameter, which we assume is smoothly distributed in the population. Such a model predicts a smooth distribution of earnings in the absence of the AET, and bunching, or excess mass, at the exempt amount  $z^{AET}$ , when the AET is introduced (Saez 2010; Gelber et al. 2020). To capture realistic patterns of movement in and out of employment, we posit a fixed cost of positive earnings (Cogan, 1981; Eissa et al., 2008). These fixed costs could be financial costs associated with working (transportation, child care, etc.) or fixed commuting times to work, or the loss of non-divisible leisure (*e.g.* blocks of time with one’s family or grandchildren).<sup>9</sup> In this case, the extensive margin decision of whether or not to have positive earnings will be a function of the average net-of-tax rate (ANTR):

$$\begin{aligned} ANTR &\equiv \frac{[z - T(z) + B(z)] - [-T(0) + B(0)]}{z} \\ &= 1 - \frac{[T(z) - B(z)] - [T(0) - B(0)]}{z} \\ &= 1 - a \end{aligned} \tag{2}$$

This measures the fraction of earnings that are kept when moving from earning 0 to earning  $z$ , when taking into account taxes, transfers, any benefit reduction.

To demonstrate the impact of a kinked budget on the decision to work, we illustrate the extensive margin incentives created by a kink in Panel A of Figure 1. Here we plot the ANTR, as a function of counterfactual earnings, *i.e.* the earnings level in state 0, in the absence of the AET. With a constant benefit schedule, *i.e.* state 0, the ANTR is constant. We also plot the ANTR of a kinked benefit schedule, *i.e.* like that of state 1, *while holding income fixed at the state 0 level*. The ANTR now decreases above  $z^{AET}$ , and the slope of the ANTR decreases discontinuously at  $z^{AET}$ , *i.e.* there is a kink in the ANTR. Panel B of Figure 1 shows potential employment responses to the two different benefit schedules. The  $x$ -axis again measures counterfactual earnings in state 0 conditional on having positive earnings and will serve as a constant index of individuals across different states. On the  $y$ -axis, the dashed line plots an illustrative employment rate in the case of a constant benefit schedule.

We plot two additional patterns. First, the dotted line plots the employment rate in state 1 under a

---

<sup>9</sup>See Eissa et al. (2008) for a discussion of how the the empirical distribution of annual hours worked in the US, with mass points at zero and 2,000 hours, is inconsistent with a convex model of labor supply, and more in line with a model featuring fixed costs of employment.

kinked benefit schedule, when individuals are unrestricted in their intensive margin earnings choices. We see that the employment function is unchanged at counterfactual earnings levels below  $z^{AET}$ , as the benefit schedule remains the same in the two states. Above  $z^{AET}$  we see a gradual decrease in the probability of positive earnings in state 1 relative to state 0, due to a decrease in the ANTR, which in turn dampens the incentive to work. Nonetheless, the kink in the ANTR in Panel A does not translate into a kink in the employment rate. Intuitively, the ability to adjust on the intensive margin smoothes the changes in the slope of the ANTR at  $z^{AET}$ . Next, the solid line depicts the relationship between counterfactual earnings in state 0 and the probability of positive earnings when a kink is present in state 1 but earnings conditional on working are held fixed to their state 0 value. That is, we shut down the ability to adjust on the intensive margin. Now that we have closed one of the channels through which individuals respond to benefit reduction, the slope of the employment rate decreases discontinuously at  $z^{AET}$ . In practice, the observed pattern may lie somewhere between these two extremes, depending on the extent of intensive margin frictions in the population.

We formally derive these two results below in Section 6. The model in Section 6 allows us to map our reduced form estimates to key earnings supply elasticities, and in Section 6.6 we further extend the model in a number of ways to establish the robustness of our results to a set of increasingly general cases. For now, we proceed to estimate the parameters  $\alpha$  and  $\beta$  from Figure 1. These parameters measure the kink in the ANTR and in the employment rate. Estimating them lets us test whether the kink in ANTR indeed passes through to the employment rate. This primary test remains relevant in all of these extended cases.

### 3 Empirical Strategy

Our conceptual framework provides a path forward in quantifying the extensive margin effect of the AET by estimating the change in the slope of employment at the kink created by the AET. This can be accomplished with a regression kink design (RKD). In this section, we describe the assumptions, data requirements, and specifications we use to implement an RKD. Below in Section 6, we microfound a model of earnings and employment from which an RKD estimation naturally arises. In Section 6.7 we specifically describe how the features of that model map directly to the assumptions underlying the RKD procedure presented in this section.

#### 3.1 Regression Kink Design

Recent work on RKDs provides conditions under which a change in the slope of the treatment intensity can be used to identify local treatment effects by comparing the relative magnitudes of the kink in the treatment variable and the induced kink in the outcome variable (Nielsen *et al.*, 2010). Under some smoothness assumptions, the estimates can be interpreted as a treatment-on-the-treated or “local average response” parameter (Card *et al.*, 2015).

In our context, the treatment intensity is the effective ANTR for individual  $i$  at time  $t$ ,  $ANTR_{it}$ , and



the primary outcome variable is the probability of positive earnings after becoming subject to the AET,  $\Pr(z_{it} > 0)$ , where  $z_{it}$  are realized earnings. The  $ANTR_{it}$  is a function of earnings conditional on working, which we take as desired earnings in the absence of the AET,  $\tilde{z}_{i0t}$ . Consistent with Figure 1, the AET creates a discontinuity at the exempt amount in the slope of the ANTR as a function of  $\tilde{z}_{i0t}$ . In this case, we can estimate the causal effect of the AET on the employment rate by estimating the change at the exempt amount in the slope of the employment rate as a function of  $\tilde{z}_{i0t}$ , in comparison to the change at the exempt amount in the slope of the ANTR. Specifically, using a “fuzzy RKD” we can estimate the marginal effect of the ANTR on the employment probability as:

$$\left. \frac{\partial \Pr(z_{it} > 0)}{\partial ANTR_{it}} \right|_{\tilde{z}_{i0t} = z^{AET}} = \frac{\lim_{\tilde{z}_{i0t} \rightarrow z^{AET+}} \frac{\partial \Pr(z_{it} > 0 | \tilde{z}_{i0t})}{\partial \tilde{z}_{i0t}} - \lim_{\tilde{z}_{i0t} \rightarrow z^{AET-}} \frac{\partial \Pr(z_{it} > 0 | \tilde{z}_{i0t})}{\partial \tilde{z}_{i0t}}}{\lim_{\tilde{z}_{i0t} \rightarrow z^{AET+}} \frac{\partial ANTR_{it}(\tilde{z}_{i0t})}{\partial \tilde{z}_{i0t}} - \lim_{\tilde{z}_{i0t} \rightarrow z^{AET-}} \frac{\partial ANTR_{it}(\tilde{z}_{i0t})}{\partial \tilde{z}_{i0t}}} = \frac{\beta}{\alpha} \quad (3)$$

That is, the marginal effect we can estimate is the change at the exempt amount in the slope of the employment probability as a function of desired earnings ( $\beta$ ), divided by the change at the exempt amount in the slope of the ANTR as a function of desired earnings ( $\alpha$ ). These kinks correspond to the theoretical kinks depicted in Figure 1.

The numerator of (3) can be estimated by relating  $\Pr(z_{it} > 0)$  to desired earnings:

$$\Pr(z_{it} > 0) = \sum_{j=0}^J \delta_j (\tilde{z}_{i0t} - z_t^{AET})^j + D \cdot \sum_{j=0}^J \beta_j (\tilde{z}_{i0t} - z_t^{AET})^j + \mu_t + \varepsilon_{it} \quad (4)$$

where  $D = 1\{\tilde{z}_{i0t} \geq z_t^*\}$  is an indicator for being above the exempt amount, the  $\mu_t$  are time fixed effects,  $\varepsilon_{it}$  is an error term, and  $\beta_1$  is the change in the slope of  $\Pr(z_{it} > 0)$  at the exempt amount. We calculate  $\Pr(z_{it} > 0)$  at the individual level by averaging an indicator for employment over a range of ages, specifically ages 63 and 64 in our main specification. We retain the  $t$  subscript to allow for the fact that in different specifications we may investigate employment over different time periods. This RKD yields a non-parametric estimate of the effect, because asymptotically the estimator uses only data arbitrarily close to the exempt amount and is unbiased without any functional form restriction on the underlying function to be estimated (Card *et al.*, 2015).

Identification of the effect of the AET on employment relies on two key assumptions (Card *et al.*, 2015). First, in the neighborhood of the exempt amount there is no discontinuity in the slope of age 63 to 64 employment that occurs for reasons unrelated to the AET. For example, beneficiaries’ earnings could in principle be affected by other public programs, or by their human capital or work experience as manifested in their hourly wage. Married beneficiaries’ earnings could also be affected by their spouses’ effective MTRs. We follow Saez (2010) and subsequent literature studying the effects of public programs in assuming that these factors jointly would have a smooth effect on earnings in the neighborhood of  $z^{AET}$ .<sup>10</sup> Second, conditional

<sup>10</sup>The 1978 and 1986 amendments to the Age Discrimination in Employment Act (ADEA) extended the ages at which age discrimination in employment was prohibited, but this did not have a discontinuous effect on elderly work incentives at the exempt amount. The 1977 Social Security Amendments increased the Delayed Retirement Credit for those 65 to 69 beginning in 1982, and the maximum age to which the AET applied decreased from 71 to 69 in 1983, but again neither of these should confound our RKD strategy. Other programs, such as Medicaid, Supplemental Security Income, Disability Insurance (DI), or taxes such as unemployment insurance payroll taxes distort earnings incentives at low earnings levels. We eliminate essentially all DI recipients from our sample through the restriction to those who claimed at 62 or later. The kinks created by other programs are typically at least several thousand dollars away from the AET convex kink, and we verify that the kink in employment at the AET exempt amount in particular is statistically significant

on unobservables, the distribution of the assignment variable,  $\tilde{z}_{i0t}$ , is smooth (*i.e.* the p.d.f. is continuously differentiable) in this neighborhood. These assumptions may not hold if we observe sorting in relation to the exempt amount, as indicated by a discontinuous change at the exempt amount in the slope or level of the density of the assignment variable, or in the distribution of predetermined covariates.

### 3.2 Measuring the ANTR

We would like to measure the ANTR in period  $t$  (*i.e.* ages 63 to 64) using the rules of the AET and the earnings level  $\tilde{z}_{i0t}$ , the counterfactual level of earnings the individual would choose if there were no AET in period  $t$ . We hold earnings fixed at this counterfactual level in order to isolate policy variation in the ANTR from changes driven by individual labor decisions. However, we do not observe this counterfactual level of earnings, because earnings at ages 63 to 64 are endogenous to the AET, and, furthermore, earnings are not observed for those who are not employed at those ages. Because of this endogeneity, actual observed earnings at ages 62, 63, or 64 cannot serve as adequate proxies for this counterfactual earnings level.

Many other papers have grappled with the issue of how to proxy for earnings or wages if individuals choose to work, and thus how to proxy for the incentive to work. Given that the econometrician does not directly observe counterfactual earnings, some set of assumptions must be made in order to proceed. One solution is a selection correction in the context of the effect of wages on labor supply, which generally requires functional form assumptions (Heckman, 1979) or very powerful instruments (Powell, 1994). Another solution is to use demographics to impute earnings if an individual works (*e.g.* Meyer and Rosenbaum, 2001), which is more difficult in settings such as ours with a limited number of demographic variables in our administrative data.

Our solution to this issue is to estimate a bound on the elasticity by assuming that the ANTR an individual would face at age 63 or 64 at her desired earnings level is the same as the ANTR she would face given her earnings at age 60. Thus, we proxy for desired earnings conditional on working at ages 63 and 64 in the absence of the AET,  $\tilde{z}_{i0t}$ , using actual earnings at age 60,  $z_i^{60}$ . Individuals do not face the AET kink at age 60, and therefore at age 60 their budget set is on average approximately linear in the region of the exempt amount. We use earnings at age 60, rather than 61, to avoid any anticipatory manipulation of earnings, and we show that there is no evidence for anticipatory manipulation in age 60 earnings in our sample. Age 60 is the closest age to 63-64 that our data show is not subject to concerns about manipulation, and therefore serves as the best proxy for counterfactual earnings. In our baseline RKD specification, our running variable is the distance between age 60 earnings in a given year and the exempt amount applying to those aged 62 to 64 in that year.

The use of a proxy variable and the inability to observe  $ANTR_{it}$  raise challenges in the estimation of the first stage of a “fuzzy RKD” (see Battistin and Chesher, 2014, on measurement error in the general average treatment effect framework). In our case, we rely on an analytic expression for the denominator of equation (3). Specifically, we calculate the kink in  $ANTR_{it}$  using the AET rules, *i.e.*  $\alpha \equiv db/z^{AET}$ , where

---

relative to placebo kink locations.

$db$  is the BRR. Next, we must address the use of a proxy for desired earnings,  $z_i^{60}$ , which may suffer from measurement error. Following Card *et al.* (2015), we model the measurement error in our proxy as:

$$z_i^{60} = \tilde{z}_{i0t} + p_{it} \cdot v_{it}, \quad (5)$$

where  $v_{it}$  is a continuously distributed random variable, and  $p_{it}$  is an indicator variable that equals zero with probability  $\pi(\tilde{z}_{i0t}, v_{it}) > 0$ . Thus, the error in our observed proxy is a composite of a continuously distributed variable and a mass point at zero. This implies that with some positive probability,  $\pi$ , earnings at age 60 equals desired earnings conditional on working in period  $t$  under a budget set with no AET. Note that we need not assume that the measurement error is mean zero, only that the distribution of  $v_{it}$ , conditional on  $z_i^{60}$ , is a smooth function of  $z^{60}$  and that  $\pi(\tilde{z}_{i0t}, v_{it}) > 0$ .

Importantly, if  $\pi(\tilde{z}_{i0t}, v_{it}) < 1$ , so that desired earnings are variable from age 60 to ages 63 to 64, then our procedure should overestimate the absolute value of the change in the ANTR at the exempt amount in the first stage, and therefore we should estimate a lower bound on the treatment effect on employment. This is an implication of Card *et al.* (2015) and we show this formally in Appendix A. Since the lower bound we estimate will be large, our conclusion will be that the elasticity is large. If there is no such persistence in desired earnings, then we expect our estimate of the kink to be significantly attenuated, further driving the interpretation of our estimate as a lower bound. It is possible to put assumptions on the structure of the measurement error to gain further traction. As an extension, we model the measurement error process in Section 5.5 and use simulations to quantify the extent of attenuation. This is in turn used to adjust our estimates upward to account for measurement bias.<sup>11</sup>

### 3.3 Additional Specification Considerations

Implementation of an RKD requires a number of choices regarding specification. For our main results, we implement the Calonico *et al.* (2014) data-driven method for bandwidth selection. We report confidence intervals corrected for bias following Calonico *et al.* (2014), and we use a triangular kernel to weight the data near the exempt amount. Card *et al.* (2015) use both linear and quadratic specifications in their analysis. Calonico *et al.* (2014) propose an RKD estimator in which the quadratic specification can be used to correct for the bias in the linear estimator, while Ganong and Jäger (2015) argue in favor of a cubic specification. We implement linear, quadratic, and cubic versions of (4), to investigate the robustness of our results. We use the linear specification without additional controls as our baseline because it minimizes the corrected Akaike Information Criterion (AICc) and the Bayesian Information Criterion (BIC) for our main outcome (the employment rate at ages 63 to 64), as well as for nearly every other outcome.

We must also decide whether to allow a discontinuity at the exempt amount in the level of the outcome variable and whether to control for covariates (Ando, 2017). We allow for a discontinuity, although the results are virtually unchanged if we do not. We present results with and without controlling for covariates. Thus, for each sample or outcome we can produce estimates using six regressions: the linear, quadratic, and

---

<sup>11</sup>We provide extensive details on our simulation-based adjustment for measurement error in Online Appendix D.

cubic regressions; and a version of each further adding predetermined covariates using Calonico *et al.* (2016).

Finally, we focus primarily on employment outcomes at ages 63 and 64, even though the AET begins to apply as early as age 62 and remains in effect at ages 65 and older in the period we study. We make this choice due to the following considerations. The AET first applies to claimants when they reach Social Security eligibility at age 62, but it does not make sense to examine the effect of the AET on whether an individual has positive earnings in the calendar year she turns age 62. The reason is that we observe calendar year earnings, and we measure “age” in a given calendar year coarsely as the highest age a person attains during that calendar year. If an individual claims OASI during the calendar year she turns age 62, the AET only applies to earnings in the months after the individual claims. If the claimant earns at all during this calendar year—even during months prior to claiming OASI—then she will have positive earnings in that year. Thus, a person who is induced by the AET to stop earning after claiming would appear in the data with positive earnings during this calendar year, and therefore would appear to have no measured response to the AET. As a result, it appears highly unlikely that we would observe a measurable employment response to the AET at age 62 as we measure it in the data, even if the AET has a substantial impact on employment. We expect measurable effects only to appear as early as age 63. We choose 64 as the oldest age at which to examine employment effects because age 60 earnings are a better proxy for desired earnings at ages 63 to 64 than for older ages. Moreover, at age 65 individuals with earnings near the under-NRA exempt amount are only exposed to this exempt amount for part of the year, as they transition to the much higher exempt amount applying to those at NRA and above on their 65th birthday.

## 4 Data

We implement our RKD estimation strategy using the restricted-access Social Security Administration Master Earnings File (MEF) linked to the Master Beneficiary Record (MBR) (Social Security Administration 2014a; Social Security Administration 2014b). The data contain Federal Insurance Contributions Act (FICA) pre-tax earnings for all Social Security Numbers (SSNs) in the U.S. in each calendar year.<sup>12</sup> Separate information is available on self-employment and non-self-employment earnings. The data are from W-2s, mandatory forms filed with the Internal Revenue Service for each employee for whom the firm withholds taxes and/or to whom remuneration exceeds a modest threshold. Thus, we have data on earnings regardless of whether an employee files taxes. The data longitudinally follow individuals over time.

The MBR contains information on exact date of birth, exact date of death, month and calendar year of claiming Social Security, race, and sex. In the calendar year after an individual dies, earnings and employment appear in the dataset as zeroes. Thus, some of an effect on employment could in principle be mediated through an effect on mortality, which would affect the interpretation—but not the validity—of our

---

<sup>12</sup>Data are also available on total W-2 earnings, as opposed to FICA earnings that are capped at the maximum taxable earnings level. We use FICA earnings because in 1978 to 1980 the data on total W-2 earnings are subject to measurement error that is not an issue in the FICA earnings data (Utendorf, 2001/2002; Kopczuk, Saez, and Song, 2010). The cap on FICA earnings at the maximum taxable earnings level does not affect our results: the AET exempt amount was far below the maximum taxable earnings level, so the FICA earnings data are capped at a level far above the top of the range included in our bandwidth.

results. The effects we estimate are relevant to policy, in the sense that they reflect the overall effect on employment.

The AET is applied to pre-tax Social Security benefits; this affects the results negligibly relative to measuring after-tax benefits, because benefits only became taxable in 1984, and after 1984 benefits were only taxable above an income threshold much higher than the AET exempt amount. By examining pre-tax benefits, we answer a policy-relevant question: how a given change in the AET BRR would affect employment. After-tax benefits are slightly smaller than pre-tax benefits—and the benefit reduction rate associated with after-tax benefits should change by slightly less at the exempt amount—suggesting that the effect of the pre-tax BRR should if anything reflect a lower bound on the effect of the after-tax BRR, and again strengthening our case that we will estimate a lower bound on the elasticity.

Due to institutional considerations and computational constraints, we focus on a specific set of cohorts observed over certain calendar years. We focus on a sample that is subject to the AET in 1978 and after. The reason is that we observe only calendar year earnings, and the AET has effectively applied to calendar year earnings beginning in 1978 (Gelber *et al.*, 2013, Appendix A). Before 1978, the AET was applied to quarterly earnings, which we do not observe in the data. This limits our focus to individuals born in 1918 or later, who turned age 60 in 1978 or later. Due to computational constraints, we were able to obtain data on individuals born from 1918 until 1923, and due to these constraints it was also necessary to draw a 25 percent random sample of this group. This leaves 3,629,183 sampled individuals; we included data on these individuals for all calendar years in the panel. We observe these individuals until they turn 64, which occurs as late as 1987 (in the 1923 birth cohort).

Next, from this initial sample, we drop individuals who claimed Social Security OASI or Disability Insurance prior to age 62—usually those claiming on a spouse’s record—so that our measure of earnings prior to 62 is not affected by potential contemporaneous impacts of the AET (19.5 percent of the sample); those with missing values for gender (0.049 percent of the remaining sample); those with positive self-employment earnings at age 60 (7.82 percent of the remaining sample), as these earnings are often subject to manipulation (*e.g.* Chetty *et al.*, 2013); and those with negative measured earnings at some point between ages 50 and 70 (0.035 percent of the remaining sample). Since the key running variable is earnings at age 60, we restrict attention to individuals with positive earnings at age 60 (47.05 percent of the remaining sample). We include all remaining individuals, including those who collect benefits as retired workers, auxiliary beneficiaries, or survivors. This leaves us with a sample of 1,424,667 individuals. We use data on individuals from ages 57 to 64, or eight calendar years per person, so that the total number of individual-year observations is 11,397,336. We do not limit the sample in any given year to individuals claiming in that year, since claiming OASI is potentially endogenous to the AET. Our data extract does not link spouses to one another.<sup>13</sup>

---

<sup>13</sup>A non-convex kink in the budget set occurs where the AET has completely phased out Social Security benefits, and therefore the BRR jumps from positive to zero. We are not able to investigate the effects of this non-convex kink in a reliable way. First, we do not directly observe an individual’s benefits in each year in our data extract. Second, the effect of the non-convex kink should be spread throughout the earnings distribution, because the non-convex kink occurs at a different earnings level for each individual. Thus, we would not expect any sharp change in aggregate behavior at any particular average earnings level due to this non-convex kink. It is also not possible to use individual-level data, because we are not able to link spouses. If a spouse is claiming on a primary beneficiary’s

Several features of the data merit discussion. First, these administrative data are subject to little measurement error. Second, earnings as measured in the dataset are not subject to manipulation through tax deductions, credits, or exemptions, and they are subject to third-party reporting (among the non-self-employed). Third, like most other administrative datasets, the data do not contain information on hours worked, hourly wage rates, amenities at individuals' jobs, underground earnings, assets, savings, or consumption. They also do not contain data on unearned income or marital status.

Table 1 shows summary statistics for our sample. The main sample shown is restricted to the 95,960 individuals who have positive age 60 earnings within \$2,797 of the exempt amount and who satisfy the other sample restrictions. This bandwidth, \$2,797, is the one we choose in our regressions for our main outcome using the procedure of Calonico *et al.* (2014). Our main outcome in our regressions is the mean annual employment rate among 63 to 64 year-olds, *i.e.* the percent of the corresponding calendar years when the individual has positive earnings; the mean level of this variable in the main sample is 56.53 percent.

The exempt amounts during our period, ranging from \$9,787 to \$11,517, are close to mean annual earnings, \$10,977, at ages 63 to 64 in the full sample of those with positive or zero age 60 earnings (not restricted to being close to the exempt amount); and close to median earnings, \$9,166, at ages 63 to 64 among those with positive age 60 earnings (not restricted to being close to the exempt amount). Thus, our estimates apply to a group near the exempt amount that is close to this mean and median in these broader populations; this is relevant for interpreting the regression results. Mean earnings (including zeroes) at ages 63 to 64 is \$5,814 in our main sample. The mean age of claiming Social Security benefits is 63.13. This sample is 67.99 percent female; the large female percentage is a consequence of focusing on individuals near the exempt amount at age 60. For comparison we also show the full sample of those with positive earnings at age 60, not restricted to those around the exempt amount but satisfying the other sample restrictions. Throughout the paper, dollar figures are expressed in real 2010 terms. We use the Consumer Price Index for All Urban Consumers (CPI-U) to adjust dollar figures across years (Bureau of Labor Statistics, 2011).

## 5 Empirical Results

### 5.1 Preliminary Analysis

We begin our analysis with initial validity checks of our empirical method, demonstrating that individuals do not appear to sort around the exempt amount when choosing their earnings at age 60. First, Figure 2 shows that the density of earnings at age 60 (solid circles) appears continuous around the exempt amount. The regressions in Table 2 confirm that the density of observations is smooth in the region of the exempt amount. Following Landais (2015), in these regressions we use the procedure of McCrary (2008) to estimate discontinuities in the level and slope of the density function at the exempt amount, allowing for linear, quadratic, or cubic functions of the running variable. To test further for sorting, we also test for bunching at

---

record, then the non-convex kink occurs at the earnings level at which the family's entire benefit is completely phased out; thus, we would need information on whether a spouse is claiming on the primary beneficiary's record to observe the non-convex kink accurately.

the exempt amount. Figure 2 shows that the density of earnings at age 60 appears smooth near the exempt amount, and that the amount of bunching, calculated using a method similar to Chetty *et al.* (2011) — and described in detail in Appendix C — is statistically insignificant.<sup>14</sup>

Similarly, Figure 3 shows that there is no clear visual discontinuity in the slope of predetermined covariates (sex, birth year, and probability of being white). In this figure and subsequent figures, the  $x$ -axis measures our running variable, earnings at age 60. The range of the  $x$ -axis is  $[-\$3,000, \$3,000]$ , which corresponds roughly to the Calonico *et al.* (2014) bandwidth of \$2,797. On the  $y$ -axis is the outcome in question, taking means within \$500 bins. The data are pooled across all years of the sample. Table 2 confirms that in the baseline linear specification, there is no significant discontinuity in the slope of the fraction white or female, although there is a small, significant discontinuity in the slope of the year of birth. Nonetheless, there is no significant discontinuity in year of birth in the cubic specification, in which we will still find a significant discontinuity in our main outcome, the age 63 to 64 annual employment rate. Moreover, the significance of the discontinuity in the slope of year of birth does not survive a Bonferroni correction that is relevant here since we have no theoretical presumption of a change in slope. Finally, controlling for these covariates, including year of birth, will have no material effect on the results. Note that although we show our data aggregated by bin means, our regressions here and elsewhere are run using individual-level data as in (4). All of this evidence of covariate smoothness shows that agents act as if they do not anticipate the imposition of the AET at age 60 (and consistent with the evidence in Gelber *et al.*, 2020, and Gelber *et al.*, 2016, of no reaction to Social Security eligibility, and the AET in particular, until closer to when individuals are affected by it).

## 5.2 Main Results Documenting a Kink in Employment

Having passed these validity tests, we proceed to our main results. In Figure 4 we see a sharp decrease at the exempt amount in the slope of the annual employment rate at ages 63 to 64: under the exempt amount the slope appears positive and steep, whereas above the exempt amount the slope appears to change immediately to nearly flat. Appendix Figure E.5 shows that this visual pattern—a sharp decrease in slope at the exempt amount—is also clear with a wider bandwidth of \$6,000. Table 3, Column 1, reports the estimated coefficient  $\beta_1$  from (4), confirming that the change in the slope is statistically significant and substantial in the baseline linear specification that uses a Calonico *et al.* (2014) bandwidth of \$2,797.

Even without estimating an elasticity, these results constitute a new source of evidence that the AET causes an extensive margin effect. In the context of our framework in Section 2.2, evidence of a kink in the employment rate also implies that some individuals face frictions in adjusting earnings on the intensive margin, which is a prerequisite to carrying out our elasticity estimates below. Our estimates represent the effects of the AET while holding other factors constant. We do not interpret the discontinuity in the slope of

<sup>14</sup>This is in contrast to age 62 where, as shown in Gelber *et al.* (2020), individuals in this sample do begin to bunch at the exempt amount. Bunching at age 62, however, is not a threat to our identification strategy, as we only require the absence of sorting at age 60. Moreover, as shown in Gelber *et al.* (2020), bunching at age 62 is not sufficient to rule out the presence of intensive margin frictions among a subset of the population. They use variation in bunching over time to show evidence of economically significant frictions.

earnings at the exempt amount as reflecting changes in demand by firms, since such demand changes should have been materially similar on either side of the exempt amount, as should any general equilibrium effects of the policy change more broadly. Like other papers based on local variation, including others in the AET literature, our identification strategy does not attempt to address general equilibrium impacts of the AET. We do, however, interpret our measured responses as potentially including the effects of misperception of substitution incentives or other frictions that could affect adjustment.

While the most salient pattern in Figure 4 is the difference in slopes to the left or right of the exempt amount, another subtle feature is that the data are noisier to the right. The probability of positive earnings to the right of the exempt amount has a much more muted slope, but also more variation above or below its regression line. In particular, the mean employment for our bin just to the right of the exempt amount is lower than just to the left. While our estimates of the change in slope are not sensitive to whether or not we allow for a level shift at the exempt amount in our regression kink specification, this potential drop could itself be of interest. Our theoretical model only predicts a change in slope, but if there is indeed a shift in the level of employment, it could mean that older workers perceive or experience some discontinuous drop in payoffs on the right side, *i.e.* a notch. For now, we proceed with a focus on the change in slope, but return to a discussion of this potential misperception in Section 7.

### 5.3 Robustness and Validity Checks

Our baseline results pass several robustness and validity checks. First, our method suggests a natural “placebo” test in the period prior to being subject to the AET. If the kink in employment we observe during ages 63 to 64 were driven by nonlinear patterns in unobservables for those whose age 60 earnings are near the exempt amount, and not the AET, we might also expect to detect these nonlinearities at ages 56 and 57. Figure 5, Panel A shows no clear visible change in the slope of employment at ages 56 and 57, as a function of age 60 earnings. Like ages 63 and 64, ages 57 and 56 are respectively three and four years away from age 60, the age at which the running variable is measured. In Table 3, Column 4 we present the estimates from a corresponding regression in which the dependent variable is the annual employment rate at ages 56 to 57. We maintain the same running variable as earlier, *i.e.* age 60 earnings, since we would like to test for a spurious relationship between earnings at age 60 and lagged employment. Across all three polynomial specifications, mean annual employment at ages 56 to 57 shows an insignificant change in slope.

Furthermore, in Figure 5, Panel B and Table 4 we show that the kink arises precisely at ages 63 and 64—not earlier—which we consider our single most convincing evidence that we have uncovered a true effect of the AET. Figure 5, Panel B shows the employment probability by single year of age from 61 to 64. There is no clear kink at ages 61 and 62, and a visible kink only arises at ages 63 and 64. Recall that no kink should appear at age 62 because the AET only applies for part of the year. Table 4 confirms that the kink estimates are insignificant at ages 61 and 62, but they are substantial and highly significant at ages 63 and 64. We are particularly reassured by the small and insignificant kink at ages 61 and 62. This placebo check



provides further evidence that nonlinearities in the probability of employment (conditional on earnings at age 60) do not drive our kink estimates. The absolute value of the kink grows substantially from age 63 to age 64, from 1.60 to 1.98, only a modest fraction of which can be explained by the higher fraction claiming at 64 (79.5 percent) than at 63 (74.5 percent). This is consistent with gradual adjustment to the AET, which Gelber *et al.* (2020) document at the intensive margin.

Our estimates remain significant in Table 3 under the quadratic and cubic specifications, with moderately larger estimates. The results are also robust to controls for demographics in Column 2 (dummies for sex, year of birth, and race groups). In Figure 6 we demonstrate that the size of the estimated kinks remains relatively stable across different bandwidths. As another assessment of the validity of our approach, we conduct a permutation test in the spirit of Ganong and Jäger (2015). In particular, we estimate a set of placebo changes in slope in the age 63 to 64 employment rate, using the same specification as our main estimates, except that we examine the change in slope at placebo locations of the exempt amount away from the true exempt amount. In Figure 7, we show that the point estimate recovered from the actual exempt amount is located well below the distribution of placebo estimates, reinforcing the view that we are detecting a true effect of the AET. This permutation test shows a significant kink ( $p = 0.025$ ).<sup>15</sup> Finally, following Landais (2015), we show in Online Appendix Figure E.2 that the R-squared of the regression in the baseline linear specification is maximized at the actual location of the kink, rather than at placebo kink locations, again consistent with our finding a true effect of the AET.<sup>16</sup>

## 5.4 Estimating an Extensive Margin Elasticity

Our fuzzy RKD estimator in equation 3 allows us to recover the marginal effect of ANTR on employment. This in turn can be rescaled into the policy-relevant employment elasticity with respect to the average-net-of-tax rate (ANTR):  $\eta \equiv \partial \Pr(z > 0) / \partial ANTR \times (1 - a) / \Pr(z > 0)$ . This parameter is useful, for example, in calculations of deadweight loss and optimal income tax schedules. We use a net-of-tax measure that incorporates the AET BRR as well as the average federal and state income and payroll tax rates. These rates are calculated using the TAXSIM calculator of the National Bureau of Economic Research and information on individuals within the bandwidth distance from the kink in the Statistics of Income data in the years we examine. The baseline probability of working,  $\Pr(z > 0)$  is measured by the realized probability of working during ages 63 and 64 for those with age 60 earnings just below the exempt amount. Table 5 shows that the elasticity estimated using the full sample in the baseline linear specification is 0.49. As discussed above, we interpret this as a lower bound on the elasticity. Since the lower bound is large, our

<sup>15</sup>Relative to our baseline specification, we also find similar, statistically significant results when we run our regressions using data at the \$50 bin mean level (or within bins of other sizes). This is an alternative way of demonstrating that the results are significant and robust to filtering out within-bin correlations. However, these results are superseded by the permutation test results as the latter demonstrate that the kink is statistically unusual relative to the distribution of placebo estimates.

