

# Earnings Adjustment Frictions: Evidence from the Social Security Earnings Test

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

September 2013

## Abstract

Using a panel of Social Security Administration microdata on 1 percent of the U.S. population from 1961 to 2006, we study frictions in adjusting earnings to changes in the Social Security Annual Earnings Test (AET). Individuals continue to "bunch" at the convex kink the AET creates even when they are no longer subject to the AET, consistent with the existence of earnings adjustment frictions in the U.S.. We specify a model consistent with the evidence that allows us to estimate the elasticity and adjustment cost. Intuitively, inertia due to an adjustment cost leads to residual bunching after a kink in the budget set becomes less pronounced (or disappears altogether). Our primary estimation method uses the degree of such inertia (in combination with the initial amount of bunching at the kink) in estimating the size of the adjustment cost (and elasticity). We estimate in a baseline case that the elasticity of earnings with respect to the implicit net-of-tax share is 0.23, and the fixed cost of adjustment is \$152.08. Our results imply that eliminating the AET would cause a substantial rise in the earnings of many elderly and near-elderly individuals.

---

\*Gelber: UC Berkeley Goldman School of Public Policy and NBER, [agelber@berkeley.edu](mailto:agelber@berkeley.edu); Jones: Harris and NBER, [damonjones@uchicago.edu](mailto:damonjones@uchicago.edu); Sacks: Wharton, [dansacks@wharton.upenn.edu](mailto:dansacks@wharton.upenn.edu). We thank Raj Chetty, Jim Cole, Jim Davis, Mark Duggan, Jonathan Fisher, Richard Freeman, John Friedman, Bill Gale, Adam Isen, Henrik Kleven, Emmanuel Saez, and numerous seminar participants. We are extremely grateful to David Pattison for generously running the code on the data. We acknowledge financial support from the Wharton Center for Human Resources and the Wharton Risk and Decision Processes Center, from NIH grant #1R03 AG043039-01, from a National Science Foundation Graduate Research Fellowship, and from support from the U.S. Social Security Administration through grant #5RRC08098400-05-00 to the National Bureau of Economic Research (NBER) as part of the SSA Retirement Research Consortium. 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. The research uses data from the Census Bureau's Longitudinal Employer Household Dynamics Program, which was partially supported by the following National Science Foundation Grants: SES-9978093, SES-0339191 and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation. All errors are our own.

# 1 Introduction

In a traditional model of workers’ earnings or labor supply choices, individuals optimize their behavior frictionlessly in response to policies that affect their incentives. However, several recent papers have suggested that individuals face frictions in adjusting behavior to policy (Chetty *et al.*, 2009, 2011, 2012a,b; Chetty, 2012; Kleven and Waseem, forthcoming). Costs of adjusting behavior help to govern the welfare cost of taxation (Chetty *et al.*, 2009), and they also help to explain heterogeneity across contexts in the observed elasticity of earnings with respect to net-of-tax rates (Altonji and Paxson, 1992; Chetty *et al.*, 2011, 2012b; Chetty, 2012).<sup>1</sup>

This paper develops evidence on the existence, nature and size of frictions in adjusting earnings in response to policy. The U.S. Social Security Annual Earnings Test (AET) represents a promising environment for studying these questions. This setting provides a useful illustration of many issues—such as the development and application of a methodology for estimating elasticities and adjustment costs simultaneously—that are applicable to studying earnings responses to policy more broadly. The AET reduces Social Security Old Age and Survivors Insurance (OASI) claimants’ current OASI benefits as a proportion of earnings, once an individual earns in excess of an exempt amount. For example, for OASI claimants aged 62-65 in 2013, current OASI benefits are reduced by 50 cents for every extra dollar earned above \$15,120. The AET may lead to very large effective benefit reduction rates (BRRs) on earnings above the exempt amount, creating a strong incentive for many individuals to “bunch” at the convex kink in the budget constraint located at the exempt amount (Burtless and Moffitt 1985; Friedberg, 1998, 2000).<sup>2</sup> Reductions in current benefits due to the AET sometimes lead to increases in later benefits; nonetheless, as we discuss in detail in Section 2, several factors may explain why individuals’ earnings still respond to the AET.

The AET is an appealing context for studying earnings adjustment for at least three reasons. First, bunching at the AET kink is easily visible on a graph, allowing credible documentation of behavioral responses.<sup>3</sup> Second, the AET represents one of the few known kinks at which bunching occurs in the U.S.; indeed, our paper represents the first study to find robust evidence

---

<sup>1</sup>The net-of-tax rate is defined as one minus the marginal tax rate (MTR).

<sup>2</sup>Other papers on the AET include Gruber and Orszag (2003) and Song and Manchester (2007).

<sup>3</sup>Other papers have examined bunching in the earnings schedule, including Blundell and Hoynes (2004) and Saez (2010). Saez shows that the amount of bunching can be related to the elasticity of earnings with respect to the net-of-tax rate.

of bunching among the non-self-employed at any kink in the U.S.<sup>4</sup> Third, the AET is important to policy-makers in its own right, as it is a significant factor that affects the earnings of the elderly in the U.S.

We make three main contributions to understanding adjustment frictions. First, we document that earnings adjustment frictions exist in the U.S., by showing that in some contexts individuals do not adjust immediately to changes in AET. We focus particularly on cases in which a kink in the effective tax schedule disappears, either because individuals reach an age at which they are no longer subject to the ET, or because legislative changes remove the AET for some groups.<sup>5</sup> We focus on the disappearance of kinks because in the absence of adjustment frictions, removal of a convex kink in the effective tax schedule should immediately lead to no bunching at the earnings level associated with the former kink; thus, any observed delay in reaching zero bunching should reflect adjustment frictions. We observe clear evidence of delays in some contexts, consistent with adjustment frictions. Nonetheless, across several contexts—including both anticipated and unanticipated changes in policy—the vast majority of individuals' adjustment occurs within at most three years. Adjustment appears even faster in certain contexts.

Second, we assess the mechanisms that underlie the patterns of adjustment we observe, in order to build a model consistent with these descriptive patterns. We assess the extent to which employers play a role in coordinating individual responses to the AET by offering jobs with earnings at the AET exempt amount.<sup>6</sup> In our main period of study, we find little evidence that individuals not subject to the AET—those too young to claim benefits—bunch at the kink, suggesting that the primary responses to the AET are mediated by employees' choices. We also find evidence that the individuals who respond to the removal of the AET are primarily those locating at the kink prior to its removal, suggesting that these individuals are particularly

---

<sup>4</sup>The lack of bunching at other kinks is consistent with the existence of adjustment costs, although this finding could also be explained by other factors such as a low elasticity of earnings with respect to the net-of-tax rate. As we discuss in greater detail in Section 6, Chetty *et al.* (2012) do find evidence of more diffuse earnings responses to the Earned Income Tax Credit among the non-self-employed.

<sup>5</sup>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," while recognizing that the AET reduces Social Security benefits and is not administered through the tax system. The "effective" marginal tax rate is affected by the AET BRR, among other factors.

<sup>6</sup>Due to interactions between adjustment costs for workers and hours constraints set by firms, some individuals may bunch at a kink even though they are not directly subject to the policy that creates the kink. Chetty *et al.* (2011) document that employers play such a role in Denmark.

responsive. Others subject to the AET appear to be unresponsive, suggesting heterogeneity in adjustment costs or elasticities in the population.

Third, we specify a model of earnings adjustment consistent with the descriptive evidence that allows us to estimate a fixed adjustment cost and the elasticity of earnings with respect to the effective net-of-tax rate. Recent work demonstrating the importance adjustment costs has raised the question of how to estimate both the elasticity and adjustment cost simultaneously. We develop tractable methods that allow estimation of elasticities and adjustment costs with kinked budget sets. This complements Kleven and Waseem (2013), who develop a method to estimate related parameters in the presence of a notch in the budget set. Our method relies on the fact that the amount of bunching at a kink increases with the elasticity but decreases with the adjustment cost. This prevents estimation of both parameters using a single cross-section—since a small amount of bunching, for example, could be consistent with either a low elasticity or a high adjustment cost—but with two or more cross-sections of individuals facing different tax rates in the region of the kink, we can specify two or more equations and find the values of two variables (the elasticity and the adjustment cost).<sup>7</sup> The model of Saez (2010) describes how bunching should vary between two different kinks in a frictionless setting, and the extent to which observed bunching deviates from this pattern is attributed to the adjustment cost. Intuitively, inertia due to an adjustment cost leads to an excess amount of bunching after a kink in the budget set becomes less sharply bent (or disappears altogether). Our primary estimation method uses the degree of such inertia (in combination with the initial amount of bunching at the kink) in estimating the size of the adjustment cost (and elasticity).<sup>8</sup>

We apply our method to data spanning the decrease in the AET benefit reduction rate from 50 percent to 33 percent in 1990 for those aged 66 to 69, as well as two settings in which the AET no longer applies for certain groups (at age 70 in the 1990-1999 period, and for ages 66-69 beginning in the year 2000). In a baseline specification, we estimate that the adjustment cost is \$152.08 (in 2010 dollars) and that the earnings elasticity with respect to the net-of-tax share is 0.23.<sup>9</sup> Other empirical strategies show results in the same range. By contrast, when we

---

<sup>7</sup>Under certain approximations that we later specify, this can yield a system of linear equations that can be solved in closed form. However, this intuition also helps in understanding the forces that drive estimation in a more general framework that we use as our primary estimation strategy.

<sup>8</sup>As we describe in detail later, this intuition applies to our primary empirical approach, the Sharp Change approach.

<sup>9</sup>This specification examines data on 67-68 year-olds in 1989 and 1990 (before and after the fall in the BRR

constrain the adjustment cost to be zero in 1990, we estimate a statistically significantly higher earnings elasticity of 0.39 in the baseline specification (69 percent larger than the unconstrained estimate).<sup>10</sup> These estimates suggest that while adjustment costs are modest in our setting, they have the potential to change elasticity estimates substantially, thus demonstrating the potential importance of incorporating adjustment costs when estimating elasticities. However, our estimates are specific to our setting, and adjustment costs and elasticities may be substantially different (larger or smaller) in other contexts.

In the course of investigating these issues relating to frictions and earnings adjustment, we build on previous literature on the AET to provide new evidence that enriches our understanding of how the AET affects earnings. First, we systematically investigate each of the major AET policy changes since 1961. Second, we use SSA administrative data, building on several previous studies that use survey data. Third, our study is the first to estimate bunching in the context of the AET through a method similar to Saez (2010). Fourth, we present evidence on individuals' earnings reaction to changes in the Delayed Retirement Credit. Fifth, we investigate whether mortality expectations help drive individuals' earnings responses to the AET. Sixth, we investigate whether individuals change earnings in response to the AET by changing jobs or by changing earnings levels within a job, as well as whether employers coordinate employees on the AET exempt amount. Finally, we investigate how the entire distribution of earnings shifts when the AET is introduced, and when it no longer applies.

The remainder of the paper is structured as follows. Section 2 describes the policies we examine. Section 3 presents a framework for analyzing the behavioral response to these policies and describes our empirical strategy for quantifying bunching. Section 4 describes the administrative data we use. Section 5 presents empirical evidence on the earnings response to the AET, particularly focusing on the context in which the AET is removed. Section 6 explores mechanisms underlying the behavioral responses, including heterogeneity in the response across the earnings distribution and the possible role of employers in mediating responses. Section 7 specifies a tractable model of earnings adjustment and estimates the fixed adjustment cost and

---

for this group from 50 percent to 33 percent in 1990). Thus, our estimated adjustment cost represents the cost of adjusting earnings in the first year after the policy change.

<sup>10</sup>In our context, it makes sense that the estimated elasticity is higher when we do not allow for adjustment costs than when we do, as adjustment costs keep individuals bunching at the kink even though tax rates have fallen.

elasticity simultaneously. Section 8 concludes with discussion and avenues for future work.

## 2 Policy Environment

Figure 1 shows key features of the AET rules over the period we study, 1961 to 2006. The AET became less stringent over this period. The dashed line and right vertical axis show the benefit reduction rate. From 1961 to 1989, every dollar of earnings above the exempt amount reduced OASI benefits by 50 cents.<sup>11</sup> In 1990 and after, the benefit reduction rate fell to 33.33 percent for beneficiaries above the Normal Retirement Age (NRA).<sup>12</sup> The figure also shows that the AET applied to a narrower set of ages over time. In 1961, the AET applied to ages 62-71; starting in 1983, the AET was eliminated for 70-71 year-olds; and starting in 2000, the AET was also eliminated for those above NRA. The solid line and left vertical axis show the real exempt amount. Between 1961 and 1971, the exempt amount rose with price inflation. Beginning in 1972, the exempt amount typically rose faster than inflation. Starting in 1978, the AET had different rules for beneficiaries younger than NRA and those of at least NRA but younger than the maximum age subject to the AET. Subsequently, the exempt amount rose much faster on average for beneficiaries NRA and older than for younger beneficiaries.<sup>13</sup>

We later model the AET as creating a positive implicit marginal tax rate for some individuals, consistent with the empirical finding that some individuals bunch at AET kinks, the theoretical presumptions we describe below, and previous literature. In the empirical section, we explore evidence relating to certain mechanisms that explain this response.<sup>14</sup>

---

<sup>11</sup>In addition to this threshold, until 1972 there was a second, higher earnings threshold over which the benefit reduction rate was 100 percent (Social Security Annual Statistical Supplement 2012). The second threshold is well above the first threshold, ranging from 25 percent to 80 percent higher depending on the year.

<sup>12</sup>The NRA, the age at which workers can claim their full OASI benefits, is 65 for those born 1937 and before, rises by two months a year for cohorts between 1938 and 1943, is constant at age 66 for cohorts between 1943 and 1954, and rises by two months a year until reaching age 67 for those born in 1960 and later.

<sup>13</sup>The exempt amount has not been a "focal" earnings level—such as \$1,000, \$5000, or \$10,000—that could lead to bunching at the exempt amount even in the absence of AET. Indeed, in our main period of study we find no evidence of bunching at the exempt amount among those younger than the ages to which the AET applies.

<sup>14</sup>In this paper, we focus on the marginal incentives created by the AET and intensive margin responses, following previous literature based on the technique of Saez (2010) in focusing on intensive margin but not extensive margin responses to kinks. Other important decisions could include the choice of whether to earn a positive amount, or the decision to claim OASI. We abstract from the claiming decision by examining a sample of OASI claimants; however, it is worth noting that there is no clear relationship between the effect of the AET on the claiming decision and the magnitude or nature of the bunching responses we document among claimants. We examine the extensive margin response in a companion paper (Gelber, Jones, and Sacks 2013). (Cogan 1981 is a classic reference on fixed costs of adjustment in the extensive margin choice.) Empirically, we do find

When current OASI benefits are lost to the AET, future scheduled benefits are increased in some circumstances. This is sometimes called "benefit enhancement." Appendix A describes benefit enhancement in detail. Prior to 1972, the AET caused a pure loss in benefits for those NRA and above, as there was no benefit enhancement for these individuals. For beneficiaries subject to the AET aged NRA and above, a one percent DRC was introduced in 1972, meaning that each year of benefits foregone led to a one percent increase in future yearly benefits. The DRC was raised to three percent in 1982 and gradually rose to eight percent from 1990-2008 (though the AET was eliminated in 2000 for those above NRA). A credit between seven and eight percent is meant to be actuarially fair on average, meaning that an individual with no liquidity constraints and average life expectancy should be indifferent between either claiming benefits now or delaying claiming and receiving higher benefits once she begins to collect OASI (as shown, for example, by Diamond and Gruber 1999 with respect to the actuarial adjustment).

Importantly, as we describe in Appendix A, future benefits are only raised due to the DRC when annual earnings are sufficiently high that the individual loses an entire month's worth of OASI benefits due to the reductions associated with the AET (Friedberg 1998; Social Security Administration 2012). In particular, an entire month's benefits are lost once the individual earns  $z^* + (MB/\tau)$  or higher, where  $z^*$  is the exempt amount,  $MB$  is the monthly benefit, and  $\tau$  is the AET benefit reduction rate. For example, with a typical monthly benefit of \$1,000 and a benefit reduction rate of 33.33 percent, one month's benefit enhancement occurs when the individual's annual earnings are \$3,000 ( $=\$1000/0.3333$ ) above the exempt amount. Benefit enhancement corresponding to one more month of reduced earnings occurs once annual earnings reaches \$6,000 above the exempt amount, and so forth. Thus, benefit enhancement is only relevant to an individual considering earning substantially in excess of the exempt amount and is therefore not relevant to marginal earnings decisions at the exempt amount. As a result, at or just above the AET threshold, earning an extra dollar does not affect subsequent OASI benefits. Indeed, this theoretical presumption is consistent with suggestive evidence we later describe that indicates little systematic bunching reaction to changes in the DRC.

Thus, the AET could affect the earnings decisions of those NRA and above for a number of reasons. As we have discussed, for those for whom benefit enhancement is effectively irrelevant (because they are only considering earning sufficiently near to the AET that they would not

---

evidence of an intensive margin response.

receive benefit enhancement through increasing earnings), the marginal incentives they face are not affected by benefit enhancement. For those for whom benefit enhancement is relevant (because they are considering earning in a region well above the AET exempt amount, thus triggering benefit enhancement), the AET could also affect decisions, for several reasons. Those whose expected lifespan is shorter than average should expect to collect OASI benefits for less long than average, implying that the AET is more financially punitive (though we find no evidence consistent with this hypothesis in our empirical results). Liquidity-constrained individuals or those who discount faster than average could also reduce work in response to the AET. Prior to 1972, the AET represented a pure loss in benefits for those NRA and above, and we examine this period in our empirics. Finally, many individuals may also not understand many features of the AET or other aspects of OASI (Liebman and Luttmer 2011); we explore potential misperceptions of the AET in a separate paper (Gelber, Jones, and Sacks 2013).