<sup>16</sup>Recall that our running variable measures earnings at age 60 relative to the exempt amount applying to OASI claimants at that time, i.e. those 62 and older. Since the exempt amount is rising modestly over time (Figure E.1), the exempt amount individuals will actually face during ages 63 and 64 is slightly higher than the prevailing exempt amount when they are age 60. However, the exempt amount rises by only \$75.77 on average in these cases, far smaller than the bin size of \$500 that we use in our figures. Thus, our statistical and graphical results are relatively unchanged if we instead measure our running variable as age 60 earnings relative to the exempt that individuals will face at ages 63 to 64.

main conclusion is that the elasticity is large.

Separating the regressions by demographic group, the elasticity for women is 0.49, and the elasticity for men is 0.25, consistent with the typical stylized fact that women have larger employment elasticities than men.<sup>17</sup> One interpretation of this result is that women’s employment is more responsive to these incentives, which is consistent with an overview of the literature by Chetty et al. (2012), which finds that (Hicksian) extensive margin elasticities tend to be larger when estimated among women. On the other hand, as discussed in Section 2.2 and fleshed in our more detail below in Section 6 and Online Appendix Section B.2.1 (see equation (B.21)), the observed extensive margin elasticity is likely to increase in the share of agents facing intensive margin frictions. It could be the case that these frictions are more prominent among women, which is consistent with the fact that men appear to bunch more at the exempt amount than women (Gelber et al., 2013). In Table 5, we do not detect as much heterogeneity in other dimensions, e.g. race or mean of prior earnings.

A variety of considerations indicate that these estimated elasticities are robust; that if anything they reflect lower bounds; and that these large elasticities make sense given that we are studying older workers, who have marginal attachment to employment in many cases. Appendix Table E.1 shows that the elasticity is modestly larger in the quadratic or cubic specifications, reaching 0.66 in the quadratic and 0.82 in the cubic. The results are similar with or without controls. In the baseline first stage, we calculated the exact change in slope of the ANTR, under the assumption that desired earnings are fixed at their age 60 level. To show that the results are robust to instead running a first stage regression analogous to the reduced form regression, in Appendix Table E.2 we show that the elasticities are similar when we assume that desired earnings are fixed at their age 60 level but estimate the change in the slope of the ANTR using a regression whose bandwidth is chosen using Calonico *et al.* (2014).

The elasticity calculations are mechanically affected by how we specify the first stage. Given that the stringency of the AET, and therefore the effective ANTR, is potentially affected by benefit enhancement, we consider our elasticity estimates to be illustrative, as the scaling of the first stage could be affected by accounting for benefit enhancement. However, it is also important to note that if we took benefit enhancement into account, the first stage absolute value of the change in the ANTR at the exempt amount would be smaller, and thus the estimated elasticity would be larger. As a result, if anything we would estimate larger elasticities, strengthening our conclusion that the lower bound on the elasticity is large. Moreover, the kink estimates in Table 3 and Table 4 are unaffected by the scaling, and our counterfactual estimates of the effect of policy changes on employment would be unaffected by the scaling of the first stage. Finally, if individuals (mis-)perceive the AET as a pure tax (without taking benefit enhancement into account), then we have specified the first stage in a way that accurately captures their perceptions.

The relatively large elasticities we estimate are within the range of elasticities estimated using microdata

---

<sup>17</sup>Our study of women and men differs from the pre-2003 empirical literature on the employment effects of the AET cited above, as well as Baker and Benjamin (1999) and Haider and Loughran (2008), all of which focus on studying only men. Gruber and Orszag (2003) find suggestive evidence of an impact among women but not among men.

for groups with low employment attachment, though they fall outside the range of some previous surveys of extensive margin elasticities for the full population estimated using microdata. Chetty *et al.* (2012) report Hicksian and Frisch (*i.e.* constant-marginal-utility-of-wealth) extensive margin elasticities of 0.25, in a meta-analytic synthesis of microdata estimates not specific to particular population groups.<sup>18</sup> Saez (2002) discusses that the Negative Income Tax (NIT) experiments find small extensive margin elasticities for males of about 0.2, but find participation elasticities for those groups less attached to the labor force—*e.g.* second earners, single household heads, and the young—between 0.5 and 1. Our estimates among older workers at or near retirement age are most comparable to estimates for groups with low employment attachment. Indeed, a small group of theoretical and empirical studies has found relatively large labor supply or earnings responses among the elderly in particular (French, 2005; Laitner and Silverman, 2012; Alpert and Powell, 2014), and we complement those studies by using a new method and applying it to administrative data.

## 5.5 Accounting for Measurement Error in the Running Variable

To support the validity of using age 60 earnings as a proxy for age 63 to 64 desired earnings, we show that desired earnings remain stable across a “placebo” set of ages. Specifically, we show that the distribution of real earnings growth from one period to a subsequent period exhibits a spike at zero, as postulated in (5). Figure 8, Panel A shows that from age 59 to age 60—a placebo set of ages near the EEA during which our sample is not subject to the AET—a noticeable spike in real earnings growth does occur near zero percent growth, as workers’ wages grow at about the rate of CPI in a substantial fraction of cases. To further probe this assumption, we perform another test. In Figure 8, Panel B, we plot age 60 earnings against age 59 earnings. In the case where there is a substantial share of workers with zero real earnings growth, we would expect to see a mass located along the 45 degree line. We indeed find this pattern, further supporting our assumption regarding earnings growth.<sup>19</sup>

Although the figures suggest significant persistence in earnings, they also show that many people exhibit non-zero earnings changes, and our estimates so far do not account for this fact. Specifically, non-zero earnings growth acts as a kind of measurement error, which attenuates the “first stage” relationship between the running variable and the endogenous regressor: the ANTR at desired, age 63 earnings. This attenuation in turn implies that our baseline elasticity of 0.49 is too low.<sup>20</sup>

In this section, we present a range of estimates that account for measurement error in the form of non-zero earnings changes. The essence of the approach is to calibrate an earnings growth process, obtain a correct first stage that accounts for this earnings growth, and then use this first stage to appropriately rescale our

<sup>18</sup>In a survey of macroeconomic literature, Chetty *et al.* (2012) find average Hicksian elasticities of 0.17 and average Frisch elasticities of 2.77.

<sup>19</sup>We reproduce this plot in Appendix Figure E.3, omitting the 45 degree line, which allows the reader to further assess the evidence. We note that both figures are constructed using the publicly available Earnings Public Use File, which contains annual earnings histories for a 1 percent sample of Social Security Numbers. The earnings amounts are randomly rounded to the nearest \$100, which means there is more year-to-year variability in this figure than in our estimation sample.

<sup>20</sup>The reduced form, on the other hand, is not attenuated because we obtain the reduced form from a regression of positive earnings at age 63-64 on the mismeasured regressor. The first stage is biased, however, because in our baseline estimates we use age 60 earnings to calculate the age 63 ANTR. If age 60 earnings are an imperfect proxy for age 63 desired earnings, this procedure overstates the first stage kink.

reduced form estimate. We provide a detailed description of the approach in Appendix D. The appendix also shows that this approach yields approximately unbiased estimates. We perform inference in the presence of this measurement error using either (1) an approximation to the delta method, or (2) the simulated distribution of estimates under an imposed null hypothesis (also described in more detail in Appendix D). There are other approaches to addressing measurement error in related models, such as Pei and Shen (2017). However, their approach would be difficult to adapt to our setting, because it requires researchers to observe a correctly measured binary treatment variable, but our endogenous regressor is continuous and unobserved.

We consider three ways of modelling the earnings growth process. All models use the growth distribution at ages 59-60 to predict desired earnings changes from age 60 to age 63, and hence desired earnings at 63. We focus on these ages because later ages are potentially subject to the earnings test, and earlier ages (before retirement nears) may involve very different earnings processes. In our first model, we assume that earnings growth rates are perfectly correlated from one year to the next, and we draw one-year growth rates from the distribution shown in Figure 8, Panel A. Drawing from this figure lets us preserve the evident mass point at zero. As a complementary second model, we instead use independent draws from the growth rate distribution. Our third model strikes a middle ground between perfect correlation and perfect independence. Specifically, it assumes that earnings growth rates are normally distributed and conditionally independent given sex and permanent income quartile.

We report the estimated elasticities under these different approaches in Appendix Table E.4. The elasticities, which are statistically significant, range from 2.2 to 2.9, and thus are not terribly sensitive to the assumed earnings growth process. Adjusting for measurement error does, however, imply elasticities that are substantially higher than our baseline estimates. The elasticities are larger than in the baseline case because the first stage kink implied by the earnings growth process is about -1, rather than about -5 in the baseline case.

## 5.6 Accounting for Claiming Behavior

Among those with age 60 earnings within our baseline bandwidth relative to the exempt amount, 75 and 80 percent of the sample has claimed by ages 63 and 64, respectively. The fact that some have not claimed Social Security benefits affects the interpretation of the elasticity we estimate. In particular, the elasticity we have estimated so far should be interpreted as the elasticity *inclusive* of the impact of the fact that some individuals have not yet claimed. We estimate this elasticity in the baseline because this is policy-relevant, in the sense that it informs us how a given change in policy would affect employment in this group, given the effects of the percentage of our sample that claims.

At the same time, an alternative elasticity concept is also of interest: the “conditional” elasticity among those in the sample claiming Social Security (*i.e.* we “condition” on claiming). In estimating this conditional elasticity, the smaller the fraction that claims by 63 or 64, the smaller the implied first stage will be, and thus the larger the conditional elasticity estimate will be. Since we are estimating a lower bound on the

unconditional elasticity above, perhaps the simplest way to address estimation of the conditional elasticity—given endogenous claiming behavior—is to note that if anything the estimated conditional elasticity would be still larger than the unconditional elasticity, if the first stage were attenuated due to the fact that some individuals have not claimed by 63 or 64.

Beyond noting this, we can divide the unconditional elasticity by the share claiming, as shown in Appendix B.5, to calculate a conditional elasticity. Here we can use the claiming behavior of those with  $z_i^{60}$  just below the kink to impute the percentage claiming among those at the kink, *i.e.* those to whom the elasticity estimates apply. Using the percent claiming at ages 63 and 64, this method implies that to estimate the conditional elasticity, we would inflate the elasticity estimates by 29.9 percent relative to those shown in Table 5.<sup>21</sup> Thus, for example, the baseline elasticity would be 0.63 rather than 0.49. Appendix Table E.3 shows the elasticities and standard errors calculated in this way.

In principle it would also be possible to adjust the conditional elasticity estimates further to account for the fact that the fraction that claims may be directly affected by the AET itself, as in Online Appendix B.5. However, Appendix Figure E.6 shows that the probability of claiming shows no clear change in slope at the exempt amount. Corresponding regressions show no robust effect of the AET on claiming at these ages: a placebo test parallel to our previous test in the spirit of Ganong and Jäger (2015) results in a  $p = 0.15$  for the two-sided test of equality of the coefficient with zero.<sup>22</sup> Because there is no evidence for an effect on claiming, there is little case for implementing our most flexible adjustment for endogenous claiming in Online Appendix B.5.3.

## 5.7 Counterfactual Simulations

We can use our elasticity estimates to understand the effect of changing the parameters of the AET, under the assumption that our extensive margin elasticity estimates apply throughout the earnings range shown in our graphs, from age 60 earnings \$3,000 under the exempt amount to \$3,000 over. One key issue for policy-makers is the impact of eliminating the AET (Tergesen, 2016). We simulate the increase in the employment rate among 63 and 64 year-olds that would result from eliminating the AET entirely, using our baseline elasticity estimate of 0.49 in Table 5. As before we assume that at ages 63 and 64, desired earnings conditional on employment in the absence of the AET is given by earnings at age 60. To calculate the baseline probability of participation at ages 63 to 64 at different earnings levels, we bin the earnings distribution into \$500 bins as in our figures and calculate this probability,  $\Pr(PE_{0i})$ . For each age 60 earnings bin, we calculate the mean percent change in the average net-of-tax rate associated with moving from the AET’s current 50 percent BRR to the ANTR associated no AET (but keeping other taxes in place),  $(ANTR_{1i} - ANTR_{0i})/ANTR_{0i}$ . We then use our estimated elasticity  $\hat{\eta}$  to calculate the predicted change in the employment rate if the AET were eliminated,  $\hat{\eta}[(ANTR_{1i} - ANTR_{0i})/ANTR_{0i}] \Pr(PE_{0i})$ , and we aggregate across bins (weighting by the fraction of the population in each bin).

<sup>21</sup>We calculate 29.9 percent as  $100 * (1/[(0.795 + 0.745)/2] - 1)$ .

<sup>22</sup>Our evidence is not incompatible with Gruber and Orszag’s (2003) evidence on claiming: our results pertain to 63-64 year olds, whereas Gruber and Orszag (2003) find an effect of the AET on claiming among older groups.

This calculation shows that eliminating the AET would increase the employment rate by 1.4 percentage points in the group with age 60 earnings within \$3,000 of the exempt amount, or a 2.5 percent increase. We consider this a lower bound for the reasons described earlier. This counterfactual exercise illustrates that the observed elasticity can be useful for counterfactual predictions about employment levels under alternative AET parameters.

We can also use our elasticity estimates to calculate the change in earnings due to the AET, as well as the fraction of the change that is due to extensive margin responses. We use our baseline extensive margin elasticity estimate of 0.49. To calculate the intensive margin response, we assume quasilinearity as in much recent literature on earnings responses to kinked budget sets (*e.g.* Saez, 2010; Chetty *et al.*, 2011; Kleven and Waseem, 2013; Kleven *et al.* 2014). For simplicity we assume that the intensive margin elasticity with respect to the marginal net-of-tax share is a constant 0.35, and that the intensive margin adjustment cost is \$278, corresponding to the Gelber *et al.* (2020) baseline point estimates. We calculate the increase in earnings associated with the elimination of the AET through each channel, intensive and extensive. For an individual  $i$ , the estimated total earnings response  $\Delta E_i$  is:

$$\begin{aligned}\Delta E_i &= \Pr(PE_{no,i})z_{no,i} - \Pr(PE_{AET,i})z_{AET,i} \\ &= \Pr(PE_{AET,i})(z_{no,i} - z_{AET,i}) + [\Pr(PE_{no,i}) - \Pr(PE_{AET,i})]z_{no,i}\end{aligned}\quad (6)$$

where  $\Pr(PE_{AET})$  and  $z_{AET}$  are the probability of positive earnings under the AET and earnings under the AET, respectively, and  $\Pr(PE_{no})$  and  $z_{no}$  refer to the levels of these variables once the AET has been eliminated. The second line shows that the total response can be decomposed into an intensive margin response (the first term) and an extensive margin response (the second term).

Our results show that in the group within \$3,000 of the exempt amount, mean annual earnings at ages 63-64 (including zeroes) increase by 9.8 percent (\$579) due to eliminating the AET, and 27.6 percent of the increase in earnings is due to the extensive margin effect on the employment rate. Focusing only on the intensive margin impact would overlook over one quarter of the total earnings response.

Our specific predictions are influenced by multiple factors: the elasticity is sizeable; the change in the BRR due to eliminating the AET is large; but for those near the AET exempt amount that we study, the change in the ANTR is relatively small. The elasticity could also be different at other points in the earnings distribution, and moreover some individuals may tend to be unresponsive (*e.g.* those with zero earnings at age 60) while others might be more responsive (*e.g.* those with age 60 earnings farther above the exempt amount who experience a larger change in the ANTR, but not so far that the effect of the AET on the ANTR greatly diminishes). Ultimately, our calculations do demonstrate that eliminating the AET can cause large increases in earnings and employment, even using our lower bound point estimates.

## 6 Interpreting the Empirical Estimates

In Section 5.4 we convert the average treatment effect estimated with an RKD into an employment elasticity. In this section we provide an underlying model of earnings and employment that clarifies the interpretation of this elasticity, the assumptions under which the parameter has a structural meaning, and for which subpopulation is the parameter identified. A key result is that the RKD estimator we implement in Section 3 naturally emerges from the model. Thus, our model provides a new microfoundation for a widely used method, the RKD, in a labor supply setting.

Building on the framework in Section 2.2, the budget set faced by individual with ability  $n$  in state  $j$  is now:

$$c_{nj} = z_{nj} - T(z_{nj}) + B_j(z_{nj}), \quad (7)$$

where  $B_j(\cdot)$  is defined as above in equation 1. As is empirically relevant to our setting, we assume the tax function does not vary based on the presence of the AET and has a constant slope of  $T'(z) \equiv \tau_0$  in the neighborhood of the AET exempt amount. As before, individuals maximize utility  $u(c, z; n)$ , where  $u(\cdot)$  is a function of class  $C^2$ . We do not make assumptions constraining the nature of income effects, *e.g.* we do not assume that utility is quasi-linear, that leisure is a normal good. Our index of “ability” is  $n$ ; the marginal rate of substitution of  $c$  for  $z$  is decreasing in  $n$  at all levels of  $c$  and  $z$ .<sup>23</sup>

Given this setup, we briefly review the intensive margin effect of a kink in the budget set created by the AET. As shown in Saez (2010), a kinked budget set leads to a discontinuity in the earnings density at the kink due to intensive margin responses. Assuming a smooth distribution of ability  $n$ , a range of individuals who would earn between  $z^{AET}$  and  $z^{AET} + \Delta z$  in state 0 will respond in state 1 by reducing earnings to the kink at  $z^{AET}$ . This is referred to as “bunching” at the kink. The reduction in earnings  $\Delta z$  of the “marginal buncher”—*i.e.* the buncher who earns  $z^{AET} + \Delta z$  in state 0—can be related to the size of the change in the effective marginal tax rate at the kink,  $db$  and used to estimate an intensive margin elasticity (Saez, 2010), although Blomquist et al. (2019) show that the elasticity is not identified without either placing restrictions on unobserved heterogeneity or leveraging variation in budget sets. In prior work (Gelber et al., 2020) we extend the Saez (2010) method to allow for optimization frictions and, by using variation in budget sets, apply the method to the AET and estimate an intensive margin elasticity. We estimate positive intensive margin elasticities and economically significant frictions on intensive margin adjustment, which will be relevant for our theoretical results below.

### 6.1 Extensive Margin Responses

In addition to the intensive margin response that many recent studies have focused on, individuals may also respond at the extensive margin. In the many models of earnings supply, such as that of Saez (2010), preferences and budget sets are convex, which restricts a small tax change to only affecting the choice

<sup>23</sup>This implies a standard single-crossing property assumed in these models, which generates rank preservation in earnings, conditional on earning a positive amount.

between zero and infinitesimally small earnings supply (*e.g.* Kleven and Kreiner, 2005). To capture the realistic pattern of potential entry to or exit from non-trivial levels of earnings, we introduce a fixed cost of employment (Cogan, 1981; Eissa *et al.*, 2008). Utility conditional on working is now given by:

$$u(c_{nj}, z_{nj}; n) = v(c_{nj}, z_{nj}; n) - q_{nj} \cdot 1\{z_{nj} > 0\} \quad (8)$$

where  $j \in \{0, 1\}$  indexes the state of the world, and the state-specific, additively separable fixed cost of employment,  $q_{nj}$ , is drawn from a distribution with CDF  $G(q|n)$  and pdf  $g(q|n)$ . If an agent does not work, she receives a reservation level of utility of  $u(c^0, 0; n) = v^0$  in either state of the world.<sup>24,25</sup>

We pay special attention to whether or not the individual locates at a corner, *i.e.*  $z_{nj} = 0$ . Let  $\tilde{z}_{nj}$  denote the optimal level of earnings in state  $j$  conditional on working (corresponding to  $\tilde{z}_{0,it}$  and  $\tilde{z}_{1,it}$  above in Section 3). This is chosen by maximizing  $u(c, z; n)$  subject to (7). The individual therefore works in state  $j$  if:

$$v(\tilde{z}_{nj} - T(\tilde{z}_{nj}) + B_j(\tilde{z}_{nj}), \tilde{z}_{nj}; n) - q_{nj} > v^0 \quad (9)$$

This is demonstrated in Figure 9. We illustrate the indifference curves governing the extensive margin decision under the alternative tax and benefit schedules. We model the fixed cost of working visually by allowing agents to choose a level of earnings along the prevailing tax schedule, or to earn zero and receive a level of consumption equal to benefits  $b$  and an additional term  $q$  that is functionally equivalent to the fixed cost of entry employment.<sup>26</sup> In Panel A the agent's optimal level of earnings conditional on working is  $z^{AET}$ . In this case she prefers earning  $z^{AET}$  to earning zero. Her response to the kink is simply a reduction in earnings. In Panel B the agent similarly has optimal earnings of  $z^{AET}$  conditional on having positive earnings. In this case, however, the individual's preferences lead her to earn zero rather than earning at the kink, in part driven by a larger value of  $q$ . Next we formally explore such responses.

Our behavioral response of interest is the extensive margin response to the presence of a kink. Here we define an individual's type by their optimal interior earnings conditional on working in state 0, *i.e.*  $\tilde{z}_{n0}$ . An isomorphism exists between this earnings amount and ability  $n$ , and for empirical purposes using an earnings amount is natural to implement. The probability of working in state  $j$  conditional on type  $\tilde{z}_{n0}$  is:

$$\Pr(z_{nj} > 0 | \tilde{z}_{n0}) = \Pr(q_{nj} \leq v(\tilde{z}_{nj} - T(\tilde{z}_{nj}) + B_j(\tilde{z}_{nj}), \tilde{z}_{nj}; n) - v^0 | \tilde{z}_{n0}) = G(\bar{q}_{nj} | n) \quad (10)$$

where:

$$\bar{q}_{nj} \equiv v(\tilde{z}_{nj} - T(\tilde{z}_{nj}) + B_j(\tilde{z}_{nj}), \tilde{z}_{nj}; n) - v^0 \quad (11)$$

is the critical value for the fixed cost of employment that leaves the agent indifferent between working

<sup>24</sup>Without loss of generality, the outside option,  $v^0$ , does not vary with  $n$  or  $j$ . This is because cross-sectional and state-specific variation in the outside option is not separately identified from the fixed cost of entry  $q_{nj}$ . We therefore collapse all such variation into the fixed cost of entry.

<sup>25</sup>Writing the fixed cost  $q$  as separable from  $v$ , as in (8) above, simplifies the exposition. Without loss of generality, this model is equivalent to a model in which these are not separable *per se*, and instead we express utility simply as  $v(c, z; n)$ . Letting  $c_n^0$  be consumption when not working, we can then posit a discontinuity in  $v(c, z; n)$  at the boundary of the support of  $z$  that reflects the fixed cost. Thus, we can define a fixed cost  $q_n$  as:  $q_n \equiv \lim_{z \rightarrow 0^+} [v(c_n^0, 0; n) - v(c_n^0, z; n)]$ .

<sup>26</sup>Formally, the parameter  $q$  in this figure is still defined as  $q \equiv \lim_{z \rightarrow 0^+} [v(c_n^0, 0) - v(c_n^0, z)]$ .



and not working.<sup>27</sup> We allow the  $G(\cdot)$  function to vary across individuals so that we have two sources of heterogeneity: (1) preferences captured by the  $v(\cdot)$  function pin down intensive margin heterogeneity but also affect the extensive margin through  $\bar{q}_{n1}$ , and (2) the unrestricted heterogeneity in the  $G(\cdot)$  function allows for differences in extensive margin responses independent of the  $v(\cdot)$  function. We make four key assumptions regarding smoothness in heterogeneity.

**Assumption 1 (Smoothness)** *The fixed cost distribution  $G(\cdot)$  and the ability parameter  $n$  have the following properties:*

1.  $G(q_{nj}|n)$  is continuous
2. The partial derivative of  $G(q_{nj}|n)$  with respect to  $q_{nj}$ ,  $g(q_{nj}|n)$ , is continuous in  $q_{nj}$  and  $n$
3. The partial derivative of  $G(q_{nj}|n)$  with respect to  $n$ ,  $\partial G(q_{nj}|n)/\partial n$ , is continuous in  $q_{nj}$  and  $n$
4. The CDF of  $n$  is continuously differentiable

Note, the smoothness of the ability and fixed costs distributions imply the smoothness assumptions required for implementation of the RKD estimator in Section 3.<sup>28</sup>

## 6.2 Employment Probability with a Kink and Unconstrained Intensive Margin Responses

We first consider the standard context in which individuals are free to adjust their earnings anywhere on the intensive margin. In other words, individuals' earnings, conditional on having positive earnings  $\tilde{z}_{nj}$ , may differ across the two tax schedules, and earnings choices are subject to no constraints other than the budget constraint  $c = z - T(z) + B(z)$ . Let the employment function in state 1, conditional on counterfactual, interior earnings in state 0, be  $\Pr(z_{n1} > 0 | \tilde{z}_{n0})$ . This is the probability of having zero earnings in state 1 as a function of the level of earnings in state 0. We have shown that  $\Pr(z_{n1} > 0 | \tilde{z}_{n0}) = G(\bar{q}_{n1}|n)$ . We now explore how this function changes as  $\tilde{z}_{n0}$  changes. In Appendix A, we prove the following result:

**Proposition 1** *Under Assumption 1, if the individuals can freely adjust their earnings on the intensive margin, then the employment probability, as a function of earnings in state 0, will exhibit no change in slope at  $z^{AET}$ . That is:*

$$\beta \equiv \lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}} - \lim_{\tilde{z}_{n0} \rightarrow z^{AET-}} \frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}} = 0. \quad (12)$$

Figure 1 illustrates this result. The  $x$ -axes measure counterfactual earnings in state 0 conditional on having positive earnings, *i.e.*  $\tilde{z}_{n0}$ . In Panel B, the  $y$ -axis plots an illustrative employment rate. The dashed line represents a presumed smooth relationship between the employment rate in state 0 under a linear tax

<sup>27</sup>This probability is analogous to  $\Pr(z_{it} > 0)$  in Section 3.

<sup>28</sup>Blomquist et al. (2019) show that without placing restrictions on unobserved heterogeneity, the bunching method is not able to identify a taxable income elasticity. Likewise, a completely unrestricted distribution of fixed costs,  $G(\cdot)$ , would confound identification in our setting, if, for example, that distribution also exhibited a kink in the ability parameter  $n$ .

schedule, *i.e.*  $\Pr(z_{n0} > 0 | \tilde{z}_{n0})$ , and earnings conditional on having positive earnings, *i.e.*  $\tilde{z}_{n0}$ . The dotted line plots the employment rate in state 1 under a kinked tax schedule,  $\Pr(z_{n1} > 0 | \tilde{z}_{n0})$ , while the  $x$ -axis continues to plot  $\tilde{z}_{n0}$ . In this case, we assume that individuals are unrestricted in their earnings choices. We see that the employment function is unchanged at counterfactual earnings levels below  $z^{AET}$ , as the tax schedule remains the same in the two states. Above  $z^{AET}$  we see a gradual decrease in the probability of positive earnings in state 1 relative to state 0, due to the decrease in the ANTR (recall Figure 1 Panel A). Nonetheless, the kink in the ANTR does not translate into a kink in the employment rate. Intuitively, the ability to adjust on the intensive margin smoothes the changes in the slope of the ANTR at  $z^{AET}$ .

### 6.3 Employment Probability with a Kink and Constrained Intensive Margin Responses

Although the kink in the budget set does not lead to a kink in employment when individuals are free to adjust to any earnings level, a kink in the employment rate arises when frictions impede intensive margin adjustment. To illustrate the idea as simply and transparently as possible, we begin with the case that individuals are completely restricted from earning other amounts at the intensive margin, so that  $\tilde{z}_{n1} \equiv \tilde{z}_{n0}$ . Individuals are still allowed to vary their extensive margin choices across the two states. Numerous papers have found evidence for such restrictions on labor supply or earnings, for example due to constraints on hours or earnings choices, or fixed costs of adjustment that would prevent adjustment to the kink for those in this region (*e.g.* Altonji and Paxson, 1988; Dickens and Lundberg, 1993; Chetty *et al.*, 2007; Chetty *et al.*, 2011; Gelber *et al.*, 2020). Modeling and estimating frictions that could give rise to such restrictions is the focus of other work (Gelber *et al.* 2020); for the purpose of this paper it is not necessary to take a stand on what specific process gives rise to such restrictions, as the existence of such restrictions is sufficient to generate our results. In Appendix A, we prove the following result:

**Proposition 2** *Under Assumption 1, if individuals are not able to adjust earnings on the intensive margin, *i.e.*  $\tilde{z}_{n1} \equiv \tilde{z}_{n0}$ , then the employment probability, as a function of desired earnings conditional on employment in state 0, will exhibit a change in slope (*i.e.* a kink) at  $z^*$ . This kink is given by:*

$$\beta \equiv \lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}} - \lim_{\tilde{z}_{n0} \rightarrow z^{AET-}} \frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}} = -db \cdot \lambda_{n^{AET}} \cdot g(\bar{q}_{n^{AET}1} | n^{AET}) \quad (13)$$

where  $\lambda_n \equiv v_c$  is the marginal utility of consumption, and  $\bar{q}_{n^{AET}1}$ ,  $n^{AET}$ , and  $\lambda_{n^{AET}}$  all refer to the individual for whom  $\tilde{z}_{n0} = z^{AET}$ .

Returning to Figure 1 Panel B, the solid line depicts the relationship between counterfactual earnings in state 0 and the probability of positive earnings when a kink is present in state 1 and  $\tilde{z}_{n1} \equiv \tilde{z}_{n0}$ . The slope of the employment rate now discontinuously changes at  $z^{AET}$ , where the ANTR also changes slope. We have closed one of the channels through which individuals respond to the increase in tax liability, and thus the slope of the employment rate decreases discontinuously at  $z^*$ . Equation (13) has an intuitive interpretation:

the kink in the employment rate is proportional to  $db$ , the size of the kink in the tax schedule;  $\lambda_{n^{AET}}$ , the marginal utility of after-tax income; and  $g(\bar{q}_{n^{AET}1} | n^{AET})$ , the density of workers who are on the margin of entering employment in state 0. These parameters apply to the individual earning  $z^{AET}$  in state 0. Because the kink in the employment rate we model is only detectable in the presence of frictions in intensive margin adjustment, our method therefore also provides an incidental test of whether intensive margin frictions exist.

In Appendix B.1, we extend the model to allow for an arbitrary set of discrete earnings choices (other than  $z^{AET}$  for the constrained types), rather than the parsimonious and transparent—but more restrictive—assumption of this section that  $\tilde{z}_{n1} \equiv \tilde{z}_{n0}$ .<sup>29</sup> We show that there is a discontinuity in the slope of the employment rate as long as individuals who earn near  $z^{AET}$  in state 0 are constrained from making small adjustments exactly to  $z^{AET}$  in state 1. We also demonstrate in this case that the extensive margin behavior for those near  $z^{AET}$  is not affected by the ability to make discrete adjustments to other earnings levels, *e.g.* a part-time job.

## 6.4 Estimating an Extensive Margin Elasticity with a Kink

When a kink is created in the employment rate due to a kink in the budget set, we may use this behavioral response to estimate an extensive margin elasticity  $\eta$  for a given average net-of-tax-and-transfer rate  $1 - a \equiv 1 - [(T(z) - B(z)) - (T(0) - B(0))] / z$ . Using the fact that  $\Pr(z > 0) = G(\bar{q})$ , we have:

$$\eta \equiv \frac{d\Pr(z > 0)}{d(1 - a)} \frac{1 - a}{\Pr(z > 0)} = g(\bar{q}) \frac{\partial \bar{q}}{\partial(1 - a)} \frac{1 - a}{\Pr(z > 0)}$$

For now we maintain the assumption of  $\tilde{z}_{n1} \equiv \tilde{z}_{n0}$  as in Section 6.3. Denote  $\beta$  as the estimated discontinuity at  $z^{AET}$  in the slope of the employment probability, *i.e.* the parameter estimated in Equation 3 in Section 3. Given (13) and the fact that  $\partial \bar{q} / \partial(1 - a) = \lambda z$ , we have:

$$\eta^{AET} = g(\bar{q}_{n^{AET}1} | n^{AET}) \frac{\partial \bar{q}_{n^{AET}1}}{\partial(1 - a)} \frac{1 - a}{\Pr(z_{n1} > 0 | \tilde{z}_{n0} = z^{AET})} = \frac{\beta}{\alpha} \cdot \frac{1 - a}{\Pr(z_{n1} > 0 | \tilde{z}_{n0} = z^{AET})} \quad (14)$$

where  $\eta^{AET}$  is the extensive margin elasticity for the individual earning  $z^{AET}$  in state 0, and  $\alpha = -db/z^{AET}$  is the magnitude of the kink at  $z^{AET}$  in the slope of the ANTR,  $1 - a$ .<sup>30</sup> In other words, the RKD approach from Section 3 for estimating the elasticity of extensive margin participation with respect to the ANTR naturally emerges from our model. Although this approach yields a single elasticity, if the function  $g(\cdot)$  is heterogeneous across different sub-groups, then the elasticity may also vary across these sub-groups. Note, our approach does not allow us to estimate the extent to which the marginal utility of income  $\lambda_{n^{AET}}$  and the fixed cost distribution  $g(\bar{q}_{n^{AET}1} | n^{AET})$  separately contribute to  $\eta^{AET}$ . Nonetheless, the extensive margin elasticity  $\eta^{AET}$  on its own can allow us to perform counterfactual simulations of the implications of different tax schedules (*e.g.* Eissa *et al.*, 2008).

<sup>29</sup>In fact, one can allow for some continuity in the choice set as long as, loosely speaking, individuals cannot make adjustments exactly to  $z^{AET}$ . See Appendix B.1 for details.

<sup>30</sup>To see this, note that below  $z^{AET}$ ,  $(1 - a) = 1 - \tau_0$  and  $\partial(1 - a)/z = 0$ , while above  $z^{AET}$ ,  $(1 - a) = 1 - (\tau_0 z + db(z - z^{AET})) / z = 1 - \tau_0 - db(1 - z^{AET}/z)$  and  $\partial(1 - a)/z = -dbz^{AET}/z^2$ . When  $z = z^{AET}$ , this reduces to  $-db/z^{AET}$ .

## 6.5 Interpreting the Observed Elasticity

Thus far we have derived an expression for the extensive margin elasticity under a kink assuming that individuals cannot adjust on the intensive margin across tax schedules. In fact, some individuals may be able to adjust on the intensive margin to  $z^{AET}$ . Suppose that under a kink we observe both the intensive margin bunching described above and a discontinuity at counterfactual earnings  $z^{AET}$  in the slope of the employment rate. In this case we discuss how to interpret the “observed” extensive margin elasticity,  $\hat{\eta} = \left(\hat{\beta}/\alpha\right) \cdot (1 - a) / \Pr(z_{n1} > 0 | \tilde{z}_{n0} = z^*)$ , where  $\hat{\beta}$  is the estimated kink at  $z^{AET}$  in the employment rate. In Online Appendix B.2 we explore two separate frameworks that allow for both bunching among some individuals and a kink in the employment rate among others. We show that in either case, the observed elasticity,  $\hat{\eta}$ , can be interpreted as a weak lower bound on the “structural” elasticity  $\eta^{AET}$ , *i.e.*  $\hat{\eta} \leq \eta^{AET}$ . Here “observed” and “structural” are used in a sense analogous to Chetty (2012). The “observed” elasticity refers to what we observe due to an increase in effective marginal tax rates above the kink among those earning near  $z^{AET}$ , which is in part affected by the lack of an extensive margin response among those unconstrained at the intensive margin. The “structural” elasticity refers to the elasticity we would hypothetically observe in response to a change in average tax rates everywhere, among those earning near  $z^{AET}$  and including both those constrained and those unconstrained on the intensive margin.<sup>31</sup> Thus, the presence of bunching that is observed at ages 62 and older in our data (see Gelber et al., 2020) need not be inconsistent with our model, and the observed extensive margin elasticity estimated in that case can be interpreted as a lower bound on the structural elasticity.

## 6.6 Extensions to the Model

In our Online Appendix, we consider several extensions to the model. In Online Appendix B.5 we extend our model to show under mild assumptions that when claiming can respond jointly with working, the kink in the probability of earning a positive amount will be less negative than in the case in which claiming is exogenous. Nonetheless, we can recover the elasticity among those claiming Social Security by dividing the observed elasticity by the probability of claiming at  $z^{AET}$ . In the most general case in which we relax these mild assumptions, we can still obtain a lower bound on this employment elasticity by subtracting the kink in claiming from the kink in the employment probability. In Online Appendix B.4 we present a fully dynamic, multi-period model with a joint decision over saving and earnings, and we show that under a set of empirically relevant assumptions our results for kinks still hold: if intensive margin frictions exist, a change in the slope of the employment rate occurs at the budget set kink, and we can still recover the elasticity using (14). This appendix also discusses the interpretation of the parameter we estimate, depending on whether the tax change is anticipated or unanticipated. Finally, in Online Appendix B.4, we extend this model to tailor it to our particular policy setting by showing that our method still applies when reductions in current

<sup>31</sup>Although we use the term “structural elasticity,” the extensive margin elasticity depends on the number of individuals who are just indifferent between working and not working—*i.e.*  $g(\bar{q}|n^{AET})$ —which may vary depending on the employment level.

benefits due to the AET can lead to increases in later benefits, as under benefit enhancement.

## 6.7 Mapping the RKD Estimator to the Model

The RKD statistical model from Section 3 can be related directly to our theoretical model. First, the smoothness conditions imposed on preferences and heterogeneity in Assumption 1 imply the smoothness assumptions in Card *et al.* (2015) that allow us to interpret the RKD estimate as a treatment-on-the-treated parameter. Second, the RKD coefficients map directly to the parameters of our model. The parameter  $\beta$  in equation (14) corresponds to the  $\beta_1$  from (4), and we can calculate  $\alpha$  analytically, as described above. We calculate the remaining elements of equation (14),  $1 - a$  and  $\Pr(z_{n1} > 0 | \tilde{z}_{n0} = z^*)$ , using data on individuals who have age 60 earnings near  $z^*$ . Moreover, just as the RKD returns an estimate that is local to agents located at the kink, our theoretical model identifies parameters that apply to agents located at the kink.

As described in Section 3.2, we use earnings at age 60 to proxy for desired earnings at ages 63 and 64 in the absence of the AET, positing equation (5). Thus, in terms of our model, we require that the ability distribution remain relatively stable between ages 60 and 63 to 64. We can still allow  $\pi(\tilde{z}_{i0t}, v_{it}) < 1$  in Section 3.2, *i.e.* we do not require this distribution to be absolutely static across this time span. Furthermore, this assumption places no restrictions on the evolution of the fixed cost distribution,  $G(q|n)$ , over time. In the data, we observe a declining employment rate between ages 60 and 63 to 64, which can be accommodated by a rightward shift in the distribution  $G(q|n)$  at older ages. While we use age 60 earnings to proxy for desired earnings at ages 63 to 64 as above, our RKD effectively uses the employment rate at ages 63 to 64 for those with age 60 earnings just under the exempt amount to reflect state 0 employment rates, while the employment rate at ages 63 to 64 for those with age 60 earnings just over the exempt amount reflects state 1 employment rates. In our model, we require that the distribution  $G(q|n)$  is the same across states 0 and 1. The empirical analogue is thus a requirement that the distribution of fixed costs  $G(q|n)$  is same for those with age 60 earnings just below and just above the exempt amount.

## 7 Conclusion

In this study, we bring new evidence to bear on the effect of the Social Security Annual Earnings Test (AET) on the employment of older workers. Previous studies have found mixed or no effect on the extensive margin decision of whether or not to work, and current policy discussions likewise take a lack of a labor force participation effect as a given. To the contrary, we show that the AET has a large impact on extensive margin earnings decisions. Under our preferred specification, the point estimate shows that the elasticity of the employment rate with respect to the ANTR is at least 0.49. We interpret this as a lower bound on the elasticity for several reasons. Estimating this large lower bound is useful because it clearly distinguishes our study from previous literature that largely finds the AET has little effect on employment. The source of discrepancy between ours and prior estimates in the literature is likely due to the new method we develop for estimating extensive margin responses to nonlinear budget sets, combined with the use of large scale,

administrative Social Security Administration data on earnings. In particular, we leverage a discontinuity in the slope of incentives created by the AET to obtain regression kink design (RKD) estimates of the employment effect of the policy.

This large response suggests that the AET can have important effects on the employment of older workers. Our point estimates imply that eliminating the AET would increase the employment rate among affected older workers by 1.4 percentage points (2.5 percent) in the group we study, and would increase earnings by 9.8 percent. We calculate that 27.6 percent of the aggregate earnings increase due to eliminating the AET is associated with extensive margin decisions, demonstrating that focusing on intensive margin responses can abstract from an important component of the earnings impacts of tax and transfer programs. These estimates also suggest that the impact of eliminating the AET on the budget is more positive than the budget-neutral estimates of the Congressional Budget Office (2000), which abstracted from extensive margin effects anywhere (including in the group around the kink we study). It should be noted, however, that in order to fully evaluate the welfare effects of the AET, one must compare its labor supply effects with any potential consumption smoothing benefits the policy may produce. That is, by retiming Social Security benefits to later in life, the AET may benefit those whose early claiming decisions increase the risk of navigating late life expenses with low monthly benefits (see, for example, Figinski and Neumark, 2018).

We subject our results to a variety of robustness checks and placebo tests. We do not detect manipulation in the running variable key to our RKD estimates, and likewise do not find any robust kinks in pre-determined covariates. Our results are generally insensitive to adding demographic controls. Perhaps our cleanest test involves estimating placebo kinks in employment among the same workers, but at ages where the AET does not apply (ages 56, 57, 61, and 63). In contrast to our main estimates, we do not find kinks in employment at those ages. In nearly all cases, adjustments for common threats to validity lead us to conclude that our estimates likely represent a lower bound on the extensive margin elasticity in our sample. In that case, the employment response to the AET appears to have been even more underappreciated than appears at first glance.

In addition to exploring the effects of the AET in this particular setting, our study makes a number of broader contributions. First, we have developed a new method for identifying extensive margin elasticities using a kinked budget set. The estimation method requires: (1) a kinked budget set; (2) data on earnings or labor supply; and (3) a proxy for desired earnings under a counterfactual linear budget set. Thus, it appears that our method is applicable in many other contexts in which individuals make extensive margin employment decisions, even outside of the AET, particularly because lagged earnings can proxy for current desired earnings in a broad range of ages as we show in Appendix Figure E.4. In fact, much like our study, a recent working paper by Escobar et al. (2019) uses a kink in Swedish inheritance taxes to motivate a regression kink design and estimate an extensive margin elasticity in bequests.

While we consider the case of a kink, our method could be extended to budget sets that feature discontinuities in the level of incentives, i.e. “notches.” It might also be possible to apply our methods to other

non-linear pricing settings, where extensive margin demand decisions are made. In fact, as we mention in Section 5.2, there is a possibility that agents in our setting perceive kink created by the AET as notch. In future work, we plan to explore both what additional evidence might shed light on this potential behavior, e.g. asymmetric bunching on one side of the “notch”, and also what parameters can still be recovered in that case.

More generally, the employment responses we estimate could be useful in simulating and evaluating tax reforms when older workers are among the group affected. This requires knowledge of both intensive margin and extensive margin behavior (Essa *et al.*, 2008). However, standard approaches would have to accommodate the presence of frictions on earnings adjustment, which we show theoretically are likely to be present in cases like ours where kinks in the average net-of-tax rate (ANTR) translate into kinks in the employment rate. The presence of frictions is consistent with recent literature including Chetty (2012), Kleven and Waseem (2013), or Gelber *et al.* (2020). It is also consistent with Abraham and Houseman’s (2005) findings from a self-reported survey that suggest frictions among older workers in particular, and Baker and Benjamin’s (1999) findings consistent with frictions in Canada. Our model points out that as a result of these frictions, the observed extensive margin elasticity in our setting need not reflect the structural elasticity, but the observed elasticity generally provides a lower bound. Overall, our findings suggest that ongoing discussions regarding employment responses to policies such as the AET may need to evolve in light of new evidence of both an extensive margin behavior and also of frictions in adjustment of earnings.

## References

- Aaronson, S., Cajner, T., Fallick, B., Galbis-Reig, F., Smith, C., and Wascher, W. (2014). "Labor Force Participation: Recent Developments and Future Prospects," *Brookings Papers on Economic Activity*, 197-275.
- Abraham, K., and Houseman, S. (2005). "Work and Retirement Plans among Older Americans." In *Reinventing the Retirement Paradigm*, R. Clark and O. Mitchell, eds. Oxford, UK; New York: Oxford University Press, 70-91.
- Abraham, K., and Kearney, M. (2017). "The Factors Behind the Decline in U.S. Employment Rates: A Review of the Evidence." University of Maryland Working Paper.
- Alpert, A., and Powell, D. (2014). "Estimating Intensive and Extensive Tax Responsiveness: Do Older Workers Respond to Income Taxes." RAND Working Paper.
- Altonji, J. G. and Paxson, C. H. (1988). "Labor Supply Preferences, Hours Constraints, and Hours-Wage Trade-Offs," *Journal of Labor Economics*, 6, 254-276.
- Brown, S. K., and Perron, R. (2011). "Assessing Current and Future Beneficiaries' Knowledge of Social Security Benefits," American Association of Retired Persons research report.
- Ando, M. (2017). "How Much Should We Trust Regression-Kink-Design Estimates?" *Empirical Economics*, 53, 1287-1322.
- Baker, M., and Benjamin, D. (1999). "How do retirement tests affect the labor supply of older men?" *Journal of Public Economics*, 71, 27-51.
- Battistin, E., and Chesher, A. (2014). "Treatment Effect Estimation with Covariate Measurement Error." *Journal of Econometrics*, 178(2), 707-715.
- Bernheim, B. D., Fradkin, A., and Popov, I. (2015). "The Welfare Economics of Default Options in 401(k) Plans," *American Economic Review*, 105(9), 2798-2837.
- Bitler, M. P., Gelbach, J. B., and Hoynes, H. W. (2006). "What Mean Impacts Miss: Distributional Effects Of Welfare Reform Experiments," *American Economic Review*, 96(4), 988-1012.
- Blinder, A. S., Gordon, R. H., and Wise, D. E. (1980). "Reconsidering the Work Disincentive Effects of Social Security." *National Tax Journal*, 33(4), 431-442.
- Blomquist, S., Newey, N., Kumar, A., and Liang, C-Y. (2019). "On Bunching and Identification of the Taxable Income Elasticity." NBER Working Paper No. 24136.
- Blundell, R., French, E., and Tetlow, G. (2016). "Retirement Incentives and Labor Supply." In *Handbook of the Economics of Population Aging*, John Piggott and Alan Woodland, eds., 457-566. Amsterdam: Elsevier.
- Blundell, R., and MaCurdy, T. (1999). "Labor Supply: A Review of Alternative Approaches." In Ashenfelter, O. and Card, D., eds., *Handbook of Labor Economics* 3A, 1560-1695. Amsterdam: Elsevier.
- Blundell, R. and Shephard, A. (2012). "Employment, Hours of Work and the Optimal Taxation of Low-Income Families." *Review of Economic Studies*, 79, 481-510.
- Brown, J., Kapteyn, A., Mitchell, O., and Mattox, T. (2013). "Framing the Social Security Earnings Test." Wharton Pension Research Council Working Paper.
- Bureau of Labor Statistics. (2011). "Historical Consumer Price Index for All Urban Consumers (CPI-U): U.S. city average, all items." Available at <http://www.bls.gov/cpi/tables/historical-cpi-u-201709.pdf>
- Burkhauser, R. V., and Turner, J. A. (1978). "A Time Series Analysis on Social Security and the Market Work of Men at Younger Ages." *Journal of Political Economy*, 86(4), 701-715.
- Burtless, G., and Hausman, J. A. (1978). "The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment." *Journal of Political Economy*, 86(6), 1103-1130.
- Burtless, G., and Moffitt, R. A. (1985). "The Joint Choice of Retirement Age and Postretirement Hours of Work." *Journal of Labor Economics*, 3, 209-236.
- Calonico, S., Cattaneo, M., and Titiunik, R. (2014). "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica*, 82(6), 2295-2326.
- Calonico, S., Cattaneo, M., Farrell, M., and Titiunik, R. (2016). "Regression Discontinuity Designs Using Covariates." University of Michigan Working Paper.
- Card, D., Lee, D., Pei, Z., and Weber, A. (2015). "Inference on Causal Effects in a Generalized Regression Kink Design." *Econometrica*, 83(6), 2463-2483.



- Card, D., Lee, D., Pei, Z., and Weber, A.** (2017). "Regression Kink Design: Theory and Practice." In Cattaneo, M. D. and, Escanciano, J. C., eds. *Regression Discontinuity Designs: Theory and Applications* (Advances in Econometrics, Volume 38). (pp. 341-382). May.
- Chetty, R.** (2012). "Bounds on Elasticities with Optimization Frictions." *Econometrica*, 80(3), 969-1018.
- 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, 749-804.
- Chetty, R., Friedman, J. N., and Saez, E.** (2013). "Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings." *American Economic Review*, 103(7), 2683-2721.
- Chetty, R., Guren, A., Manoli, D., and Weber, A.** (2012). "Does Indivisible Labor Explain the Difference Between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities." In Acemoglu, D., Parker, J., and Woodford, M., eds., *NBER Macroeconomics Annual 2012* (Vol. 27, pp. 1-56). Chicago: University of Chicago Press.
- Chetty, R., Looney, A., and Kroft, K.** (2007), "Salience and Taxation: Theory and Evidence." NBER Working Paper No. 13330.
- Christensen, K.** (1990). "Bridges Over Troubled Water: How Older Workers View the Labor Market," in P. Doeringer, ed., *Bridges to Retirement: Trends in the Labor Market for Older Workers* (pp. 175-207). Ithaca, NY: ILR Press, Cornell University.
- Cogan, J.** (1981). "Fixed Costs and Labor Supply." *Econometrica*, 49, 945-963.
- Congressional Budget Office.** (2000). *Pay-As-You-Go Estimate: H.R. 5, Senior Citizens' Freedom to Work Act of 2000*. Washington, D.C.: U.S. Government Printing Office.
- Congressional Research Service.** (2019). *Social Security Retirement Earnings Test: How Earnings Affect Benefits*. Washington, D.C.: Congressional Research Service. <https://fas.org/sgp/crs/misc/R41242.pdf>
- Diamond, P.** (1980). "Income Taxation with Fixed Hours of Work." *Journal of Public Economics*, 13(1), 101-110.
- Diamond, P., and Gruber, J.** (1999). "Social Security and Retirement in the United States." In Gruber, J. and Wise, D., eds., *Social Security and Retirement around the World* (pp. 437-473). Chicago: University of Chicago Press.
- Dickens, W., and Lundberg, S.** (1993). "Labor Supply with Hours Restrictions." *International Economic Review*, 34(1), 169-192.
- Disney, R., and Smith, S.** (2002). "The labour supply effect of the abolition of the earnings rule for older workers in the United Kingdom." *Economic Journal*, 112, C136-C152.
- Eissa, N., and Hoynes, H. W.** (2004). "Taxes and the Labor Market Participation of Married Couples: the Earned Income Tax Credit." *Journal of Public Economics*, 88(9-10), 1931-1958.
- Eissa, N., Kleven, H., and Kreiner, C.** (2008). "Evaluation of four tax reforms in the United States: Labor supply and welfare effects for single mothers." *Journal of Public Economics*, 92(3-4), 795-816.
- Eissa, N., and Liebman, J.** (1996). "Labor Supply Response to the Earned Income Tax Credit." *Quarterly Journal of Economics*, 111(2), 605-637.
- Engelhardt, G., and Kumar, A.** (2014). "Taxes and the Labor Supply of Older Americans: Recent Evidence from the Social Security Earnings Test." *National Tax Journal*, 67(2), 443-458.
- Escobar, S., Ohlsson, H., and Selin, H.** (2019). "Giving to Children or the Taxman? Lessons from a Swedish Inheritance Tax Loophole." Unpublished Manuscript
- Figinski, T., and Neumark, D.** (2018). "Does Eliminating the Earnings Test Increase the Incidence of Low Income Among Older Women?" NBER Working Paper No. 21601.
- French, E.** (2005). "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour." *Review of Economic Studies*, 72, 395-427.
- Friedberg, L.** (1998). "The Social Security Earnings Test and Labor Supply of Older Men." In Poterba, J. M., ed., *Tax Policy and the Economy* (pp. 121-150). Chicago: University of Chicago Press.
- Friedberg, L.** (2000). "The Labor Supply Effects of the Social Security Earnings Test." *Review of Economics and Statistics*, 82, 48-63.
- Friedberg, L., and Webb, A.** (2009). "New Evidence on the Labor Supply Effects of the Social Security Earnings Test." In Brown, J.R., and Poterba, J. M., eds., *Tax Policy and the Economy* (pp. 1-35).

Chicago: University of Chicago Press.

**Ganong, P., and Jäger, S.** (2015). “A Permutation Test for the Regression Kink Design.” Harvard University Working Paper.

**Gelber, A., and Mitchell, J.** (2012). “Taxes and Time Allocation: Evidence from Single Women and Men.” *Review of Economic Studies*, 79(3), 863-897.

**Gelber, A., Isen, A., and Song, J.** (2016). “The Effect of Pension Income on Elderly Earnings: Evidence from Social Security and Full Population Data.” UC Berkeley Working Paper.

**Gelber, A., Jones, D., and Sacks, D.** (2013). “Earnings Adjustment Frictions: Evidence from the Social Security Earnings Test.” NBER Working Paper No. 19491.

**Gelber, A., Jones, D., and Sacks, D.** (2020). “Estimating Adjustment Frictions Using Nonlinear Budget Sets: Method and Evidence from the Earnings Test.” *American Economic Journal: Applied Economics*. Vol. 12, No. 1, 1-31.

**Gelber, A., Jones, D., Sacks, D., and Song, J.** (2017). “The Labor Force Participation Effects of the Social Security Earnings Test.” Indiana University Working Paper.

**Gelber, A., Moore, T., and Strand, A.** (Forthcoming). “The Effect of Disability Insurance Payments on Beneficiaries’ Earnings.” *American Economic Journal: Economic Policy*.

**Gruber, J., and Orzag, P.** (2003). “Does the Social Security Earnings Test Affect Labor Supply and Benefits Receipt?” *National Tax Journal*, 56, 755-773.

**Gruber, J., and Wise, D.** (1999). “Introduction and Summary.” In Gruber, J., and Wise, D., eds., *Social Security and Retirement Around the World* (pp. 1-35). Chicago: University of Chicago Press.

**Gustman, A., and Steinmeier, T.** (1985). “The 1983 Social Security Reforms and Labor Supply Adjustment of Older Individuals in the Long Run.” *Journal of Labor Economics*, 3(2), 237-253.

**Gustman, A., and Steinmeier, T.** (1991). “Changing the Social Security Rules for Work after 65.” *Industrial Labor and Relations Review*, 44(4), 733-745.

**Haider, S., and Loughran, D.** (2008). “The Effect of the Social Security Earnings Test on Male Labor Supply: New Evidence from Survey and Administrative Data.” *Journal of Human Resources*, 43(1), 57-87.

**Hausman, J.** (1981). “Labor Supply.” In Aaron, H. J. and Pechman, J. A., eds., *How Taxes Affect Economic Behavior* (pp. 27-71). Washington, DC: Brookings Institution.

**Heckman, J.** (1979). “Sample Selection Bias as a Specification Error.” *Econometrica*, 47(1), 153-161.

**Honig, M., and Reimers, C.** (1989). “Is it Worth Eliminating the Retirement Test?” *American Economic Review*, 79(2), 103-107.

**Hoynes, H. W.** (1996). “Welfare Transfers in Two-Parent Families: Labor Supply and Welfare Participation Under AFDC-UP.” *Econometrica*, 64, 295-332.

**Ito, K.** (2014). “Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing.” *American Economic Review*, 104(2), 537-563.

**J. K. Lasser Institute** (1998). *Your Income Tax: 1998*. New York: Macmillan General Reference.

**Keane, M. P.** (2011). “Labor Supply and Taxes: A Survey.” *Journal of Economic Literature*, 49(4), 961-1075.

**Keane, M. P. and Moffitt, R.** (1998). “A Structural Model of Multiple Welfare Program Participation and Labor Supply.” *International Economic Review*, 39, 553-589.

**Kleven, H.** (2016). “Bunching.” *Annual Review of Economics*, 8, 435-464.

**Kleven, H., and Kreiner, C.** (2005). “Labor Supply Behavior and the Design of Tax and Transfer Policy.” *Danish Journal of Economics*, 143, 321-358.

**Kleven, H., Landais, C., Saez, E., and Schultz, E.** (2014). “Migration and Wage Effects of Taxing Top Earners: Evidence from the Foreigners’ Tax Scheme in Denmark.” *Quarterly Journal of Economics*, 129(1), 333-378.

**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.

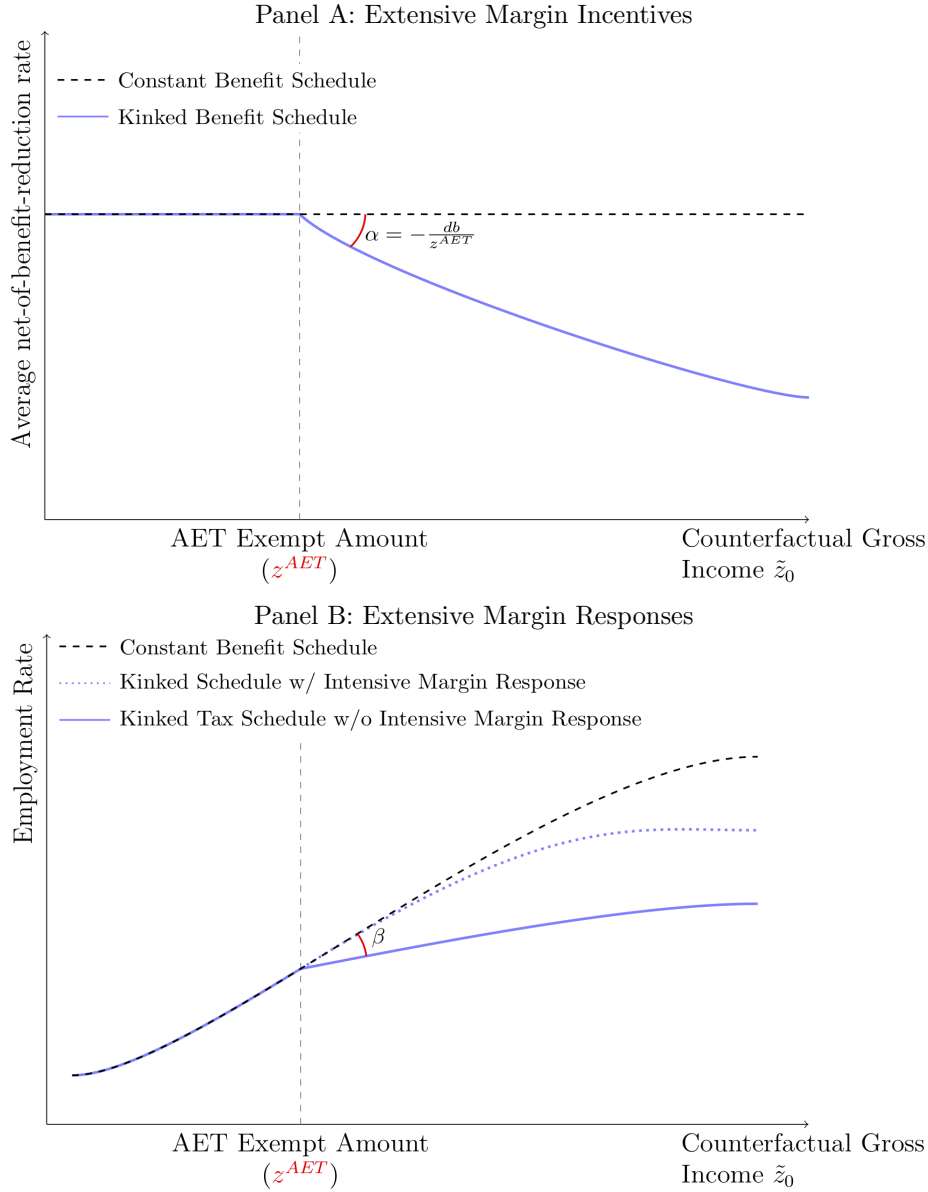
**Kline, P., and Tartari, M.** (2016). “Bounding Labor Supply Responses to a Randomized Welfare Experiment: A Revealed Preference Approach.” *American Economic Review*, 106, 972-1014.

**Kopczuk, W., and Munroe, D.** (2015). “Mansion Tax: The effect of Transfer Taxes on the Residential Real Estate Market.” *American Economic Journal: Economic Policy*, 7(2), 214-257.

- Kopczuk, W., Saez, E., and Song, J.** (2010). "Earnings Inequality and Mobility in the United States: Evidence from Social Security Data since 1937." *Quarterly Journal of Economics*, 125(1), 91-128.
- Laitner, J., and Silverman, D.** (2012). "Consumption, Retirement, and Social Security: Evaluating the Efficiency of Reform that Encourages Longer Careers." *Journal of Public Economics*, 96, 615-634.
- Landais, C.** (2015). "Assessing the Welfare Effects of Unemployment Benefits Using the Regression Kink Design." *American Economic Journal: Economic Policy*, 7(4), 243-278.
- Lee, D. S., and T. Lemieux** (2010). "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48, 281-355.
- Leonesio, M. V.** (1990). "The Effects of the Social Security Earnings Test on the Labor-Market Activity of Older Americans: A Review of Empirical Evidence." *Social Security Bulletin*, 53(5), 2-21.
- Liebman, J., and Luttmer, E.** (2015). "Would People Behave Differently If They Better Understood Social Security? Evidence From a Field Experiment." *American Economic Journal: Economic Policy*, 7(1), 275-99.
- Manoli, D., and Weber, A.** (2016). "The Effects of the Early Retirement Age on Retirement Decisions." NBER Working Paper 22561.
- Manoli, D., and Weber, A.** (Forthcoming). "Nonparametric Evidence on the Effects of Financial Incentives on Retirement Decisions." *American Economic Journal: Economic Policy*.
- Manski, C.** (2014). "Identification of Income-Leisure Preferences and Evaluation of Income Tax Policy." *Quantitative Economics*, 5, 145-174.
- McCrary, J.** (2008). "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics*, 142(2), 698-714.
- McDonald, J. F., and Moffitt, R. A.** (1980). "The Uses of Tobit Analysis." *Review of Economics and Statistics*, 62(2), 318-321.
- Meyer, B. D., and Rosenbaum, D. T.** (2001). "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers." *Quarterly Journal of Economics*, 116(3), 1063-1114.
- Moffitt, R.** (1983). "An Economic Model of Welfare Stigma." *American Economic Review*, 73(5), 1023-1035.
- Nielsen, H. S., Sorensen, T., and Taber, C.** (2010). "Estimating the Effect of Student Aid on College Enrollment from a Government Grant Policy Reform." *American Economic Journal: Economic Policy*, 2(2), 185-215.
- Pei, Z., and Shen, Y.** (2017). "The Devil is in the Tails: Regression Discontinuity Design with Measurement Error in the Assignment Variable." In Cattaneo, M. D. and Escanciano, J. C., eds. *Regression Discontinuity Designs: Theory and Applications* (Advances in Econometrics, Volume 38). (pp. 455-502). May.
- Powell, J. L.** (1994). "Estimation of Semiparametric Models." In Engle, R. F., and McFadden, D., eds., *Handbook of Econometrics* (pp. 2443-2521). Amsterdam: Elsevier.
- Rubin, D. B.** (1974). "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology*, 66.5, 688-701.
- Saez, E.** (2002). "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses." *Quarterly Journal of Economics*, 117(3), 1039-1073.
- Saez, E.** (2010). "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy*, 2, 180-212.
- Saez, E., Slemrod, J. and Giertz, S.** (2012). "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature*, 50(1), 3-50.
- Social Security Administration.** (2011). "Earnings Public Use File." Social Security Database, available at <https://www.ssa.gov/policy/docs/microdata/epuf/index.html>, last accessed November 13, 2015.
- Social Security Administration.** (2012). *Social Security Handbook*. Washington, D.C.: SSA.
- Social Security Administration.** (2014). "Master Earnings File." Social Security Database, last accessed May 7, 2014.
- Social Security Administration.** (2014). "Master Earnings File." Social Security Database, last accessed May 7, 2014.
- Song, J. G.** (2004). "Evaluating the Initial Impact of Eliminating the Retirement Earnings Test." *Social Security Bulletin*, 65(1), 1-15.