For beneficiaries under NRA, the actuarial adjustment raises future benefits for each month of *withheld* benefits, which are triggered by an individual earning over the AET threshold (by 0.55 percent per month withheld for the first three years of AET assessment).<sup>15</sup> Importantly, the individual receives the increase in future benefits when the individual earns any amount over the AET threshold.<sup>16</sup> This creates a notch in the budget set at the AET threshold—as opposed to a simple kink, whose properties we explore in our theory sections. Our discussion of the effects of kinks therefore does not directly apply to estimates in the pre-NRA years. Thus, in our estimates of elasticities and adjustment costs, we limit the sample to ages NRA and above, for which the budget set (in the region of the exempt amount) is a kink rather than a notch.

### 3 Initial Bunching Framework

As a preliminary step, we begin with a model with no frictions. This model is well-known and described in detail elsewhere, but we briefly describe it here and in more detail in Appendix

---

<sup>15</sup>The difference here is subtle. Formally, the number of months worth of benefit enhancement enjoyed by OASI recipients will be  $\text{floor}(\tau \cdot (z - z^*)/MB)$  for those NRA and above, and  $\text{ceiling}(\tau \cdot (z - z^*)/MB)$  for those below NRA. See Appendix (A) for more details.

<sup>16</sup>Social Security Administration (2012), Section 728.2; Gruber and Orszag (2003).



D.<sup>17</sup> After we have presented our empirical results, we specify a model with frictions that is consistent with the descriptive patterns we document.

Appendix Figure E.1 shows the budget constraint and incentives created by the AET for those NRA and above in the frictionless case. Start with a linear tax (Panel A) at a rate of  $\tau$ . Now, suppose the AET is introduced (on top of pre-existing taxes), so that the marginal net-of-tax rate decreases to  $1 - \tau - d\tau$  for earnings above a threshold  $z^*$  (Panel B). For small  $d\tau$ , individuals earning in the neighborhood above  $z^*$  will reduce their earnings. If ability is smoothly distributed, a range of individuals initially locating between  $z^*$  and  $z^* + \Delta z$  (as depicted in the density in Panel C) will instead locate exactly at  $z^*$ , due to the discontinuous jump in the marginal net-of-tax rate at  $z^*$ . Panel D shows that once the AET is introduced, these individuals locate in the neighborhood of  $z^*$ .

To measure the amount of bunching, we use a technique similar to Chetty, Friedman, Olsen, and Pistaferri (2011) and Kleven and Waseem (2013), which we illustrate in Appendix Figure E.2. The x-axis measures before-tax income,  $z$ , while the y-axis measures the density of earnings. In Panel A, we show that the ex-post density of earnings in the presence of a kink is comprised of a number of groups. Those in the region labeled X in the figure ("bunchers") have optimal earnings above  $z^*$  under the lower rate of  $\tau$  and at  $z^*$  under the higher rate of  $\tau + d\tau$ . Those in the region labeled Y in the figure consist of individuals whose optimal earnings are below  $z^*$  under a lower marginal tax rate of  $\tau$ , and individuals whose optimal earnings are above  $z^*$  under the higher marginal tax rate of  $\tau + d\tau$ . Panel B shows that to estimate the size of region X, we must estimate the *ex post* density and subtract the mass associated with Group Y.

In our estimation strategy, we divide the data into \$800 bins and estimate a seven-degree polynomial through the densities associated with the bins. In estimating this polynomial, we control for dummies for being in the seven bins nearest to the kink,<sup>18</sup> to capture the bunching near the kink that we wish to ignore when we estimate the counterfactual polynomial density.

---

<sup>17</sup>We review this framework in the Appendix since its results are thoroughly covered in previous literature. Saez (2010) describes this model in greater detail. This work follows earlier work on estimation of labor supply responses on nonlinear budget sets, including Burtless and Hausman (1978) and Hausman (1981). Moffitt (1986, 1990) surveys these methods and their applications.

<sup>18</sup>This implies that our estimate of excess bunching is driven by individuals locating within \$2800 of the kink (as the central bin runs from \$400 under the kink to \$400 above the kink). We discuss this issue further in the Appendix. As we show in the Appendix, we have also experimented with other bandwidths, which yield similar results.

Our estimate of bunching,  $B$ , is the difference between the mass in these seven bins and the area under the polynomial in this \$5600-wide region. We estimate confidence intervals through a bootstrap procedure that we describe further in Appendix D.8 (and the results are similar under the delta method). We report our bunching amount,  $B$ , normalized by the share of individuals in the neighborhood  $[z^* - \delta, z^* + \delta]$  who belong to Group Y (which we approximate as the area under our polynomial over this range).<sup>19</sup>

Some apparent limitations of our approach are worth discussion. First, following previous literature on earnings responses to kinks, we do not take into account other choices that could affect earnings in the long run, such as human capital accumulation. However, human capital accumulation is likely to be less important for the older workers we study than it is for the population as a whole. Second, other programs—such as Medicaid, Supplemental Security Income, Disability Insurance, or taxes such as unemployment insurance payroll taxes—create earnings incentives near the bottom of the earnings distribution. While we acknowledge that other incentives represent a concern in principle—applicable to most of the literature on bunching at kinks—we also note that the kinks created by these programs are typically inapplicable or safely far away from the AET convex kink.<sup>20</sup> The results show very clear evidence of bunching at the AET kink and no visible, systematic evidence of bunching in other regions close to the AET kink. Third, we follow the previous work and do not distinguish among the potential reasons for a response to the AET. Following previous literature, our bunching framework presumes that consistent with the empirical evidence documenting clear responses to the incentives created by the AET, certain individuals treat the AET as creating *some* effective marginal tax rate above the exempt amount.

Finally, the results are specific to the AET and may not generalize outside of this context.

---

<sup>19</sup>While we show this excess bunching at the kink as arising in a frictionless model here, this technique is also suited to measuring the excess bunching at the kink arising in a model with frictions (as the key in either setting is that there is excess bunching at the kink, which this technique can measure in either case).

<sup>20</sup>We have found that many other incentives, including income tax rates, are smooth on average around the AET convex kink. The AET also potentially creates other distortions that we discuss in Appendix A, including: a slight notch for those NRA age and above every time an entire month’s worth of benefits are lost due to the Delayed Retirement Credit; an additional non-convex kink in the budget constraint at the point at which OASI benefits are fully phased out; and a notch for those below the NRA for every month of withheld benefits that triggers the actuarially adjustment described above. However, in the case of those NRA and above that we focus on, these incentives are not likely to be relevant for potential bunchers and do not appear to be empirically relevant (as we discuss elsewhere). For these reasons, we abstract from these additional features but discuss such incentives further when we present our empirical evidence and estimation exercise.

We estimate the speed of adjustment among those initially bunching at a kink, a group that our empirical results suggest is more responsive to the AET than other groups. We therefore believe it is all the more interesting that we still find evidence of modest adjustment frictions among this group whose initial bunching indicates a substantial degree of flexibility (enough to locate at the kink initially). Furthermore, our estimation procedure relies on estimating bunching at more than one kink, and therefore it has the potential to incorporate information on the responses of individuals across a fairly wide range of the income distribution (across multiple kinks).

## 4 Data

We primarily rely on the restricted-access Social Security Administration Master Earnings File (MEF) and Master Beneficiary Record (MBR), described more fully in the Appendix. The data contain a complete longitudinal earnings history with yearly information on earnings since 1951; the type and amount of yearly Social Security benefits an individual receives; year of birth; the year (if any) that claiming began; and sex (among other variables). Separate information is available on self-employment earnings and non-self-employment earnings. Prior to 1978, the data measure annual Federal Insurance Contribution Act (FICA) earnings. Starting in 1978, the measure of earnings in the MEF reflects total wage compensation, as reported on Internal Revenue Service (IRS) forms (but continues to measure capped earnings). Our dataset is a one percent random sample of all Social Security numbers in the MEF, keeping all available years of data for each individual sampled.

Several features of the dataset are worth discussion. First, these administrative data allow large sample sizes and are subject to little measurement error. Second, earnings (as measured in the dataset) are the base for FICA taxes and are not subject to manipulation through tax deductions, credits, or exemptions. Third, because earnings are taken from the W-2 form, they are subject to third-party reporting for the non-self-employed; third-party reporting has been found in the literature to greatly reduce evasion (Kleven et al. 2011). This limits the degree to which observed bunching among the non-self-employed—to whom we limit our sample—could reflect reporting issues. Fourth, the data do not contain information on hours worked or amenities at individuals' jobs.

Table 1 shows summary statistics for the sample of individuals aged 18-75, and for the sample that we typically focus on, all those aged 62-69 who claimed by age 65.<sup>21</sup> The larger (smaller) sample has 15,840,014 (1,684,122) observations on 651,938 (577,973) individuals. 45 percent (41 percent) of the observations have positive earnings. Mean earnings in the sample (conditional on having positive earnings) is \$35,352 (\$28,843). Note that median earnings among our group of interest (\$17,543) is not far from the AET exempt amount; the population our study examines is in a range with a thick density of earnings that is not far from the median. 57 percent of the sample is male. In the main sample we use (62-69 year-olds), 17.8 percent of observations (including both the self-employed and non-self-employed) have positive self-employment earnings.<sup>22</sup>

The second dataset we use is the Longitudinal Employer Household Dynamics (LEHD) dataset of the U.S. Census (McKinney and Vilhuber, 2008; Abowd et al., 2009), described further in the Appendix. The data are based on unemployment insurance earnings records and longitudinally follow workers' earnings over time. The data have information on around nine-tenths of workers in covered states and their employers, though we are only able to use data on a 20 percent random subsample. We use these data primarily in order to link employees to employers, as the SSA data that we have access to have no information about individuals' employers. We secondarily use these data because the sample size we are able to obtain in the LEHD is much larger than the (large) sample size we obtain in the SSA data.<sup>23</sup>

---

<sup>21</sup>We focus on this fixed sample to hold the sample constant across ages or years; as we describe later, we also investigate a number of other samples as robustness checks, including a sample in which we examine only those who have claimed by the time we observe them.

<sup>22</sup>In calculating this self-employment percentage, we use the period 1978-2005, since the self-employment earnings data are more reliable during this period. Over the full 1961-2005 period, 17.6 percent of observations in the 62-69 year-old age group have positive self-employment earnings. In the full age sample, 9.8 percent of observations have positive self-employment earnings in the 1978-2005 period.

<sup>23</sup>The LEHD lacks information on whether a given individual is claiming Social Security. Nonetheless, the ultimate importance of this shortcoming is limited. The primary individuals that we consider in the LEHD are those over age 65. In our SSA data, 97 percent of people claim by age 69, so it is a safe assumption that the great majority of the individuals observed in the LEHD data of the ages we are interested in (primarily ages 69 and 70) have claimed Social Security. The magnitude of the bunching we observe is likely to be slightly understated relative to the magnitude we would measure in the population of Social Security claimants, as the results include non-claimants in the sample. However, our primary interest concerns the patterns of responses to the AET across ages and over time, which prove to be visually and statistically clear in the LEHD.

## 5 Earnings Response to Policy

### 5.1 Adjustment Frictions: Descriptive Evidence from Policy Variation Across Ages

We first examine the pattern of excess bunching across ages, in order to determine how quickly individuals respond to changes in policy across ages and whether they face delays (consistent with the existence of adjustment frictions). Empirical work often estimates only short-run responses to changes in policy (see Saez *et al.*, 2012, for a review of literature on earnings responses to taxation). If individuals are able to respond more (less) in the long run than in the short run, then this large body of work would under-estimate (over-estimate) long-run responses. Moreover, most empirical specifications have related an individual's tax rate in a given year to the individual's earnings in that year. In order to choose the appropriate time horizon over which to study behavioral responses to policy, it is necessary to establish how long it takes to respond to policy changes.

Subsequent to 1982, the AET applies to ages 62-69. The policy changes at ages 62 and 70—the imposition and removal of the AET—are "anticipated," by which we refer to changes that would be anticipated by those who have knowledge of the relevant policies. We begin by examining the period 1990-1999.<sup>24</sup> Figure 2 plots earnings histograms for each age from 59 to 73. Earnings are measured along the x-axis, relative to the exempt amount, which is shown by a vertical line.<sup>25</sup>

Figure 2 Panel A shows clear visual evidence of substantial excess bunching from ages 62-71.<sup>26</sup> Figure 2 Panel B plots the point estimates and 95 percent confidence intervals for bunching at each age. Bunching is statistically significantly different from zero at each age in the 62-71 range ( $p < 0.01$  at all ages). We find no robust evidence of adjustment in anticipation

---

<sup>24</sup>The 1990-1999 period is relatively recent and represents a substantial period over which the AET parameters were stable (aside from the change in the Delayed Retirement Credit over this period).

<sup>25</sup>For ages younger than 62, we define the (placebo) kink in a given year as the kink that applies to pre-NRA individuals in that year. For individuals 70 and above, we define the (placebo) kink in a given year as the kink that applies to post-NRA individuals in that year.

<sup>26</sup>As discussed above, in this period individuals aged 62 to 64 faced a notch in the budget set (due to the actuarial adjustment of benefits) at the exempt amount; thus, the incentives they faced were different than those for individuals aged 65 to 69. However, the histograms show no evidence of a spike in earnings just above the kink (as one might predict on the basis of this notch); indeed, as Gelber, Jones and Sacks (2013) discuss, individuals tend to locate just below the kink much more frequently than just above the kink.

of future changes in policy, as those younger than 62 do not bunch.<sup>27</sup> We do find evidence that unbunching takes more than one year, however, as those at ages 70 and 71 show evidence of bunching.<sup>28</sup> We find it striking that even among the group bunching prior to age 70—that (the data reveal) are able to adjust earnings to the kink—we still find evidence of modest adjustment frictions.

Figure 2 Panel B shows that excess bunching is substantially lower at age 65 than surrounding ages. One explanation is that the location of the kink changes substantially from age 64 to age 65; as Figure 1 shows, during this period the exempt amount is much higher for individuals NRA and above than for individuals below NRA. Individuals may have difficulty adjusting their earnings to the new, higher kink within one year. This provides suggestive evidence that individuals also face delays in adjusting in this context.<sup>29</sup>

Similar patterns of adjustment occur when looking at the periods 1972-1982, 1983-1989 and 2000-2006 (Appendix Figures E.4 to E.6). We find evidence of adjustment delays, as individuals continue to bunch at the kink at ages older than the highest age to which the AET applies (except in the 2000-2006 period).<sup>30</sup> However, in all cases, adjustment appears to take at most three years.

---

<sup>27</sup>If the cost of adjustment in each year rose with the size of adjustment and this relationship were convex, we would expect anticipatory adjustment.

<sup>28</sup>The AET is assessed on earnings until the month in which the individual turns age 70. For simplicity, in our baseline sample we measure age as calendar year minus year of birth. Thus, if an individual turns age 70 later in the year—in the extreme case, on December 31—she will have had an incentive to bunch at the kink during nearly the entire year when she is classified as age 70 in our data. As a result, her yearly earnings may appear to be located at or near the kink even though she is bunching at the kink applicable to 69-year-olds through almost all of the calendar year over which her earnings are observed. However, the figure shows that significant bunching occurs at age 71, which cannot be due to this coarse measure of birth dates. Thus, the results do show a delay in complete adjustment. We have also found substantial and significant ( $p < 0.01$ ) bunching at age 70 among those born in January, who no longer face the AET immediately in January of the year they turn 70 and therefore should not show excess bunching at this age in the absence of adjustment frictions.

<sup>29</sup>In our context, the only "appearance" of a new kink that we observe is the appearance of a kink at age 62. The amount of time since the appearance of the kink at age 62 is correlated with age, and elasticities and adjustment costs could also be correlated with age—thus confounding analysis of the time necessary to adjust. While recognizing these caveats, it is worth noting that the amount of bunching slowly rises from age 62 to 63, which could suggest gradual adjustment. In principle, this could also relate to the fact that these graphs show the sample of those who have claimed by age 65, and the probability of claiming at a given age (conditional on claiming by age 65) rises from age 62 to 63. To address this issue, in Appendix Figure E.3 we show the results when the sample at a given age consists of those who have claimed by that age, which still shows a substantial increase in excess bunching from age 62 to 63 (though less steep than in Figure 2).

<sup>30</sup>This is also true in the sample that only contains those born in January.

## 5.2 Adjustment Frictions: Descriptive Evidence from Changes in Policy Across Time

We next examine adjustment to a legislated change in AET policy. As shown in Figure 1, the AET was eliminated for those over 65 in 2000. This policy change was unanticipated prior to the year 2000, in the sense that the legislation enacting the policy change was passed in April 2000 and applied to workers' earnings in October 2000, and discussions prior to 2000 did not widely anticipate these changes.<sup>31</sup>

Figure 3 shows the results for those aged 66-69. Bunching in the earnings distribution is easily visible in the years prior to 2000. In 2000, however, there is immediately little bunching visible, and this lack of bunching persists after 2000.<sup>32</sup> A very small bump in the earnings histogram is visible near the kink, but this proves to be insignificant in these data. Figure 4 also shows the amount of excess bunching estimated by year, along with 95 percent confidence intervals.<sup>33</sup> The amount of bunching is significantly greater than zero in all years prior to 2000, and estimates for 2000 and subsequent years show no significant bunching.

Appendix Figure E.8 shows bunching in 1999, 2000, and 2001 in the LEHD. A spike at the kink is easily visible in 1999, and a small amount of bunching is visible in 2000 as the two bins on either side of the exempt amount are raised relative to the rest of the density (as in the small bump in the earnings histogram in 2000 in the MEF). In fact, excess normalized bunching proves to be significantly different from zero in 2000 in the LEHD ( $p < 0.01$ ). By 2001, there is no clear visual evidence of bunching at the kink, and normalized excess bunching

---

<sup>31</sup>The AET was also eliminated in 1983 for individuals aged 70 and 71. However, our results across ages show that individuals bunch at ages 70 and 71 in the 1990-99 period, so that persistent bunching at these ages in the 1983-89 cannot cleanly be interpreted as delayed adjustment to the 1983 change, as opposed to delayed reaction to the disappearance of the kink at age 70.