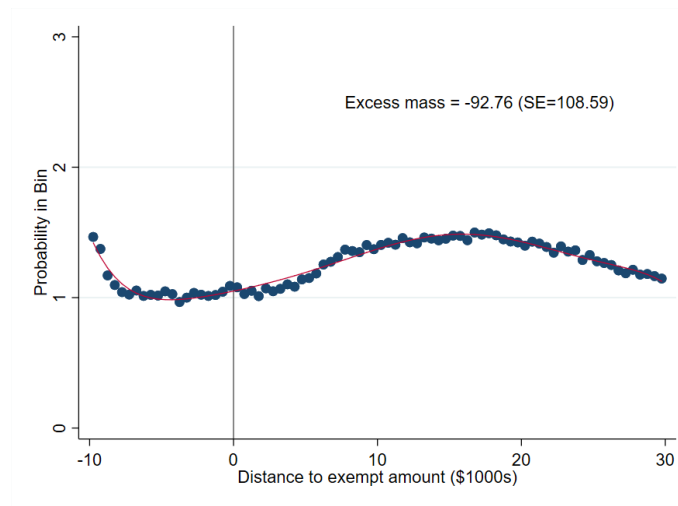
- Song, J. G., and Manchester, J.** (2007). “New Evidence on Earnings and Benefit Claims Following Changes in the Retirement Earnings Test.” *Journal of Public Economics*, 91, 669-700.
- Tergesen, A.** (2016). “How to Get People to Delay Retirement.” *Wall Street Journal*, March 21.
- Turner, L.** (2014). “The Road to Pell is Paved with Good Intentions: The Economic Incidence of Federal Student Grant Aid.” University of Maryland Working Paper.
- Utendorf, K.** (2001/2002). “The Upper Part of the Earnings Distribution in the United States: How Has It Changed?” *Social Security Bulletin*, 64 (3), 1-11.
- Viscusi, K.** (1979). *Welfare of the Elderly: An Economic Analysis and Policy Prescription*. New York, NY: John Wiley and Sons.
- Vroman, W.** (1985). “Some Economic Effects of the Retirement Test.” In Ehrenberg, R., ed., *Research in Labor Economics* (Vol. 7, pp. 31-89). Greenwich, CT: Jai Press, Inc.

Figure 1: Extensive Margin Response by State 0 Counterfactual Earnings



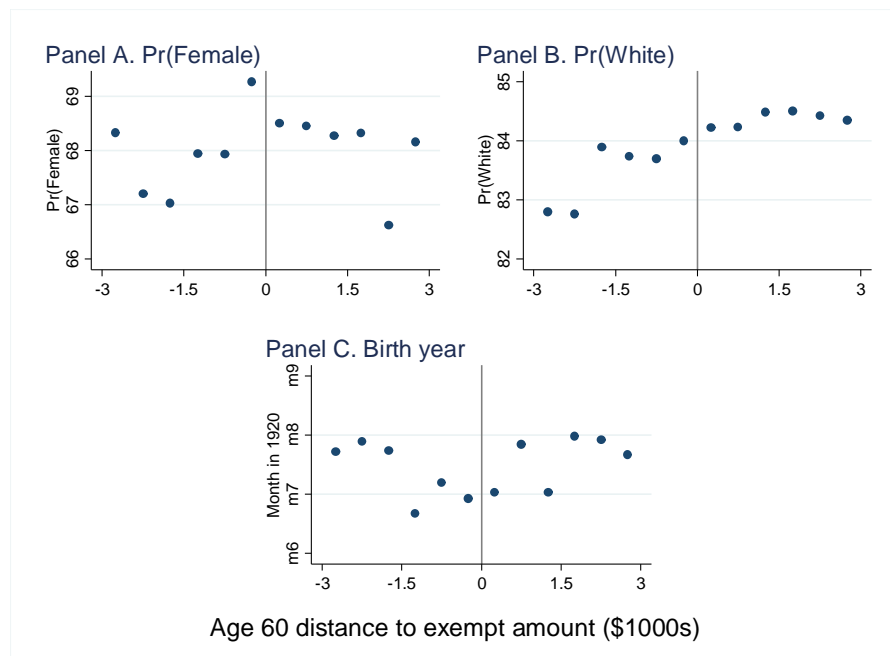
Notes: The figures illustrate features of the framework described in Section 2.2. The  $x$ -axes show desired income if employed constant benefit budget set in state 0, i.e.  $\tilde{z}_0$ . The  $y$ -axis in Panel A shows the ANTR. Under a constant benefit schedule the ANTR is flat everywhere and represented by a dashed line. Under a kinked tax schedule, the ANTR is equal to  $1 - db$  above the kink point  $z^{AET}$ , and is represented by a solid line. Panel A shows that under a kink, the slope of the ANTR discontinuously decreases at the kink point  $z^{AET}$  by a value of  $a$ , due to the imposition of effectively higher marginal tax on earnings above  $z^{AET}$ . Panel B shows a hypothetical probability of employment, under three scenarios: a constant benefit schedule in state 0 (dashed line); a kinked tax schedule in state 1 when individuals can make intensive margin adjustments (dotted line); and a kinked tax schedule in state 1 when individuals cannot make intensive margin adjustments (solid line).  $\beta$  refers to the change in slope at the exempt amount of the employment probability as a function of desired earnings when a kink is present and there are no intensive margin adjustments.

Figure 2: Pre-AET Earnings Distribution



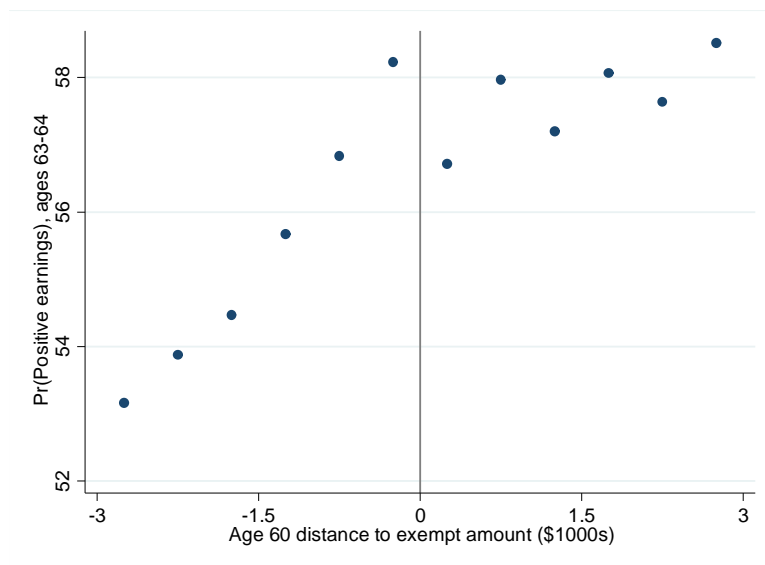
Notes: The figure shows the actual earnings density (dotted plot) and smooth fit (dashed line). The data shown are means within bins of width \$500. The exempt amount is normalized to zero and shown with the vertical line at zero. Following Chetty *et al.* (2011), the smooth density is estimated using a seventh-order polynomial, excluding the region within a \$3,000 bandwidth of the exempt amount. The source is a 25 percent random sample of the Social Security Administration Master Earnings File linked to the Master Beneficiary Record. The sample consists of individuals with age 60 earnings that are positive and within \$40,000 of the exempt amount, born 1918 to 1923, with no self-employment income at age 60, and excluding individuals who ever have negative earnings at ages 50-57 or 63-70. See Appendix C for a description of how we estimate excess mass.

Figure 3: Predetermined Covariates



Note: The figure shows the bin means of predetermined covariates as a function of the distance to the age 60 exempt amount. The figure demonstrates that there are no clear visual changes in slope in any of these covariates at the age 60 exempt amount, consistent with the assumptions necessary for the validity of the regression kink design.

Figure 4: Employment Rate Near the Exempt Amount, Ages 63 to 64

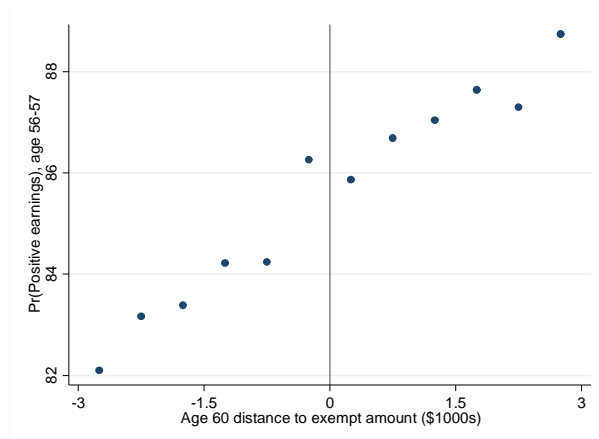


Notes: The figure plots the mean annual employment rate, *i.e.* the probability of positive earnings, at ages 63 to 64 averaged, as a function of the distance of age 60 earnings from the exempt amount. The sample is individuals with positive age 60 earnings and no age 60 self-employment income, born 1918 to 1923. See other notes to Figure 2.

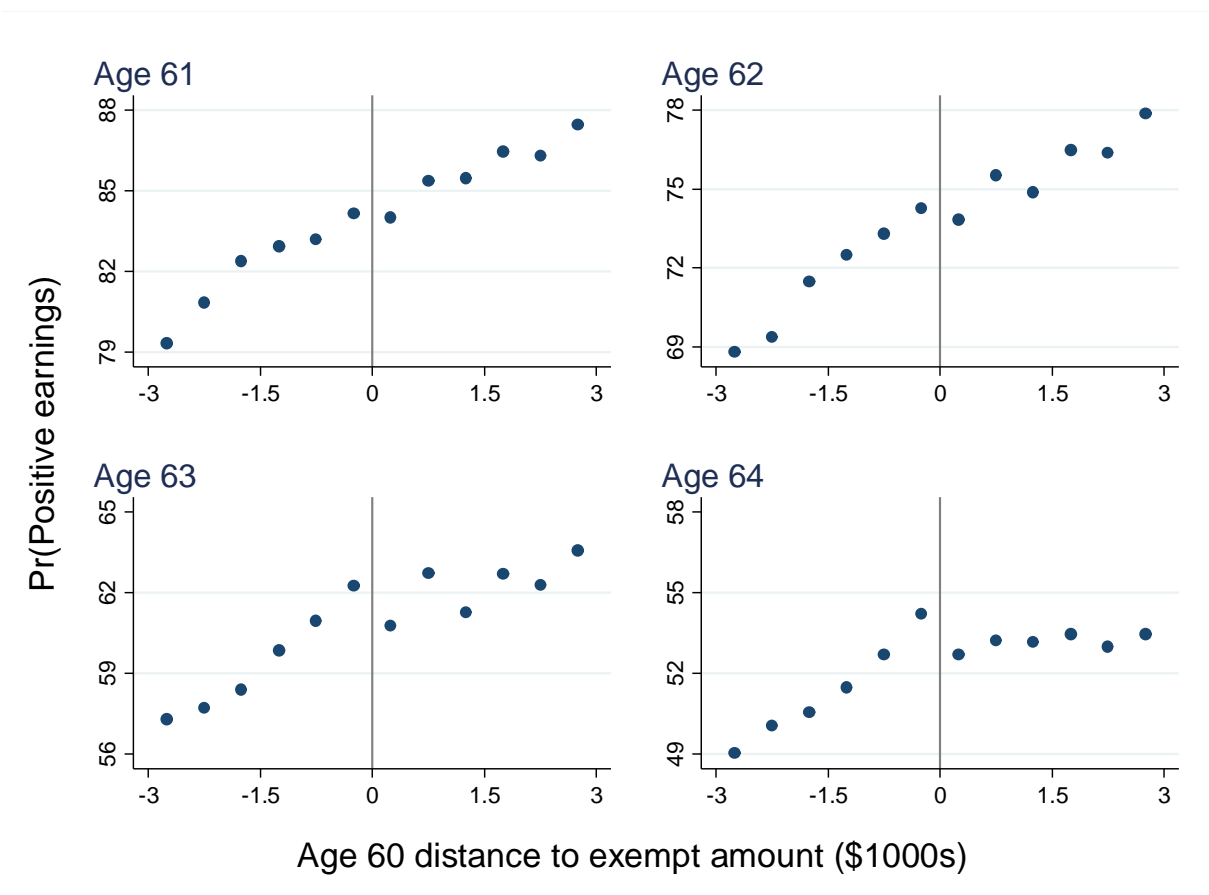


Figure 5: Placebo Tests

Panel A: Probability of Positive Earnings, Ages 56 to 57

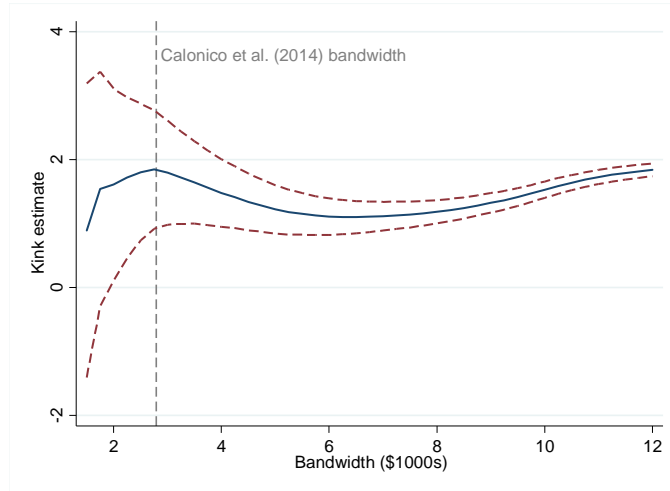


Panel B: Probability of Positive Earnings by Single Year of Age, Ages 61 to 64



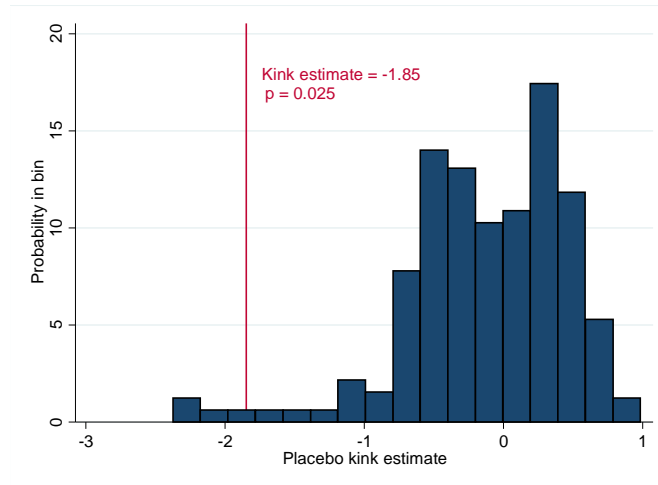
Notes: Panel A plots the annual probability of positive earnings at ages 56 to 57, as a function of the distance of age 60 earnings from the exempt amount. In Panel B, each figure plots the probability of positive earnings for each single year of age from 61 to 64. See other notes to Figure 2.

Figure 6: Kink Estimate by Bandwidth



Notes: The figure shows that as a function of the bandwidth, the estimated change in slope at the exempt amount of the mean annual probability of age 63 to 64 employment is relatively stable. The solid line shows the point estimates, and the dashed lines show the 95 percent confidence intervals. See other notes to Figure 2.

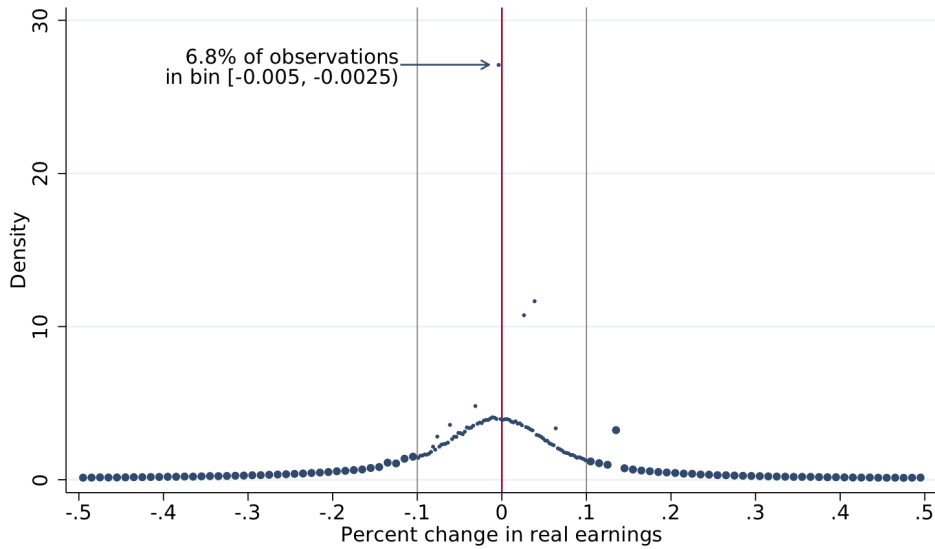
Figure 7: Permutation Test



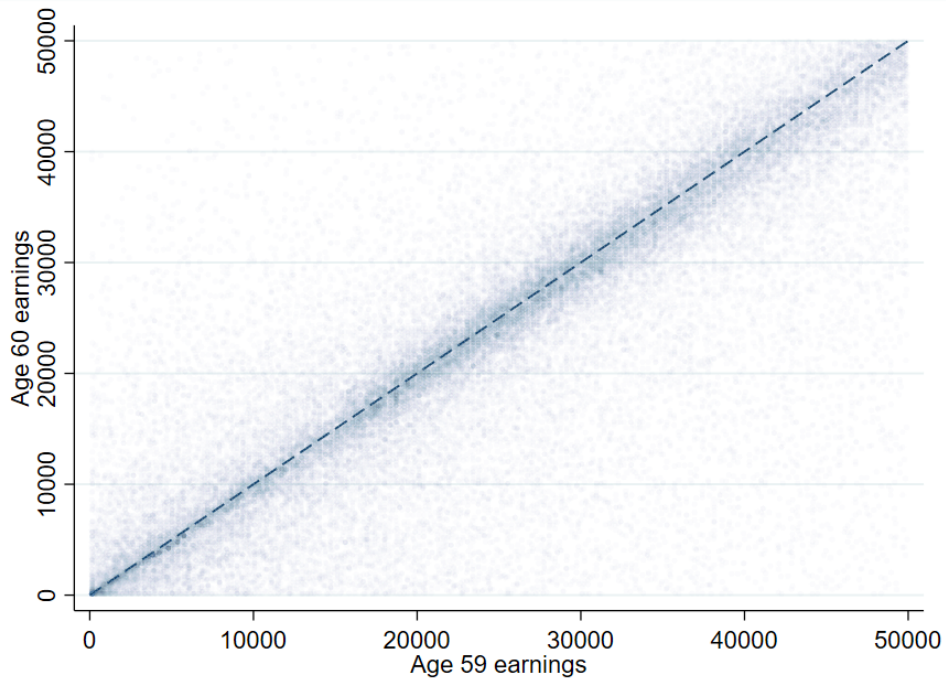
Notes: The figure plots the histogram of placebo kink estimates from a permutation test in the spirit of Ganong and Jäger (2015). We estimate a set of placebo changes in slope in the mean annual age 63 to 64 employment rate, using the same specification as our main estimates except that we examine the change in slope at placebo locations of the exempt amount away from the true exempt amount. The figure shows that the point estimate recovered from the actual exempt amount, shown by the vertical line in the figure, is located well below the significant majority of the distribution of placebo estimates, reinforcing the view that we are detecting a true effect of the AET. This permutation test shows a significant kink ( $p=0.025$ ). See other notes to Figure 2.

Figure 8: Extensive Margin Response by State 0 Counterfactual Earnings

Panel A: Distribution of Real Growth in Earnings, Age 59 to 60

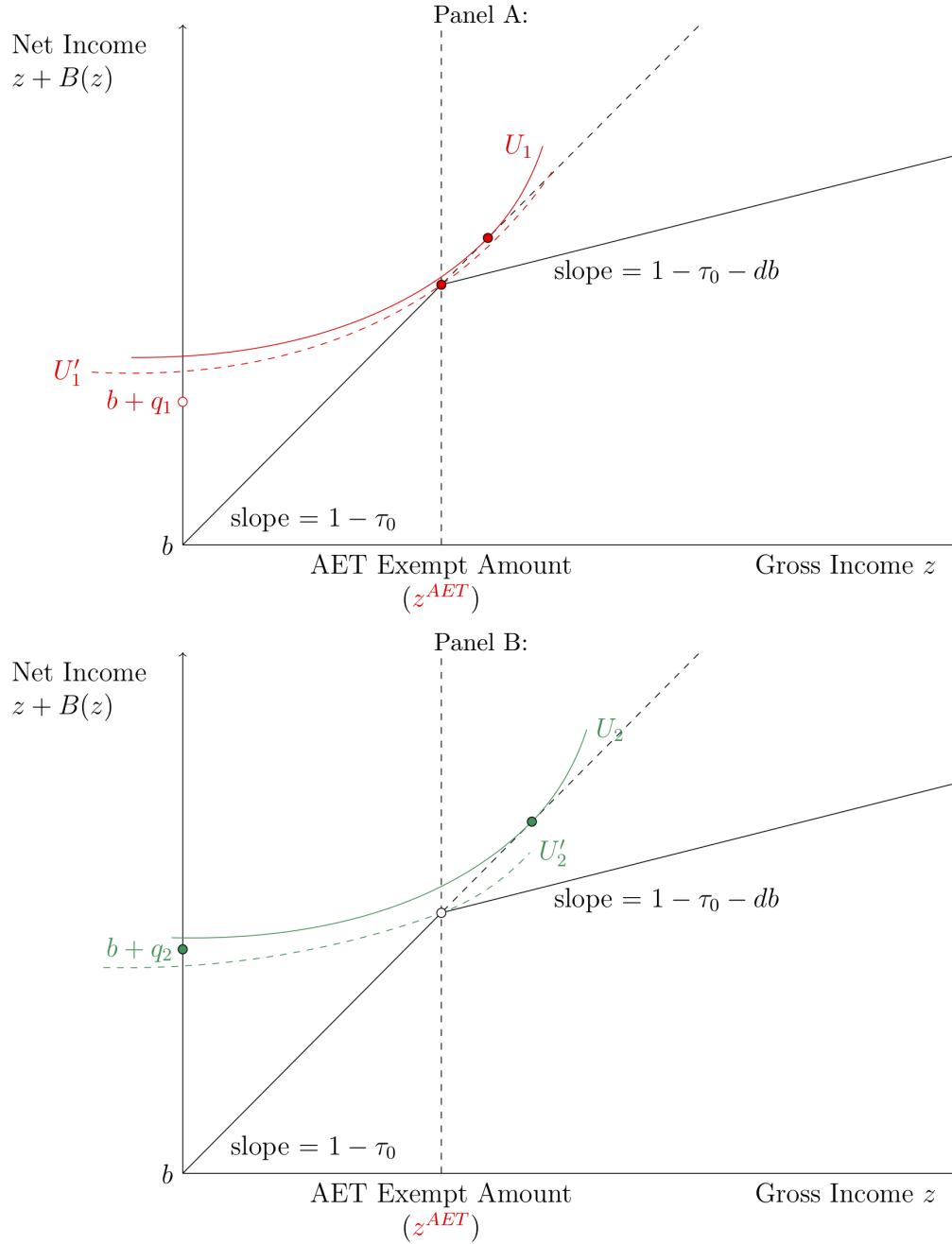


Panel B: Age 59 Earnings versus Age 60 Earnings



Notes: Panel A shows a density graph with a large mass near zero percent real earnings growth across a “placebo” set of ages, 59 and 60, when individuals do not face the AET. This indicates that a substantial mass of individuals have no growth in desired real earnings, consistent with the assumptions necessary for our RKD to estimate a lower bound on the elasticity as described in the main text. Real earnings in each year are calculated using the CPI-U. Data within -0.01 and 0.01 percent are averaged within bins of width 0.002 percent, while data outside of that range are grouped into bins of width 0.01 percent. Panel B plots age 59 earnings against age 60 earnings. The data are clustered around the dashed line, which represents the 45 degree line, again, consistent with a mass of individuals for whom there is no real earnings growth. See other notes to Figure 2. Panel B is constructed using the publicly available Earnings Public Use File, a 1 percent sample of the Master Earnings’s File, with annual earnings randomly rounded to the nearest \$100.

Figure 9: Intensive and Extensive Margin Responses to a Kink



Notes: The figure depicts potential responses to a kinked budget set. In Panel A the agent reduces earnings to the kink at  $z^{AET}$ , preferring this to the outside level of consumption. In Panel B the agent prefers the outside option of earning zero to the optimal level of earnings, conditional on being employed,  $z^{AET}$ .

Table 1: Summary Statistics

	(1)	(2)
	Within \$2,797 of Exempt Amount	Full Earnings Range
Percent of years with positive earnings, ages 63 to 64	56.53 (45.92)	63.23 (44.06)
Annual earnings, ages 63 to 64 (including zeroes)	\$5,813 (\$8,134)	\$20,087 (\$24,582)
Claim age (if claimed)	63.13 (2.61)	63.67 (2.13)
Percent female	68.00 (46.65)	41.78 (49.32)
Percent white	83.96 (36.70)	89.32 (30.88)
Year of birth	1920.62 (1.68)	1920.52 (1.69)
<i>N</i>	95,960	1,424,667

Notes: The table shows the means of each variable shown in the row headings, with the standard deviation in parentheses, from a 25 percent random sample of the Social Security Administration Master Earnings File and Master Beneficiary Record. The sample is all individuals born 1918-1923 with positive earnings at age 60, and satisfying the other sample restrictions described in the main text. In Column 1, the bandwidth relative to the AET exempt amount, \$2,797, corresponds to the bandwidth used in our main regressions. For comparison, Column 2 shows individuals from the full range of earnings. Earnings are in thousands of real 2010 dollars. *N*'s refer to the number of individuals.

Table 2: Initial Tests of Smoothness

	(1)	(2)	(3)	(4)
	Estimated Kink			
	Density	White (Indicator)	Female (Indicator)	Year of Birth
Linear	-0.057 (0.060)	-0.0005 (0.36)	-0.35 (0.34)	0.052 (0.018)***
<i>N</i>	141,110	172,383	330,556	102,640
Quadratic	-0.26 (0.37)	-0.74 (1.48)	-0.0069 (0.43)	0.074 (0.017)***
<i>N</i>	141,110	696,139	348,920	304,306
Cubic	-0.26 (0.37)	-0.98 (0.79)	-2.25 (0.72)***	0.038 (0.028)
<i>N</i>	141,110	282,460	362,174	308,690

Notes: The table presents tests for the smoothness of the earnings density, and of three pre-determined covariates (indicators for being white or female, and year of birth). The table reports coefficients corresponding to  $\beta_1$  in equation (4), reflecting the change in slope in the outcome at the age 60 exempt amount, as a function of age 60 earnings relative to the exempt amount. The column headings indicate the outcome variable, while the row headings indicate the specification of the running variable (linear, quadratic, or cubic). The number of individuals *N* in each regression is shown below the standard error. In Columns 2 to 4, robust standard errors following Calonico *et al.* (2014) are reported in parentheses, and the bandwidth is also chosen using Calonico *et al.* (2014). The resulting difference in bandwidths is the reason for the discrepancy across specifications in the number of individuals included in these regressions. The *N*'s are the same across the specifications for the density outcome in Column 1 because following previous literature, we use the McCrary (2008) procedure to calculate the bandwidth, bin size, and standard error for the density outcome. Here and in the following tables, \*\*\* indicates  $p < 0.01$ ; \*\* indicates  $p < 0.05$ ; and \* indicates  $p < 0.10$ , all from two-tailed tests of equality with zero. See other notes to Table 1.

Table 3: Main Results Documenting Kink in Employment

	(1)	(2)	(3)
	Main RKD	With Controls	Age 56-57 (Placebo)
Linear	-1.85 (0.72)***	-1.68 (0.66)***	-0.59 (0.37)
<i>N</i>	95,960	104,665	106,241
Quadratic	-2.47 (1.02)***	-2.26 (0.89)***	0.11 (0.53)
<i>N</i>	160,785	172,979	277,187
Cubic	-3.11 (1.07)***	-2.57 (1.04)***	-0.27 (0.60)
<i>N</i>	273,241	326,762	407,773

Notes: The table shows regression kink estimates showing the change at the age 60 exempt amount in the slope of the mean annual probability of positive earnings at ages 63 and 64 as a function of age 60 distance to the exempt amount, corresponding to  $\beta_1$  in equation (4). Column 1 is our main specification described in the main text. In Column 1, the Calonico *et al.* (2014) bandwidth is \$2,797 for the linear specification, \$3,047 for the quadratic, and \$3,093 for the cubic. Column 2 adds indicators for demographic categories (sex, race groups, and year of birth). Column 3 shows a placebo test in which the mean probability of positive earnings at ages 56 and 57 is the dependent variable. Robust standard errors, following Calonico *et al.* (2014), are in parentheses. The number of individuals in each regression is shown below the standard error. See other notes to Table 2.

Table 4: RKD Estimates by Single Year of Age

	(1)	(2)	(3)	(4)
	Age 61	Age 62	Age 63	Age 64
Estimate of slope discontinuity	-0.33 (0.81)	-0.90 (0.91)	-1.60 (0.72)**	-1.98 (0.74)***
<i>N</i>	80,148	244,465	105,871	98,502

Notes: The table presents regression kink estimates showing the change at the age 60 exempt amount in the slope of the probability of positive earnings at each age separately from 61 to 64, as a function of the age 60 distance to the exempt amount, corresponding to  $\beta_1$  in equation (4). We use our baseline linear specification. The age in question is shown in each column heading. See other notes to Table 2.

Table 5: Elasticity Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full Sample	Men	Women	White	Non- White	High Prior Earnings	Low Prior Earnings
Elasticity	0.49 (0.19)***	0.24 (0.19)	0.49 (0.21)**	0.47 (0.17)***	0.41 (0.43)	0.48 (0.48)	0.41 (0.16)***
<i>N</i>	95,960	68,971	66,251	93,722	39,271	19,574	101,709

Notes: The table presents the estimates for the full sample and various subgroups, from the baseline linear specification. The “high prior earnings” and “low prior earnings” subgroups refer to individuals for whom mean earnings from ages 40 to 59 were above the median and below the median, respectively. The elasticity is calculated as described in the text, using the AET rules to calculate analytically the first stage change in slope of the ANTR at the exempt amount; averaging across calendar years, the mean first stage change in the slope of the ANTR at the exempt amount is -4.88. *N*’s refer to the number of individuals included in each “reduced form” regression (4). The *N*’s differ across specifications both because the samples differ, and because the the Calonico *et al.* (2014) bandwidths differ across samples. See other notes to Table 2.