<sup>32</sup>For comparison, Appendix Figure E.7 shows that bunching stayed relatively constant for the 62-64 year-old group that experienced no policy change in 2000. While this group faces a notch at the exempt amount rather than a kink (as explained above), the relative comparison is instructive and suggests that the fall in bunching in the 66-69 year-old group in 2000 and subsequent years was due to the removal of the AET for this group in 2000.

<sup>33</sup>We have estimated this amount of excess bunching using three ways of calculating "placebo" kinks in 2000 and after: 1) by adjusting the exempt amount in 1999 using the CPI-U; 2) by adjusting the exempt amount in 1999 using the Employment Cost Index; 3) by using the exempt amount applicable to individuals in the year of attaining NRA in a given year (which is the same as the exempt amount that had been scheduled prior to the 2000 legislation to apply in each year to those NRA and above). Figures 3 and 4 show the first of these methods, but all of these methods show no significant bunching in these years (which is unsurprising given the lack of bunching visible in the histograms in 2000 and after).

is insignificantly different from zero in the LEHD.<sup>34</sup>

In 2000, we find weaker evidence of a delay in adjustment—it is only statistically significant in the LEHD, and only among a small number of individuals. Moreover, bunching in the LEHD in 2000 is not necessarily immediately apparent in the earnings density and is therefore substantially less convincing than the residual bunching in the SSA data at ages 70-71. Thus, we do not wish to rely on the finding of residual bunching in the year 2000; instead, we consider this evidence to be merely suggestive of a small amount of residual bunching (but not definitive).

However, it is clear that both when changes are anticipated (*i.e.* the changes in policy across age) and unanticipated (*i.e.* the policy change in 2000), adjustment occurs fairly rapidly, with the vast majority occurring within a maximum of three years. It is interesting to note that adjustment appears to be faster in the case in which the change is unanticipated than in the anticipated case. While this may be surprising, many other differences between the two sets of changes—including differences in the degree to which the changes are publicized, the ages affected, the calendar year, and the distribution of individuals' earnings relative to their desired earnings in the absence of adjustment costs—could be responsible for the discrepancy in the speed of adjustment. As we observe only a small number of changes in AET policy and confront several candidate explanations for heterogeneity in the speed of adjustment, we do not explicitly try to distinguish among these explanations. Because the change was passed in April 2000 and implemented in October 2000—both after most salaried workers would have learned about their pay—the fairly fast reaction suggests that bunching is driven by workers with substantial flexibility in their earnings.<sup>35</sup>

---

<sup>34</sup>Since the sample size is much larger in the LEHD than in the MEF, it makes sense that we could estimate a small but statistically significant amount of bunching in the LEHD but not in the MEF in 2000. We also investigated the speed of adjustment from quarterly earnings data in the LEHD. These data show a finer level of temporal disaggregation than the yearly SSA data. However, we do not primarily rely on these data because the AET is assessed yearly, and thus individuals can appear to bunch at the quarterly kink—defined as one-quarter of the earnings level associated with the kink in each year—even though their yearly earnings does not put them at the kink, or vice versa, muddying our interpretation of bunching in these data. These data show a small but significant ( $p < 0.01$ ) amount of bunching in each quarter of 2000 and in the first two quarters of 2001 but no significant bunching in subsequent quarters. Thus, these data again suggest that adjustment is not immediate but that adjustment is still fast in the sense that the majority occurs within a small number of quarters of the policy change.

<sup>35</sup>Due to changes that raised the scheduled exempt amount beginning in 1996, the ET had been scheduled to increase from \$15,500 in 1999 to \$30,000 by 2002. In principle, this could have affected the amount of bunching in 2000, even absent the elimination of the AET in this year. Nonetheless, bunching is unlikely to have been zero in the absence of the AET elimination, as the quarterly LEHD data show substantial evidence of bunching in quarterly earnings data prior to the fourth quarter of 2000, when the AET was eliminated.



### 5.3 Other Evidence on Changes in Bunching

Figure 4 shows that there is no sharp change in the amount of bunching around the increases in the Delayed Retirement Credit in 1972 or 1982. We consider this suggestive—but not definitive—evidence of little discernable reaction to policy changes in benefit enhancement (particularly in light of our other results suggesting fast adjustment). A general downward trend in the amount of excess bunching is discernable in the 1990s—with the notable exception of a number of years, including 1995—which is coincident in the rise in the DRC through this period. However, even if this does represent a trend, we cannot conclusively attribute this trend to the influence of the DRC, as it could be due to other factors that changed over this period.<sup>36</sup> A final piece of evidence concerns adjustment to the decrease in the AET marginal tax rate from 50 percent to 33.33 percent in 1990. As shown in Figure 4, the amount of bunching fell from the years before 1990 to the years after 1990. We discuss this evidence in detail later when we use this policy change to estimate elasticities and adjustment costs.

We also conduct a variety of robustness tests. Appendix Figure E.9 uses a bandwidth of \$500 instead of \$800, which changes our estimates little (as have other bandwidths we have chosen). In Appendix Figure E.10, we vary the degree of the polynomial we use between 6 and 8, which produces similar results; other sufficiently rich variations of the order of polynomial we have tried have also shown similar results. In Appendix Figure E.11, we vary the region near the kink we exclude when estimating the amount of excess bunching (from \$2,000 to \$3,000 to \$4,000) and again estimate similar results. Limiting the sample to those who have substantial benefits (such as those with \$1,000 or higher in benefits)—so that they are safely far from the concave kink in the budget set created when the AET reduces OASI benefits to zero—also yields very similar results.<sup>37</sup>

---

<sup>36</sup>For example, the AET threshold amount rose much faster in the 1996-1999 period than in the previous period. It is possible that this helps to explain the decrease in the amount of bunching observed in these years, as individuals may find it difficult to adjust earnings to a rapidly-increasing kink (or elasticities and adjustment costs may be different in this segment of the earnings distribution). Meanwhile, as we discuss later, the fall in excess bunching after 1990 may relate to adjustment to the reduction in the benefit reduction rate in 1990.

<sup>37</sup>Since 1978, the earnings test has been assessed on yearly earnings, implying that we analyze the appropriate time period, *i.e.* earnings in a calendar year. Prior to 1978, the earnings test was assessed on quarterly earnings. While there is likely some error in measuring the amount of bunching pre-1978, we believe that this is not a major issue: the patterns of bunching in the pre-1978 period that we discuss below are visually clear and appear unlikely to be changed in a qualitative sense by an examination of quarterly data. Moreover, Figure 4 shows that the amount of excess bunching falls from 1977 to 1978 and subsequent years, rather than rising as we might expect if we hypothetically measured bunching more accurately starting in 1978.

In Appendix Figure E.12, we show that both men and for women bunch at the kink (though interestingly, men show more bunching than women). Previous work has demonstrated very different patterns of bunching among the self-employed and non-self-employed (Chetty *et al.* 2011) and has shown that bunching at the kink in response to the Earned Income Tax Credit is primarily driven by the self-employed (e.g. Chetty *et al.* 2012b). In Appendix Figure E.13 we show histograms for those with self-employment income in 1990-1999, who also show a spike in the earnings distribution near the kink.

## 6 Mechanisms

This section probes the mechanisms that underlie patterns of adjustment to the AET. Specifically, we examine which parts of the earnings distribution adjust to AET changes, whether adjustment relates to age at death, and whether employers or employees drive responses to the AET.

### 6.1 Who Adjusts?

We investigate who adjusts to the AET using the large sample sizes in the LEHD data, which allow us to estimate parameters precisely in relatively small population groups. Specifically, we examine how earnings change as the AET is removed from age 69 to age 70 during 1990-1999, when the AET applied to individuals aged 62-69. Figure 5 shows the mean percentage change in earnings (expressed as a fraction of initial earnings)<sup>38</sup> from age 69 to age 70 (y-axis), against earnings at age 69 (x-axis). The graph shows a large spike at the kink: individuals locating near the kink at age 69 on average increase their earnings substantially from age 69 to age 70.<sup>39</sup>

This finding suggests that individuals locating near the kink at age 69 are different than other individuals at the same age. Indeed, the AET applies not only to individuals locating at the kink, but also to individuals initially locating above the kink (and below the concave kink). Thus, if individuals initially locating at the kink had the same elasticity and adjustment cost as others, we might have expected to see a large increase in earnings in a substantial range of

---

<sup>38</sup>Thus, for example, 0.15 on the y-axis of the graph corresponds to a 15 percent mean increase in earnings.

<sup>39</sup>The increase near the kink is significantly higher than that in adjacent bins ( $p < 0.01$ ). This spike in earnings growth in Figure 5 is interesting in part because it directly documents responses to policy along the intensive margin, which is often found to be very inelastic (e.g. Eissa and Liebman, 1996; Meyer and Rosenbaum, 2001).

earnings above the kink, as well. The fact that we do not observe this pattern is suggestive of heterogeneity in adjustment costs or elasticities.<sup>40</sup> For example, those initially locating at the kink may have low adjustment costs and react to the AET removal quickly, but those who do not bunch at the kink to begin with may have higher adjustment costs.<sup>41</sup>

In Appendix Figure E.14, we show that individuals at the kink tend to follow the kink from year to year. We graph the probability of being at the kink in year  $t + 1$ , as a function of earnings in year  $t$ . There are clear spikes at the kink for ages 62-63 and ages 65-68, showing that individuals at the kink in year  $t$  are disproportionately likely to be at the kink again the next year.<sup>42</sup> We interpret this as further suggestive evidence that certain individuals are particularly responsive to the incentives created by the kink, in the sense that they serially bunch at the kink.

To understand where in the earnings distribution bunchers come from and go to, we examine more closely how the distribution of earnings differs across adjacent ages that are substantially different in terms of AET incentives. Appendix Figure E.15 stacks the estimated distributions of earnings at ages 60, 61, and 62, as well as 69, 70, and 71. The distribution of earnings changes modestly from year to year due to factors unrelated to the AET, such as from ages 60 to 61 as shown in the Figure. However, the age-62 distribution shows a sharply different pattern than the age-60 or 61 distributions, with a sharp spike at the kink (particularly to the left of the kink), a higher density immediately to the right of the kink, and generally a lower density at earnings levels starting around \$5,000 greater than the kink until meeting the age-61 distribution well above this level—consistent with the hypothesis that those ultimately bunching in the vicinity of the kink tend to come from a broad region well to the right of the

---

<sup>40</sup>The income effect of the AET also rises with income, which would also lead the mean percentage earnings change to fall as income rises (under the assumption that leisure is a normal good). However, the income effect rises only gradually, whereas the mean percent earnings increase quickly falls just to the right of the kink level of earnings and remains relatively constant at this lower level as earnings rises—consistent with the hypothesis that those initially locating at the kink are more responsive to the AET. In fact, the data suggest that income effects (if any) are sufficiently small that they do not cause a noticeable systematic decrease in the mean percentage change in earnings as we move increasingly far to the right of the kink.

<sup>41</sup>The observed pattern is also consistent with such heterogeneity in elasticities or income effects.

<sup>42</sup>The probability of being near the kink in year  $t + 1$  is significantly higher ( $p < 0.01$ ) for those near the kink in year  $t$  than in adjacent bins of the year  $t$  earnings distribution. For those aged 58-60, who should not be affected by the incentives to bunch at the kink, no such spike occurs—demonstrating that the spike at the kink for ages subject to the AET is not simply an artefact of the natural evolution of the earnings distribution. We define the "placebo" kink for individual aged 58-60 as the kink affecting those aged 62-64.

kink.<sup>43</sup> Similarly, the age-69 earnings distribution shows a sharply higher earnings density than the age 70 distribution in the immediate region of the kink (particularly to its left) but shows a lower density than at age 70 at higher earnings levels, eventually reaching a similar earnings density starting around \$6,000 above the kink.<sup>44</sup> We return to this pattern of adjustment when discussing our model of fixed adjustment costs below.

Finally, it is possible that those with short expected lifespan should disproportionately bunch near the kink: the DRC should increase lifetime benefits more for claimants with longer life expectancy, which could lead the AET to be a larger effective tax on those with shorter lifespans (though as we note, the DRC only takes effect at earnings substantially above the exempt amount). In Appendix Figure E.16, we show graphs illustrating that life expectancy is smooth near the kink (not significantly different from adjacent bins), suggesting little evidence for such a mechanism.

Recent literature has documented responses to kinks not captured by bunching, including Chetty, Friedman, and Saez (2012) in the context of the EITC and Kline and Tartari (2013) in the context of the Connecticut Jobs First program. In the context of the AET, our evidence shows that responses to incentives appear to be concentrated among a group of bunchers at the kink, with little apparent response among others (though we cannot rule out that such changes occur in ways we do not capture, such as responses over a longer time frame).<sup>45</sup>

## 6.2 Employers and the AET

We use the LEHD data to investigate whether employers play a role in mediating responses to the AET. Chetty *et al.* (2011) argue that employers drive a significant share of the bunching at kink points observed in Denmark. In their context, some individuals bunch at kinks even though they are not directly subject to the policy that creates the kink. Chetty *et al.* conclude that these individuals bunch at the kink because employers create jobs that have those earnings levels. In other words, some individuals bunch at kinks because their employers present them a

---

<sup>43</sup>The age-61 distribution of earnings conditional on locating in the vicinity of the kink at 62, and the age-70 distribution of earnings conditional on locating in the vicinity of the kink at age 69, are also consistent with this.

<sup>44</sup>Some adjustment takes place after age 70; the evolution of the income distribution from 70 to 71 shows similar patterns.

<sup>45</sup>We further partially addressed the possibility of bunching over a different range by varying the bandwidth that we chose for estimating excess bunching.

limited equilibrium menu of earnings levels (including the kink earnings level), and they would face costs of adjusting earnings to a different level.

We explore this possibility by testing for bunching among workers who are too young to claim OASI benefits and are therefore unaffected by the AET. Above, we have presented evidence indicating that in 1990-1999, individuals at ages earlier than those subject to the AET show little evidence of bunching at the AET kink. Thus, during this period, the evidence is consistent with the hypothesis that responses to the AET are driven by employees' choices.<sup>46</sup>

We extend this analysis by estimating bunching over the entire age distribution in the pre-1972 period, when the DRC did not exist, as Appendix Figure E.17 shows. Appendix Figure E.17 shows a small amount of statistically significant excess bunching at some ages younger than those subject to the AET (though not at other ages), suggesting that some employers do coordinate employment responses in this way in the pre-1972 period—though this behavior is small in the aggregate.<sup>47</sup>

In Appendix Figure E.18, we graph the probability that individuals change at least one employer from age  $t$  to age  $t + 1$ , against earnings at age  $t$  (during 1990-1999 period). At ages when people face the AET, the probability that individuals change jobs across employers is sharply lower for individuals locating near the kink at age  $t$  than for individuals with other initial earnings levels.<sup>48</sup> The probability of changing employers is also sharply lower at the kink when individuals transition from being subject to the AET at age 69 to being no longer subject at age 70, and this is true when we limit the sample to those who increase their earnings from age 69 to age 70. Those initially locating at the kink evidently have sufficiently flexible pay arrangements that they can change their earnings from age 69 to age 70 while typically staying at the same employer.<sup>49</sup>

---

<sup>46</sup>It is possible that employers drive some of the bunching at ages older than those subject to the AET, but we might then expect some degree of employer earnings coordination on the AET exempt amount for ages younger than 62 (and older than 70, including ages 72 and above).

<sup>47</sup>Figure E.17 bins three adjacent years of ages (e.g. 18 to 20); doing so in the post-1971 period shows no statistically significant bunching in pre-62 age bins. While individuals could in principle be locating at the kink prior to age 62 in anticipation of facing the AET later, it seems unlikely that they would do so as early as their late 30s, when a small amount of statistically significant bunching first appears—around 25 years before they are first eligible for OASI.

<sup>48</sup>The probability of changing employers near the kink is significantly lower than that in adjacent bins ( $p < 0.01$ ).

<sup>49</sup>Those locating at the kink might be different from other individuals for reasons—such as different demographics—that lead them to switch across employers less frequently. It is worth emphasizing that we attempt only to document the descriptive pattern that they change employers less frequently. We have also

Overall, we interpret this evidence as suggesting that in the post-1972 period, employees drive the primary responses to the AET, and they do so through changing earnings at the same job (through work arrangements that allow them the flexibility to do so in some way). These findings are reflected in the estimation framework that follows.

## 7 Estimating Elasticities and Adjustment Costs

The results thus far suggest a role for adjustment frictions in individuals' earnings choices in some contexts. As a first step in incorporating frictions into an estimable model of earnings supply, we build on the Saez (2010) model (described briefly in the first sections of Appendix D), which uses normalized excess bunching to identify the elasticity of (taxable) earnings with respect to the net-of-tax rate.<sup>50</sup> We extend this model to allow for a cost of adjusting to tax changes.<sup>51</sup> We first develop the theory graphically to show how adjustment costs affect bunching. Next, we show that using data on bunching at multiple kinks associated with different jumps in the net-of-tax rate, we can jointly estimate elasticities and adjustment costs. As discussed in Chetty *et al.* (2009), these parameters—adjustment costs and the elasticity that characterizes individuals' behavior in the absence of adjustment costs—are jointly sufficient for welfare calculations in many applications.

Our model relies on features of the empirical results that we have documented in the previous two sections. We find evidence of adjustment frictions, which we model through a cost of adjusting earnings. The empirical results also suggest that employees' choices are primarily responsible for patterns of bunching in the main period we study; this motivates a model in which employees choose their earnings rather than a model in which employers coordinate responses.