# Using Non-Linear Budget Sets to Estimate Extensive Margin Responses: Evidence and Method from the Earnings Test

Alexander M. Gelber, Damon Jones, Daniel W. Sacks, and Jae Song

## Online Appendix

### A Proofs of Propositions and Other Claims

**Proposition 1:** In general the slope of  $\frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}}$  will be given by:

$$\frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}} = g(\bar{q}_{n1} | n) \frac{d\bar{q}_{n1}}{d\tilde{z}_{n0}} + \frac{\partial G(\bar{q}_{n1} | n)}{\partial n} \frac{dn}{d\tilde{z}_{n0}} \quad (\text{A.1})$$

Focusing on the first term in the expression for  $d\Pr(z_{n1} > 0 | \tilde{z}_{n0})/d\tilde{z}_{n0}$  in (A.1), we have:

$$\begin{aligned} \frac{d\bar{q}_{n1}}{d\tilde{z}_{n0}} &= \frac{\partial v(\tilde{z}_{n1} - T(\tilde{z}_{n1}) + B_1(\tilde{z}_{n1}), \tilde{z}_{n1}; n)}{\partial \tilde{z}_{n1}} \frac{d\tilde{z}_{n1}}{d\tilde{z}_{n0}} + \frac{\partial v(\tilde{z}_{n1} - T(\tilde{z}_{n1}) + B_1(\tilde{z}_{n1}), \tilde{z}_{n1}; n)}{\partial n} \frac{dn}{d\tilde{z}_{n0}} \\ &= \frac{\partial v(\tilde{z}_{n1} - T(\tilde{z}_{n1}) + B_1(\tilde{z}_{n1}), \tilde{z}_{n1}; n)}{\partial n} \frac{dn}{d\tilde{z}_{n0}} \end{aligned} \quad (\text{A.2})$$

When agents are unrestricted in their intensive margin earnings choice, we can set the first term on the right side of (A.2) to zero. For those with  $\tilde{z}_{n0} < z^{AET}$  or  $\tilde{z}_{n0} > z^{AET} + \Delta z$  we have:

$$\partial v(\tilde{z}_{n1} - T(\tilde{z}_{n1}) + B_1(\tilde{z}_{n1}), \tilde{z}_{n1}; n) / \partial \tilde{z}_{n1} = 0$$

due to the envelope theorem.<sup>32</sup> For those with  $z^{AET} \leq \tilde{z}_{n0} \leq z^{AET} + \Delta z$ , we have  $d\tilde{z}_{n1}/d\tilde{z}_{n0} = 0$ , since  $\tilde{z}_{n1} = z^{AET}$  for everyone in this set—*i.e.* these agents bunch at  $z^{AET}$ . Substituting for  $d\bar{q}_{n1}/d\tilde{z}_{n0}$  in (A.1) using (A.2), we have:

$$\frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}} = g(\bar{q}_{n1} | n) \frac{\partial v(\tilde{z}_{n1} - T(\tilde{z}_{n1}) + B_1(\tilde{z}_{n1}), \tilde{z}_{n1}; n)}{\partial n} \frac{dn}{d\tilde{z}_{n0}} + \frac{\partial G(\bar{q}_{n1} | n)}{\partial n} \frac{dn}{d\tilde{z}_{n0}} \quad (\text{A.3})$$

when individuals are able to adjust on both the intensive and extensive margins.

Our smoothness assumptions imply that this slope is continuous, and in particular it is continuous at  $z^{AET}$  since  $n$ ,  $\bar{q}_{n1}$ ,  $\tilde{z}_{n1}$ ,  $T(\cdot)$  and  $\partial G(\bar{q}_{n1} | n) / \partial n$  are all continuous in  $\tilde{z}_{n0}$  at  $z^{AET}$ . Furthermore,  $g(\cdot)$  and  $\partial v / \partial n$  are likewise continuous in their arguments. Thus, we have:

$$\lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}} = \lim_{\tilde{z}_{n0} \rightarrow z^{AET-}} \frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}} \quad (\text{A.4})$$

That is, the employment probability does not exhibit any change in slope at  $z^{AET}$ , even though the ANTR does feature such a discontinuity.

**Proposition 2:** If  $\tilde{z}_{n1} \equiv \tilde{z}_{n0}$ , the general expression for  $d\Pr(z_{n1} > 0 | \tilde{z}_{n0})/d\tilde{z}_{n0}$  from (A.1) still holds. However, we now have a slightly different expression for the critical level of fixed costs, which is now evaluated at  $\tilde{z}_{n0}$ , implying  $\bar{q}_{n1} \equiv v(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_1(\tilde{z}_{n0}), \tilde{z}_{n0}; n) - v^0$ . Accordingly, we have a different expression for  $d\bar{q}_{n1}/d\tilde{z}_{n0}$  relative to (A.2). Since  $\tilde{z}_{n1} = \tilde{z}_{n0}$  for everyone, we have:

$$\frac{d\bar{q}_{n1}}{d\tilde{z}_{n0}} = \frac{\partial v(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_1(\tilde{z}_{n0}), \tilde{z}_{n0}; n)}{\partial \tilde{z}_{n0}} + \frac{\partial v(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_1(\tilde{z}_{n0}), \tilde{z}_{n0}; n)}{\partial n} \frac{dn}{d\tilde{z}_{n0}} \quad (\text{A.5})$$

where the key difference is that  $\partial v / \partial z$  and  $\partial v / \partial n$  are evaluated at  $\tilde{z}_{n0}$  instead of  $\tilde{z}_{n1}$ . For those with  $\tilde{z}_{n0} < z^*$ , since  $B_1(\cdot) = B_0(\cdot)$  and  $\tilde{z}_{n1} = \tilde{z}_{n0}$ , it is still the case that  $\partial v(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_1(\tilde{z}_{n0}), \tilde{z}_{n0}; n) / \partial \tilde{z}_{n0} = 0$

<sup>32</sup>In this and other similar expressions elsewhere we evaluate the partial derivative of  $v$  with respect to  $z$  allowing both earnings and consumption to change via the budget constraint, but holding  $n$  constant.

due to the envelope theorem. However, the first term in (A.5) for those with  $\tilde{z}_{n0} > z^{AET}$  is now:

$$\begin{aligned} \frac{\partial v(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_1(\tilde{z}_{n0}), \tilde{z}_{n0}; n)}{\partial \tilde{z}_{n0}} &= (1 - \tau_0 - db) v_c + v_z \\ &= \lambda_n \left[ (1 - \tau_0 - db) + \frac{v_z}{v_c} \right] \end{aligned} \quad (\text{A.6})$$

where  $\lambda_n \equiv v_c$ , and  $v_c$  and  $v_z$  are the partial derivatives of  $v(\cdot)$  with respect to  $c$  and  $z$ , respectively, evaluated at  $(\tilde{z}_{n0} - T_1(\tilde{z}_{n0}) + B_1(\tilde{z}_{n0}), \tilde{z}_{n0}; n)$ , and  $db$  is the benefit reduction rate above  $z^{AET}$ .

Thus, we now have:

$$\frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}} = \begin{cases} g(\bar{q}_{n1} | n) \frac{\partial v(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_1(\tilde{z}_{n0}), \tilde{z}_{n0}; n)}{\partial n} \frac{dn}{d\tilde{z}_{n0}} + \frac{\partial G(\bar{q}_{n1} | n)}{\partial n} \frac{dn}{d\tilde{z}_{n0}}, & \text{if } \tilde{z}_{n0} < z^{AET} \\ g(\bar{q}_{n1} | n) \left[ \lambda_n \left[ (1 - \tau_0 - db) + \frac{v_z}{v_c} \right] + \frac{\partial v(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_1(\tilde{z}_{n0}), \tilde{z}_{n0}; n)}{\partial n} \frac{dn}{d\tilde{z}_{n0}} \right] + \frac{\partial G(\bar{q}_{n1} | n)}{\partial n} \frac{dn}{d\tilde{z}_{n0}}, & \text{if } \tilde{z}_{n0} \geq z^{AET} \end{cases} \quad (\text{A.7})$$

Note also that:

$$\lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \frac{v_z(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_1(\tilde{z}_{n0}), \tilde{z}_{n0}; n)}{v_c(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_1(\tilde{z}_{n0}), \tilde{z}_{n0}; n)} = -(1 - \tau_0) \quad (\text{A.8})$$

where we have used the first order condition for  $\tilde{z}_{n,0}$  and the fact that  $\lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} B_1(\tilde{z}_{n0}) = B_0(z^{AET})$ . We now have the following expression for the difference in slopes at  $z^{AET}$ :

$$\begin{aligned} \lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}} - \lim_{\tilde{z}_{n0} \rightarrow z^{AET-}} \frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}} &= \lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} g(\bar{q}_{n1} | n) \cdot \lambda_n \left[ (1 - \tau_0 - db) + \frac{v_z}{v_c} \right] \\ &= g(\bar{q}_{n^{AET1}} | n^{AET}) \cdot \lambda_{n^{AET}} [(1 - \tau_0 - db) - (1 - \tau_0)] \\ &= -db \cdot \lambda_{n^{AET}} \cdot g(\bar{q}_{n^{AET1}} | n^{AET}) \end{aligned} \quad (\text{A.9})$$

where  $\bar{q}_{n^{AET1}}$ ,  $n^{AET}$ , and  $\lambda_{n^{AET}}$  all refer the individual for whom  $\tilde{z}_{n0} = z^{AET}$ .

**Estimation of Kink with Measurement Error:** In Section 3.2, we argue that in the case of measurement error, i.e.  $\pi(\tilde{z}_0, v_{it}) < 1$ , we estimate a lower bound on our elasticity. Here we formally demonstrate this result. Recall that our relationship between earnings at age 60 and desired earnings at ages 63-64 is  $z^{60} = \tilde{z}_{0it} + p_{it}v_{it}$ . We assume that the joint distribution of  $(p, v)$ , conditional on  $z^{60}$ , is continuous and continuously differentiable in  $z^{60}$  at  $z^{60} = z^{AET}$ . Consider the estimated kink in employment, as a function of earnings at age 60, i.e. the numerator in equation (3), using  $z^{60}$  as the running variable. Denote this as  $\beta^{RKD}$ :

$$\beta^{RKD} \equiv \lim_{z_i^{60} \rightarrow z^{AET+}} \frac{\partial \Pr(z_{it} > 0 | z_i^{60})}{\partial z_i^{60}} - \lim_{z_i^{60} \rightarrow z^{AET-}} \frac{\partial \Pr(z_{it} > 0 | z_i^{60})}{\partial z_i^{60}}$$

Let the potential deviation in desired earnings,  $v$ , have a distribution with the CDF  $M(v|z)$  and pdf  $m(v|z)$ , conditional on  $z^{60} = z$ . The sample with  $z^{60} = z^{AET}$  that we use to estimate our kink is comprised of two groups. The first group draws  $p_{it} = 0$  with probability  $\pi(z^{AET}, v)$ , and therefore have  $\tilde{z}_{0it} = z^{AET}$ . The second draw  $p_{it} = 1$  with probability  $1 - \pi(z^{AET} - v, v)$  and therefore have  $\tilde{z}_{0it} = z^{AET} - v_{it}$ . We



therefore have the following observed kink in employment:

$$\begin{aligned}
\beta^{RKD} &= \int \pi(z^{AET}, v) \left[ \lim_{z_i^{60} \rightarrow z^{AET+}} \frac{\partial \Pr(z_{it} > 0 | z_i^{60})}{\partial z_i^{60}} - \lim_{z_i^{60} \rightarrow z^{AET-}} \frac{\partial \Pr(z_{it} > 0 | z_i^{60})}{\partial z_i^{60}} \right] m(v | z^{AET}) dv \\
&+ \int (1 - \pi(z^{AET} - v, v)) \left[ \lim_{z_i^{60} \rightarrow z^{AET+}} \frac{\partial \Pr(z_{it} > 0 | z_i^{60})}{\partial z_i^{60}} - \lim_{z_i^{60} \rightarrow z^{AET-}} \frac{\partial \Pr(z_{it} > 0 | z_i^{60})}{\partial z_i^{60}} \right] m(v | z^{AET}) dv \\
&= \int \pi(z^{AET}, v) \underbrace{\left[ \lim_{\tilde{z}_{i0t} \rightarrow z^{AET+}} \frac{\partial \Pr(z_{it} > 0 | \tilde{z}_{i0t})}{\partial \tilde{z}_{i0t}} - \lim_{\tilde{z}_{i0t} \rightarrow z^{AET-}} \frac{\partial \Pr(z_{it} > 0 | \tilde{z}_{i0t})}{\partial \tilde{z}_{i0t}} \right]}_{\hat{\beta}(z^{AET})} m(v | z^{AET}) dv \\
&+ \int \left\{ (1 - \pi(z^{AET} - v, v)) \cdot \right. \\
&\quad \left. \underbrace{\left[ \lim_{\tilde{z}_{i0t} \rightarrow z^{AET+} - v_{it}} \frac{\partial \Pr(z_{it} > 0 | \tilde{z}_{i0t})}{\partial \tilde{z}_{i0t}} - \lim_{\tilde{z}_{i0t} \rightarrow z^{AET-} - v_{it}} \frac{\partial \Pr(z_{it} > 0 | \tilde{z}_{i0t})}{\partial \tilde{z}_{i0t}} \right]}_{\hat{\beta}(z^{AET} - v)} m(v | z^{AET}) \right\} dv \\
&= \bar{\pi} \cdot \hat{\beta}(z^{AET}) + (1 - \bar{\pi}) \int \frac{1 - \pi(z^{AET} - v, v)}{1 - \bar{\pi}} \hat{\beta}(z^{AET} - v) m(v | z^{AET}) dv
\end{aligned}$$

where

$$\begin{aligned}
\bar{\pi} &= \int \pi(z^{AET}, v) m(v | z^{AET}) dv \\
\hat{\beta}(z) &= \lim_{\tilde{z}_{i0t} \rightarrow z} \frac{\partial \Pr(z_{it} > 0 | \tilde{z}_{i0t})}{\partial \tilde{z}_{i0t}} - \lim_{\tilde{z}_{i0t} \rightarrow z} \frac{\partial \Pr(z_{it} > 0 | \tilde{z}_{i0t})}{\partial \tilde{z}_{i0t}}
\end{aligned}$$

That is,  $\bar{\pi}$  is the probability that  $p_{it} = 0$  among individuals with  $z_i^{60} = z^{AET}$  and  $\hat{\beta}(z)$  is the kink in employment at  $z$  that would be estimated if desired earnings,  $\tilde{z}_{i0t}$ , were observed.

Given our assumptions on smoothness in heterogeneity and assuming the only locally relevant kink in the budget is at  $z^{AET}$ , we have:

$$\hat{\beta}(z) = \begin{cases} \beta & \text{if } z = z^{AET} \\ 0 & \text{if } z \neq z^{AET} \end{cases}$$

Here  $\beta$  is the kink in employment among those who face a kink, as defined in equation (14). Thus, we have:

$$\beta^{RKD} = \bar{\pi} \cdot \beta$$

It follows that if  $\pi(z, v) < 1$ , *i.e.* we have measurement error in desired earnings when using  $z^{60}$  as a proxy, then we estimate a lower bound on our elasticity when using the formula in equation (14). In other words, the estimated kink among those with age 60 earnings near  $z^{AET}$  reflects a weighted average of the subset of individuals who will actually face the kink at ages 63-64 and the subset who will not. Furthermore, when  $\pi(z, v) = 0$ , *i.e.* there is no persistence in earnings from age 60 to ages 63-64, we expect no employment kink. Note, *i.e.* we only require that the distribution of  $v$  be smooth as a function of  $z^{60}$ . Thus, we can allow for cases where  $v$  is not mean zero, *e.g.* where there is a systematic change in mean desired earnings from age 60 to ages 63-64. Note as well that we can relax the assumption that the only kink in the tax schedule is at  $z^{AET}$  and can instead assume that the set of other possible kinks is of measure zero.

## B Model Extensions

### B.1 Extension to Two or More Discrete Job Choices

We begin by focusing on two key options within the menu of positive earnings: one at the interior optimum in state 0,  $\tilde{z}_{n0}$ , and another “next-best” job at an alternative level of earnings,  $\tilde{z}_n^{nb}$ . This allows for an arbitrary number (whether finite or infinite) of discrete choices that are less preferred than the “next-best” job. The model can also be easily extended to allow for the possibility that multiple earnings levels give the

same “next best” utility level as  $\tilde{z}_n^{nb}$ .

Before showing this result more formally, we briefly illustrate the intuition. Consider an individual whose optimal earnings is just above  $z^{AET}$  under a linear tax and no benefit reduction above  $z^{AET}$ , i.e. in the absence of the kink. For the moment, suppose this person can either earn just above  $z^{AET}$  or exit. Now, we introduce a kink at  $z^{AET}$ . The effect of the kink on the average net of tax rate for the person just above  $z^{AET}$  is vanishingly small. So, if this person now decides to exit, they must have been virtually indifferent between  $z^{AET}$  and exiting prior to the kink. Now, let's suppose instead that a part-time job is available at some earnings partway between 0 and  $z^{AET}$ . Also, suppose that in the absence of the kink, the person continues to earn just above  $z^{AET}$ , rather than exiting or taking the part-time job. By revealed preference, we learn that the job earning just above  $z^{AET}$  is preferred to the part-time job. Furthermore, since we have established that this person is virtually indifferent between earning just above  $z^{AET}$  or exiting, it follows that exiting is also preferred to the part-time job. Thus, when we again introduce the kink, the features of the part-time job are irrelevant for the decision to exit or not. If, in the absence of the kink, exiting is preferred to the part-time job, it will continue to be preferred once the kink is introduced.

Formally, let  $v_n^{nb}$  be the utility level associated with the next-best level of earnings:

$$v_n^{nb} \equiv v(\tilde{z}_n^{nb} - T(\tilde{z}_n^{nb}) + B(\tilde{z}_n^{nb}), \tilde{z}_n^{nb}; n) \quad (\text{B.10})$$

As it is possible that the kink lowers utility at  $\tilde{z}_{n0}$  while leaving utility at  $\tilde{z}_n^{nb}$  unaffected, the probability of working in state 1 is one minus the probability that non-employment is preferred to both the earnings level  $\tilde{z}_{n0}$  and the earnings level  $\tilde{z}_n^{nb}$  of the next-best job:

$$\begin{aligned} \Pr(z_{n1} > 0 | \tilde{z}_{n0}) &= 1 - \Pr(v^0 \geq v_n^{nb} - q_{n1} \text{ and } v^0 \geq v(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_1(\tilde{z}_{n0}), \tilde{z}_{n0}; n) - q_{n1}) \\ &= 1 - \Pr(v^0 \geq v_n^{nb} - q_{n1} | \bar{q}_{n1} < q_{n1}) \cdot \Pr(\bar{q}_{n1} < q_{n1}) \\ &= 1 - \Pr(v^0 \geq v_n^{nb} - q_{n1} | \bar{q}_{n1} < q_{n1}) \cdot [1 - G(\bar{q}_{n1} | n)] \end{aligned} \quad (\text{B.11})$$

The slope of the employment function is now a more complex expression:

$$\begin{aligned} \frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}} &= - \frac{d\Pr(v^0 \geq v_n^{nb} - q_{n1} | \bar{q}_{n1} < q_{n1})}{d\tilde{z}_{n0}} [1 - G(\bar{q}_{n1} | n)] \\ &\quad + \Pr(v^0 \geq v_n^{nb} - q_{n1} | \bar{q}_{n1} < q_{n1}) \cdot \left[ \frac{dG(\bar{q}_{n1} | n)}{d\tilde{z}_{n0}} \right] \end{aligned} \quad (\text{B.12})$$

We now explore under what conditions this slope reduces to that of our earlier model in Section 6.3, in which intensive margin earnings in state 1 are constrained at their state 0 level. We will show in general that this is true for those with state 0 earnings below  $z^{AET}$ . Next, we show that for those with state 0 earnings just above  $z^{AET}$ , the slope is likewise unaffected relative to the model in Section 6.3, as long as the next-best job offers a level of earnings that is discretely different than the new interior optimum  $z^{AET}$ .

Consider individuals earning  $\tilde{z}_{n0} < z^{AET}$  in state 0. We first focus on the term  $\Pr(v^0 \geq v_n^{nb} - q_{n1} | \bar{q}_{n1} < q_{n1})$ . We can show the following for the agents in this set for whom  $\bar{q}_{n1} < q_{n1}$ :

$$\begin{aligned} v^0 &> v(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_1(\tilde{z}_{n0}), \tilde{z}_{n0}; n) - q_{n1} \\ &= v(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_0(\tilde{z}_{n0}), \tilde{z}_{n0}; n) - q_{n1} \\ &> v_n^{nb} - q_{n1} \end{aligned} \quad (\text{B.13})$$

where in the first line we used the fact that  $\bar{q}_{n1} < q_{n1}$  and the definition of  $\bar{q}_{n1}$  in equation (11). In the second line we used the fact that  $B_1(\tilde{z}_{n0}) = B_0(\tilde{z}_{n0})$  for individuals with  $\tilde{z}_{n0} < z^{AET}$ . In the third line we used the fact that  $v(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_0(\tilde{z}_{n0}), \tilde{z}_{n0}; n) \geq v_n^{nb}$  due to revealed preference in state 0. It follows that:

$$\begin{aligned} \Pr(v^0 \geq v_n^{nb} - q_{n1} | \bar{q}_{n1} < q_{n1}, \tilde{z}_{n0} < z^{AET}) &= 1 \\ \Rightarrow \frac{d\Pr(v^0 \geq v_n^{nb} - q_{n1} | \bar{q}_{n1} < q_{n1}, \tilde{z}_{n0} < z^{AET})}{d\tilde{z}_{n0}} &= 0 \end{aligned} \quad (\text{B.14})$$

In other words, if an individual with state 0 earnings below  $z^{AET}$  prefers the outside option in the absence of the next-best job, she would continue to prefer it in the presence of the next-best job.

Using the results in (B.14), we can simplify the expression in (B.12), for those with  $\tilde{z}_{n0} < z^{AET}$ :

$$\begin{aligned} \frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}} &= \frac{dG(\bar{q}_{n1} | n)}{d\tilde{z}_{n0}} \\ &= g(\bar{q}_{n1} | n) \frac{\partial v(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_1(\tilde{z}_{n0}), \tilde{z}_{n0}; n)}{\partial n} \frac{dn}{d\tilde{z}_{n0}} + \frac{\partial G(\bar{q}_{n1} | n)}{\partial n} \frac{dn}{d\tilde{z}_{n0}} \end{aligned} \quad (\text{B.15})$$

where the second line follows from equation (A.7). Thus, the presence of a menu of discrete options does not affect the results for those with  $\tilde{z}_{n0} < z^{AET}$ . Intuitively, after the introduction of a kink, the individual's state 0 optimal earnings amount  $\tilde{z}_{n0}$  is still available, at the same level of utility, and thus is the only positive earnings level relevant for extensive margin decisions.

Now consider individuals with  $\tilde{z}_{n0} \geq z^{AET}$  and recall that we are ultimately interested in the change in slope of the employment rate at  $z^{AET}$ . Any change in the slope of the employment function at  $z^{AET}$  will depend on the following limit:

$$\begin{aligned} \lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \frac{d\Pr(z_{n1} > 0 | z_{n0})}{dz_{n0}} &= \left[ - \lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \frac{d\Pr(v^0 \geq v_n^{nb} - q_{n1} | \bar{q}_{n1} < q_{n1})}{d\tilde{z}_{n0}} \right] [1 - G(\bar{q}_{n^{AET1}} | n^{AET})] \\ &\quad + \left[ \lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \Pr(v^0 \geq v_n^{nb} - q_{n1} | \bar{q}_{n1} < q_{n1}) \right] \left( \frac{dG(\bar{q}_{n^{AET1}} | n^{AET})}{d\tilde{z}_{n0}} \right) \end{aligned} \quad (\text{B.16})$$

Note the following:

$$\begin{aligned} \lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \Pr(v^0 \geq v_n^{nb} - q_{n1} | \bar{q}_{n1} < q_{n1}) &= \lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \Pr(v^0 \geq v_n^{nb} - q_{n1} | \\ &\quad v^0 \geq v(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_1(\tilde{z}_{n0}), \tilde{z}_{n0}; n) - q_{n1}) \\ &= \Pr(v^0 \geq v_{n^{AET}}^{nb} - q_{n^{AET1}} | \\ &\quad v^0 \geq v(z^{AET} - T(z^{AET}) + B_0(z^{AET}), z^{AET}; n^{AET}) - q_{n^{AET1}}) \\ &= 1 \end{aligned} \quad (\text{B.17})$$

In the second line, we used the fact that  $\lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \tilde{z}_{n0} = z^{AET}$  and  $B_1(z^{AET}) = B_0(z^{AET})$ , and the final line follows from revealed preference:  $z^{AET}$  was initially chosen over the next-best job.

We require that in the neighborhood of  $z^{AET}$  the earnings level offered at the next-best job be discretely different than that of the state 0 job; thus, we assume that:

$$\lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \tilde{z}_n^{nb} \neq z^{AET} \quad (\text{B.18})$$

This assumption rules out alternative jobs that can be made arbitrarily close to the level of state 0 earnings. Intuitively, if this were not so then individuals earning just above  $z^{AET}$  in state 0 would be able to replicate intensive margin adjustment, which we have shown smooths out any kink in the employment function that would otherwise exist.

The assumption in (B.18) implies that the limit in (B.17) is reached at some level of state 0 earnings strictly above  $z^{AET}$ . That is, as we approach  $z^{AET}$  from above, the probability that preferring the outside option without an alternative job implies preferring it in the presence of the next-best job plateaus at 1 at some point before reaching  $z^{AET}$ . Thus, we have:

$$\lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \frac{d\Pr(v^0 \geq v_n^{nb} - q_{n1} | \bar{q}_{n1} < q_{n1})}{d\tilde{z}_{n0}} = 0 \quad (\text{B.19})$$

As before, (B.17) and (B.19) can be used to simplify (B.16) as follows:

$$\begin{aligned}
\lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}} &= \lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \frac{dG(\bar{q}_{n1} | n)}{d\tilde{z}_{n0}} \\
&= g(\bar{q}_{n^{AET1}} | n^*) \left[ -db \cdot \lambda_{n^*} \right. \\
&\quad \left. + \frac{\partial v(\tilde{z}_{n^{AET0}} - T(\tilde{z}_{n^{AET0}}) + B_1(\tilde{z}_{n^{AET0}}), \tilde{z}_{n^{AET0}}; n^{AET})}{\partial n} \frac{dn}{d\tilde{z}_{n0}} \right] \\
&\quad + \frac{dG(\bar{q}_{n^{AET1}} | n^{AET})}{dn} \frac{dn}{d\tilde{z}_{n0}} \tag{B.20}
\end{aligned}$$

where the second line follows from equation (A.7). Combining (B.20) and (B.15), we have the same result as equation (13) of Section 6.3, *i.e.* our earlier model with  $\tilde{z}_{n1} = \tilde{z}_{n0}$ . Note that this result features as a special case the scenario in which the agent has the choice among a full-time job at  $\tilde{z}_{n0}$ , a part-time job at some lower level of earnings, or not working.

## B.2 Allowing for both Bunching at the Kink in the Budget Set and a Kink in the Employment Rate

As discussed in Section 6.5, our baseline models with either completely constrained earnings levels or discrete but limited earnings options do not allow for bunching, which is in clear violation of the empirical evidence (See Gelber et al., 2020). In this section we outline two separate cases where it is possible to observe bunching among a subset of individuals and also a kink in the employment rate overall. In both cases, we establish that the observed elasticity that is estimated is a lower bound for the structural elasticity among everyone near the kink in the budget set.

### B.2.1 Model with Mixture of Types

One approach to capturing both bunching and a kink in the employment probability is to posit a model with two types of individuals: Type *A* that can adjust on the intensive margin, and Type *NA* that cannot (*e.g.* Kleven and Waseem, 2013). We have shown that among Type *A* agents, the employment function has a continuous slope. Among Type *NA* agents, the slope is discontinuous at  $z^{AET}$ . Let  $\pi_{NA}^{AET} = \Pr(NA | \tilde{z}_{n0} = z^{AET})$  be the probability of being Type *NA* conditional on having earnings at  $z^{AET}$  in state 0. It follows that:

$$\lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}} - \lim_{\tilde{z}_{n0} \rightarrow z^{AET-}} \frac{d\Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d\tilde{z}_{n0}} = -\pi_{NA}^{AET} \cdot db \cdot g_{NA}(\bar{q}_{n^{AET1}} | n^{AET}) \tag{B.21}$$

where  $g_{NA}(\cdot)$  is the pdf of fixed costs among Type *NA* agents. In this sense, our estimate of the extensive margin elasticity is attenuated by a factor  $\pi_{NA}^{AET}$  and can therefore be considered a weak lower bound on the elasticity among Type *NA* agents with state 0 earnings  $z^{AET}$ . Among Type *A* agents who earn above  $z^{AET}$  in state 0, there may also be a response to the kink that increases gradually above  $z^{AET}$  as in Figure 1 Panel B, but our method only picks up responses among Type *NA* agents. Nonetheless, the observed elasticity is a lower bound on the elasticity among all of those earning  $z^{AET}$  in state 1:

$$\hat{\eta} = \frac{\hat{\beta}}{\alpha} \cdot \frac{1 - a}{\Pr(z_{n1} > 0 | \tilde{z}_{n0} = z^{AET})} = \pi_{NA}^{AET} \cdot \eta_{NA}^{AET} \leq \pi_A^{AET} \cdot \eta_A^{AET} + \pi_{NA}^{AET} \cdot \eta_{NA}^{AET} = \eta^{AET} \tag{B.22}$$

In principle it would be possible to use the observed elasticity  $\hat{\eta}$ , together with an estimate of the fraction constrained  $\pi_{NA}^{AET}$ , to estimate the structural elasticity among constrained agents,  $\eta_{NA}^{AET} = \hat{\eta} / \pi_{NA}^{AET}$ . However, estimating  $\pi_{NA}^{AET}$  requires more restrictive assumptions, including assuming that types *A* and *NA* have the same distribution of  $q$  and  $n$ , as we explain in Appendix B.3. Our observed elasticity remains of interest regardless of the underlying proportion of agents displaying different types of behavior, both in the sense that policy-makers are interested in the raw employment effects of changing policy parameters like the average tax rate, and in the sense that it reflects a lower bound on the structural elasticity.

## B.2.2 Model with a Fixed Cost of Intensive Margin Adjustment

In an alternative model of intensive margin frictions, individuals face a fixed cost of adjusting earnings on the intensive margin in response to variation in the tax schedule (see Gelber *et al.* (2020), for a detailed exposition of this model). Such frictions could reflect a variety of factors, including lack of knowledge of a tax regime, the cost of negotiating a new contract with an employer, or the time and financial cost of job search. With a fixed intensive margin adjustment cost individuals will only adjust if the utility gain of intensive margin adjustment exceeds the fixed cost. Recall that for individuals earning  $\tilde{z}_{n0} < z^{AET}$  there is no change in the tax schedule from state 0 to state 1, and therefore  $\tilde{z}_{n1} = \tilde{z}_{n0}$ . Gelber *et al.* (2020) show that due to the fixed cost of intensive margin adjustment individuals with  $\tilde{z}_{n0} > z^{AET}$  for whom  $\tilde{z}_{n0}$  is sufficiently close to  $z^{AET}$  will also prefer to keep earnings fixed across the two tax schedules, *i.e.*  $\tilde{z}_{n1} = \tilde{z}_{n0}$ . The reason is that the utility gain from adjusting on the intensive margin converges to zero as  $\tilde{z}_{n0}$  approaches  $z^{AET}$ : the optimal level of earnings is  $z^{AET}$  in state 1 for this group. In this case, in a close enough neighborhood around  $z^{AET}$ , individuals behave as in Section 6.3, and our results from Section 6.3 follow. In other words, a fixed cost of intensive margin adjustment can rationalize the assumption that some individuals do not adjust to  $z^{AET}$  in state 1, and it follows that the observed elasticity reflects the structural elasticity.<sup>33</sup>

## B.3 Jointly Estimating the Structural Elasticity among Constrained Types and $\pi_B^{AET}$

In the model presented in Section B.2, we may wish to estimate the structural elasticity among constrained types,  $\eta_{NA}^{AET}$ . We may be able to use data on extensive margin responses between states 0 and 1 along with evidence on intensive margin responses in state 1 to perform this decomposition. In particular, we continue to draw on the kink in employment in state 1. We also use the amount of bunching in the second period, which is related to intensive margin response among those who are not constrained. In addition, we estimate a second kink, this time in the average net-of-tax rate in state 1. The idea is that we have an analytical expression for this kink under complete frictions. The extent to which the observed kink in the average net-of-tax rate deviates from that quantity is a function of the share of the sample that faces intensive margin frictions. Finally, this method will require more restrictive assumptions on the underlying primitives, as explained below.

For notational convenience, define the set  $R = \{n | z^{AET} < z_{n,0} < z^{AET} + \Delta z\}$  as the set of individuals in state 0 with earnings in the range that bunches under a kink in the absence of intensive margin frictions. Define  $N_{R,0}$  as the number of individuals in this range in state 0. As before, denote Type *A* earners as those who can adjust on the intensive margin and Type *NA* earners as those who cannot, in state 1. Define  $N_{RA,0}$  and  $N_{RNA,0}$  as the number of Type *A* and Type *NA* earners in the set  $R$ , respectively. It follows that:

$$N_{R,0} = N_{RA,0} + N_{RNA,0}$$

Similarly, define  $N_{R,1}$  as the number of individuals in the set  $R$  that are still employed in state 1. That is  $\{n | z^{AET} < z_{n,0} < z^{AET} + \Delta z, z_{n,1} > 0\}$ . Again, define  $N_{RA,1}$  and  $N_{RNA,1}$  as the number of Type *A* and Type *NA* earners in the set  $R$  that are also employed in state 1. Again, we have:

$$N_{R,1} = N_{RA,1} + N_{RNA,1}$$

Finally, define  $N_{Bunch}$  as the number of individuals in the set  $R$  that bunch in state 1. Note that:

$$N_{Bunch} = N_{RA,1}$$

We can show the following:

$$\begin{aligned} \Pr(z_{n,1} > 0 | z^{AET} < z_{n,0} < z^{AET} + \Delta z) &= \Pr(z_{n,1} > 0 | n \in R) \\ &= (1 - \pi_{NA}^{AET}) \Pr(z_{n,1} > 0 | n \in R, A) \\ &\quad + \pi_{NA}^{AET} \Pr(z_{n,1} > 0 | n \in R, NA) \end{aligned}$$

<sup>33</sup>If there is heterogeneity in the fixed cost of intensive margin adjustment, then as long as the fixed cost is strictly positive then the above still holds in a neighborhood near  $z^{AET}$ .

Define  $N_0$  as the total number of people in the labor force in state 0 and define  $N_1$  as the number of these people also in the labor force in state 1. If we define  $B$  as the share of all such earners bunching in state 1, then we have:

$$\begin{aligned}
B &= \frac{N_{Bunch}}{N_1} \\
&= \frac{N_{RA,1}}{N_1} \\
&= \frac{N_{RA,1}}{N_{RA,0}} \frac{N_{RA,0}}{N_{R,0}} \frac{N_{R,0}}{N_0} \frac{N_0}{N_1} \\
&= \frac{\Pr(z_{n,1} > 0 | n \in R, A) (1 - \pi_{NA}^{AET}) \times \Pr(z^{AET} < z_{n,0} < z^{AET} + \Delta z)}{\Pr(z_{n,1} > 0 | z_{n,0} > 0)}
\end{aligned}$$

In addition, if we assume that Type  $A$  and Type  $NA$  individuals have the same preferences and differ only in the ability to adjust to  $z^{AET}$ , then the only difference in employment exit between the two groups is due to the lack of intensive margin adjustment on the part of the Type  $NA$  agents. In this case, we can show the following:

$$\begin{aligned}
\Pr(z_{n,1} > 0 | n \in R, B) &= \int_{z^{AET}}^{z^{AET} + \Delta z} \frac{d\Pr(z_{n,1} > 0 | \tilde{z}_{n0} = \zeta, NA)}{d\tilde{z}_{n0}} h_0(\zeta) d\zeta \\
&= \int_{z^{AET}}^{z^{AET} + \Delta z} \left[ \frac{d\Pr(z_{n,1} > 0 | \tilde{z}_{n0} = \zeta, A)}{d\tilde{z}_{n0}} + \lambda_n \left[ (1 - \tau_0 - db) + \frac{v_z}{v_c} \right] g(\bar{q}_{n,1} | n) \right] h_0(\zeta) d\zeta \\
&= \Pr(z_{n,1} > 0 | n \in R, A) + \int_{z^{AET}}^{z^{AET} + \Delta z} \left[ \lambda_n \left[ (1 - \tau_0 - db) + \frac{v_z}{v_c} \right] g(\bar{q}_{n,1} | n) \right] h_0(\zeta) d\zeta,
\end{aligned}$$

where  $h_0(\cdot)$  is the density of earnings  $\tilde{z}_{n0}$  in state 0. The second line follows from equation (A.7). We will use a first-order approximation, assuming  $\lambda_n \left[ (1 - \tau_0 - db) + \frac{v_z}{v_c} \right] g(\bar{q}_{n,1} | n)$  is constant in the set  $R$ . Then we have:

$$\begin{aligned}
\int_{z^{AET}}^{z^{AET} + \Delta z} \left[ \lambda_n \left[ (1 - \tau_0 - db) + \frac{v_z}{v_c} \right] g(\bar{q}_{n,1} | n) \right] h_0(\zeta) d\zeta &\approx -db \cdot \lambda_{n^{AET}} \cdot g(\bar{q}_{n^{AET},1} | n = n^{AET}) \int_{z^{AET}}^{z^{AET} + \Delta z} h_0(\zeta) d\zeta \\
&= \frac{\beta^{RKD}}{\pi_{NA}^{AET}} \cdot \Pr(z^{AET} < z_{n,0} < z^{AET} + \Delta z)
\end{aligned}$$

Generally,  $\lambda_n$  and  $v_z/v_c$  are decreasing over this range. If  $g(\bar{q}_{n,1} | n)$  is also weakly decreasing, then our first-order approximation will overstate the difference between  $\Pr(z_{n,1} > 0 | n \in R, NA)$  and  $\Pr(z_{n,1} > 0 | n \in R, A)$ .

Now, consider estimating the following:

$$\begin{aligned}
\beta^{ATR} &\equiv \lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \frac{d\mathbb{E} [1 - [T(z_{n,1}) - B_1(z_{n,1})] - [T(0) - B_1(0)]] / z_{n,1} | \tilde{z}_{n0}, z_{n,1} > 0]}{d\tilde{z}_{n0}} \\
&\quad - \lim_{\tilde{z}_{n0} \rightarrow z^{AET-}} \frac{d\mathbb{E} [1 - [T(z_{n,1}) - B_1(z_{n,1})] - [T(0) - B_1(0)]] / z_{n,1} | \tilde{z}_{n0}, z_{n,1} > 0]}{d\tilde{z}_{n0}}
\end{aligned}$$

which is the difference at the exempt amount  $z^*$  in the slope of the average net-of-tax rate in state 1, as a function of state 0 earnings, among the set of individuals who do not exit employment between states 0 and 1. Note that since Type  $A$  individuals bunch, the average tax rate is constant for this group. Thus, the

difference in the slope will be zero for this group. For the Type  $NA$  individuals, the difference will be:

$$\begin{aligned}
\beta_{NA}^{ATR} &= \left[ \frac{d(1 - [[T(z_{n,1}) - B_1(z_{n,1})] - [T(0) - B_1(0)]] / z_{n,1})}{dz_{n,1}} \right. \\
&\quad \left. - \frac{d(1 - [[T(z_{n,1}) - B_0(z_{n,1})] - [T(0) - B_0(0)]] / z_{n,1})}{dz_{n,1}} \right] \Big|_{z_{n,1}=z^{AET}} \\
&= \left[ \frac{d([[B_0(z_{n,1}) - B_1(z_{n,1})] - [B_0(0) - B_1(0)]] / z_{n,1})}{dz_{n,1}} \right] \Big|_{z_{n,1}=z^{AET}} \\
&= \frac{B'_0(z^{AET})}{z^{AET}} - \frac{B_0(z^{AET})}{(z^{AET})^2} - \frac{B'_1(z^{AET})}{z^{AET}} + \frac{B_1(z^{AET})}{(z^{AET})^2} \\
&= -\frac{B'_1(z^{AET}) - B'_0(z^{AET})}{z^{AET}} \\
&= -\frac{db}{z^{AET}}
\end{aligned}$$

As a result, the average in the difference in slopes for the total group will be:

$$\begin{aligned}
\beta^{ATR} &= \frac{N_{RA,1}}{N_{R,1}} \cdot 0 + \frac{N_{RNA,1}}{N_{R,1}} \cdot \left( -\frac{db}{z^{AET}} \right) \\
&= \frac{N_{R,0}}{N_{R,1}} \frac{N_{RNA,0}}{N_{R,0}} \frac{N_{RNA,1}}{N_{RNA,0}} \cdot \left( -\frac{db}{z^{AET}} \right) \\
&= \frac{\Pr(z_{n,1} > 0 | n \in R, NA) \pi_{NA}^{AET}}{\Pr(z_{n,1} > 0 | z^{AET} < \tilde{z}_{n,0} < z^{AET} + \Delta z)} \left( -\frac{db}{z^{AET}} \right)
\end{aligned}$$

We thus have four equations:

$$\Pr(z_{n,1} > 0 | z^{AET} < z_{n,0} < z^{AET} + \Delta z) = (1 - \pi_{NA}^{AET}) \Pr(z_{n,1} > 0 | n \in R, A) + \pi_{NA}^{AET} \Pr(z_{n,1} > 0 | n \in R, NA)$$

$$B = \frac{\Pr(z_{n,1} > 0 | n \in R, A) (1 - \pi_{NA}^{AET}) \times \Pr(z^{AET} < z_{n,0} < z^{AET} + \Delta z)}{\Pr(z_{n,1} > 0 | z_{n,0} > 0)}$$

$$\beta^{RKD} = \frac{\Pr(z_{n,1} > 0 | n \in R, NA) - \Pr(z_{n,1} > 0 | n \in R, A)}{\Pr(z^{AET} < z_{n,0} < z^{AET} + \Delta z^{AET}) / \pi_{NA}^{AET}}$$

$$\beta^{ATR} = \frac{\Pr(z_{n,1} > 0 | n \in R, B) \pi_{NA}^{AET}}{\Pr(z_{n,1} > 0 | z^{AET} < z_{n,0} < z^{AET} + \Delta z)} \left( -\frac{db}{z^{AET}} \right)$$

and four unknowns:  $\Delta z$ ,  $\pi_{NA}^{AET}$ ,  $\Pr(z_{n,1} > 0 | n \in R, A)$  and  $\Pr(z_{n,1} > 0 | n \in R, NA)$ . We do not have a closed form solutions for either  $\Pr(z_{n,1} > 0 | z^{AET} < z_{n,0} < z^{AET} + \Delta z)$  or  $\Pr(z^{AET} < z_{n,0} < z^{AET} + \Delta z^{AET})$ .

However, if we estimate a flexible polynomial for the employment rate and for the density in state 0, we can numerically solve for  $\pi_{NA}^{AET}$ . This can be combined with  $\hat{\eta}$  to recover  $\eta_{NA}^{AET}$ . This requires additional assumptions relative to our estimate of the observed elasticity, which is a non-parametric lower bound on the structural elasticity.

## B.4 Fully Dynamic Extension of the Model

In this section we briefly demonstrate under what conditions our results continue to hold once our model is extended to a multi-period setting with forward-looking agents. We again have two states of the world, state 0 and state 1. In this multi-period model, the tax and benefit schedule is the same across the two states for

periods  $1, \dots, t-1$ . However, in period  $t$ , there is no benefit reduction rate in state 0 and a kink at  $z^{AET}$  in state 1 created by a benefit reduction rate on earnings above  $z^{AET}$ . The tax and benefit schedules are once again the same across the two states during periods  $t+1, \dots, T$ . (For simplicity we assume here that the tax and benefit schedules across the two states are once again the same during periods  $t+1, \dots, T$ , but the model can be easily extended to assume that in each of these periods there is no benefit reduction rate in state 0 and a kink at  $z^{AET}$  in state 1.)

We assume that preferences and the economic environment yield a dynamic programming problem as follows. In each period, individuals maximize:

$$u_t(c_{njt}, z_{njt}; n) = v(c_{njt}, z_{njt}; n) - q_{njt} \cdot 1\{z_{njt} > 0\} + V_t(A_{njt}, z_{njt}; n) \quad (\text{B.23})$$

subject to a dynamic budget constraint:

$$c_{njt} = (1 + r_{t-1}) A_{nj,t-1} + z_{njt} - T(z_{njt}) + B_j(z_{njt}) - A_{njt} \quad (\text{B.24})$$

where  $A_{njt}$  is the level of assets at the end of period  $t$ . The value function for the next period,  $V_t(\cdot)$ , may depend on the level of assets passed forward and potentially the level of earnings in the current period. For example, working today may have some effect on the choice set in the next period.

We once again index individuals by their counterfactual earnings in period  $t$  in state 0, and we focus on the probability of having positive earnings in period  $t$  in state 1, conditional on the counterfactual earnings level in state 0:  $\Pr(z_{n1t} > 0 | \tilde{z}_{n0t})$ . In addition to the assumptions we have made above in Section 6.2, we assume that the value function  $V_t(\cdot)$  is  $C^1$  in  $A$ ,  $z$ , and  $n$ . Agents choose  $c$ ,  $z$ , and  $A$  to maximize utility. The outside value of not working in period  $t$ ,  $v^0(A_{nj,t-1})$ , depends on the current level of assets and includes the continuation value of future periods. Finally, the distribution of fixed costs of working,  $G(q|n, t)$ , now depends on the time period as well.

The first-order conditions when earnings are positive are now:

$$\begin{aligned} v_z + V_z &= -\lambda(1 - T'(z) + B'(z)) \\ v_c &= V_A = \lambda \end{aligned} \quad (\text{B.25})$$

where  $\lambda$  is the marginal utility of wealth. Using these conditions, we can show that there will still be bunching in response to a kink among those who can adjust on the intensive margin. As before, individuals will work if the utility conditional on working exceeds that of not working:

$$v(\tilde{c}_{njt}, \tilde{z}_{njt}; n) + V_t(\tilde{A}_{njt}, \tilde{z}_{njt}; n) - v^0(A_{nj,t-1}) - q_{njt} > 0 \quad (\text{B.26})$$

where the “ $\sim$ ” denotes optimal levels conditional on working. The probability of working in period 1 is still:

$$\Pr(z_{n1t} > 0 | \tilde{z}_{n0t}) = G(\bar{q}_{n1t} | n, t) \quad (\text{B.27})$$

where now:

$$\bar{q}_{n1t} \equiv v(\tilde{c}_{n1t}, \tilde{z}_{n1t}; n) + V_t(\tilde{A}_{n1t}, \tilde{z}_{n1t}; n) - v^0(A_{n1,t-1}) \quad (\text{B.28})$$

We now show under what conditions our main results still hold in this dynamic setting. First, the slope of the employment rate will still be:

$$\frac{d\Pr(z_{n1t} > 0 | \tilde{z}_{n0t})}{d\tilde{z}_{n0t}} = g(\bar{q}_{n1t} | n, t) \frac{d\bar{q}_{n1t}}{d\tilde{z}_{n0t}} + \frac{dG(\bar{q}_{n1t} | n, t)}{dn} \frac{dn}{d\tilde{z}_{n0t}} \quad (\text{B.29})$$

We will have a new expression for the first term on the right of equation (B.29). After substituting for  $\tilde{c}_{n1t}$



in (B.28) using the dynamic budget constraint in (B.24) we have:

$$\begin{aligned}
\frac{d\bar{q}_{n1t}}{d\tilde{z}_{n0t}} &= \left[ \frac{\partial v \left( (1+r_{t-1})\tilde{A}_{n1,t-1} + \tilde{z}_{n1t} - T(\tilde{z}_{n1t}) + B_1(\tilde{z}_{n1t}) - \tilde{A}_{n1t}, \tilde{z}_{n1t}; n \right)}{\partial \tilde{z}_{n1t}} + \frac{\partial V_t \left( \tilde{A}_{n1t}, \tilde{z}_{n1t}; n \right)}{\partial \tilde{z}_{n1t}} \right] \frac{d\tilde{z}_{n1t}}{d\tilde{z}_{n0t}} \\
&+ \left[ \frac{\partial v \left( (1+r_{t-1})\tilde{A}_{n1,t-1} + \tilde{z}_{n1t} - T(\tilde{z}_{n1t}) + B_1(\tilde{z}_{n1t}) - \tilde{A}_{n1t}, \tilde{z}_{n1t}; n \right)}{\partial \tilde{A}_{n1t}} + \frac{\partial V_t \left( \tilde{A}_{n1t}, \tilde{z}_{n1t}; n \right)}{\partial \tilde{A}_{n1t}} \right] \frac{d\tilde{A}_{n1t}}{d\tilde{z}_{n0t}} \\
&+ \left[ \frac{\partial v \left( (1+r_{t-1})\tilde{A}_{n1,t-1} + \tilde{z}_{n1t} - T(\tilde{z}_{n1t}) + B_1(\tilde{z}_{n1t}) - \tilde{A}_{n1t}, \tilde{z}_{n1t}; n \right)}{\partial n} + \frac{\partial V_t \left( \tilde{A}_{n1t}, \tilde{z}_{n1t}; n \right)}{\partial n} \right] \frac{dn}{d\tilde{z}_{n0t}} \\
&+ \left[ \frac{\partial v \left( (1+r_{t-1})\tilde{A}_{n1,t-1} + \tilde{z}_{n1t} - T(\tilde{z}_{n1t}) + B_1(\tilde{z}_{n1t}) - \tilde{A}_{n1t}, \tilde{z}_{n1t}; n \right)}{\partial \tilde{A}_{n1,t-1}} - \frac{\partial v^0 \left( \tilde{A}_{n1,t-1} \right)}{\partial A_{n1,t-1}} \right] \frac{dA_{n1,t-1}}{d\tilde{z}_{n0t}} \\
&= [(1 - T'(z) + B_1'(z))v_c + v_z + V_z] \frac{d\tilde{z}_{n1t}}{d\tilde{z}_{n0t}} + [-v_c + V_A] \frac{d\tilde{A}_{n1t}}{d\tilde{z}_{n0t}} + [v_n + V_n] \frac{dn}{d\tilde{z}_{n0t}} \\
&+ [(1 + r_{t-1})v_c - v_A^0] \frac{dA_{n1,t-1}}{d\tilde{z}_{n0t}}
\end{aligned} \tag{B.30}$$

For those with  $\tilde{z}_{n0t} < z^{AET}$  or  $\tilde{z}_{n0t} > z^{AET} + \Delta z$ , we can use the first-order conditions in (B.25) to show that the first term in (B.30) equals zero when agents are able to adjust on the intensive margin as in Section 6.2. For those with  $z^{AET} \leq \tilde{z}_{n0t} \leq z^{AET} + \Delta z^*$ , the first term in (B.30) equals zero:  $d\tilde{z}_{n1t}/d\tilde{z}_{n0t} = 0$  since  $\tilde{z}_{n1t} = z^{AET}$  for everyone in this latter set due to bunching. Additionally, the second term in (B.30) equals zero for everyone, due to the first-order condition in (B.25) for saving.

Thus, when agents are able to adjust on both the intensive and extensive margin we have:

$$\begin{aligned}
\frac{d\Pr(z_{n1t} > 0 | \tilde{z}_{n0t})}{d\tilde{z}_{n0t}} &= g(\bar{q}_{n1t} | n, t) \left( \left[ \frac{\partial v(\tilde{c}_{n1t}, \tilde{z}_{n1t}; n)}{\partial n} + \frac{\partial V_t(\tilde{A}_{n1t}, \tilde{z}_{n1t}; n)}{\partial n} \right] \frac{dn}{d\tilde{z}_{n0t}} \right. \\
&\quad \left. + \left[ (1 + r_{t-1}) \frac{\partial v(\tilde{c}_{n1t}, \tilde{z}_{n1t}; n)}{\partial \tilde{c}_{n1t}} - \frac{\partial v^0(\tilde{A}_{n1,t-1})}{\partial A_{n1,t-1}} \right] \frac{dA_{n1,t-1}}{d\tilde{z}_{n0t}} \right) + \frac{dG(\bar{q}_{n1t} | n, t)}{dn} \frac{dn}{d\tilde{z}_{n0t}}
\end{aligned} \tag{B.31}$$

Note that  $n$ ,  $\tilde{q}_{n,1}$ ,  $\tilde{z}_{n,1}$ ,  $\tilde{A}_{n,1}$ , and  $T_1(\cdot)$  are all continuous in  $\tilde{z}_{n0t}$  at  $z^{AET}$ . Furthermore, our smoothness assumptions imply that  $g(\cdot)$ ,  $\partial v/\partial n$ ,  $\partial V/\partial n$ ,  $\partial n/\partial \tilde{z}_{n0t}$ , and  $\partial G(\bar{q}_{n,1} | n, t)$  are likewise continuous in their arguments. We additionally assume that  $\partial A_{n1,t-1}/\partial \tilde{z}_{n0t}$  is continuous at  $z^*$ ; we discuss below the conditions under which this assumption holds and argue that they are satisfied in our setting. Given these assumptions, our original result follows when there are intensive margin adjustments:

$$\lim_{\tilde{z}_{n0t} \rightarrow z^{AET+}} \frac{d\Pr(z_{n1t} > 0 | \tilde{z}_{n0t})}{d\tilde{z}_{n0t}} - \lim_{\tilde{z}_{n0t} \rightarrow z^{AET-}} \frac{d\Pr(z_{n1t} > 0 | \tilde{z}_{n0t})}{d\tilde{z}_{n0t}} = 0 \tag{B.32}$$

We now turn to the case in which we make the same assumptions, except that individuals are not able

to adjust on the intensive margin, as in Section 6.3. We now have:

$$\begin{aligned}
\frac{d\bar{q}_{n1t}}{d\tilde{z}_{n0t}} &= \left[ \frac{\partial v \left( (1+r_{t-1})\tilde{A}_{n1,t-1} + \tilde{z}_{n0t} - T(\tilde{z}_{n0t}) + B_1(\tilde{z}_{n0t}) - \tilde{A}_{n1t}, \tilde{z}_{n0t}; n \right)}{\partial \tilde{z}_{n0t}} + \frac{\partial V_t \left( \tilde{A}_{n1t}, \tilde{z}_{n0t}; n \right)}{\partial \tilde{z}_{n0t}} \right] \\
&+ \left[ \frac{\partial v \left( (1+r_{t-1})\tilde{A}_{n1,t-1} + \tilde{z}_{n0t} - T(\tilde{z}_{n0t}) + B_1(\tilde{z}_{n0t}) - \tilde{A}_{n1t}, \tilde{z}_{n0t}; n \right)}{\partial \tilde{A}_{n1t}} + \frac{\partial V_t \left( \tilde{A}_{n1t}, \tilde{z}_{n0t}; n \right)}{\partial \tilde{A}_{n1t}} \right] \frac{d\tilde{A}_{n1t}}{d\tilde{z}_{n0t}} \\
&+ \left[ \frac{\partial v \left( (1+r_{t-1})\tilde{A}_{n1,t-1} + \tilde{z}_{n0t} - T(\tilde{z}_{n0t}) + B_1(\tilde{z}_{n0t}) - \tilde{A}_{n1t}, \tilde{z}_{n0t}; n \right)}{\partial n} + \frac{\partial V_t \left( \tilde{A}_{n1t}, \tilde{z}_{n0t}; n \right)}{\partial n} \right] \frac{dn}{d\tilde{z}_{n0t}} \\
&+ \left[ \frac{\partial v \left( (1+r_{t-1})\tilde{A}_{n1,t-1} + \tilde{z}_{n0t} - T(\tilde{z}_{n0t}) + B_1(\tilde{z}_{n0t}) - \tilde{A}_{n1t}, \tilde{z}_{n0t}; n \right)}{\partial \tilde{A}_{n1,t-1}} - \frac{\partial v^0 \left( \tilde{A}_{n1,t-1} \right)}{\partial A_{n1,t-1}} \right] \frac{dA_{n1,t-1}}{d\tilde{z}_{n0t}} \\
&= [(1 - T'(z) + B'_1(z))v_c + v_z + V_z] + [-v_c + V_A] \frac{d\tilde{A}_{n1t}}{d\tilde{z}_{n0t}} + [v_n + V_n] \frac{dn}{d\tilde{z}_{n0t}} \\
&+ [(1 + r_{t-1})v_c - v_A^0] \frac{dA_{n1,t-1}}{d\tilde{z}_{n0t}}
\end{aligned} \tag{B.33}$$

where the primary difference from (B.30) is that earnings are fixed at  $\tilde{z}_{n0t}$ . We still have the second term dropping out of the expression in (B.33), as assets,  $\tilde{A}_{n1t}$ , are optimally chosen even when earnings are fixed. Furthermore, for those with  $\tilde{z}_{n0t} \leq z^{AET}$  we have  $T_1 = T_0$ , and thus the first term also drops out among this group due to the envelope theorem. Note that for those with  $\tilde{z}_{n0t} \geq z^{AET}$  we can rewrite:

$$(1 - T'(z) + B'_1(z))v_c + v_z + V_z = \lambda \left[ (1 - \tau_0 - db) + \frac{v_z + V_z}{v_c} \right] \tag{B.34}$$

Similar to our less dynamic model above in Section 6.3, we therefore have:

$$\frac{d\Pr(z_{n1t} > 0 | \tilde{z}_{n0t})}{d\tilde{z}_{n0t}} = \begin{cases} g(\bar{q}_{n1t}|n, t) \left( [v_n + V_n] \frac{dn}{d\tilde{z}_{n0t}} + [(1 + r_{t-1})v_c - v_A^0] \frac{dA_{n1,t-1}}{d\tilde{z}_{n0t}} \right) & \text{if } \tilde{z}_{n0t} < z^{AET} \\ g(\bar{q}_{n1t}|n, t) \left( \lambda_n \left[ (1 - \tau_0 - db) + \frac{v_z + V_z}{v_c} \right] + [v_n + V_n] \frac{dn}{d\tilde{z}_{n0t}} + [(1 + r_{t-1})v_c - v_A^0] \frac{dA_{n1,t-1}}{d\tilde{z}_{n0t}} \right) + \frac{dG(\bar{q}_{n1t}|n, t)}{dn} \frac{dn}{d\tilde{z}_{n0t}} & \text{if } \tilde{z}_{n0t} \geq z^{AET} \end{cases} \tag{B.35}$$

We note the following limit, making use of the first-order condition in (B.25):

$$\lim_{\tilde{z}_{n0t} \rightarrow z^{AET+}} \frac{v_z \left( (1+r_{t-1})\tilde{A}_{n1,t-1} + \tilde{z}_{n0t} - T(\tilde{z}_{n0t}) + B_1(\tilde{z}_{n0t}) - \tilde{A}_{n1t}, \tilde{z}_{n0t}; n \right) + V_z \left( \tilde{A}_{n1t}, \tilde{z}_{n0t}; n \right)}{v_c \left( (1+r_{t-1})\tilde{A}_{n1,t-1} + \tilde{z}_{n0t} - T(\tilde{z}_{n0t}) + B_1(\tilde{z}_{n0t}) - \tilde{A}_{n1t}, \tilde{z}_{n0t}; n \right)} = -(1 - \tau_0) \tag{B.36}$$

Maintaining our smoothness assumptions for this section, we thus have our original result when earnings in

state 1 are constrained ( $\tilde{z}_{n1t} = \tilde{z}_{n0t}$ ):

$$\begin{aligned} \lim_{\tilde{z}_{n0t} \rightarrow z^{AET+}} \frac{d\Pr(z_{n1t} > 0 | \tilde{z}_{n0t})}{d\tilde{z}_{n0t}} - \lim_{\tilde{z}_{n0t} \rightarrow z^{AET-}} \frac{d\Pr(z_{n1t} > 0 | \tilde{z}_{n0t})}{d\tilde{z}_{n0t}} &= \lim_{\tilde{z}_{n0t} \rightarrow z^{AET+}} g(\bar{q}_{n1t} | n, t) \cdot \lambda_n \left[ (1 - \tau_0 - db) \right. \\ &\quad \left. + \frac{v_z + V_z}{v_c} \right] \\ &= -db \cdot \lambda_{n^*} \cdot g(\bar{q}_{n^{AET}1t} | n^{AET}, t) \end{aligned} \tag{B.37}$$

As in Appendix B.1, this result easily generalizes to the case of multiple, discrete job options away from  $z^{*AET}$ . The kink in this case can be used to calculate an extensive margin elasticity, as in equation (14).

Returning to our assumption that  $\partial A_{n1,t-1} / \partial \tilde{z}_{n0t}$  is continuous at  $\tilde{z}_{n0t} = z^{AET}$ , we view this as a natural assumption in our setting. First, if agents act as if the tax change is unanticipated, then this is assumption holds: in this case the slope of expected lifetime wealth is continuous at the kink. Many contexts will feature unanticipated changes in taxes. In our empirical application, Gelber *et al.* (2013) show that prior to being subject to the AET, individuals do not appear to act as if they anticipate the later imposition of the AET (consistent with the myopia suggested in Gelber, Isen, and Song, 2016). If the tax change were anticipated they would bunch at  $z^{AET}$  in anticipation of the later imposition of the kink (due to the fixed cost of adjustment), but empirically we do not observe such behavior. Thus, they do not appear to anticipate its imposition, so our empirical application appears consistent with this case. Moreover, in our setting, we test for kinks in predetermined or placebo outcomes, including measures of the employment rate at ages 56, 57, 60, and 61, as well as demographic variables. We show in Section 5.3 that predetermined variables as a function of age 60 earnings do not exhibit systematic discontinuities in their slopes at the exempt amount, consistent with our assumption of continuity in  $\partial A_{n1,t-1} / \partial \tilde{z}_{n0t}$  at  $\tilde{z}_{n0t} = z^{AET}$ .

Second, even if some individuals in our context acted as if the imposition of the AET were anticipated, if the AET is also actuarially fair and other sources of lifetime wealth are also smooth, then  $\partial A_{n1,t-1} / \partial \tilde{z}_{n0t}$  should be smooth at  $\tilde{z}_{n0t} = z^{AET}$ . It is commonly understood that the AET is approximately actuarially fair (*e.g.* Diamond and Gruber, 1999), and thus this assumption should approximately hold. Note that even if the AET is actuarially fair on average—but better than actuarially fair in expectation for some types and worse for others—it is possible that  $\partial A_{n1,t-1} / \partial \tilde{z}_{n0t}$  is continuous at  $\tilde{z}_{n0t} = z^{AET}$ , even though we also we observe a non-zero substitution effect of the incentives created by the AET when individuals are later subject to it. For example, if the AET is worse than actuarially fair for those who are particularly responsive to the substitution incentives created by the AET, then we could see a reduction in earnings or employment due to a bunching or extensive margin response to the substitution incentives created by the AET once individuals have claimed, even though  $\partial A_{n1,t-1} / \partial \tilde{z}_{n0t}$  is continuous at  $\tilde{z}_{n0t} = z^{AET}$ .

Third, in the case that the imposition of the tax change is anticipated—which could be consistent with the data if the AET is actuarially fair on average—we could interpret our estimated elasticity as a marginal-utility-of-wealth-constant elasticity. In this case our paper then can fit into a broader literature estimating Frisch elasticities (see Chetty *et al.*, 2013, for a review), and provides a novel estimate of a Frisch elasticity using transparent variation from the RKD. (Following Chetty *et al.* (2012) we call this extensive margin elasticity, holding the marginal utility of lifetime wealth constant, a Frisch elasticity, while recognizing that extensive margin Frisch elasticities are technically not defined because non-participants do not locate at an interior optimum.)

Fourth, even in the case in which the AET is not actuarially fair, we can estimate an elasticity that represents the response to a parametric shift in the entire wage profile; as Blundell and MaCurdy (1999) and others point out, this is the most relevant for policy evaluation of the impacts of permanent shifts in the entire wage profile and will reflect a combination of income and substitution effects. We have run several numerical simulations to gauge the quantitative importance of anticipatory savings for the continuity of  $\partial A_{n1,t-1} / \partial \tilde{z}_{n0t}$  at  $\tilde{z}_{n0t} = z^{AET}$ , and we find that extensive margin estimates using this method are little affected by a lack of smoothness in  $A_{n1,t-1}$ .<sup>34</sup> Furthermore, since predetermined variables do not show systematic discontinuities

<sup>34</sup>Results from our simulations are available upon request. We continue to recover the extensive margin elasticity in a dynamic model when intensive margin adjustment is constrained and we observe the correct, counterfactual running variable. We estimate an attenuated version of the true elasticity when intensive margin adjustment is constrained and we use lagged

in their slopes as a function of age 60 earnings, this issue again does not appear to affect the validity of the smoothness assumption underlying our empirical design. We view our method as most easily applicable in settings in which the data are consistent with such interpretations of the parameters and the smoothness assumptions appear to be satisfied.