Agents maximize utility  $u(c, z; n)$  over consumption and earnings (where greater earnings are associated with greater disutility at the margin), subject to a budget constraint

---

found that the graph of the probability of changing employers at age 59-61 against earnings at age 69 is smooth near the kink, suggesting that absent the Earnings Test incentives these individuals do not display noticeably different behavior in this regard.

<sup>50</sup>Formally, the elasticity of earnings with respect to the net-of-tax rate is defined as  $\varepsilon \equiv -(\partial z / z) / (\partial \tau / (1 - \tau))$ .

<sup>51</sup>Following recent literature on bunching—including Saez (2010), Chetty *et al.*, 2009, 2011, 2012a,b, Chetty (2012), and Kleven and Waseem (forthcoming)—we specify a static model of earnings choice in each period. As we discuss further in the Conclusion, dynamic considerations represent an important topic for future research.

$c = (1 - \tau)z + R$ , where  $R$  is virtual income and the parameter  $n$  reflects the tradeoff between consumption and earnings supply.<sup>52</sup> We assume that in order to change earnings from an initial level, individuals must pay a fixed utility cost of  $\phi^*$ . This cost could represent the information costs associated with navigating a new tax regime if, for example, individuals only make the effort to understand their earnings incentives when the utility gains from doing so are sufficiently large (as in Simon 1955 or Chetty *et al.* 2007). Alternatively, this cost may represent frictions such as the cost of negotiating a new contract with an employer or the time and financial cost of job search, assuming that these costs do not depend on the size of the desired earnings change.

We model a fixed cost in order to build on recent literature that has focused on fixed costs (*e.g.* Chetty *et al.*, 2011; Chetty, 2012). The distribution of earnings at ages 62 or 69 is higher in a region surrounding the kink but lower in a region substantially above the kink than at ages 61 or 70, respectively, which is consistent with a simple model with fixed adjustment costs that lead to a region of inaction and a region of adjustment.<sup>53</sup> In Appendix D.6.2, we extend our model to a case in which the cost of adjustment is linear in the size of the adjustment (and may include a fixed cost component).

We develop two different approaches for estimating elasticities and adjustment costs. Our first approach, which we call the "Comparative Static" approach, relies on comparing bunching in two separate cross-sections of data. Our second approach, which we call the "Sharp Change" approach, additionally relies on attenuation (due to adjustment costs) in the change in bunching among individuals who face a decrease in the marginal tax rate over time. The Sharp Change method relates more directly to our observation that bunching persists among individuals who formerly faced a kink. Nonetheless, the Comparative Static approach is potentially applicable in many settings, as we describe later. We begin with the Comparative Static approach because it introduces concepts that the Sharp Change method builds upon.

---

<sup>52</sup>We describe the model in more detail in Appendix D.

<sup>53</sup>However, note that even with a fixed adjustment cost, the AET could in principle cause some individuals to reduce their earnings to levels just above the kink, which in principle could lead to a rise in the density to the right of the kink due to the imposition of the AET. Moreover, the shape of the distribution of earnings at age 70 conditional on locating at the kink at age 69 cannot be predicted *a priori*, as it should depend among other things on the correlation of the fixed cost of adjustment with the elasticity of earnings with respect to the net-of-tax rate. For example, if individuals with low fixed costs of adjustment tend to have low elasticities, then the earnings distribution at age 70 should stay closer to the kink on average than if individuals with low fixed costs of adjustment tend to have high elasticities. As a result of these factors, it is difficult to use the effect of the AET on such moments of the earnings distribution in estimating elasticities and adjustment costs.

## 7.1 Estimation: Comparative Static Approach

The Comparative Static approach is best suited to estimating elasticities and adjustment costs from two cross-sections of different individuals who face different policies. Figure 6 Panel A illustrates how a fixed adjustment cost attenuates the level of bunching. Recall that our frictionless model predicts that bunchers have initial earnings (i.e. earnings in the absence of a kink) in the range  $[z^*, z^* + \Delta z_1]$ . Consider the person with initial earnings  $\underline{z}_1$  (on the linear budget constraint with tax rate  $\tau_0$ ). This individual faces a higher marginal tax rate  $\tau_1$  after the kink is introduced, which increases the marginal tax rate to  $\tau_1$  above earnings level  $z^*$ . Because she faces an adjustment cost, she could decide to keep her earnings at  $\underline{z}_1$  and locate at point 1. Alternatively, with a sufficiently low adjustment cost, she would like to pay the adjustment cost and reduce her earnings to the kink at  $z^*$  marked by point 2. We assume that the benefit of relocating to the kink is increasing in distance from the kink for initial earnings in the range  $[z^*, z^* + \Delta z_1]$ .<sup>54</sup> These assumptions imply that above a threshold level of initial earnings,  $\underline{z}_1$ , individuals adjust their earnings to the kink, and below this threshold individuals remain inert. We have drawn this individual as the marginal buncher who is indifferent between staying at the initial level of earnings  $\underline{z}_1$  (at point 1) and moving to the kink earnings level  $z^*$  (point 2) by paying the adjustment cost  $\phi^*$ .

In Panel B, we show that the level of bunching is attenuated due to the adjustment cost: only individuals with initial earnings in the range  $[\underline{z}_1, z^* + \Delta z_1]$  bunch at the kink (areas *ii*, *iii*, *iv*, and *v*)—whereas in the absence of an adjustment cost, individuals with initial earnings in the range  $[z^*, z^* + \Delta z_1]$  bunch (areas *i*, *ii*, *iii*, *iv*, and *v*). The amount of bunching is equal to the integral of the initial earnings density over the range  $[\underline{z}_1, z^* + \Delta z_1]$ :

$$B_1(\boldsymbol{\tau}_1, z^*; \varepsilon, \phi^*) = \int_{\underline{z}_1}^{z^* + \Delta z_1} h(\zeta) d\zeta. \quad (1)$$

Bunching therefore depends on the preference parameters  $\varepsilon$  and  $\phi^*$ , the tax rates below and above the kink,  $\boldsymbol{\tau}_1 = (\tau_0, \tau_1)$ , and the exempt amount  $z^*$ . The lower limit of the integral,  $\underline{z}_1$ ,

---

<sup>54</sup>In general, this requires that the size of the optimal adjustment in earnings increases in  $n$  at a rate faster than the decrease in the marginal utility of consumption. This is true, for example, if utility is quasilinear. We explore the implications of this assumption in Appendix D.5.



is implicitly defined by the indifference condition drawn in Figure 6, Panel A:

$$\phi^* \equiv u((1 - \tau_0)z^* + R_0, z^*; \underline{n}) - u((1 - \tau_1)\underline{z}_1 + R_1, \underline{z}_1; \underline{n}) \quad (2)$$

where  $R_0$  and  $R_1$  are virtual income above and below the earnings level  $z^*$ , respectively, and  $\underline{n}$  is the "ability" level of this marginal buncher. If the marginal tax rate above  $z^*$  were instead  $\tau_2$ , where  $\tau_0 < \tau_2 < \tau_1$ , then bunchers would be comprised only of individuals with initial earnings in the range  $[\underline{z}_2, z^* + \Delta z_2]$  (area *iii*), which is again attenuated relative to bunching under a frictionless model (areas *i*, *ii*, and *iii*). This generates a second expression for bunching and indifference condition analogous to 1 and 2, respectively.

When we later perform our estimates, we make use of a minimum distance estimator described in Appendix D.8 to solve this nonlinear system of equations. The key assumptions underlying that method are (a) that the density of initial earnings  $h(z)$  is uniform over the range  $[z^*, z^* + \Delta z^*]$  and (b) that utility is quasi-linear and isoelastic. The first assumption is common in the bunching literature (see Chetty et al. (2011) or Kleven and Waseem (2013), for example), as is the functional form for utility (see Saez 2010, Chetty et al. 2011, Kleven, Landais, Saez, and Schultz 2012 and Kleven and Waseem 2013, for example). We also relax the first assumption and instead estimate a lognormal distribution of earnings using individuals aged 61 (who are similarly aged but do not face the AET), as explained in Appendix D.8, which shows similar results. If we were to relax the assumption of quasilinearity, we would need to specify outside wealth, which we do not observe in the data.

### 7.1.1 Intuition and Tractable Approximation

To build intuition regarding our minimum distance estimation procedure, and to derive an expression relating the elasticity and adjustment cost to the level of bunching that can be easily solved in closed form, we can use a series of approximations to specify a simple system of linear equations. Let  $b \equiv B/h(z^*)$ , *i.e.* the amount of bunching scaled by the density of earnings at  $z^*$  when there is no kink. Also assume that  $h(z)$  is uniform and equal to  $h(z^*)$  in the range between  $\underline{z}$  and  $[z^* + \Delta z_1]$ . We show in Appendix D.6 that scaled bunching is approximately:

$$b_1(\tau_1, z^*; \varepsilon, \phi) = \varepsilon \left( z^* \frac{d\tau_1}{1 - \tau_0} \right) - \phi \left( \frac{1}{d\tau_1} \right), \quad (3)$$

where  $d\tau_1 = \tau_1 - \tau_0$  and  $\phi = \phi^*/u_c$  is the dollar equivalent of the adjustment cost. This equation shows several intuitive comparative statics: All else equal, bunching is increasing in the elasticity, decreasing in the adjustment cost, and increasing in the size of the tax change at the kink. This generalizes and nests the formula developed in Saez (2010), which is equivalent in the case in which there is no adjustment cost. Because the amount of bunching is decreasing in the adjustment cost, constraining  $\phi = 0$  and using the Saez (2010) will in general weakly underestimate the elasticity, since attenuation in bunching is attributed to a small elasticity rather than to the adjustment cost. Note that the expository derivation in (3) does not impose quasilinearity but uses the uniform density assumption and a first-order approximation for utility in the neighborhood of the kink.

Equation (3) also shows the features of the data that allow us to identify  $\varepsilon$  and  $\phi$ . We need to observe bunching at two or more kinks, with variation in the change in tax rate  $d\tau_1$ .<sup>55</sup> If we observe bunching at exactly two kinks (*i.e.* two cross-sections of data in which the tax rate differs), then we can solve for  $\varepsilon$  and  $\phi$  exactly, as we then have a system of two equations and two variables. More generally, we could estimate a regression of  $b$  on  $z^*(d\tau_1/(1 - \tau_0))$  and  $-1/d\tau_1$ , with the constant omitted. The coefficient on the first term is  $\varepsilon$ , and the coefficient on the second term is  $\phi$ .

Intuitively, with only a single cross-section of data, the amount of excess bunching increases in the elasticity and decreases in the adjustment cost, and thus it is not possible to identify both. Suppose that instead we have two cross-sections of data featuring different changes in marginal tax rates at the kink. The difference in the amount of bunching from one cross-section to the other will also depend on the elasticity and adjustment cost.<sup>56</sup> Phrased differently, the Saez (2010) formula describes how bunching should vary between two different kinks in a frictionless model, and the extent to which observed bunching deviates from this pattern is attributed to the adjustment cost.

---

<sup>55</sup>Variation in  $z^*$  alone would let us identify  $\varepsilon$  and  $\phi$ , but this identification would come solely from functional form.

<sup>56</sup>Under the approximations above, (3) implies that  $\frac{\partial b^2}{\partial \phi \partial (d\tau)} = \frac{1}{\tau^2} > 0$ ; as  $\phi$  increases, the marginal impact of  $d\tau$  on  $b$  increases.

### 7.1.2 Heterogeneity in Elasticities and Fixed Costs of Adjustment

Our empirical results suggest heterogeneity in the elasticity and the fixed cost of adjustment, as some individuals are more responsive to removal of the AET than others. Let  $(\varepsilon_i, \phi_i, n_i)$  be jointly distributed according to a smooth CDF  $G(\cdot)$ , which translates into a smooth, joint distribution of elasticities, fixed costs and earnings  $\tilde{H}(z, \varepsilon, \phi)$ . As shown in Appendix D.6.1, assuming that the density of earnings,  $\tilde{h}(z, \varepsilon, \phi)$ , is again constant over the interval  $[z, z^* + \Delta z^*]$ , we derive a generalized formula analogous to equation (3):

$$b = \bar{\varepsilon} \left( z^* \frac{d\tau_1}{1 - \tau_0} \right) - \bar{\phi} \left( \frac{1}{d\tau_1} \right),$$

where  $\bar{\varepsilon}$  and  $\bar{\phi}$  are the average elasticity and adjustment cost for those who bunch at the kink.<sup>57</sup>

## 7.2 Estimation: Sharp Change Approach

The Sharp Change approach is best suited to estimating elasticities and adjustment costs when we are examining a constant population that experiences a change in the marginal tax rate at a kink (which may involve the kink disappearing), as we observe in our empirical applications. Let  $K_1$  and  $K_2$  be two kinks that involve jumps at  $z^*$  in the marginal tax rate of  $d\tau_1 = \tau_1 - \tau_0$  and  $d\tau_2 = \tau_2 - \tau_0$ , respectively. Now consider a transition from  $K_1$  to  $K_2$  and assume  $d\tau_2 < d\tau_1$ . The Comparative Static approach would predict that the range of bunchers would decrease from  $[z_1, z^* + \Delta z_1]$  (areas *ii*, *iii*, *iv*, and *v* in Figure 6 Panel B) to  $[z_2, z^* + \Delta z_2]$  (area *iii*). However, since we are moving from a larger kink to a smaller kink, adjustment costs prevent some individuals from "un-bunching" from the kink, even though they would prefer to move away from the kink in the absence of an adjustment cost. The fixed adjustment cost therefore not only attenuates bunching in a cross-section (as under the Comparative Static approach) but also attenuates the change in bunching between two cross-sections in response to a reduction in the size of the kink.<sup>58</sup>

The first source of attenuation in the change is driven by individuals in area *ii* of Panel B. They bunch under  $K_1$  and continue to bunch after transitioning to  $K_2$ . The reason is that the

<sup>57</sup>We are grateful to Henrik Kleven for suggesting the approach that led to this derivation.

<sup>58</sup>This result is not symmetric. If  $d\tau_2 > d\tau_1$  instead – *i.e.* the kink becomes larger – then additional individuals will be induced to bunch (while facing an adjustment cost), and the change in bunching will be the same as in the Comparative Static model.

frictionless optimum under  $K_2$  is  $z^*$  for everyone initially earning in the range  $[z^*, z^* + \Delta z_2]$ .

The second source of attenuation in the change is driven by individuals in area *iv* of Panel B. Panel C of Figure 6 demonstrates this. At point 0, we show an individual's initial earnings  $\bar{z}_0 \in [z^*, z^* + \Delta z_1]$  under a constant marginal tax rate of  $\tau_0$ . We now introduce the first kink,  $K_1$ . The individual responds by bunching at  $z^*$  at point 1. Next, we transition to the muted kink  $K_2$ . Note that since  $\bar{z}_0 > z^* + \Delta z_2$ , this individual would have chosen earnings  $\bar{z}_2 > z^*$  (marked as point 2) under  $\tau_2$ , if we had gone directly from no kink to  $K_2$ . However, in order to move to point 2, this individual must pay a fixed cost of  $\phi^*$ . We have drawn this individual as a marginal buncher who is indifferent between staying at  $z^*$  and moving to  $\bar{z}_2$ . All individuals with initial earnings in the range  $[z^* + \Delta z_2, \bar{z}_0]$  will remain at the kink.

Thus, bunching under  $K_2$  is:

$$B_2(\tau_2, z^*; \varepsilon, \phi^*) = \int_{z_1}^{\bar{z}_0} h(\zeta) d\zeta. \quad (4)$$

It follows that the absolute value of the change in bunching from  $K_1$  to  $K_2$  under the Sharp Change approach (area *v* in Panel B) will be smaller than under the Comparative Static approach (areas *ii*, *iv* and *v*). As discussed in Appendix D.7,  $\varepsilon$  is still identified by the adjustment of the top-most buncher:  $\varepsilon = \frac{\bar{z}_0 - \bar{z}_2}{\bar{z}_2} \frac{(1 - \tau_0)}{d\tau_2}$ . The critical earnings level  $\bar{z}_2$  is defined implicitly by the indifference condition in Panel C:

$$\phi^* \equiv u((1 - \tau_2)\bar{z}_2 + R_2, \bar{z}_2; \bar{n}) - u((1 - \tau_0)z^* + R_0, z^*; \bar{n}). \quad (5)$$

As before, our estimates use the minimum distance estimator described in Appendix D.8 to solve the system of nonlinear equations defined by (1), (4), and (5). Intuitively, as before, with a single cross-section of data, the amount of excess bunching increases in the elasticity and decreases in the adjustment cost, and thus it is not possible to identify both. We rely on a before-and-after comparison of bunching at the same kink, once the jump in marginal tax rates has been reduced. In this case, inertia generates an excess amount of bunching in the period after the policy change. In the extreme case in which a kink has been eliminated, we can attribute any residual bunching to adjustment costs. The amount of residual bunching at the kink, in combination with the amount of bunching prior to the change in the jump in MTRs

at the kink, therefore helps to identify both the elasticity and the adjustment cost.<sup>59</sup>

### 7.3 Estimates of Elasticity and Adjustment Cost

To estimate  $\varepsilon$  and  $\phi$ , we examine several changes in the AET benefit reduction rate, including a reduction in the rate in 1990; the elimination of the AET from ages 69 to 70; and the elimination of the AET in 2000. Our Sharp Change method is applicable in all of these contexts, as we observe a group from before to after these changes in policy. By contrast, the Comparative Static method is more applicable to analyzing changes in bunching when comparing two different groups each with positive marginal tax rates in different steady states, and is inapplicable to estimating elasticities and adjustment costs using residual bunching when a kink has disappeared.

Estimating  $\varepsilon$  and  $\phi$  requires estimates of the implicit marginal tax rate that individuals face. This requires estimates of both the "baseline" marginal tax rate,  $\tau$ —the rate that individuals near the AET threshold face in the absence of the AET due to federal and state taxes—and estimates of the implicit marginal tax rate associated with the AET. We begin by using a marginal tax rate that incorporates the effects of the AET BRR, as well as the average marginal income and FICA tax rates (including federal and state taxes).<sup>60</sup> These estimates are predicated on correctly specifying the marginal tax rate (net of benefit enhancement), but recall that Appendix A shows that benefit enhancement is not relevant to an individual's marginal incentives for earning an extra dollar near the AET exempt amount. This assumption is also consistent with the methodology of Friedberg (1998, 2000), who treats the AET as a pure tax. Moreover, as we have noted, the evidence shows little systematic bunching reaction to changes in the DRC. We vary these assumptions in various dimensions, which show similar results to the baseline: we exclude FICA taxes in the calculation of the baseline tax rate,<sup>61</sup> and we alternatively assume that a portion of the benefit enhancement corresponds to a reduction in the effective marginal tax rate.

---

<sup>59</sup>An approximation explained in the Appendix also helps to build intuition.

<sup>60</sup>We calculated the mean of the sum of federal and state marginal income and FICA tax rates for people with positive Social Security benefits and earnings within \$2000 of the kink in 1988-1992, using TAXSIM and the Statistics of Income individual tax return files. The results are not sensitive to other such choices.

<sup>61</sup>Liebman, Luttmer, and Seif (2009) show that labor supply reacts to OASI benefit rules, suggesting that individuals may not perceive FICA taxes as pure taxes.

Before turning to our empirical estimates, we begin with graphical depictions of the patterns driving the estimation. Figure 4 shows excess bunching among 66-69 year-olds, for whom the BRR fell from 50 percent to 33.33 percent in 1990. It is clear that subsequent to 1990, excess bunching generally fell, corresponding to the fall in the marginal tax rate.<sup>62</sup> Meanwhile, Appendix Figure E.7 shows that bunching stayed relatively constant for the 62-64 year-old group that experienced no policy change in 2000. While this group faces a notch at the exempt amount rather than a kink (as explained above), the relative comparison is instructive and suggests that the fall in bunching in the 66-69 year-old group after 1990 was due to the policy change in 1990.

Appendix Figure E.19 shows that the elasticity we estimate using the Saez method—constraining the adjustment cost to be zero—rises substantially from 1989 to 1990, as we might expect given that adjustment in other contexts appears to take longer than a year. This relates directly to our Sharp Change theory, which predicts that following a reduction in the change in the MTR at the kink, there may be excess bunching due to inertia (corresponding to area *iv* in Figure 6, Panel B).<sup>63</sup> Once we allow for an adjustment cost, this excess bunching is attributed to optimization frictions. Note that this finding is in some sense the opposite of the usual presumption that adjustment frictions should attenuate elasticity estimates. In the context of measuring an elasticity when individuals have not yet had a chance to adjust (and the effective marginal tax rate has fallen), frictions can increase elasticity estimates.

Table 2 presents our results from the Sharp Change method, examining the 1990 change. We estimate an elasticity of 0.23 in Column (1) and a positive adjustment cost of \$152.08 in Column (2), both significantly different from zero ( $p < 0.01$ ). This specification examines data in 1989 and 1990; thus, our estimated adjustment cost represents the cost of adjusting earnings in the first year after the policy change. When we constrain the adjustment cost to zero using 1990 data in Column (3), as most previous literature has implicitly done, we estimate a substantially larger elasticity of 0.39.<sup>64</sup> Consistent with our discussion above, it makes sense

---

<sup>62</sup>The Delayed Retirement Credit changed from 3 percent to 4 percent over this period, though this should not have affected the AET MTR for those sufficiently close to the kink, as explained above.

<sup>63</sup>In the Conclusion, we discuss dynamic considerations that might subsequently cause residual bunching to disappear. We consider dynamic considerations to be an important topic for future research. Our focus in the Sharp Change approach is limited to the period just after a policy change, before residual bunching has dissipated.

<sup>64</sup>Note that despite the differences in methodology between our study and that of Friedberg (2000), our results are in the same range as her uncompensated elasticity estimates of 0.22 and 0.32 in the different samples she

that the estimated elasticity is higher when we do not allow for adjustment costs than when we do, as adjustment costs keep individuals bunching at the kink even though tax rates have fallen. The difference in the constrained and unconstrained estimates of the elasticity is substantial (69 percent higher in the constrained case) and statistically significant ( $p < 0.01$ ).

We also consider alternative specifications. In the next row, using a lognormal earnings density rather than a uniform density (as described in the Appendix) changes the results little. In the following row, adjusting the marginal tax rate to take account of benefit enhancement (relevant to those individuals to whom benefit enhancement is relevant to their earnings choices) raises the estimated elasticity but yields similar qualitative patterns across the constrained and unconstrained estimates. This makes sense: for the same behavioral response, if we assume a less pronounced percentage change in the net-of-tax rate, we infer a larger elasticity. The next rows show other specifications: excluding FICA taxes from the baseline tax rate; other bandwidths; and other years of analysis. Our results are similar under these and other variations.

In Appendix Table F.1, we apply the Sharp Change method to the disappearance of the kink at age 70 (in which context we find residual bunching in Figure 2) and find similar (slightly higher) elasticity estimates and somewhat lower adjustment costs (though still in the same range). It is worth noting that the constrained estimate of the elasticity is smaller than the unconstrained estimate. This makes sense, because we use data from age 69 to perform the constrained estimate, and adjustment costs attenuate the constrained estimate at those ages because they reduce bunching (just as they do in Table 2, Panel A, in which we perform the constrained estimate using only data from before the 1990 policy change).<sup>65</sup>

Appendix Tables F.2, F.3, and F.4 present further specifications. In Appendix Table F.2, we apply the Sharp Change approach to the 1990 policy change but assume that bunching in 1989 is not attenuated by adjustment frictions (under the rationale that bunching could have reached a "steady state" in 1989 that is not attenuated by adjustment frictions). We estimate results similar to the baseline. In Appendix Table F.3, we use the Sharp Change approach to estimate

---

considers. However, differences in the estimation strategies imply that these results are not directly comparable to ours.

<sup>65</sup>We cannot estimate the constrained elasticity using only data on age 70 because the benefit reduction rate is zero at that age. We show these results in the Appendix, rather than the main text, because the estimates of excess bunching at age 70 are potentially affected by the coarse measure of age that we use, as explained above. To address this issue, we use age 71 in estimating these results. Using only age 70 shows similar results, which is unsurprising because Figure 2 shows that normalized excess bunching is similar at ages 70 and 71.

elasticities and adjustment costs using the disappearance of the kink in the year 2000 for those NRA and above. Given our small point estimate of residual bunching in 2000, it is unsurprising that we find small adjustment costs (though they are still marginally significant). We find smaller elasticities than in the baseline specification in part because normalized bunching was relatively small (though still positive and significant) in 1999, as shown in Figure 4.

In Appendix Table F.4, we use the Comparative Static approach to perform estimates around the 1990 policy change. As noted above, the Comparative Static method is most applicable to performing estimates on unrelated cross-sections of individuals—rather than just before and just after a change in policy as in our results on residual bunching.<sup>66</sup> We use years 1989 and 1993 to make sure that none of the 66-69 year-olds in 1993 were observed in 1989 (to avoid the possibility of residual bunching at the kink among those initially bunching there). We again find estimates in the same range (with higher though less significant adjustment costs). The linear approximation (3) shows higher (but still modest) adjustment costs, and higher elasticities than in the baseline specification (which makes sense because higher elasticities are needed to reconcile the higher estimate adjustment cost with the observed change in excess normalized bunching).

## 8 Conclusion

In the context of the Social Security Annual Earnings Test, we investigate the nature and size of earnings adjustment frictions. We develop several related findings. First, we examine the speed of adjustment to the disappearance of convex kinks in the effective tax schedule. We document evidence of delays in adjustment in certain contexts, consistent with the existence of earnings adjustment frictions in the U.S. The evidence showing delayed adjustment in some cases suggests that regressing earnings or taxable income on contemporaneous tax rates is not the correct specification. Nonetheless, we find that adjustment to both anticipated and unanticipated policy changes is quite rapid, as the vast majority of adjustment occurs within at most three years of budget set changes. This suggests that long-run elasticities are similar

---

<sup>66</sup>As noted above, the Comparative Static approach is inapplicable to years in which we observe residual bunching at a former kink; in assuming zero marginal tax rate after a kink disappears, the Comparative Static method effectively cannot explain this residual bunching. Thus, the Sharp Change method is more directly motivated by our primary empirical observations.



to those estimated in a medium-run time frame of a few years in this context.

Second, we investigate the mechanisms that underlie the patterns of adjustment. Adjustment to removal of kinks occurs primarily through substantial earnings growth among those initially locating at the kink, suggesting that they are more responsive than others (due to some combination of different elasticities or adjustment costs). We additionally investigate the extent to which firms may help coordinate bunching. The responses appear to be driven mainly by employees, as only those subject to the AET bunch at the AET kink during the primary period we examine. This combination of evidence suggests that the bunching primarily results from the choices of certain particularly responsive employees who choose themselves to vary their earnings.

Third, we specify a model of employees' earnings adjustment consistent with these findings and use it to estimate the earnings elasticity and the fixed adjustment cost. When we consider the change in bunching associated with the reduction in the AET benefit reduction rate from 50 percent to 33.33 percent for those above NRA from 1989 to 1990, we estimate that the elasticity is 0.23 and the adjustment cost is \$152.08. The results are similar with other populations, time periods, and methods. When we constrain adjustment costs to zero in the baseline specification, the elasticity we estimate in 1990 (0.39) is substantially (69 percent) larger, demonstrating the potential importance of taking account of adjustment costs.

Our estimates demonstrate the applicability of the methodology and the potential importance of allowing for adjustment costs when estimating elasticities. The modest adjustment cost we estimate parallels our empirical observation that bunching responds rapidly to changes in policy. It is interesting that even among the group locating at the kink, that is able to adjust earnings initially to the level of the exempt amount, we still find evidence of modest adjustment frictions—raising the possibility that adjustment frictions could be larger in other contexts.

The analysis leaves open a number of avenues of further inquiry. First, it would further enrich the framework to extend the static analysis here to a dynamic model. We consider our static framework for estimating elasticities and adjustment costs to be a first step in understanding estimation of these parameters (in the spirit of other static papers such as Saez 2010, Chetty *et al.*, 2009, 2011, 2012a,b, Chetty 2012, and Kleven and Waseem 2013), but incorporating dynamic considerations is an important next step. The speed at which individuals respond to changes in the AET, the nature of income effects, and the distinction between anticipated and

unanticipated changes—all of which we have begun exploring in this paper—are three of several possible pieces of evidence that could help in specifying a model of this sort (perhaps including a stochastic wage arrival process).

Second, further work distinguishing among the possible reasons for reaction to the AET—such as liquidity constraints or misperceptions—remains an important outstanding issue. Third, further investigation of extensive margin responses to the AET—including both the decision of whether to earn a positive amount and the OASI claiming decision—would be valuable. Fourth, following most previous literature, we have modeled the adjustment cost as a "black box," without modeling the process that underlies this cost, such as information acquisition or job search. Future research could model such processes and distinguish these explanations using data. Finally, the AET policy environment provides a useful illustration of many issues—such as a methodology for estimating elasticities and adjustment costs simultaneously—that should be applicable more broadly to studying adjustment to policy. As elasticities and adjustment costs may be substantially different in other contexts, studying earnings adjustment to other policies is a high priority.

## References

- Abowd, John M., Bryce E. Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L. McKinney, Marc Roemer, and Simon Woodcock (2009), "The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators." In *Producer Dynamics: New Evidence from Micro Data* (Timothy Dunne, J. Bradford Jensen, and Mark J. Roberts, eds.), 149–230, University of Chicago Press.
- Altonji, Joseph G. and Christina H. Paxson (1992), "Labor Supply, Hours Constraints, and Job Mobility." *Journal of Human Resources*, 27, 256–278.
- Blundell, Richard, and Hilary Hoynes (2004), "Has 'In-Work' Benefit Reform Helped the Labor Market?" In *Seeking a Premier Economy: The Economic Effects of British Economic reforms, 1980–2000*, (David Card, Richard Blundell, and Richard B. Freeman, eds.), 411–59, University of Chicago Press.
- Burtless, Gary and Robert A. Moffitt (1985), "The Joint Choice of Retirement Age and Postretirement Hours of Work." *Journal of Labor Economics*, 3, 209–236.
- Caplin, Andrew S. and Daniel F. Spulber (1987), "Menu Costs and the Neutrality of Money." *Quarterly Journal of Economics*, 102, 703–726.

- Chetty, Raj (2012), "Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." *Econometrica*, 80, 969–1018.
- Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri (2011), "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." *Quarterly Journal of Economics*, 126, 749-804.
- Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber (2012), "Does Indivisible Labor Explain the Difference Between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities." In *NBER Macroeconomics Annual 2012* (Daron Acemoglu, Jonathan Parker, and Michael Woodford, eds.), vol. 27, University of Chicago Press.
- Chetty, Raj, Adam Looney, and Kory Kroft (2007), "Salience and Taxation: Theory and Evidence." NBER Working Paper No. 13330.
- Chetty, Raj, Adam Looney, and Kory Kroft (2009), "Salience and Taxation: Theory and Evidence." *American Economic Review*, 99, 1145–1177.
- Chetty, Raj, Emmanuel Saez, and John N. Friedman (2012), "Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings." Forthcoming, *American Economic Review*.
- Cogan, John (1981), "Fixed Costs and Labor Supply." *Econometrica*, 49, 945-963.
- Coile, Courtney, and Jonathan Gruber (2001), "Social security incentives for retirement." In *Themes in the Economics of Aging* (David Wise, Ed.), 311-354, University of Chicago Press.
- Diamond, Peter, and Jonathan Gruber (1999), "Social Security and Retirement in the United States." In *Social Security and Retirement around the World* (Jonathan Gruber and David Wise, eds.), 437-473, University of Chicago Press.
- Eissa, Nada, and Jeffrey B. Liebman (1996), "Labor Supply Response to the Earned Income Tax Credit." *Quarterly Journal of Economics*, 111, 605-37.
- Friedberg, Leora (1998), "The Social Security earnings test and labor supply of older men." In *Tax Policy and the Economy* (James M. Poterba, ed.), 121–150, University of Chicago Press.
- Friedberg, Leora (2000), "The Labor Supply Effects of the Social Security Earnings Test." *Review of Economics and Statistics*, 82, 48–63.
- Gelber, Alexander, Damon Jones, and Daniel Sacks (2013), "Misperceptions of the Social Security Earnings Test." University of Chicago Working Paper.
- Gruber, Jonathan and Peter Orszag (2003), "Does the Social Security Earnings Test Affect Labor Supply and Benefits Receipt?" *National Tax Journal*, 56, 755–773.

- Hausman, Jerry A. (1981), "Labor Supply." In *How Taxes Affect Economic Behavior* (Henry J. Aaron and Joseph A. Pechman, Eds.), 27–71, Brookings Institution.
- Kleven, Henrik, Martin Knudsen, Claus Kreiner, Søren Pedersen, and Emmanuel Saez (2011), "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark." *Econometrica*, 79, 651-692.
- Kleven, Henrik, Camille Landais, Emmanuel Saez, and Esben Schultz (2012), "Taxation and International Migration of Top Earners: Evidence from the Foreigner Tax Scheme in Denmark." Forthcoming, *American Economic Review*.
- Kleven, Henrik and Mazhar Waseem (2013), "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan." *Quarterly Journal of Economics*, 128, 669-723.
- Kline, Patrick, and Melissa Tartari (2013), "What Distributional Impacts Mean: Welfare Reform Experiments and Competing Margins of Adjustment," UC Berkeley Working Paper.
- Liebman, Jeffrey B., and Erzo F.P. Luttmer (2011), "Would People Behave Differently If They Better Understood Social Security? Evidence From a Field Experiment." NBER Working Paper 17287.
- Liebman, Jeffrey B., Erzo F.P. Luttmer, and David Seif (2009), "Labor Supply Responses to Marginal Social Security Benefits: Evidence from Discontinuities," *Journal of Public Economics*, 93, 1208-1223.
- McKinney, Kevin L. and Lars Vilhuber (2008), "LEHD Infrastructure Files in the Census RDC - Overview Revision: 219." U.S. Census Bureau, LEHD Program.
- Meyer, Bruce D and Dan T. Rosenbaum (2001), "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers." *Quarterly Journal of Economics*, 116, 1063–1114.
- Moffitt, Robert (1986), "The Econometrics of Piecewise-Linear Budget Constraints: A Survey and Exposition of the Maximum Likelihood Method." *Journal of Business and Economic Statistics*, 4, 317–28.
- Moffitt, Robert (1990), "The Econometrics of Kinked Budget Constraints." *Journal of Economic Perspectives*, 4, 119–39.
- Saez, Emmanuel (2010), "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy*, 2, 180–212.
- Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz (2012), "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature*, 50, 3–50.

Simon, Herbert A. (1955), "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics*, 69, 99–118.

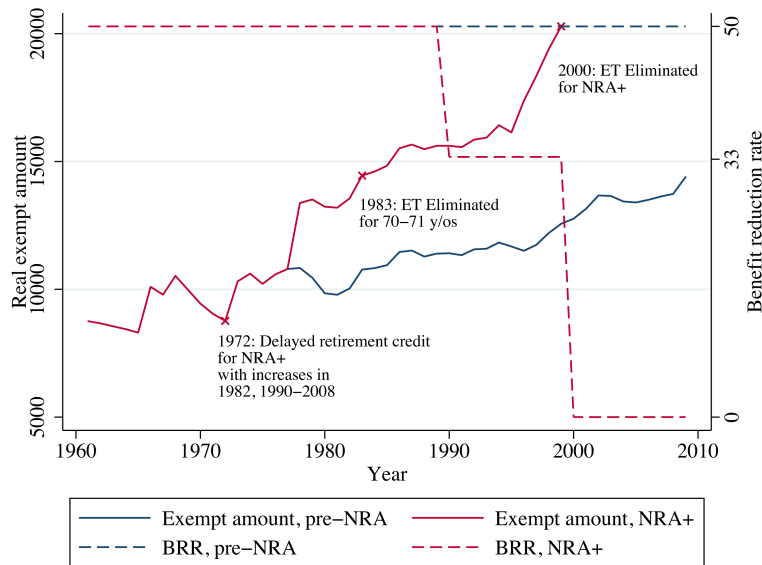
Social Security Administration (2012a), *Annual Statistical Supplement*. Washington, D.C.