#### B.4.1 Explicitly Modeling Actuarial Adjustment in the Context of the Dynamic Model

Thus far, we have abstracted from the impact of actuarial adjustment of future benefits on the behavioral response to the AET. In this section, we explicitly incorporate this feature of the AET and highlight how our estimator is affected. We maintain the setup of the dynamic model above with the following alterations. First, the dynamic program now involves choosing earnings to maximize:

$$u_t(c_{njt}, z_{njt}; n) = v(c_{njt}, z_{njt}; n) - q_{njt} \cdot 1\{z_{njt} > 0\} + V_t(A_{njt} + B_{njt}^{PDV}(z_{njt}), z_{njt}; n) \quad (\text{B.38})$$

Relative to our previous case, the value function,  $V_t(\cdot)$ , is now a function of total assets, *i.e.* the sum of savings,  $A_{njt}$ , and the present discounted value of all future Social Security benefits,  $B_{njt}^{PDV}(z_{njt})$ . These benefits are potentially affected by the current level of earnings. The dynamic budget constraint is as before:

$$c_{njt} = (1 + r_{t-1})A_{nj,t-1} + z_{njt} - T(z_{njt}) + B_j(z_{njt}) - A_{njt} \quad (\text{B.39})$$

The flow of benefits,  $B_{jt}(z_{njt})$ , does potentially vary across states, due to the presence of the AET in state 1. The first-order conditions are now:

$$\begin{aligned} v_z + V_z &= -\lambda(1 - T'(z) + [B'(z) + B^{PDV'}(z)]) \\ v_c &= V_A = \lambda \end{aligned} \quad (\text{B.40})$$

where  $B'_{jt}(z)$  is the marginal effect of earning more on the current flow of benefits, *i.e.* the benefit reduction rate (BRR). It takes the following form in state 1:

$$B'(z) = \begin{cases} 0 & \text{if } z < z^{AET} \\ -db & \text{if } z \geq z^{AET} \end{cases} \quad (\text{B.41})$$

Finally,  $B^{PDV'}(z)$  is the effect of increasing current earnings on the future stream of benefits from OASI.  $B^{PDV'}(z)$  similarly takes the form:

$$B^{PDV'}(z) = \begin{cases} 0 & \text{if } z < z^{AET} \\ db^{PDV} & \text{if } z \geq z^{AET} \end{cases} \quad (\text{B.42})$$

Note that when the adjustment is actuarially fair, we have  $db = db^{PDV}$ . In this version of the model, we also express the outside utility as a function of savings and OASI benefits:

$$v^0 = v^0(A_{nj,t-1}, B_{nj,t-1}^{PDV}) \quad (\text{B.43})$$

Using similar steps as in the basic dynamic model above, we have the following expression for the (potential) kink in the employment rate at  $z^*$ , when intensive margin earnings are constrained:

$$\lim_{\tilde{z}_{n0t} \rightarrow z^{AET+}} \frac{d\Pr(z_{n1t} > 0 | \tilde{z}_{n0t})}{d\tilde{z}_{n0t}} - \lim_{\tilde{z}_{n0t} \rightarrow z^{AET-}} \frac{d\Pr(z_{n1t} > 0 | \tilde{z}_{n0t})}{d\tilde{z}_{n0t}} = -(db - db^{PDV}) \cdot \lambda_{n^{AET}} \cdot g(\bar{q}_{n^{AET}1t} | n^{AET}, t)$$

Here we make explicit that an additional necessary condition for finding a kink in the employment rate is that  $db \neq db^{PDV}$ , either because the AET is not actually fair, or perhaps, because individuals do not pay attention to actuarial adjustment. The sign of  $db^{PDV}$  is non-negative, as the actuarial adjustment never reduces future benefits, but  $db^{PDV}$  can be smaller than  $db$  if adjustment is not full. It can also be larger than  $db$  for those with high life expectancies, in which case there will be a positive kink in the slope of the employment

earnings as a proxy for the running variable, though our simulations show that the extent of attenuation is slight. If anything, this slight attenuation again strengthens our case that the elasticity is large, as the lower bound we estimate is itself large.

function. In our applications above, we have effectively assumed  $db^{PDV} = 0$  for illustrative purposes and following previous literature, although our method can easily accommodate alternative assumptions. Since we do find a kink in the employment rate empirically (and also find intensive margin bunching), this assumption is validated.

## B.5 Model Extension: Joint Claiming and Employment Decision

A specific feature of our empirical context is that whether or not individuals face a kink is endogenous. That is, if an agent does not claim, she faces a linear budget set. We now extend the model to our particular empirical application, so that the model includes this trade-off between facing a kink or delaying claiming. To model claiming, we introduce additional notation. We model claiming in a somewhat “reduced form” fashion, but our model easily generalizes to the previous dynamic setting, in which we can explicitly incorporate the effects of claiming Social Security on the timing and magnitude of Social Security benefits in different periods, as well as the resulting effects on wealth and savings. In our case, when  $j = 0$  an individual faces a linear tax schedule whether or not she claims. However, when  $j = 1$  an individual faces a kinked tax schedule when claiming Social Security or a linear schedule when not claiming Social Security. We focus on our version of the model in which intensive margin adjustments are constrained, *i.e.*  $\tilde{z}_{n1} \equiv \tilde{z}_{n0}$ , and we will determine whether a kink in the probability of working still occurs when claiming is endogenous.

First, we will denote  $v_n^0 \equiv v(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_0(\tilde{z}_{n0}), \tilde{z}_{n0}, n)$  as the flow utility (net of the fixed cost of working) when earning  $\tilde{z}_{n0}$ , facing no benefit reduction rate, and claiming in state 0. Next, we denote  $v_n^1 \equiv v(\tilde{z}_{n0} - T(\tilde{z}_{n0}) + B_1(\tilde{z}_{n0}), \tilde{z}_{n0}, n)$ , *i.e.* the net flow utility when earning  $\tilde{z}_{n0}$ , facing the benefit reduction rate  $db$  above  $z^{AET}$ , and claiming in state 1. As before,  $v^0$  is the utility received when claiming Social Security and not working, in either state 0 or state 1. Next, we specify utility when claiming is delayed. The payoff when not working and not claiming Social Security is  $v^0 + \theta_n$ , in both state 0 and state 1. Likewise, the payoff when working and not claiming Social Security is  $v_n^0 + \gamma_n$ , again in either state 1 or state 0. Thus, the parameters  $\theta$  and  $\gamma$  capture the relative change in utility when claiming is delayed. These can be considered to capture the income effect of delaying claiming. We leave these parameters relatively unrestricted, and thus they may represent an increase in lifetime wealth when delaying claiming increases lifetime benefits (for example due to the actuarial adjustment being better than actuarially fair), or they may represent a utility decrease for those facing liquidity constraints, for example.

In addition to  $n$  we now have three parameters that capture unobserved heterogeneity,  $(\gamma, \theta, q)$ . We allow these variables to have a relatively unrestricted, joint distribution, which we represent with a conditional, joint density of  $\gamma$  and  $\theta$ ,  $m(\gamma, \theta|q, n)$  and our previous marginal density of  $q$ ,  $g(q|n)$ . The joint density of  $(\gamma, \theta, q)$  is therefore  $m(\gamma, \theta|q, n) \cdot g(q|n)$ . We extend our assumption of smoothness in unobserved heterogeneity to the joint distribution of  $(\gamma, \theta, q)$ , conditional on  $n$ .

In State 1, our model will now feature comparisons among four discrete choices: (1) working while claiming,  $v_n^1 - q_n$ ; (2) not working while claiming,  $v^0$ ; (3) working while not claiming  $v_n^0 + \gamma_n - q_n$ ; and (4) not working while not claiming,  $v^0 + \theta_n$ . We will additionally denote critical values of our unobservable parameters, which arise when comparing the various discrete options. As before, when comparing working while claiming to not working while claiming, the expression for indifference is  $v_n^1 - q_n = v^0$ , which implies a critical value for  $q$ ,  $\bar{q}_n \equiv v_n^1 - v^0$ . Similarly, when comparing working while not claiming to not working while not claiming, the expression for indifference is  $v_n^0 + \gamma_n - q_n = v^0 + \theta_n$ . We can write this in terms of  $\gamma_n$  as  $\gamma_n = \theta_n + q_n - (v_n^0 - v^0) = \theta_n + q_n - \bar{q}_n^0$ , where  $\bar{q}_n^0 \equiv v_n^0 - v^0$  is an analogous critical value for  $q$  in state 0 when claiming.

Furthermore, for those who are indifferent between working while claiming and not working while not claiming, we have  $v_n^1 - q_n = v^0 + \theta_n$ . This implies a critical value for  $\theta_n$ :  $\theta_n = \bar{q}_n - q_n$ . Symmetrically, for those who are indifferent between not working while claiming and working while not claiming, we have  $v^0 = v_n^0 + \gamma_n - q_n$ , with a critical value for  $\gamma_n$ :  $\gamma_n = \bar{q}_n^0 - q_n$ . Indifference between working while claiming and working while not claiming implies  $v_n^1 - q_n = v_n^0 + \gamma_n - q_n$ , and a critical value for  $\gamma_n$  is:  $\gamma_n = v_n^1 - v_n^0 \equiv \Delta v_n$ . Finally, indifference between not working while claiming and not working while not claiming implies  $v^0 = v^0 + \theta_n$  or  $\theta_n = 0$ .

Using our previous results, the envelope theorem implies the following:

$$\frac{\partial \bar{q}_n^0}{\partial \tilde{z}_{n0}} = \frac{\partial v_n^0}{\partial \tilde{z}_{n0}} = 0 \quad (\text{B.44})$$

Additionally, we have:

$$\frac{\partial \bar{q}_n}{\partial \tilde{z}_{n0}} = \frac{\partial \Delta v_n}{\partial \tilde{z}_{n0}} = \frac{\partial v_n^1}{\partial \tilde{z}_{n0}} = \begin{cases} \lambda_n \left[ (1 - \tau_0) + \frac{v_z}{v_c} \right] & \text{if } \tilde{z}_{n0} < z^{AET} \\ \lambda_n \left[ (1 - \tau_0 - db) + \frac{v_z}{v_c} \right] & \text{if } \tilde{z}_{n0} \geq z^{AET} \end{cases} \quad (\text{B.45})$$

### B.5.1 Claiming Decision

We now characterize the claiming probability and derive an expression for a kink in the claiming probability at  $\tilde{z}_{n0} = z^{AET}$ . The probability of claiming can be expressed as follows:

$$\begin{aligned} \Pr(\text{claim} | \tilde{z}_{n0}) &= \int_{-\infty}^{\bar{q}} \int_{-\infty}^{\bar{q}-q} \int_{-\infty}^{\Delta v} m(\gamma, \theta | q, n) g(q | n) d\gamma d\theta dq \\ &+ \int_{\bar{q}}^{\infty} \int_{-\infty}^0 \int_{-\infty}^{q-\bar{q}^0} m(\gamma, \theta | q, n) g(q | n) d\gamma d\theta dq \end{aligned} \quad (\text{B.46})$$

The first term integrates over the values of  $q$  for which working while claiming is preferred to not working while claiming,  $q \in [-\infty, \bar{q}]$ . The next two integrals restrict attention to values of  $\theta$  that render not working while not claiming dominated by working while claiming, *i.e.*  $\theta \in [-\infty, \bar{q} - q]$  and values of  $\gamma$  that similarly render working while not claiming dominated by working while claiming, *i.e.*  $\gamma \in [-\infty, \Delta v]$ . The second term integrates over values of  $q$  for which not working while claiming is preferred to working while claiming,  $q \in [\bar{q}, \infty]$ . Over this range we restrict analysis to  $\theta \in [-\infty, 0]$  and  $\gamma \in [-\infty, \bar{q} - q]$ , *i.e.* values that render not claiming dominated by not working while claiming.

Consider the slope of  $\Pr(\text{claim} | \tilde{z}_{n0})$ , which can now be expressed as:

$$\frac{d \Pr(\text{claim} | \tilde{z}_{n0})}{d \tilde{z}_{n0}} = \frac{\partial \Pr(\text{claim} | \tilde{z}_{n0})}{\partial \tilde{z}_{n0}} + \frac{\partial \Pr(\text{claim} | \tilde{z}_{n0})}{\partial n} \frac{dn}{d \tilde{z}_{n0}} \quad (\text{B.47})$$

Given our assumptions regarding smoothness and our results above, we know that the second term will be continuous at  $\tilde{z}_{n0} = z^{AET}$ . We therefore focus on the first term in this expression. Leibniz's rule implies:

$$\begin{aligned} \frac{\partial \Pr(\text{claim} | \tilde{z}_{n0})}{\partial \tilde{z}_{n0}} &= \int_{-\infty}^{\bar{q}} \int_{-\infty}^{\bar{q}-q} \frac{\partial \Delta v_n}{\partial \tilde{z}_{n0}} m(\Delta v_n, \theta | q, n) g(q | n) d\gamma d\theta dq \\ &+ \int_{-\infty}^{\bar{q}} \frac{\partial \bar{q}_n}{\partial \tilde{z}_{n0}} \left[ \int_{-\infty}^{\Delta v} m(\gamma, \bar{q} - q | q, n) d\gamma \right] g(q | n) dq \\ &+ \frac{\partial \bar{q}_n}{\partial \tilde{z}_{n0}} \left[ \int_{-\infty}^0 \int_{-\infty}^{\Delta v} m(\gamma, \theta | \bar{q}, n) d\gamma d\theta \right] g(\bar{q} | n) \\ &- \int_{\bar{q}}^{\infty} \int_{-\infty}^0 \frac{\partial \bar{q}_n^0}{\partial \tilde{z}_{n0}} m(q - \bar{q}^0, \theta | q, n) g(q | n) d\theta dq \\ &- \frac{\partial \bar{q}_n}{\partial \tilde{z}_{n0}} \left[ \int_{-\infty}^0 \int_{-\infty}^{\Delta v} m(\gamma, \theta | \bar{q}, n) d\gamma d\theta \right] g(\bar{q} | n). \end{aligned} \quad (\text{B.48})$$

Rearranging terms, and using the fact that  $\partial \bar{q}^0 / \partial \tilde{z}_{n0} = 0$  and  $\partial \bar{q} / \partial \tilde{z}_{n0} = \partial \Delta v / \partial \tilde{z}_{n0} = \lambda \left[ (1 - \tau_0 + B'(z) + \frac{v_z}{v_c}) \right]$ , we have:

$$\begin{aligned} \frac{\partial \Pr(\text{claim} | \tilde{z}_{n0})}{\partial \tilde{z}_{n0}} &= \lambda \left[ (1 - \tau_0 + B'(z) + \frac{v_z}{v_c}) \right] \times \\ &\left\{ \int_{-\infty}^{\bar{q}} \left[ \int_{-\infty}^{\bar{q}-q} m(\Delta v, \theta | q, n) d\theta + \int_{-\infty}^{\Delta v} m(\gamma, \bar{q} - q | q, n) d\gamma \right] g(q | n) dq \right\} \end{aligned} \quad (\text{B.49})$$

Finally, the kink in the probability of claiming can be expressed as:

$$\lim_{\tilde{z}_{n0} \rightarrow z^{*+}} \frac{\partial \Pr(\text{claim} | \tilde{z}_{n0})}{\partial \tilde{z}_{n0}} - \lim_{\tilde{z}_{n0} \rightarrow z^{*-}} \frac{\partial \Pr(\text{claim} | \tilde{z}_{n0})}{\partial \tilde{z}_{n0}} = \delta \quad (\text{B.50})$$

where

$$\begin{aligned} \delta &= \lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \lambda \left[ (1 - \tau_0 - db) + \frac{v_z}{v_c} \right] \cdot \left\{ \int_{-\infty}^{\bar{q}} \left[ \int_{-\infty}^{\bar{q}-q} m(\Delta v, \theta | q, n) d\theta + \int_{-\infty}^{\Delta v} m(\gamma, \bar{q} - q | q, n) d\gamma \right] g(q | n) dq \right\} \\ &= -db \cdot \lambda_{n^{AET}} \cdot \left\{ \int_{-\infty}^{\bar{q}} \left[ \int_{-\infty}^{\bar{q}-q} m(0, \theta | q, n^{AET}) d\theta + \int_{-\infty}^0 m(\gamma, \bar{q} - q | q, n^{AET}) d\gamma \right] g(q | n^{AET}) dq \right\} \quad (\text{B.51}) \end{aligned}$$

and where we used the results from Section 6.3 and the fact that  $\lim_{\tilde{z}_{n0} \rightarrow z^{AET}} \Delta v_n = 0$ .

The two integrals in the expression can be interpreted as joint probabilities. The first is the joint probability that (1) working while claiming is preferred to not working while claiming, (2) working while claiming is preferred to not working while not claiming, and (3) the individual is indifferent between working while claiming and working while not claiming, *i.e.*  $\gamma_{n^*} = \Delta v_{n^*} = 0$ . The second term is the joint probability that (1) working while claiming is preferred to not working while claiming, (2) working while claiming is preferred to working while not claiming, and (3) the individual is indifferent between working while claiming and not working while not claiming, *i.e.*  $\theta_{n^{AET}} = \bar{q}_{n^{AET}} - q_{n^{AET}}$ . Note that the probabilities are conditional on  $n = n^{AET}$ .

Thus, there will be a downward kink in the probability of claiming, which increases with the size of the kink and the size of two key sets of marginal claimants. The first set are on the margin of moving from working while claiming to working while not claiming, and the second set are on the margin of moving from working while claiming to not working while not claiming. Finally, note that the kink only affects claiming among those for whom working while claiming is optimal. Therefore, the only relevant shifting is from working while claiming to either state of not claiming.

### B.5.2 Extensive Margin Response with Endogenous Claiming

We now consider the extensive margin choice of whether to work, allowing for an endogenous claiming response. The probability of having positive earnings in state 1 is now:

$$\begin{aligned} \Pr(z_{n1} > 0 | \tilde{z}_{n0}) &= \int_{-\infty}^{\bar{q}} \left[ \int_{-\infty}^{\bar{q}-q} \int_{-\infty}^{\infty} m(\gamma, \theta | q, n) d\gamma d\theta + \int_{\bar{q}-q}^{\infty} \int_{\theta+q-\bar{q}^0}^{\infty} m(\gamma, \theta | q, n) d\gamma d\theta \right] g(q | n) dq \\ &+ \int_{\bar{q}}^{\infty} \left[ \int_{-\infty}^0 \int_{q-\bar{q}^0}^{\infty} m(\gamma, \theta | q, n) d\gamma d\theta + \int_0^{\infty} \int_{\theta+q-\bar{q}^0}^{\infty} m(\gamma, \theta | q, n) d\gamma d\theta \right] g(q | n) dq \quad (\text{B.52}) \end{aligned}$$

The probability is comprised of four terms. The first two terms correspond to values of  $q$  that render working while claiming preferable to not working while claiming. The first term in this set captures individuals for whom working while claiming also dominates not working while not claiming. In this case, the individual will always work, regardless of the value of  $\gamma$ . The second term captures individuals for whom not working while not claiming dominates working while claiming. In this case, only those who prefer working while not claiming to not working while not claiming will work. The second set of terms likewise covers the two settings in which not working while claiming dominates working while claiming, but the individual still decides to work. Put another way, an agent will work when  $\max(v_n^1 - q, v_n^0 + \gamma - q) > \max(v^0, v^0 + \theta)$ .

As in the case of claiming above, we focus on the discontinuity in the partial derivative of this probability,

$\partial \Pr(z_{n1} > 0 | \tilde{z}_{n0}) / \partial \tilde{z}_{n0}$ . Leibniz's rule again implies:

$$\begin{aligned}
\frac{\partial \Pr(z_{n1} > 0 | \tilde{z}_{n0})}{\partial \tilde{z}_{n0}} &= \int_{-\infty}^{\bar{q}} \left\{ \frac{\partial \bar{q}_n}{\partial \tilde{z}_{n0}} \left[ \int_{-\infty}^{\infty} m(\gamma, \bar{q} - q | q, n) d\gamma - \int_{\Delta v}^{\infty} m(\gamma, \bar{q} - q | q, n) d\gamma \right] \right. \\
&\quad \left. + \int_{\bar{q}-q}^{\infty} \frac{\partial \bar{q}_n^0}{\partial \tilde{z}_{n0}} m(\theta + q - \bar{q}^0, \theta | q, n) d\theta \right\} g(q | n) dq \\
&\quad + \frac{\partial \bar{q}}{\partial \tilde{z}_{n0}} \left[ \int_{-\infty}^0 \int_{-\infty}^{\infty} m(\gamma, \theta | \bar{q}, n) d\gamma d\theta + \int_0^{\infty} \int_{\theta+\Delta v}^{\infty} m(\gamma, \theta | \bar{q}, n) d\gamma d\theta \right] g(\bar{q} | n) \\
&\quad + \int_{\bar{q}}^{\infty} \left\{ \frac{\partial \bar{q}^0}{\partial \tilde{z}_{n0}} \left[ \int_{-\infty}^0 m(q - \bar{q}^0, \theta | q, n) d\gamma d\theta + \int_0^{\infty} m(\theta + q - \bar{q}^0, \theta | q, n) d\gamma d\theta \right] \right\} g(q | n) dq \\
&\quad - \frac{\partial \bar{q}}{\partial \tilde{z}_{n0}} \left[ \int_{-\infty}^0 \int_{\Delta v}^{\infty} m(\gamma, \theta | \bar{q}, n) d\gamma d\theta + \int_0^{\infty} \int_{\theta+\Delta v}^{\infty} m(\gamma, \theta | \bar{q}, n) d\gamma d\theta \right] g(\bar{q} | n).
\end{aligned} \tag{B.53}$$

Once again relying on the fact that  $\partial \bar{q}^0 / \partial \tilde{z}_{n0} = 0$  and  $\partial \bar{q} / \partial \tilde{z}_{n0} = \partial \Delta v / \partial \tilde{z}_{n0} = \lambda \left[ (1 - \tau_0 - B'(z)) + \frac{v_z}{v_c} \right]$ , we can rearrange terms to yield:

$$\begin{aligned}
\frac{\partial \Pr(z_{n1} > 0 | \tilde{z}_{n0})}{\partial \tilde{z}_{n0}} &= \lambda \left[ (1 - \tau_0 - B'(z)) + \frac{v_z}{v_c} \right] \times \left\{ \left[ \int_{-\infty}^0 \int_{-\infty}^{\Delta v} m(\gamma, \theta | \bar{q}, n) d\gamma d\theta \right] g(\bar{q} | n) \right. \\
&\quad \left. + \int_{-\infty}^{\bar{q}} \int_{-\infty}^{\Delta v} m(\gamma, \bar{q} - q | q, n) g(q | n) d\gamma dq \right\}.
\end{aligned} \tag{B.54}$$

The kink in the probability of having positive earnings can now be expressed as:

$$\lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \frac{d \Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d \tilde{z}_{n0}} - \lim_{\tilde{z}_{n0} \rightarrow z^{AET-}} \frac{d \Pr(z_{n1} > 0 | \tilde{z}_{n0})}{d \tilde{z}_{n0}} = \beta, \tag{B.55}$$

where

$$\begin{aligned}
\beta &= \lim_{\tilde{z}_{n0} \rightarrow z^{AET+}} \lambda \left[ (1 - \tau_0 - db) + \frac{v_z}{v_c} \right] \times \left\{ \left[ \int_{-\infty}^0 \int_{-\infty}^{\Delta v} m(\gamma, \theta | \bar{q}, n) d\gamma d\theta \right] g(\bar{q} | n) \right. \\
&\quad \left. + \int_{-\infty}^{\bar{q}} \int_{-\infty}^{\Delta v} m(\gamma, \bar{q} - q | q, n) g(q | n) d\gamma dq \right\} \\
&= -db \cdot \lambda_{n^{AET}} \cdot g(\bar{q}_{n^{AET}} | n^{AET}) \times \underbrace{\left[ \int_{-\infty}^0 \int_{-\infty}^0 m(\gamma, \theta | \bar{q}_{n^{AET}}, n^{AET}) d\gamma d\theta \right]}_{\Pr(\text{claim} | n^{AET}, q_{n^{AET}} = \bar{q}_{n^{AET}})} \\
&\quad - db \cdot \lambda_{n^{AET}} \cdot \int_{-\infty}^{\bar{q}} \int_{-\infty}^0 m(\gamma, \bar{q} - q | q, n^{AET}) g(q | n^{AET}) d\gamma dq,
\end{aligned} \tag{B.56}$$

and where again we have used the fact that  $\lim_{\tilde{z}_{n0} \rightarrow z^{AET}} \Delta v_n = 0$ .

The first term in this expression is an attenuated version of our previous kink in the probability of working. The extra term can be interpreted as the probability of claiming among the marginal labor force participants, *i.e.* those for whom  $q_{n^{AET}} = \bar{q}_{n^{AET}}$ . The integral covers the range of values for  $\gamma$  and  $\theta$  that render working while claiming preferable to either not working while not claiming or working while not claiming. Since the individual is indifferent between working while claiming and not working while claiming, this also implies that claiming is preferable to not claiming among this set of agents. The second term may be recognized as one component of the kink in claiming. In particular, the integral captures individuals who prefer working while claiming to not working while claiming, and prefer working while claiming to working while not claiming, but are indifferent between working while claiming and not working while not claiming.



The kink in the budget set while working and claiming shifts this marginal individual to not claiming, at which point not working becomes optimal.

Thus, endogenizing claiming has two effects on the kink in the probability of positive earnings. First, it attenuates this behavioral response, because the kink only affects those who claim. Second, it amplifies the behavioral response if there are individuals who are shifted to not claiming, and prefer to not work conditional on not claiming, but to work conditional on claiming. In this case, we cannot generally conclude whether our estimate of the participation elasticity is a lower bound or an upper bound. However, we detail below two approaches for establishing an upper bound on the (negative) behavioral response of participation under certain restrictions on the parameters (*i.e.* a lower bound on the absolute value of the participation response).

### B.5.3 Bounding the Behavioral Response

One method for establishing a bound involves imposing an additional restriction on unobserved heterogeneity. In particular, we can assume that  $\gamma_n \cdot \theta_n \geq 0$ , *i.e.* the signs of  $\gamma$  and  $\theta$  are the same. In this case, delaying claiming must cause utility when working and not working to either both increase or both decrease, although the absolute value of the change in utility can be different across the two states. This assumption makes intuitive sense, as we are restricting analysis to the case in which earnings are fixed at  $\tilde{z}_{n0}$ . In this case, delaying claiming does not affect employment, and delaying claiming only affects utility through its direct effect on either current disposable income, or (in a dynamic setting) lifetime benefits. This would be the case, for example, if delaying claiming has the same effect on wealth, regardless of working status. The assumption rules out cases in which, for example, working affects life expectancy in a way that changes the sign of the effect of delaying claiming on lifetime benefits, or a case in which working relaxes a borrowing constraint and the effect of delaying claiming on utility therefore changes sign, conditional on working.

If our assumption holds, we can show the following:

$$\int_{-\infty}^{\bar{q}} \int_{-\infty}^0 m(\gamma, \bar{q} - q | q, n^{AET}) g(q | n^{AET}) d\gamma dq = 0 \quad (\text{B.57})$$

The reason why this probability becomes zero is as follows. First, we are restricting attention to values of  $q \leq \bar{q}$ . In addition, we are fixing the value of  $\theta = \bar{q} - q \geq 0$ . However, we are also restricting analysis to individuals for whom  $\gamma \leq 0$ . Given our assumption that the signs of  $\gamma$  and  $\theta$  must be the same, this set must be of measure zero.

In this case, we can simplify our two kinks above as follows:

$$\delta = -db \cdot \lambda_{n^{AET}} \times \int_{-\infty}^{\bar{q}} \int_{-\infty}^{\bar{q}-q} m(0, \theta | q, n^{AET}) g(q | n^{AET}) d\theta dq \quad (\text{B.58})$$

$$\beta = -db \cdot \lambda_{n^{AET}} \times g(\bar{q} | n^{AET}) \underbrace{\left[ \int_{-\infty}^0 \int_{-\infty}^0 m(\gamma, \theta | \bar{q}_{n^{AET}}, n^{AET}) d\gamma d\theta \right]}_{\Pr(\text{claim} | n^{AET}, q_{n^{AET}} = \bar{q}_{n^{AET}})} \quad (\text{B.59})$$

The kink in participation we estimate is now a weak upper bound on a negative kink (*i.e.* the absolute value of the kink in participation we estimate is a lower bound on the kink that would obtain absent a claiming response). This implies that our observed elasticity will be a lower bound. Note that the probability of claiming that attenuates the behavioral response is local to marginal labor force participants, *i.e.*  $q = \bar{q}$ , and thus is not the same as the population probability of claiming at  $\tilde{z}_{n0} = z^{AET}$ .

If we are not willing to impose the above restrictions on the joint distribution of  $\gamma$  and  $\theta$ , we can still achieve a lower bound. Using the results from the previous two sections, we have:

$$\begin{aligned} \beta - \delta &= -db \cdot \lambda_{n^{AET}} \times g(\bar{q}_{n^{AET}} | n^{AET}, t) \left[ \int_{-\infty}^0 \int_{-\infty}^0 m(\gamma, \theta | \bar{q}_{n^{AET}}, n^{AET}) d\gamma d\theta \right] \\ &\quad + db \cdot \lambda_{n^{AET}} \times \int_{-\infty}^{\bar{q}} \int_{-\infty}^{\bar{q}-q} m(0, \theta | q, n^{AET}) g(q | n^{AET}) d\theta dq \end{aligned} \quad (\text{B.60})$$

Again, this difference in kinks provides an upper bound on the negative kink in participation. Note that

this bound is not guaranteed to be negative. However, in cases in which the kink in claiming is relatively negligible—as appears to be the case in our empirical application—it is possible to establish a non-trivial upper bound on the kink and, by extension, a lower bound on the employment elasticity. Furthermore, if there is no detectable kink in claiming, then simply rescaling the kink in employment by the share claiming is sufficient to adjust for endogenous claiming.

## C Procedure for estimating excess mass

As we explain in Gelber, Jones, and Sacks (2013), we seek to estimate the “excess mass” at the kink, *i.e.* the fraction of the sample that locates at the kink under the kinked tax schedule but not under the linear tax schedule. Following a standard procedure in the literature (*e.g.* Saez, 2010; Chetty, Friedman, Olsen, and Pistaferri, 2011), we estimate the counterfactual density (*i.e.* the density in the presence of a linear budget set) by fitting a smooth polynomial to the earnings density away from the kink, and then estimating the “excess” mass in the region of the kink that occurs above this smooth polynomial.

Specifically, for each earnings bin  $z_i$ , we calculate  $p_i$ , the proportion of the sample with earnings in the range  $[z_i - k/2, z_i + k/2)$ . The earnings bins are normalized by distance-to-kink, so that for  $z_i = 0$ ,  $p_i$  is the fraction of all individuals with earnings in the range  $[0, k)$ . To estimate bunching, we assume that  $p_i$  can be written as:

$$p_i = \sum_{d=0}^D \beta_d \cdot (z_i)^d + \sum_{j=-k}^k \gamma \cdot 1\{z_i = j\} + \varepsilon_i \quad (\text{C.61})$$

and run this regression (where 1 denotes the indicator function and  $j$  denotes the bin). This equation expresses the earnings distribution as a degree  $D$  polynomial, plus a set of indicators for each bin within  $k\delta$  of the kink, where  $\delta$  is the bin width. In our empirical application, we choose  $D = 7$ ,  $\delta = 500$  and  $k = 6$  as our baseline (so that six bins are excluded from the polynomial estimation). We control for a baseline seventh-degree polynomial through the density following Chetty, Friedman, Olsen, and Pistaferri (2011). The parameter  $\gamma$  reflects the excess density near the kink.

Our measure of excess mass is  $\hat{M} = 2k\gamma$ , the estimated excess probability of locating at the kink (relative to the polynomial term). This measure depends on the counterfactual density near the kink, so to obtain a measure of excess mass that is comparable at the kink, we scale by the predicted density that we would obtain if there were a linear budget set. This is just the constant term in the polynomial, since  $z_i$  is the distance to zero. Thus, our estimate of normalized excess mass is  $\hat{B} = \frac{\hat{M}}{\hat{\beta}_0}$ . We calculate standard errors using the delta method. We calculate the density in each bin by dividing the number of beneficiaries in the bin by the total number of beneficiaries within the bandwidth; note that this normalization should not affect the excess normalized mass or the estimated density, because dividing by the total number of beneficiaries within the bandwidth affects the numerator (*i.e.*  $\hat{M}$ ) and denominator (*i.e.*  $\hat{\beta}_0$ ) of the expression for  $\hat{B}$  ( $= \frac{\hat{M}}{\hat{\beta}_0}$ ) in equal proportions and therefore should have no impact on  $\hat{B}$ .

## D Method of Adjusting for Measurement Error

Our baseline estimates treat 60 earnings as a perfect proxy for desired earnings at age 63. However, earnings change from year-to-year, even at ages not subject to the earnings test, so it is natural to think that desired earnings would change as well. Such changes imply that our baseline first stage estimate of the kink in ANTR (at age 63) given age 60 income is too large—*i.e.* measurement error in our running variable will lead to an attenuated first stage, and therefore a larger elasticity estimate. This appendix describes our approach to correcting for this measurement error, both in estimation and inference. We also present simulations validating the approach.

### D.1 Estimation

In general, we estimate the extensive margin elasticity with the following plug-in estimator:

$$\hat{\eta} = \frac{\hat{\beta}_{PE}}{\hat{\alpha}_{ANTR}} \cdot \frac{\hat{\mu}_{ANTR}}{\hat{\mu}_{PE}}, \quad (\text{D.62})$$

where  $\hat{\beta}_{PE}$  is the estimated kink in positive earnings (the reduced form),  $\hat{\alpha}_{ANTR}$  is the estimated kink in ANTR (the first stage kink),  $\hat{\mu}_{PE}$  is the estimated probability of employment at the threshold (in the data), and  $\hat{\mu}_{ANTR}$  is the estimated average ANTR at the threshold. In our baseline estimate we impute the ANTR at age 63 using age 60 earnings, and estimate  $\hat{\alpha}_{ANTR}$  from an RK of this imputed value on age 60 distance. We now present an alternative approach where we calibrate an earnings growth process, and use that process to simulate the first stage, which yields estimates for  $\hat{\alpha}_{ANTR}$  and  $\hat{\mu}_{ANTR}$ .

### D.1.1 Calibrating and drawing from the earnings growth distribution

Before simulating the first stage, we must calibrate earnings growth from age 60 to age 63. We consider three earnings growth distributions. All three are based on the earnings growth from age 59 to age 60. We chose these ages because data from ages even earlier than age 59 may involve very different dynamics, as most people are not yet approaching retirement. Later ages are also unsuitable because observed earnings may have responded endogenously to the earnings test.

The first earnings growth process assumes that earnings growth rates are perfectly correlated from one year to the next, with the one-year distribution given by the age 59 to 60 growth rate distribution shown in Figure 8, Panel A; call this distribution  $H$ . To draw from the earnings growth distribution, we simply draw a one-year growth rate  $r_i$  from  $H$ , and then obtain a three-year growth rate as  $(1 + r_i)^3 - 1$ . Our second earnings growth process is identical except it assumes that earnings growth is perfectly independent instead of perfectly persistent. To draw from the earnings growth distribution in this case, we therefore take three draws  $r_{i1}, r_{i2}$  and  $r_{i3}$  from  $H$ , and obtain the three-year growth rate as  $(1 + r_{i1})(1 + r_{i2})(1 + r_{i3}) - 1$ .

In the third approach, we assume that earnings growth rates are conditionally independent given a time invariant type. Specifically, we divide people up into eight categories determined by gender and quartiles of permanent income, defined as average annual income between ages 35 and 55. We use these ages because annual earnings are not recorded prior to 1950, so starting at 35 ensures that we can properly define permanent income. We assume that income growth is normally distributed conditional on type, and we estimate the mean and variance of income growth (at ages 59 to 60) given type. We do this calibration using the SSA Earnings Public Use File (EPUF), which contains a one-percent sample of SSA earnings histories (Social Security Administration 2011). To draw from this distribution, we first draw a type for each person (using the observed probabilities in the public use file) and then draw three earnings growth rates,  $r_{i1}, r_{i2}$  and  $r_{i3}$  from the type-specific distribution. The three-year growth rate is  $(1 + r_{i1})(1 + r_{i2})(1 + r_{i3}) - 1$ .

### D.1.2 Simulating the first stage

Given a calibrated earnings growth process, we simulate the first stage in the following steps:

1. Resample from the distribution of age 60 earnings, shown in Figure 2 (assuming an exempt amount of \$10,000). This yields a data set on age 60 earnings,  $z_i$ , corresponding to the running variable used in our analysis.
2. For each observation, draw from the calibrated earnings growth distribution and simulate earnings forward three years (as explained above). This yields a value of age 63 earnings and distance to the exempt amount,  $z_i^*$ , for each observation.
3. Given  $z_i^*$ , find the ANTR at age 63,  $ANTR_i^*$ , using the statutory formula.
4. Estimate an RK of age 63  $ANTR_i^*$  on  $z_i$ , using the main bandwidth in estimation (\$2,800). Record the estimated kink and the mean  $ANTR_i^*$  at  $Z_i = 0$ .

For each calibrated earnings growth process, we repeat this simulation 1,000 times. We obtain  $\hat{\alpha}_{ANTR}$  and  $\hat{\mu}_{ANTR}$  as the average kink and mean ANTR across these 1,000 iterations. We then plug these values into Equation (D.62) to obtain our elasticity estimate for each iteration.

## D.2 Inference

We consider two approaches to inference. As our elasticity estimate is a nonlinear function of other estimates, we report an approximation for delta method standard errors. Implementing delta method standard errors requires that we estimate the asymptotic covariance of the first stage and reduced form, but limited access to administrative data precludes directly calculating this. Our delta method standard errors therefore

assume that all estimates are independent. We call this approach the approximate delta method. Under the independence assumption, the delta method standard error simplifies to

$$s.e.(\eta) = \eta \sqrt{\frac{Var(\hat{\beta}_{PE})}{\hat{\beta}_{PE}^2} + \frac{Var(\hat{\alpha}_{ANTR})}{\hat{\alpha}_{ANTR}^2} + \frac{Var(\hat{\mu}_{PE})}{\hat{\mu}_{PE}^2} + \frac{Var(\hat{\mu}_{ANTR})}{\hat{\mu}_{ANTR}^2}}, \quad (D.63)$$

where  $Var(\hat{X})$  is the variance of the estimate. We estimate  $Var(\hat{\beta}_{PE})$  and  $Var(\hat{\mu}_{PE})$  using the asymptotic standard errors of  $\hat{\beta}_{PE}$  and  $\hat{\mu}_{PE}$ . We estimate  $Var(\hat{\alpha}_{ANTR})$  and  $Var(\hat{\mu}_{ANTR})$  using the variance of  $\hat{\alpha}_{ANTR}$  and  $\hat{\mu}_{ANTR}$  across simulation draws.

Because we impose independence of the estimates, the delta method standard errors may be biased, and the simulation evidence described below indicates that we over reject true null hypotheses. We therefore consider an alternative approach. Below we simulate data sets based on a statistical model that closely corresponds to observable moments in our data. Within this statistical model, we can impose the null hypothesis that  $\eta = 0$ . Simulations for this model can therefore give us an estimate of the distribution of  $\hat{\eta}$  under the null hypothesis. Call this distribution  $F_0$ . Our second approach to inference uses this distribution to calculate p-values, that is  $p = 2(1 - F_0^{-1}(\hat{\eta}))$ , where  $\hat{\eta}$  is our actual elasticity estimate. This guarantees that we reject the true null five percent of the time, assuming our statistical model of measurement error model is properly specified.

### D.3 Validation

The above procedure yields an elasticity estimate and standard error that adjust for measurement error stemming from year-to-year earnings growth. We investigate the performance of this procedure in a monte carlo simulation that closely resembles the data generating process. The basic idea is to impose an elasticity and measurement error process, simulate many data sets, apply our procedure to each, obtain a distribution of estimates, and assess the bias and power of our estimator. We find, reassuringly, that our procedure is approximately unbiased. However the approximate delta method standard errors end up overrejecting true null hypotheses. As an alternative approach, we therefore construct p-values based on the (simulated) distribution of estimates under the imposed null hypothesis.

**Preliminary Calibration** The structural objects in our simulation are the distribution of age 60 earnings, the distribution of growth rates, and the expected value of the outcome given desired earnings and ANTR. We have already described Section D.1 how we calibrate and draw from the age 60 earnings distribution and the distribution of growth rates. The remaining object to calibrate is the expected value of the outcome given  $z_i^*$  and the endogenous regressor.

We calibrate this object to match the empirical relationship between positive earnings and age 60 earnings. This relationship is shown in Figure E.7, a zoomed out version of Figure 4. Each dot shows the average probability of positive earnings in each \$100 bin of distance to the exempt amount. The simulation assumes that this probability can be written as a smooth component plus a component which is kinked due to the kink in the ANTR. That is

$$E[PE_i | z_i^*, ANTR_i^*] = f(z_i^*) + \beta \cdot ANTR_i^*, \quad (D.64)$$

where  $f$  is a degree seven polynomial, and  $\beta$  is the key structural parameter, giving the effect of the ANTR on the probability of positive earnings. In the data, we estimate a reduced form kink of -1.85, which should be interpreted to mean that for an increase in \$1,000 in the running variable, the increase in the probability of positive earnings is 1.85 percentage points less just the right of the threshold than to the left. This estimate implies that  $\beta^{Data} = 0.37$  (assuming a sharp kink of -5 in ANTR as a function of desired earnings). We use this value to estimate  $f$ . To do so, we define the adjusted variable  $\tilde{PE}$  as  $PE_i - \beta^{Data} \cdot ANTR$ . We then regress  $\tilde{PE}$  on a degree 7 polynomial. We plot the implied values in the absence of the earnings test,  $f(z) + \beta^{Data} \cdot \overline{ANTR}$ , as the dashed line in the figure (where  $\overline{ANTR} = 75$  is the average net-of-tax rate in the data in the absence of the AET BRR). We also plot  $f + \beta^{Data} \cdot ANTR$  in the solid line, to show the overall fit. Note that, because Appendix Figure E.7 is based on real data, we subtract out the estimated kink, but in simulating the data, we impose different values of  $\beta$  to obtain either a zero elasticity or an elasticity closer to the measurement-error corrected value.

**Simulation details for validation exercise** The simulation is the same as the process for simulating the first stage, except we also simulate the outcome, employment, and we let the first stage bandwidth differ

from iteration to iteration. The simulation approach is similar to that used in Card et al. (2017). The simulation works as follows.

1. Fix the structural parameter  $\beta$  (explained above) and fix the measurement error process.
2. Draw a data set on the observed running variable,  $Z_i$ ; the desired earnings at age 63 relative to the threshold,  $Z_i^*$ ; the ANTR at age 63 given desired earnings,  $ANTR_i^*$ , and the outcome  $PE_i$ , as follows.
  - (a) Resample from the distribution of age 60 earnings, shown in Figure 2 (and assuming a kink point of \$10,000). This yields a data set on age 60 earnings,  $z_i$ , corresponding to the running variable used in our analysis.
  - (b) For each observation, simulate earnings forward three years according to the calibrated earnings growth distribution (as explained above). This yields a value of age 63 earnings and distance to the exempt amount,  $z_i^*$ , for each observation.
  - (c) Given  $z_i^*$ , find the ANTR at age 63,  $ANTR_i^*$ , using the statutory formula.
  - (d) Given  $z_i^*$  and  $ANTR_i^*$ , find  $E[PE_i|z_i^*, ANTR_i^*]$  as

$$E[PE_i|z_i^*, ANTR_i^*] = f(z_i^*) + \beta \cdot ANTR_i^*. \quad (\text{D.65})$$

Where  $f$  and  $\beta$  are calibrated as described above.

- (e) For each observation, draw a uniform error  $e_i$ , and set

$$PE_i = \mathbf{1}\{e_i \leq E[PE_i|z_i^*, ANTR_i^*]\}, \quad (\text{D.66})$$

where  $\mathbf{1}\{\cdot\}$  is the indicator function.

3. Estimate an RK of  $PE_i$  on  $z_i$ , using the CCT procedure to find the bandwidth. Record the  $\hat{\beta}_{PE}$  and  $\hat{\mu}_{PE}$ , as well as their standard errors.
4. Estimate an RK of  $ANTR_i^*$  on  $z_i$  using the same bandwidth as in step (3). Record the estimated kink and the mean  $ANTR_i^*$  at  $Z_i = 0$ .
5. Repeat steps 2-4 1,000 times, yielding 1,000 sets of estimates.
6. Find the average estimated kink in ANTR and mean ANTR. Obtain the elasticity in each iteration by dividing using the iteration-specific reduced form and the *average* first stage kink and mean. Calculate delta method standard errors (using the asymptotic standard errors of the reduced form and the standard deviation of estimates in the first stage).
7. Repeat steps 1-6 for each measurement error process (the three described above, as well as a no measurement error benchmark), and for two assumed structural parameters,  $\beta = 0$  and  $\beta = 1.85$ . These parameters correspond to an elasticity of zero and 2.41, the latter being five times our baseline estimate, yielding an elasticity which is roughly our estimate after adjusting for measuring error (as reported in Appendix Table E.4).

**Discussion** This approach yields a distribution of estimates given known reduced forms and first stages, and hence we can use it to assess the bias of our estimator, as well as its statistical performance. However, some elements of this simulation may appear non-standard, and we therefore explain them further. First, note that we estimate a first stage and reduced form in each iteration, but rather than use a different first stage to obtain the elasticity in each iteration, we use the same (average) first stage across all iterations. This is to parallel our empirical approach, which also uses the average first stage to obtain an elasticity. Likewise, we use the variance of first stage estimates as our estimated variance of the estimator, as we do in our main estimator. Second, in our empirical implementation, we set the first stage bandwidth equal to our reduced form bandwidth. To mimic that approach here, we use a different first stage bandwidth in each iteration, corresponding to the reduced form bandwidth from that iteration.

### D.3.1 Results

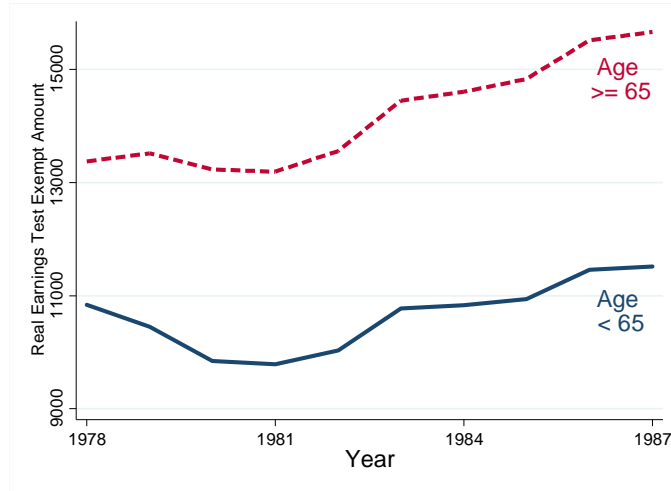
Our results from two sets of simulations are reported in Appendix Table E.5. The first table reports estimates using  $\beta = 1.85$  and the second using  $\beta = 0$ , corresponding to elasticities of 2.41 and 0. The tables report the mean and SD of the estimated reduced form and first stage kinks, as well as statistics on the estimated elasticities and the structural parameter  $\beta$ . For comparison, the first column reports on results with no measurement error. The remaining columns report the three measurement error cases.

We begin by discussing the positive elasticity case, presented in panel A of Appendix Table E.5. This case, which is based on  $\beta = 1.85$ , is closest to our empirical setting. We find that, although measurement error attenuates the reduced forms, the mean elasticities are similar across columns and roughly equal to the true value, 2.41, indicating that the estimator is roughly unbiased and effectively corrects for measurement error. We reject the truth only between 4.4 and 9.8 percent of the time (with inference conducted via the approximate delta method as described above). Some amount of bias is perhaps to be expected given that we use a nonparametric procedure, which trades bias against variance. One possible concern evident in these results is that the standard errors are fairly large, and, as a result, our power ranges from 42 to 73 percent, depending on the type of measurement error.

We now next turn to the zero elasticity case in Panel B. Across all specifications, including ones with no measurement error, the average reduced form kink is slightly negative, because our nonparametric procedure introduces some bias. We also find that our inference (based on the approximate delta method) ends up overrejecting the true null hypothesis, with rejection rates of 19-26 percent. We obtain these high rejection rates even in the absence of measurement error. We therefore conclude that they are a consequence of the bias in the nonparametric procedure or our inference approach, rather than the measurement error and its correction per se. However, the overrejection in these simulations implies that we may be likely to obtain a statistically significant elasticity estimate even if the true elasticity is zero. An alternative approach to inference can avoid this problem. Specifically, as a complementary approach, we obtain  $p$ -values by comparing our estimated elasticity to the distribution of elasticity estimates obtained in our simulation under the null hypothesis. These  $p$ -values are reported in the last column of Appendix Table E.4.

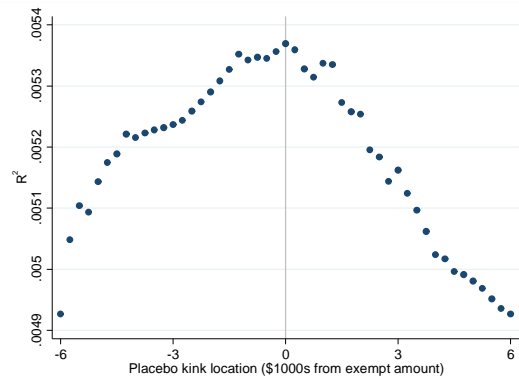
## E Appendix Figures and Tables

Appendix Figure E.1: Earnings Test Real Exempt Amount, 1978 to 1987



Notes: The figure shows the real value of the exempt amount over time among those 62-64 years old (labeled “Age<65” in the graph) and those above (labeled “Age>=65”). The AET applied to earnings of claimants from ages 62 to 71 from 1978 to 1982, but only to claimants aged 62 to 69 from 1983 to 1989. All dollar figures are expressed in real 2010 dollars.

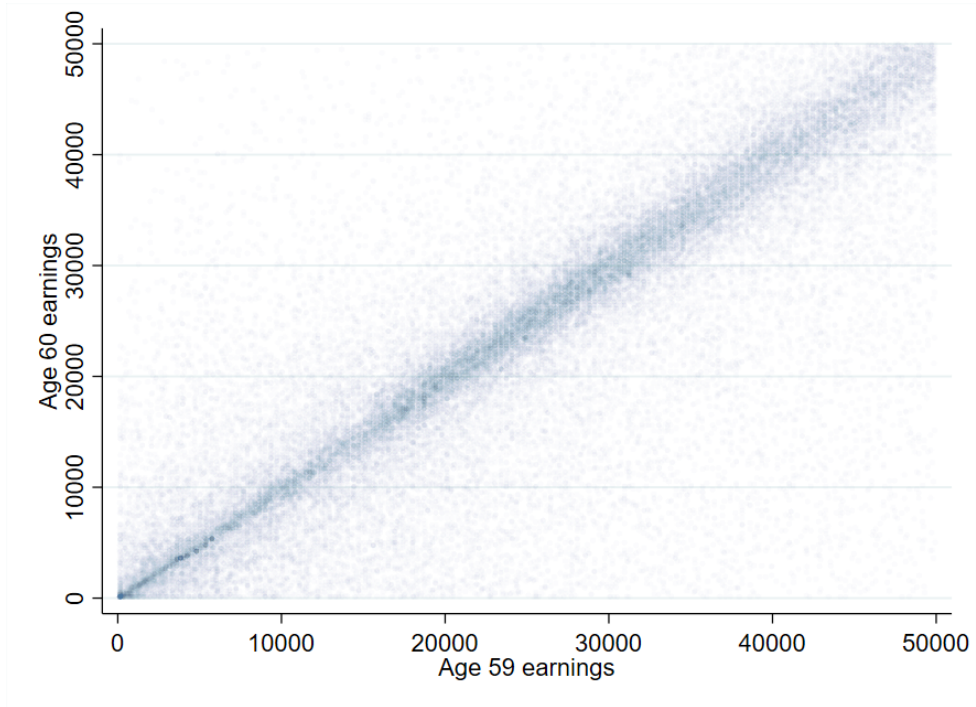
Appendix Figure E.2: R-Squared by Placebo Kink Location



Notes: The figure plots the R-squared of our baseline specification against the “placebo” kink location relative to the exempt amount, following Landais (2015). The vertical line denotes the actual location of the exempt amount. As described in the text, we estimate a set of placebo changes in slope in the mean annual age 63 to 64 employment rate, using the same specification as our main estimates except that we examine the change in slope at placebo locations of the exempt amount away from the true exempt amount. The figure shows that the R-squared is maximized at the true location of the placebo kink, supporting our hypothesis that we have found a true kink in the data rather than a spurious underlying nonlinearity in the relationship between the yearly employment rate at ages 63 to 64 and age 60 earnings. See other notes to Figure 2.

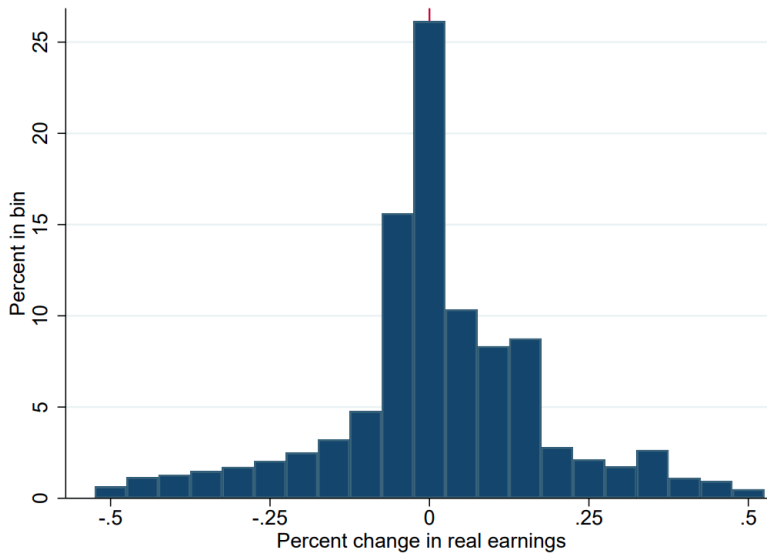


Appendix Figure E.3: Age 59 Earnings versus Age 60 Earnings, (45 Degree Line Omitted)



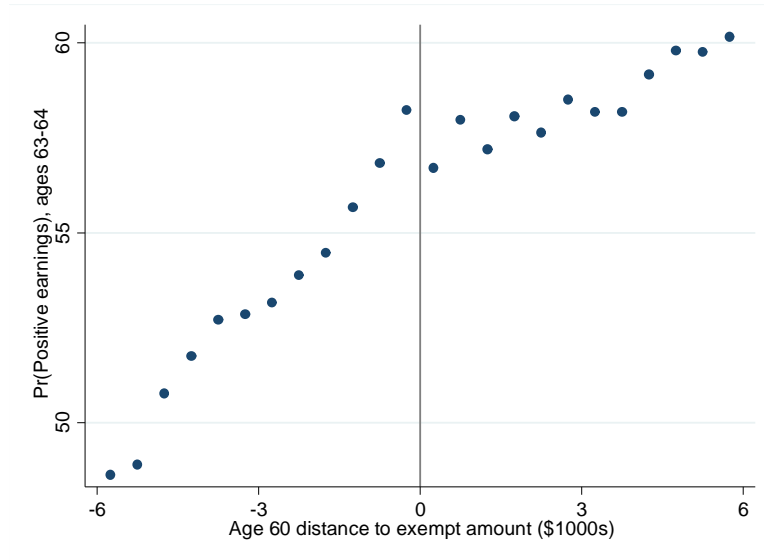
Notes: The figure reproduces Panel B of 8, but omits the 45 degree line. See Figure 8 for additional details.

Appendix Figure E.4: Histogram of Percent Earnings Growth, Ages 28 to 59



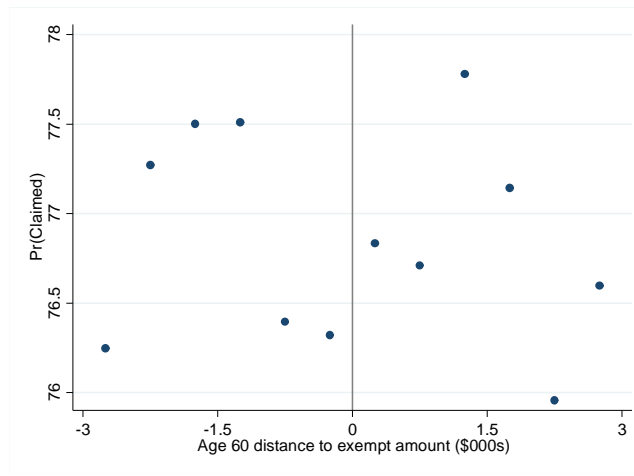
Notes: This histogram shows that there is a large mass near zero percent nominal earnings growth from one year to the next in a wider set of ages, 28 to 59, than we focus on in the main analysis. This indicates that among a broad set of ages, a substantial mass of individuals have no growth in desired nominal earnings, consistent with the assumptions necessary for our RKD to estimate a lower bound on the elasticity as described in the main text. This suggests that it should be possible to use our method, implemented in this case through using lagged earnings to proxy for desired earnings, when studying extensive margin responses to other policies applying in other age ranges. The figure uses the SSA data we have, covering the 1918 to 1923 cohorts in calendar years from 1951 to 1984. See other notes to Figure 2.

Appendix Figure E.5: Probability of Positive Earnings at Ages 63 to 64, wider x-axis



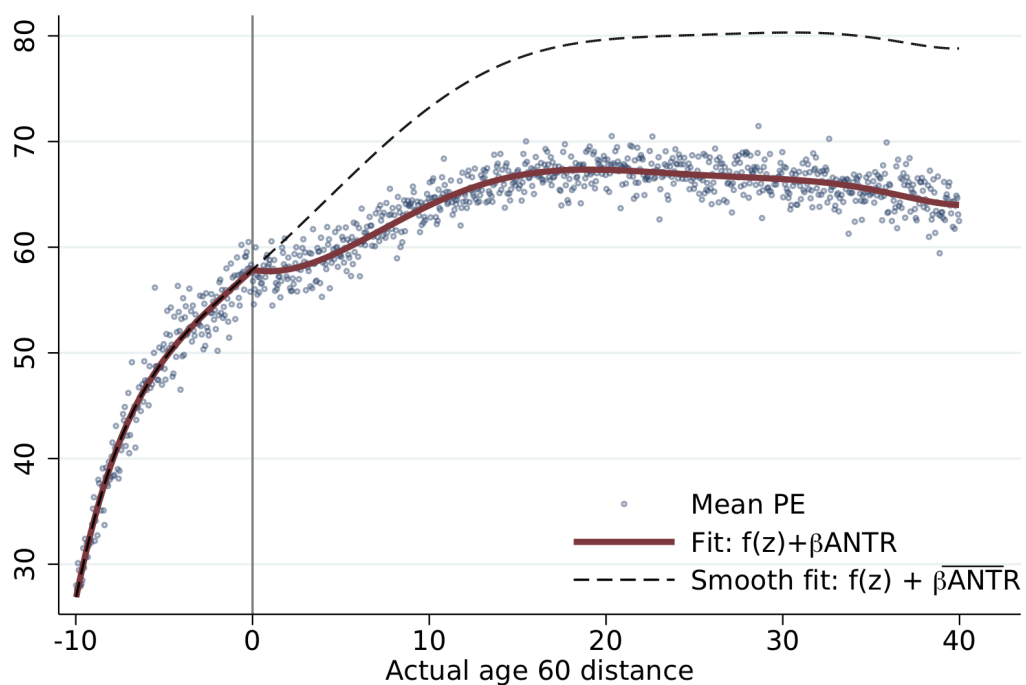
Notes: See the notes to Figure 4. This figure is identical to Figure 4, except that the range of the  $x$ -axis on this figure runs from  $-\$6,000$  to  $\$6,000$ . Like Figure 4, this figure also shows a clear, discontinuous change in slope at the exempt amount.

Appendix Figure E.6: Mean Probability of Claiming at Ages 63 to 64



Notes: The figure plots the mean claiming rate, *i.e.* the probability someone has claimed by the calendar year of reaching age  $t$ , at ages 63 to 64 averaged, as a function of the distance of age 60 earnings from the exempt amount. The figure shows that there is no clear visual change in the slope of the claiming rate, and regression evidence supports the same conclusion: a placebo test in the spirit of Ganong and Jäger (2015) shows  $p=0.15$  for the two-sided test of equality of the coefficient with zero. See other notes to Figure 2.

Appendix Figure E.7: Employment as a Function of Earnings, Actual and Simulated Counterfactual



Notes: Each dot shows the average probability of positive earnings at age 63 in each \$100 bin of age 60 distance to the exempt amount. The dashed black line is the estimated smooth fit after taking out the kink, obtained from a regression of  $PE - \beta \cdot ANTR$  on a degree seven polynomial in distance (with  $\beta = 0.37$  and the fitted values adding back in  $\beta 75$ ). The solid red line shows the fitted values including the kink.

Appendix Table E.1: Robustness of Elasticity Estimates

	(1)	(2)
	Baseline Specification	With Controls
Linear	0.49 (0.19)***	0.44 (0.16)***
<i>N</i>	95,960	104,665
Quadratic	0.64 (0.27)***	0.64 (0.27)***
<i>N</i>	160,785	172,979
Cubic	0.80 (0.27)***	0.64 (0.26)***
<i>N</i>	273,421	326,762

Notes: The table presents robustness checks on the main elasticity estimates in Table 5. The “with controls” column shows the kink in the employment probability when we control for dummies for year of birth, sex, and race. Robust standard errors, using the procedure of Calonico, Cattaneo, and Titiunik (2014), are reported in parentheses. The number of individuals included in each “reduced form” regression (4) is shown below the standard error. See other notes to Table 2.

Appendix Table E.2: RKD Elasticity Estimates using Regression-Based First Stage

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full Sample	Men	Women	White	Non- White	High Prior Earnings	Low Prior Earnings
First Stage Kink	-4.00 (0.01)***	-3.25 (0.01)***	-3.99 (0.01)***	-3.88 (0.01)***	-3.09 (0.01)***	-3.76 (0.01)***	-3.82 (0.01)***
Elasticity	0.59 (0.23)***	0.36 (0.28)	0.60 (0.26)***	0.59 (0.22)***	0.64 (0.68)	0.62 (0.62)	0.52 (0.21)***
<i>N</i>	95,960	68,971	66,251	93,722	39,271	19,574	101,709

Notes: See notes to Table 5. Relative to Table 5, the elasticities differ here because we use a linear RKD to estimate the first stage kink in the ANTR (reported in the first row), rather than using the analytic expression as in Table 5.

Appendix Table E.3: Elasticity Estimates Accounting for Claiming

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full Sample	Men	Women	White	Non- White	High Prior Earnings	Low Prior Earnings
Elasticity	0.63 (0.25)***	0.31 (0.24)	0.64 (0.28)**	0.61 (0.23)***	0.53 (0.56)	0.62 (0.62)	0.53 (0.21)***
<i>N</i>	95,960	68,971	66,251	93,722	39,271	19,574	101,709

Notes: See notes to Table 5. As explained in the main text, we calculate these elasticities by inflating the Table 5 elasticities by 29.9 percent, to account for claiming behavior. Among those with age 60 earnings below the kink, but not more than \$2,797 below the kink, 74.5 percent of the sample claims by age 63, and 79.5 by age 64. We calculate 29.9 percent as  $100 \cdot (1 / [(0.795 + 0.745) / 2] - 1)$ .

Appendix Table E.4: Elasticity Estimates, Adjusted for Measurement Error

Growth Process	First-Stage Kink	Elasticity	Standard Error	$p$ -value
Perfect persistence growth rate	-0.94	2.37	1.03	0.03
Perfect independence growth rate	-1.05	2.18	0.92	0.03
Conditionally independent growth rate	-0.81	2.86	1.46	0.06

Notes: Table reports, for the indicated measurement error process, the implied elasticity, as well as the standard error of that elasticity (obtained via the approximate delta method) and the  $p$ -value of the null hypothesis that the elasticity is zero (obtained via simulating the distribution of estimates under the null hypothesis). See Appendix D for more details.

Appendix Table E.5: Validating the Measurement Error Correction

Assumed growth process	Growth rate is:			
	Zero	Perfectly Persistent	Perfectly Independent	Conditionally Independent
A. Simulation w/ positive elasticity ( $\eta = 2.41$ )				
Mean first stage kink	-4.237	-0.914	-0.930	-0.642
Mean reduced form kink	-7.940	-1.689	-1.621	-1.141
Mean $\hat{\beta}$	1.874	1.847	1.744	1.778
Mean $\hat{\eta}$	2.425	2.618	2.408	2.472
Mean SE $\hat{\eta}$	0.240	1.049	1.269	1.830
Fraction reject $\eta = 0$ (approximate $\delta$ -method)	1.000	0.731	0.594	0.415
Fraction reject $\eta = 2.41$ (approximate $\delta$ -method)	0.098	0.073	0.062	0.044
B. Simulation w/ zero elasticity				
Mean first stage kink	-3.617	-1.094	-1.393	-1.127
Mean reduced form kink	-0.065	-0.179	-0.128	-0.141
Mean $\hat{\beta}$	0.018	0.163	0.092	0.126
Mean $\hat{\eta}$	0.021	0.191	0.110	0.151
Mean SE $\hat{\eta}$	0.110	0.365	0.285	0.373
Fraction reject $\eta = 0$ (approximate $\delta$ -method)	0.172	0.257	0.191	0.189

Notes: The table presents results from a validation exercise, where we use our method for adjusting for measurement error on simulated data. The first three rows of each panel report, for each measurement error process, the mean first stage kink (in ANTR), mean reduced form kink (in positive earnings), and mean  $\hat{\beta}$  (obtained as the ratio of the reduced form kink to the mean first stage kink), averaging over 1000 simulated data sets. The remaining rows report the mean elasticity estimate, mean standard error (obtained by the approximate delta method), fraction of iterations in which the null hypothesis  $\eta = 0$  is rejected, and fraction of iterations in which the true  $\eta$  is rejected.