Social Security Administration (2012b), *Social Security Handbook*. Washington, D.C.

Song, Jae G. and Joyce Manchester (2007), "New evidence on earnings and benefit claims following changes in the retirement earnings test in 2000." *Journal of Public Economics*, 91, 669–700.

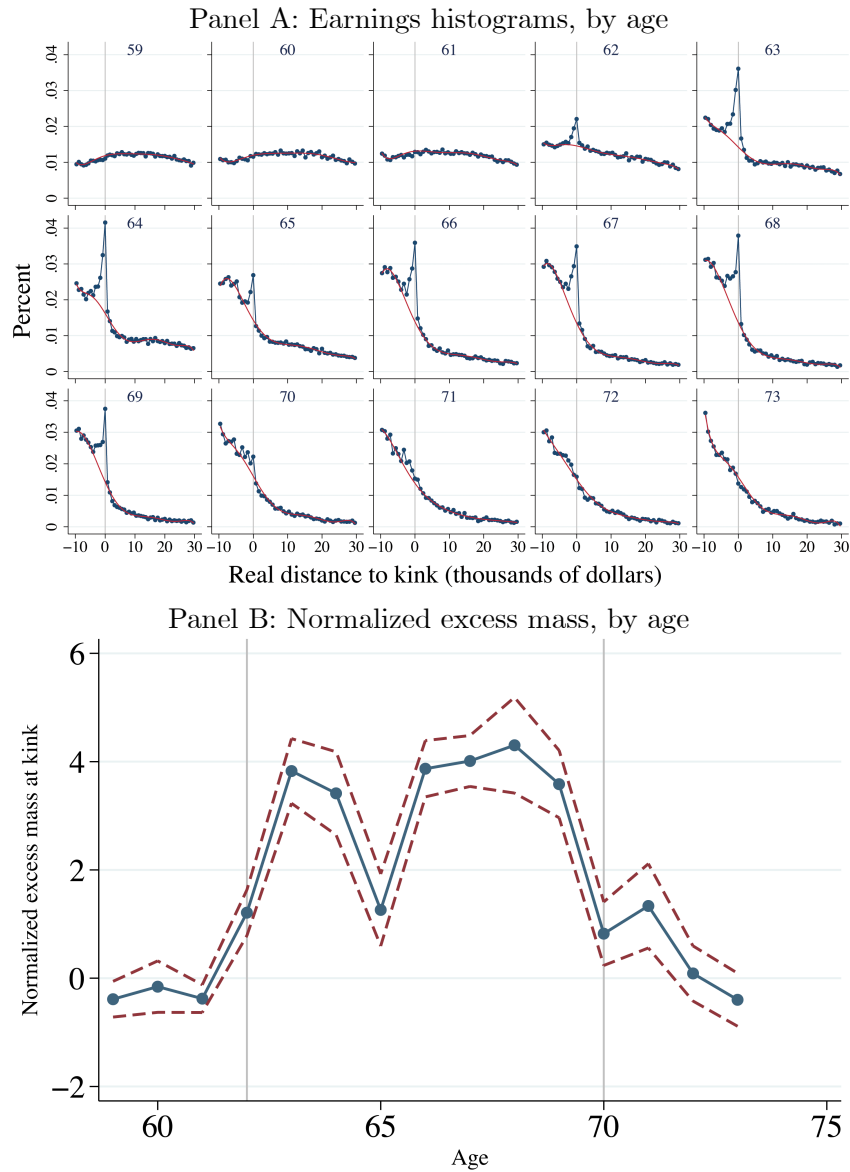
Utendorf, Kevin R. (2001/2), "The Upper Part of the Earnings Distribution in the United States: How Has It Changed?" *Social Security Bulletin*, 64, 1–11.

Figure 1: Earnings Test Rules, 1961-2009



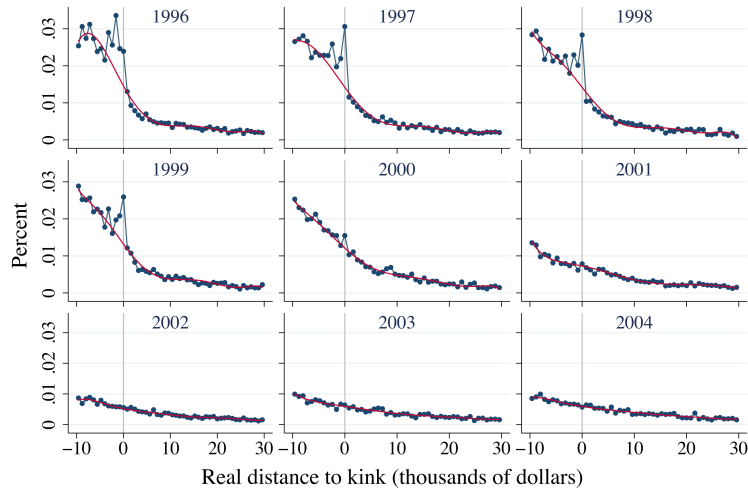
Note: The figure illustrates the primary relevant variation in AET parameters from 1961 to 2009. The right vertical axis measures the benefit reduction rate in OASI payments for every dollar earned beyond the exempt amount. The left vertical axis measures the real value of the exempt amount over time. "NRA" refers to the Normal Retirement Age. "BRR" refers to the AET benefit reduction rate above the exempt amount.

Figure 2: Adjustment Across Ages: Histograms of Earnings and Normalized Excess Mass, 59-73-year-olds Claiming OASI by Age 65, 1990-1999



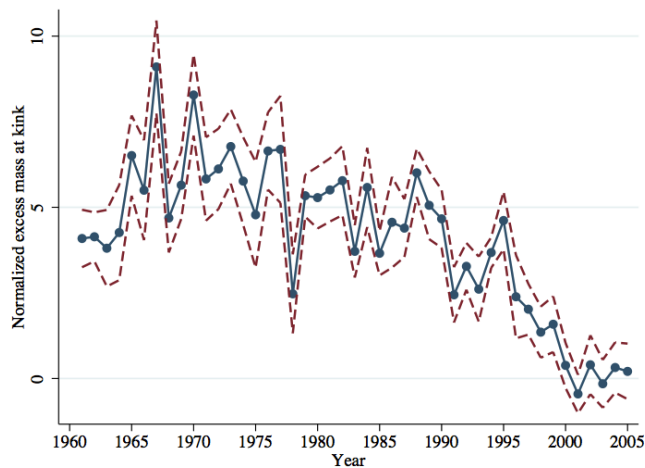
Note: The figure shows earnings histograms (Panel A) and normalized excess bunching (Panel B) from a one percent random sample of SSA administrative data on Social Security claimants aged 59-73 between 1990 and 1999 (inclusive). Normalized excess bunching is calculated as described in the text. The bin width is \$800. The earnings level zero, shown by the vertical lines in Panel A, denotes the kink. Earnings are measured in CPI-U-adjusted year 2010 dollars in each year. In this and all other figures in the main text, "claimant" refers to individuals who have claimed by age 65. Solid (dotted) lines show point estimates (95 percent confidence intervals). The vertical lines in Panel B show the ages at which the AET first applies (62) and the age at which the AET ceases to apply (70).

Figure 3: Adjustment Across Years: Histograms of Earnings and Normalized Excess Mass, 66-69 year olds Claiming OASI by Age 65, 1996-2004



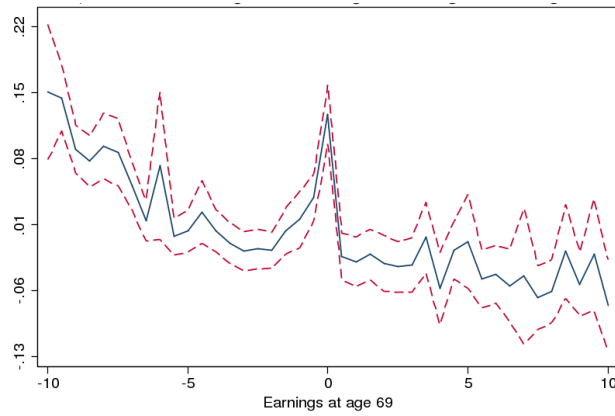
Note: The figure shows histograms of earnings from a one percent random sample of SSA administrative data on Social Security claimants aged 66-69 in each year from 1996 to 2004 (inclusive). In 2000 and after, the (placebo) kink is defined as the kink applying to those in the year of attaining NRA; as we mention in the text, the results are robust to two other ways of defining this placebo kink. See other notes to Figure 2 Panel A.

Figure 4: Normalized Excess Bunching by Year, 1961-2005



Note: The figure shows normalized excess bunching from a one percent random sample of SSA administrative data on Social Security claimants aged 66-69 in each year between 1961 and 2005. See other notes to Figure 2 Panel B.

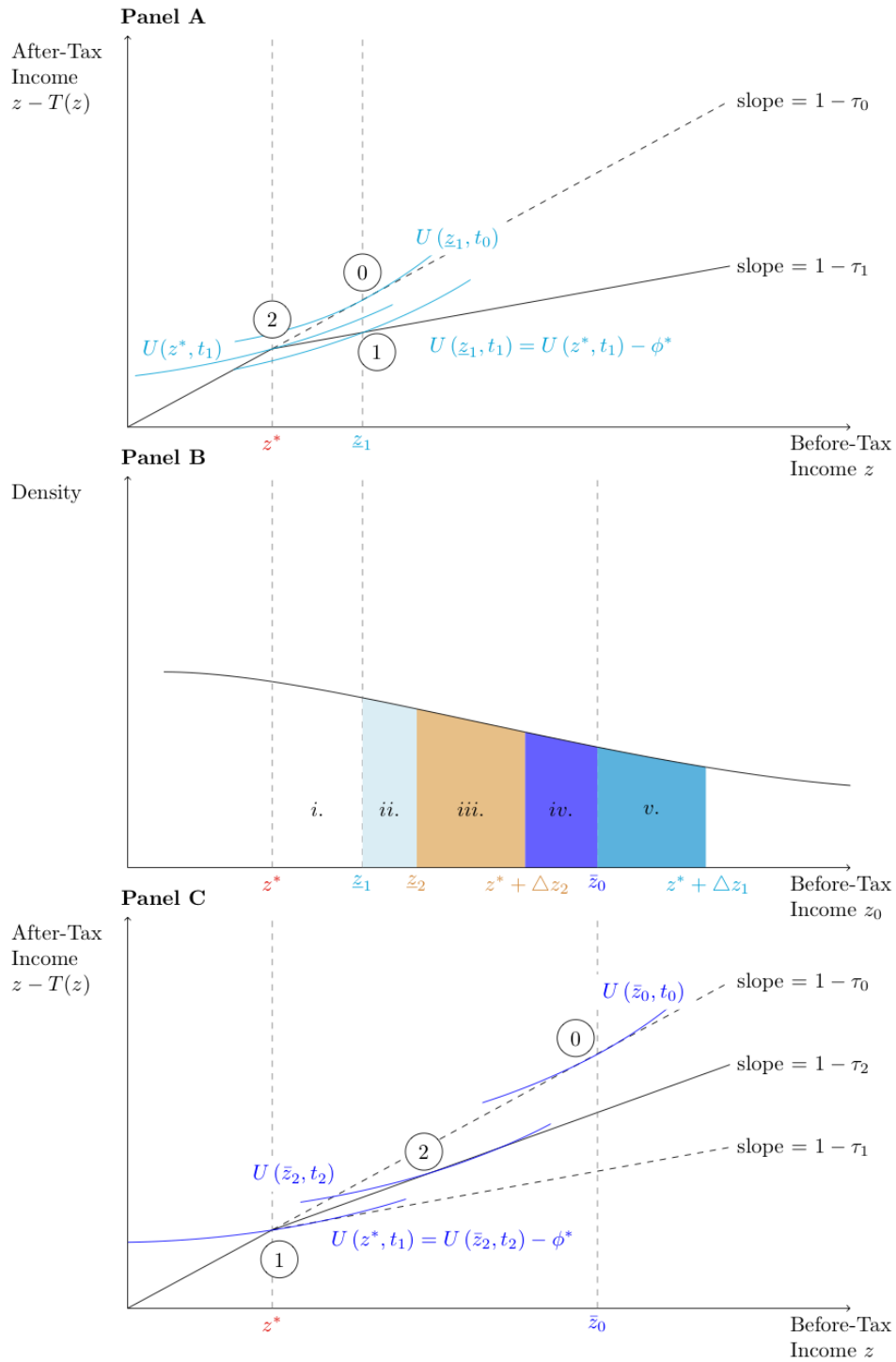
Figure 5: Mean Fraction Change in Earnings from Age 69 to 70, by Earnings at 69, 1990-1998



Note: The figure shows the mean fraction change in earnings from age 69 to age 70 (y-axis), against earnings at age 69 (x-axis). For example, 0.15 on the y-axis corresponds to a 15 percent mean increase in earnings from age 69 to age 70. Earnings are measured relative to the kink, shown at zero on the x-axis. The data are a 20 percent random sample of 69-year-olds in the LEHD in 1990-1998. We exclude 1999 as a base year in this and similar graphs because the AET is eliminated for those over NRA in 2000.



Figure 6: Bunching Responses to a Convex Kink, with Fixed Adjustment Costs



Note: See Section 7 for an explanation of the figures.

Table 1: Summary Statistics, Social Security Administration Master Earnings File

	(1)	(2)
	Ages 18-75	Ages 62-69
Mean Earnings	37,492.28 (282,940.03)	29,485.08 (783,897.87)
10th Percentile	5,234.74	1,957.78
25th Percentile	15,291.11	7,291.21
50th Percentile	32,982.17	17,739.68
75th Percentile	46,949.86	38,149.90
90th Percentile	66,370.18	58,417.99
Fraction with Positive Earnings	0.42	0.41
Fraction Male	0.56	0.57
Observations	13,612,313	1,595,139
Individuals	619,580	545,615

Note: The data in Panel A are taken from a one percent random sample of earnings records from SSA administrative data. The data for ages 18-75 cover those in 1961-2005 who claim by age 65, who do not report self-employment earnings, and who have positive earnings. (However, the fraction with positive earnings is calculated by including those who have zero earnings.) Column 2 covers the same sample but limits the ages to 62-69, the group we examine most often. Earnings are FICA (capped) earnings and are expressed in 2010 dollars. Numbers in parentheses are standard deviations. The standard deviation is large because of very rare aberrant large values of earnings (as documented in Utendorf 2001/2); these do not affect our estimates in the figures or tables because they are far above the AET exempt amount, and they affect mean earnings far less than they affect the standard deviation. (The results are robust to winsorizing.)

Table 2: Estimates of Elasticity and Adjustment Cost Using Sharp Change Method

	(1)	(2)	(3)	(4)
	$\varepsilon$	$\phi$	$\varepsilon \phi = 0$	
			1990	1989
Baseline	0.23 [0.20, 0.27]***	\$152.08 [49.54, 382.94]***	0.39 [0.33, 0.48]***	0.22 [0.18, 0.26]***
Lognormal density	0.26 [0.22, 0.30]***	\$165.76 [45.57, 473.52]***	0.43 [0.35, 0.54]***	0.24 [0.19, 0.29]***
Benefit Enhancement	0.39 [0.33, 0.45]***	81.52 [18.11, 245.88]***	0.59 [0.49, 0.73]***	0.37 [0.30, 0.43]***
Excluding FICA	0.32 [0.27, 0.36]***	\$129.14 [34.18, 360.97]***	0.50 [0.42, 0.62]***	0.30 [0.25, 0.35]***
Bandwidth = \$500	0.25 [0.21, 0.30]***	\$90.65 [6.04, 319.69]***	0.38 [0.30, 0.50]***	0.24 [0.19, 0.30]***
1988 to 1990	0.32 [0.29, 0.35]***	\$114.33 [26.49, 307.96]***	0.50 [0.42, 0.62]***	0.31 [0.27, 0.34]***

Note: The table shows estimates of the elasticity and adjustment cost using the Sharp Change method described in the text. We report bootstrapped confidence intervals shown in parentheses. We investigate the 1990 reduction in the AET BRR from 50 percent to 33.33 percent. The baseline specification assumes a uniform density, calculates the effective MTR by including the effects of the AET BRR and federal and state income and FICA taxes, and uses data from 1989 and 1990. Alternative specifications deviate from the baseline as noted. The estimates that include benefit enhancement use effective marginal tax rates due to the AET based on the authors' calculations relying on Coile and Gruber (2001) (assuming that individuals are considering earning just enough to trigger benefit enhancement). The left endpoint of the confidence interval in the "Bandwidth = \$500" row, Column 2 is zero because of the constraint  $\phi \geq 0$  from theory, as discussed in the Appendix. Columns (1) and (2) report joint estimates with  $\phi \geq 0$  imposed, while Columns (3) and (4) impose the restriction  $\phi = 0$ . The constrained estimate in Column 3 only uses data from 1990, whereas that in Column 4 uses only data from 1989 (except in the row where we investigate data from 1988 and 1990, in which Column 4 uses data from 1988). \*\*\* indicates that an estimate is significantly different from zero at the 1 percent level; \*\* at the 5 percent level; and \* at the 10 percent level.

## A Appendix: Additional Features of Annual Earnings Test

When current benefits are lost to the AET, future scheduled benefits are increased in some circumstances.<sup>67</sup> This is sometimes called benefit enhancement. This enhancement is actuarially fair if the net present value of the benefit enhancement equals the benefits lost due to the AET. Past research has sometimes interpreted benefit enhancement as implying an attenuation of the marginal work disincentives associated with the AET, with actuarially fair benefit enhancement implying no marginal work disincentive for fully informed, rational agents with average mortality expectations and discounting who are facing perfect credit markets.

We show in this Appendix that for workers NRA and older, at the AET exempt amount the effective marginal tax rate associated with AET is equal to the AET BRR, so that the AET essentially behaves as a kink for those considering earning near the AET earnings threshold. This is because benefit enhancement for this group occurs only once a month's worth of benefits are lost to the AET. As a result, no benefit enhancement occurs when an individual earns a marginal dollar at or near the AET earnings threshold. Phrased differently, the DRC is approximately actuarially fair in the sense that delaying claiming an extra year is approximately actuarially fair, but this does not imply that the DRC acts in an actuarially fair manner for an individual who has her benefits reduced due to the AET in a given year and is considering earning an extra dollar. In fact, as we describe, earning an extra dollar is not actuarially fair for (among others) those with positive OASI benefits considering earning an extra dollar above the AET exempt amount. The remainder of this Appendix describes these rules in greater detail.

The benefit enhancement rules have varied over time, and they depend on whether the beneficiary is above or below NRA. Prior to 1972, there was no benefit enhancement for people aged NRA and older. In these years, the AET represented a pure loss in benefits for those NRA and above (equivalent to a pure tax).

For beneficiaries at the NRA and older, a one percent Delayed Retirement Credit (DRC) was introduced in 1972. The DRC was intended to compensate beneficiaries who delayed claiming beyond age 65, but they also apply to earnings lost to the AET. For individuals above NRA, benefits are increased 1/12 of 1 percent for each month between ages 65 and 72 for which no benefits received after 1970 (Social Security, 2012, Table 2.A.20).<sup>68</sup> This language indicates that each month's worth of foregone benefits—either because of delayed claiming or because of the AET—results in increased future benefits. Critically, the beneficiary has to forego an *entire* month of benefits in order to receive the DRC; if, for example, she earns slightly over the exempt amount and loses a small amount of benefits to the AET, then her future benefits are not adjusted. Thus, the DRC provides no *marginal* relief from the AET.

As a result of these rules, future benefits are enhanced when the individual's yearly earnings are over  $z^* + (MB/\tau)$ , where  $z^*$  is the exempt amount,  $MB$  is the monthly benefit, and  $\tau$  is the AET benefit reduction rate. For example, with a typical monthly benefit of \$1,000 and a benefit reduction rate of 33.33 percent, benefit enhancement occurs when the individual's yearly earnings are \$3,000 ( $=\$1000/0.3333$ ) above the exempt amount.<sup>69</sup> Benefit enhancement

---

<sup>67</sup>This section is based on table 2.A.20 of the Annual Statistical Supplement of the Social Security bulletin, as well as extensive email correspondence with numerous officials at the Social Security Administration.

<sup>68</sup>The size of the DRC was increased to three percent per year in 1982, and then increased steadily throughout the 1990s, reaching eight percent for each year of foregone benefits in 2008.

<sup>69</sup>Another month's benefit enhancement would occur if the individual earns more than  $z^* + (2MB/\tau)$ ; a third

corresponding to one more month of reduced earnings occurs once annual earnings reaches \$6,000 above the exempt amount, and so forth. Thus, benefit enhancement is only relevant to an individual considering earning substantially in excess of the exempt amount and is therefore not relevant to marginal earnings decisions at the exempt amount. Indeed, this theoretical presumption is consistent with suggestive evidence we describe that indicates little systematic bunching reaction to changes in the DRC.

The AET is implemented in a number of stages. First, SSA must determine that a claimant is expected to exceed the exempt amount, or that she has already done so. Claimants can notify SSA in advance if they expect to exceed the exempt amount, or they can report their earnings *ex post facto* at any point in the year. In addition, SSA uses W-2 records at the end of the year to determine if the AET threshold has been crossed (for those who have a W-2). Second, SSA withholds OASI benefits in monthly increments, until enough benefits have been withheld to cover the AET penalty amount. For example, assume an individual aged 66 with a monthly benefit of \$1,200 earns \$1,800 dollars beyond the AET exempt amount in 1992, when the benefit reduction rate was 0.3333. This individual should receive a yearly benefit reduction of  $\$1,800 \times 0.3333 = \$600$ . SSA withholds an entire month's check, \$1,200, in order to collect the \$600. Finally, at the end of the year, SSA refunds any overwithheld benefits. In the same example, at the end of the year SSA would return \$600 in overwithheld benefits. Importantly, the DRC is not applied to future benefits in this case—less than a month's worth of benefits, \$600, was ultimately collected by SSA, after factoring overwithholding and refunds. After considering both withholding and refunds, the AET is ultimately applied at a yearly level—much in the same way that the Earned Income Tax Credit (EITC) is applied at the annual level but the receipt of the credit depends on one's withholding patterns and when the income tax return is filed.

In sum, the AET effectively acts as a kink for those earning close enough to the exempt amount, and benefit enhancement does not attenuate the marginal work disincentives associated with the AET for people NRA and older in this range of earnings. As noted in the text, however, the DRC is relevant to an individual considering earning enough to reduce her OASI benefit by at least a month's worth (*i.e.* at least  $z^* + (MB/\tau)$ ). However, this earnings level is typically a couple of thousand dollars above the AET threshold, limiting its relevance for most individuals' marginal incentives. Empirically, we find that limiting the sample to those with substantial OASI benefits—for whom this earnings level is several thousand dollars above the AET exempt amount, and for whom the notch created by the DRC is therefore less relevant—yields very similar results to those we have shown.

Note that in our empirical estimation, the region we "dummy out" near the kink  $z^*$  is  $[z^* - 2800, z^* + 2800]$ . Thus, if all of the "bunchers" arrived at the kink from initial earnings levels between  $z^*$  and  $z^* + 2800$ , we could find zero bunching despite substantial actual bunching. However, the earnings densities clearly show that the polynomial does not substantially overpredict bunching in the region above the exempt amount. Moreover, the densities also show no evidence of bunching near notches in the budget set created by the DRC. We alternatively use a bandwidth of \$500 and find similar results.

Individuals younger than NRA are subject to different rules for benefit adjustment, called "actuarial adjustment." The rule for this younger group was introduced in the legislation al-

---

month's benefit enhancement would occur if she earns more than  $z^* + (3MB/\tau)$ ; and so on. Note that this creates 12 notches in the budget set, the final one at  $z^* + (12MB/\tau)$ .

lowing people younger than 65 to claim early benefits (in 1956 for women and in 1961 for men). For those younger than NRA, future benefits are reduced 5/9 of 1 percent for each month under age 65 in which an individual claims benefits (Social Security, 2012, Table 2.A.20). SSA has interpreted this language to mean that if a beneficiary has any income withheld under the AET in a given month, then she receives a full benefit enhancement for that month. This creates a notch at the exempt amount—a discontinuous increase in future benefits when moving from just under the exempt amount to just over it—creating incentives to bunch just above the exempt amount in order to receive the monthly (or yearly) benefit adjustment. We find no evidence for this kind of behavior; in fact, as we document in Gelber, Jones, and Sacks (2013), people tend to have earnings *just below* the exempt amount, exactly opposite the behavior we would expect if people were responding to the notch just described. Because benefit adjustment occurs at exactly the same earnings level that the AET begins to apply for those under NRA, we focus on the group NRA and above.

Formally, the number of months worth of benefit enhancement enjoyed by OASI recipients is therefore  $\text{floor}(\tau \cdot (z - z^*)/MB)$  for those NRA and above, and  $\text{ceiling}(\tau \cdot (z - z^*)/MB)$  for those below NRA.

The AET applies to an individual’s earnings; spouses’ earnings do not count in the earnings total to which the AET is applied. For a retired worker beneficiary whose spouse collects spousal benefits, the AET reduces the family’s OASI benefit by the amounts we have described. The family benefit is also reduced when the spouse (separately) earns more than the AET threshold. For a retired worker beneficiary whose spouse is collecting benefits on his or her own earnings record, the AET reduces an individual’s benefits by the amounts described while not affecting the spouse’s benefits. Thus, following previous literature (e.g. Friedberg 1998, 2000), we model the AET as creating the MTRs associated with the BRRs described, because the AET reduces family benefits by these amounts (all else equal). Our data do not contain the information necessary to link spouses (except when one spouse is claiming OASI benefits on the other spouse’s record).

## B Appendix: Social Security Data

Our data are a one percent extract of the Social Security Master Earnings File (MEF), which is described more extensively in Song and Manchester (2007). The MEF is a longitudinal history of Social Security taxable earnings for all Social Security Numbers (SSNs) in the U.S. Our data are a one percent random sample of SSNs; we randomly extract SSNs from the database and follow each of these individuals over the full time period. The AET is based on earnings as measured in this dataset. Prior to 1978, the data have information on annual FICA earnings; since 1978, the data have information on uncapped wage compensation. Before 1978, the data do not clearly distinguish between earnings from self-employment and non-self-employment earnings, but we are able to distinguish them in the data starting in 1978. The data also contain information on date of birth, date of death, and sex.

We supplement the MEF with information from the Master Beneficiary Record (MBR) file, which contains information on the day, month, and year that people began to claim Social Security (and other variables). The majority of workers excluded from OASDI coverage are in four main categories: (1) federal civilian employees hired before January 1, 1984; (2) agricultural workers and domestic workers whose earnings do not meet certain minimum requirements; (3)

individuals with very low net earnings from self-employment (generally less than \$400 per year); and (4) employees of several state and local governments. However, civil service and other government workers are covered by Medicare and are therefore present in the MBR.

In choosing our main sample, we take into account a number of considerations. It is desirable to show a constant sample in making comparisons of earnings densities. Meanwhile, the AET only affects people who claim OASI, and thus we wish to focus on claimants. However, many individuals claim OASI at ages over the Early Entitlement Age (62), implying that they have not claimed at younger ages but have claimed by older ages. This implies that to investigate a constant sample, we cannot simply limit the sample to claimants at each age (because many people move from not claiming to claiming). Moreover, the claiming decision is endogenous and may depend on people’s awareness of the AET and their anticipated reaction to it. To balance these considerations, our main sample at each age and year consists of individuals who ultimately claim in the year they turn 65 or earlier. We show that the results are robust to other sample definitions. Because we focus on the intensive margin response (consistent with Saez (2010) and subsequent papers on bunching), we further limit the sample to observations with positive earnings in our main analysis.

Information on AET parameters is from table 2.A.20 and 2.A.29 of the Annual Statistical Supplement to the Social Security Bulletin. Friedberg (1998, 2000) provides a thorough description of these rules. All dollar amounts are deflated to 2010 dollars using the CPI-U.

The standard deviation is large because of very rare aberrant large values of earnings (as documented in Utendorf 2001/2); these do not affect our estimates in the figures or tables because they are far above the AET exempt amount, and they affect mean earnings far less than they affect the standard deviation. The results are robust to winsorizing.

## C Appendix: Longitudinal Employer Household Dynamics

We use the Longitudinal Employer Household Dynamics (LEHD) dataset, which contains wage data available from state-level unemployment insurance (UI) programs. These data measure uncapped quarterly earnings for employees covered by state unemployment insurance systems, estimated to cover over 95 percent of private sector employment. Although coverage laws vary slightly from state to state, UI programs do not cover federal employees, the self-employed, and many agricultural workers, domestic workers, churches, nonprofits, and state and local government employees. We examine a 20 percent random sample of the original LEHD file, as this was the largest amount of data that our available server space could handle.

These administrative earnings records are linked across quarters to create individual work histories. In addition to earnings, information on gender and date of birth are available. The data on employees are linked to data on firms. Each firm at which an individual works in a given quarter is identified through a firm identifier. We consider an employee to have changed employers from year  $t$  to year  $t + 1$  if at least one of the federal employer IDs at which the employee works in year  $t$  is different in year  $t + 1$ . However, when the individual works at one or more employer in year  $t$  and does not work at any employer in year  $t + 1$ , we drop this individual from the sample. The results are similar when we treat these individuals as if they changed employers.

We select data from 1990-1999. During this period, the AET explicit marginal tax rate was constant. 1990-1999 also represent natural years to investigate because large sample sizes are not available in the LEHD prior to 1990. When we include other years and age groups in the LEHD sample, we find similar results to those reported here. Note that the population we investigate is not constant over this period, because (among other reasons) an increasingly broad set of states is included in the LEHD over time. Data are available on 13 states in 1990, climbing to 28 states by 1999. In a given quarter, we include in our sample all states whose data are available. Holding the sample constant yields very similar results.

## D Appendix: Model of Earnings Response

### D.1 Baseline Model

We start with a baseline, frictionless model of earnings, following Saez (2010). We briefly sketch the key features of this model for comparison to our model with a fixed cost of adjustment. In Saez (2010), individuals maximize utility over consumption,  $c$ , and costly earnings,  $z$ :

$$u(c, z; n)$$

Heterogeneity is parameterized by an "ability" parameter  $n$ , which is distributed according to the smooth cdf  $F(\cdot)$ . Individuals maximize utility subject to the following budget constraint:

$$c = (1 - \tau)z + R$$

where  $R$  is virtual income. This leads to the first order condition:

$$-\frac{u_z(c, z; n)}{u_c(c, z; n)} = (1 - \tau),$$

which implicitly defines an earnings supply function:  $z(1 - \tau, R, n)$ .

When necessary, we will use a quasi-linear and iso-elastic utility function:

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

Under this assumption, the first order condition simplifies to:

$$(1 - \tau) - \left(\frac{z}{n}\right)^{\frac{1}{\varepsilon}} = 0,$$

which implies this earnings supply function:

$$z = n(1 - \tau)^\varepsilon. \tag{D.1}$$



## D.2 Linear Tax Schedule

Consider first a linear tax schedule with a constant marginal tax rate  $\tau_0$ . Observe that with a smooth distribution of skills  $n$ , we have a smooth distribution of earnings that is monotonic in skill, provided we make the typical Spence-Mirrlees assumption. Let  $H_0(\cdot)$  denote the cumulative distribution function (CDF) of earnings under the constant marginal tax rate, and let  $h_0(\cdot) = H_0'(\cdot)$  denote the density of this distribution. Under quasilinear utility, we have:

$$H_0(z) = F\left(\frac{z}{(1-\tau_0)^\varepsilon}\right).$$

Define  $H_1(\cdot)$  and  $h_1(\cdot)$  as the smooth CDF and density of earnings under a higher, constant marginal tax rate  $\tau_1$ ;  $H_1$  is defined similarly as a function of  $\tau_1$ .

## D.3 Kinked Tax Schedule

Now consider a piecewise linear tax schedule with a convex kink: the marginal tax rate below earnings level  $z^*$  is  $\tau_0$ , and the marginal tax rate above  $z^*$  is  $\tau_1 > \tau_0$ . Given the tax schedule, individuals bunch at the kink point  $z^*$ ; as explained in Saez (2010), the realized density in earnings has an excess mass at  $z^*$ . Denote the realized distribution of earnings once the kink has been introduced at  $z^*$  as  $H(\cdot)$ :

$$H(z) = \begin{cases} H_0(z) & \text{if } z < z^* \\ H_1(z) & \text{if } z \geq z^* \end{cases}$$

Denote the density of this realized distribution as  $h(\cdot) = H'(\cdot)$ . In general there is now a discrete jump in the earnings density at  $z^*$ :

$$h(z) = \begin{cases} h_0(z) & \text{if } z < z^* \\ h_1(z) & \text{if } z > z^* \end{cases}$$

The share of people who relocate to the kink is:

$$B = \int_{z^*}^{z^* + \Delta z^*} h_0(\zeta) d\zeta$$

These "bunchers" are those whose *ex ante* earnings lie in the range  $[z^*, z^* + \Delta z^*]$ , who are induced to locate at the kink by the rise in the MTR above the kink point. For relatively small changes in the tax rate, we can relate the elasticity of earnings with respect to the net-of-tax rate to the earnings change  $\Delta z^*$  for the individual with the highest *ex ante* earnings who bunches *ex post*:

$$\varepsilon = \frac{\Delta z^*/z^*}{d\tau_1/(1-\tau_0)}$$

where  $d\tau_1 = \tau_1 - \tau_0$ .<sup>70</sup>

---

<sup>70</sup>This formula holds if there is a single elasticity  $\varepsilon$  in the population. We investigate cases with heterogeneity below.

## D.4 Graphical Exposition

Appendix Figure E.1 depicts a setting in which a kinked budget set is introduced. In Panels A and B, the x-axis shows before-tax-and-transfer income,  $z$ , and the y-axis shows after-tax-and-transfer consumption,  $z - T(z)$ . Consider first a linear tax (Panel A) at a rate of  $\tau$ . An individual optimally locates at a point of tangency, where the marginal rate of substitution (MRS) between earnings and consumption equals the net-of-tax rate,  $1 - \tau$ . The figure shows indifference curves and earnings levels for low- and high-earning agents (labeled  $L$  and  $H$ , respectively). The low earner has an earnings level of  $z^*$ , while the high earner receives  $z^* + \Delta z$ .

Suppose the AET is introduced (on top of pre-existing taxes), so that the marginal net-of-tax rate decreases to  $1 - \tau - d\tau$  for earnings above a threshold  $z^*$ . For small  $d\tau$ , individuals earning in the neighborhood above  $z^*$  will reduce their earnings. If ability is smoothly distributed, a range of individuals will locate exactly at  $z^*$ , due to the discontinuous jump in the marginal net-of-tax rate at  $z^*$ . In Panels A and B, individual  $L$  has the lowest *ex ante* earnings among those who bunch at  $z^*$ , individual  $H$  has the highest *ex ante* earnings in this group, and all others previously earning between  $z^*$  and  $z^* + \Delta z$  also bunch at  $z^*$ . Those with *ex ante* earnings higher than  $z^* + \Delta z$  reduce their earnings to a level greater than  $z^*$ .

Panels C and D of Figure E.1 depict densities of earnings we would expect to observe in the absence and presence of the AET, respectively. The x-axis shows before-tax earnings,  $z$ , and the y-axis measures the density of earnings. In Panel C, the density is continuous at  $z^*$ , reflecting a smooth distribution of ability. The blue region represents the set of individuals who bunch at  $z^*$  in the presence of the AET, *i.e.* those earning in  $[z^*, z^* + \Delta z]$  in the absence of the AET. Panel D shows that once the AET is introduced, these individuals locate in the neighborhood of  $z^*$ . However, rather than depicting a mass point exactly at  $z^*$ , we have shown bunching in the region at and surrounding  $z^*$ , reflecting the fact that individuals often cannot bunch exactly at the kink point (as discussed, for example, in Saez, 2010).

## D.5 Fixed Adjustment Costs

We now extend the model to include a fixed cost of adjusting earnings. We assume that the adjustment cost reflects a disutility of  $\phi^*$  of increasing or decreasing earnings from some initial earnings level. We begin by analyzing the response to a change in the marginal tax rate from  $\tau_0$  to  $\tau_1$ , where the tax schedule is linear in both cases, in order to build intuition for the case with a kinked budget set. We assume that following a change in tax rates from  $\tau_0$  to  $\tau_1$ , the gain (absent adjustment costs) to reoptimizing is increasing in  $n$ . In general, this requires that the size of the optimal earnings adjustment increases in  $n$  at a rate faster than the decrease in the marginal utility of consumption.<sup>71</sup> This is true, for example, if utility is quasilinear.

If the gain in utility is monotonically increasing in initial earnings, and the cost of adjustment is fixed, there exists a unique level of initial earnings at which the agent is indifferent between

<sup>71</sup>To see this, note that the utility gain from reoptimizing is  $u((1 - \tau_1)z_1 + R_1, z_1; n) - u((1 - \tau_1)z_0 + R_1, z_0; n) \approx u_c \cdot (1 - \tau_1)[z_1 - z_0] + u_z \cdot [z_1 - z_0] = u_c \cdot (\tau_1 - \tau_0)[z_0 - z_1]$ , where in the first expression, we have used a first-order approximation for utility at  $((1 - \tau_0)z_0 + R_0, z_0)$  and in the second expression we have used the first order condition  $u_z = -u_c(1 - \tau_0)$ . The gain in utility is approximately equal to an expression that depends on the marginal utility of consumption, the change in tax rates, and the size of the earnings adjustment. The first term,  $u_c$ , is decreasing as  $n$  (and therefore initial earnings  $z_0$ ) increases. Thus, in order for the gain in utility to be increasing in  $n$ , we need the size of earnings adjustment  $[z_0 - z_1]$  to increase at a rate that dominates.

adjusting and staying at the initial earnings level. We formally state the implications in the following result:

**Remark 1 (Linear Tax Change and Adjustment Costs)**

*After a change in linear tax rates from  $\tau_0$  to  $\tau_1$ , if there is a constant adjustment cost of  $\phi^*$  and the size of the optimal earnings adjustment increases in  $n$  at a rate faster than the decrease in the marginal utility of consumption, then:*

1. *There is a unique threshold of initial earnings,  $z_{0,\phi}$ , above which all individuals will adjust their earnings in response to the tax change. Those initially locating below the threshold will not adjust.*
2. *The threshold level of earnings satisfies the following identity:*

$$u((1 - \tau_1) z_{1,\phi} + R_1, z_{1,\phi}) - u((1 - \tau_1) z_{0,\phi} + R_1, z_{0,\phi}) \equiv \phi^*$$

*where  $z_{1,\phi}$  is the ex post earnings level of the individual who initially locates at  $z_{0,\phi}$ . In other words, at the threshold level, the gain in  $u$  from adjusting earnings is exactly equal to the adjustment cost  $\phi^*$ .*

3. *In the case of quasilinear utility, the threshold level of earnings is:*

$$z_{0,\phi} = \frac{\phi^*}{\alpha(\varepsilon, \tau_0, \tau_1)}$$

*where*

$$\alpha(\varepsilon, \tau_0, \tau_1) \equiv \frac{1 - \tau_1}{1 + \varepsilon} \left[ \left( \frac{1 - \tau_1}{1 - \tau_0} \right)^\varepsilon - 1 + \varepsilon \left( \frac{\tau_1 - \tau_0}{1 - \tau_1} \right) \right].$$

4. *The ex post distribution of earnings is:*

$$H(z) = \begin{cases} H_0(z) & \text{if } z < z_{1,\phi} \\ H_0(z) + H_1(z) - H_0(z_{0,\phi}) & \text{if } z \in [z_{1,\phi}, z_{0,\phi}] \\ H_1(z) & \text{if } z > z_{0,\phi} \end{cases}$$

$$h(z) = \begin{cases} h_0(z) & \text{if } z < z_{1,\phi} \\ h_0(z) + h_1(z) & \text{if } z \in [z_{1,\phi}, z_{0,\phi}] \\ h_1(z) & \text{if } z > z_{0,\phi} \end{cases}$$

*where  $H_0(\cdot)$  and  $H_1(\cdot)$  are the CDFs of earnings in the presence of linear tax rates  $\tau_0$  and  $\tau_1$ , respectively.*

Next, consider choices in the presence of adjustment costs on a budget set with a convex kink. Consider again an initial linear tax schedule with marginal tax rate  $\tau_0$ . Now, introduce a higher MTR  $\tau_1 > \tau_0$  for earnings above  $z^*$ . We again assume that the gain to reoptimizing is increasing in initial earnings over the range  $[z^*, z^* + \Delta z^*]$ . Using the same logic as above—the gain in utility is monotonically increasing in initial earnings, and the cost of adjustment is fixed—there exists a unique level of initial earnings at which the agent is indifferent between adjusting and staying at the initial earnings level. Thus, we have the following result:

**Remark 2 (Non-Linear Tax and Adjustment Costs)**

When a kink is introduced in the budget set (i.e. a jump in marginal tax rates from  $\tau_0$  below  $z^*$  to  $\tau_1$  above  $z^*$ ), there is a fixed adjustment cost of  $\phi^*$ , and  $z^* \geq z_{1,\phi}$ <sup>72</sup>, then:

1. Individuals with initial earnings below a unique threshold  $\underline{z}$  do not adjust their earnings.
2. The threshold level of earnings is implicitly defined by the following:

$$\begin{aligned} u((1 - \tau_1) z^* + R_1, z^*) - u((1 - \tau_1) \underline{z} + R_1, \underline{z}) &\equiv \phi^* \\ z^* \leq \underline{z} \leq z^* + \Delta z^*. \end{aligned} \tag{D.2}$$

3. Individuals with initial earnings in  $[\underline{z}, z^* + \Delta z^*]$  bunch at the kink point  $z^*$ .
4. Individuals with initial earnings above  $z^* + \Delta z^*$  reduce their earnings to a new level of earnings higher than  $z^*$ .
5. The ex post distribution of earnings is:

$$\begin{aligned} H(z) &= \begin{cases} H_0(z) & \text{if } z < z^* \\ H_0(z) + H_1(z) - H_0(\underline{z}) & \text{if } z \in [z^*, \underline{z}] \\ H_1(z) & \text{if } z > \underline{z} \end{cases} \\ h(z) &= \begin{cases} h_0(z) & \text{if } z < z^* \\ h_0(z) + h_1(z) & \text{if } z \in [z^*, \underline{z}] \\ h_1(z) & \text{if } z > \underline{z} \end{cases} \end{aligned}$$

6. Excess bunching at  $z^*$  is given by:

$$B = \int_{\underline{z}}^{z^* + \Delta z^*} h_0(\zeta) d\zeta$$

If the kink point  $z^*$  is lower than  $z_{1,\phi}$ , then:

1. Individuals only adjust their earnings if their initial earnings level is above the threshold  $z_{0,\phi}$ .
2. There is no bunching at  $z^*$ .
3. The ex post distribution of earnings is the same as in the case of a change in a linear tax rate from  $\tau_0$  to  $\tau_1$  described in Remark (1).

---

<sup>72</sup> $z_{1,\phi}$  is again the ex post level of earnings for the individual who initially locates at  $z_{0,\phi}$  – where  $z_{0,\phi}$  is the initial earnings level over which individuals adjust their earnings defined above in Remark (1). Note that  $\underline{z}$  denotes this threshold in the non-linear budget set case, whereas  $z_{0,\phi}$  denotes this threshold in the linear budget set case.

## D.6 Derivation of Closed-Form Solution for Elasticity and Adjustment Cost (Comparative Static Approach)

As we discuss in Remark (2) above and in Section 7 in the text, the amount of bunching in the presence of a fixed adjustment cost is equal to the integral of the initial earnings density over the range  $[\underline{z}, z^* + \Delta z^*]$ :

$$B(\boldsymbol{\tau}, z^*; \varepsilon, \phi^*) = \int_{\underline{z}}^{z^* + \Delta z^*} h(\zeta) d\zeta, \quad (\text{D.3})$$

where  $\boldsymbol{\tau} \equiv (\tau_0, \tau_1)$ . If the density is locally uniform, the integral in (D.3) is:

$$B(\boldsymbol{\tau}, z^*; \varepsilon, \phi^*) \approx h(\underline{z})(z^* + \Delta z^* - \underline{z}) \quad (\text{D.4})$$

Taking a first-order Taylor approximation of  $u((1 - \tau_1)\underline{z} + R_1, \underline{z}, \underline{n})$  and  $u((1 - \tau_1)z^* + R_1, z^*, \underline{n})$  at  $((1 - \tau_0)\underline{z} + R_0, \underline{z}, \underline{n})$ , and using the first order condition for initial earnings,  $(1 - \tau_0)u_c = -u_z$ , we have from (D.2):

$$\begin{aligned} \phi^* &\approx u_c \cdot (1 - \tau_1) [z^* - \underline{z}] + u_z \cdot (z^* - \underline{z}) \\ \Rightarrow \underline{z} &\approx z^* + \frac{\phi^*/u_c}{(\tau_1 - \tau_0)} \\ &= z^* + \frac{\phi}{d\tau_1}, \end{aligned}$$

where  $d\tau_1 = \tau_1 - \tau_0$  and  $\phi = \phi^*/u_c$  is the dollar equivalent of the disutility associated with adjusting earnings. Substituting this expression for  $\underline{z}$  into (D.4), we have

$$B(\boldsymbol{\tau}, z^*; \varepsilon, \phi) = h(\underline{z})(\Delta z - \phi/d\tau_1),$$

where bunching now depends on the dollar-denominated cost of adjusting, rather than the utility cost. Finally, for small  $d\tau$ ,  $\Delta z$  is small and  $h(\underline{z}) \approx h(z^* + \Delta z) \approx h(z^*)$ . Let  $b \equiv B/h(z^*)$ , and note that  $\Delta z^* = z^*(d\tau_1/(1 - \tau_0))\varepsilon$ . The excess mass at the kink can now be expressed as a linear function of the parameters:

$$b(\boldsymbol{\tau}, z^*; \varepsilon, \phi) = \varepsilon \left( z^* \frac{d\tau_1}{1 - \tau_0} \right) - \phi \left( \frac{1}{d\tau_1} \right). \quad (\text{D.5})$$

### D.6.1 Derivation of Formula for Bunching with Heterogeneity

We derive the formula for bunching  $B$  in the presence of heterogeneity under the Comparative Static approach as follows:

$$\begin{aligned}
B &= \iiint_{\underline{z}}^{z^* + \Delta z^*} \tilde{h}(\zeta, \epsilon, \varphi) d\zeta d\epsilon d\varphi \\
&\approx \iint [z^* + \Delta z^* - \underline{z}] \tilde{h}(z^*, \epsilon, \varphi) d\epsilon d\varphi \\
&\approx \iint \left[ \epsilon \left( z^* \frac{d\tau_1}{1 - \tau_0} \right) - \varphi \left( \frac{1}{d\tau_1} \right) \right] \tilde{h}(z^*, \epsilon, \varphi) d\epsilon d\varphi \\
&= h(z^*) \cdot \left[ \left( \iint \epsilon \frac{\tilde{h}(z^*, \epsilon, \varphi)}{h(z^*)} d\epsilon d\varphi \right) \left( z^* \frac{d\tau_1}{1 - \tau_0} \right) - \left( \iint \varphi \frac{\tilde{h}(z^*, \epsilon, \varphi)}{h(z^*)} d\epsilon d\varphi \right) \left( \frac{1}{d\tau_1} \right) \right] \\
&= h(z^*) \cdot \left[ \bar{\epsilon} \left( z^* \frac{d\tau_1}{1 - \tau_0} \right) - \bar{\phi} \left( \frac{1}{d\tau_1} \right) \right],
\end{aligned}$$

where we have used the assumption of constant  $\tilde{h}(\cdot)$  and the approximations for  $\Delta z^*$  and  $\underline{z}$  in Section 7. Here  $h(z^*) = \iint \tilde{h}(z^*, \epsilon, \varphi) d\epsilon d\varphi$ , and  $\bar{\epsilon}$  and  $\bar{\phi}$  are the average elasticity and adjustment cost, respectively.

## D.6.2 Linear Adjustment Costs

We now introduce an adjustment cost that increases linearly in the size of the adjustment. Assume that given an initial level of earnings  $z_0$ , agents must pay a cost of  $\phi^* \cdot |z - z_0|$  when they change their earnings to a new level  $z$ . Utility  $\tilde{u}$  at the new earnings level can be represented as:

$$\tilde{u}(c, z; n, z_0) = u(c, z; n) - \phi^* \cdot |z - z_0|.$$

The first order condition for earnings can be characterized as:

$$\begin{aligned}
-\frac{u_z(c, z; n)}{u_c(c, z; n)} &= (1 - \tau - \phi^*/\lambda^* \cdot \text{sgn}(z - z_0)) \\
&= \begin{cases} (1 - \tau - \phi) & \text{if } z > z_0 \\ (1 - \tau + \phi) & \text{if } z < z_0 \end{cases},
\end{aligned}$$

where  $\lambda^* = u_c(c^*, z^*; n)$  is the Lagrange multiplier and  $\phi = \phi^*/\lambda^*$  is the dollar equivalent of the linear adjustment cost  $\phi^*$ .

The individual chooses earnings as if he faces an effective marginal tax rate of  $\tilde{\tau} = \tau + \phi \cdot \text{sgn}(z - z_0)$ . It follows that our predictions about earnings adjustment are similar to our previous predictions, except that the effective marginal tax rate  $\tilde{\tau}$  appears, rather than  $\tau$ . Thus, we can solve for the elasticity of earnings as a function of the change in earnings  $\Delta z^*$  due to introduction of a kink in the tax schedule and the jump in marginal tax rate  $d\tau_1$ :

$$\begin{aligned}
\varepsilon &= \frac{\Delta z^*/z^*}{d\tilde{\tau}_1/(1 - \tilde{\tau}_0)} \\
&= \frac{\Delta z^*/z^*}{(d\tau_1 - 2\phi)/(1 - \tau_0 - \phi)}.
\end{aligned}$$

Since the right-hand side is increasing in  $\phi$ , the estimate of the elasticity increases as the linear

adjustment cost increases. This makes intuitive sense: the adjustment cost attenuates bunching, so holding constant the level of bunching, the elasticity must be higher as the adjustment cost increases.

Now assume that when an individual adjusts his earnings, he incurs a linear adjustment cost  $\phi^{*L}$  for every unit of change in earnings, as well as a fixed cost  $\phi^{*F}$  associated with any change in earnings. Consider again bunching at  $z^*$ , with a tax rate jump of  $d\tau_1 = \tau_1 - \tau_0$  at earnings level  $z^*$ . We have the following set of expressions for excess mass:

$$\begin{aligned} B &= \int_{\underline{z}}^{z^* + \Delta z^*} h(\zeta) d\zeta \\ \varepsilon &= \frac{\Delta z^* / z^*}{(d\tau_1 - 2\phi^L) / (1 - \tau_0 - \phi^L)} \\ \phi^{*F} + \phi^{*L} \cdot (z - z^*) &= u\left((1 - \tau_1)z^* + R', z^*; \underline{n}\right) - u\left((1 - \tau_1)z + R', z; \underline{n}\right). \end{aligned}$$

Using a left rectangle approximation for the integral, we have:

$$\begin{aligned} b &\equiv B/h(z^*) \\ &= z^* + \Delta z^* - \underline{z} \\ &= z^* \left( \frac{d\tau_1 - 2\phi^L}{1 - \tau_0 - \phi^L} \varepsilon + 1 \right) - \underline{z}. \end{aligned}$$

We can further apply an approximation for  $\underline{z}$  similar to the approximation we used in Section 7, *i.e.*  $\underline{z} = z^* + \phi^F / (d\tau_1 - 2\phi^L)$ . Thus, the expression for bunching can be simplified to:

$$b = \varepsilon \left( z^* \frac{d\tau_1 - 2\phi^L}{1 - \tau_0 - \phi^L} \right) - \frac{\phi^F}{(d\tau_1 - \phi^L)},$$

where  $(\phi^F, \phi^L) = (\phi^{*F}/\lambda^*, \phi^{*L}/\lambda^*)$ . In this case, we need at least three kinks to separately identify  $(\varepsilon, \phi^F, \phi^L)$ . Because we do not examine a setting in which one can compare bunching under three different positive tax rates, we are not able to estimate these parameters using data (or to fruitfully estimate the parameters in the non-linear case since we do not have a credible source of variation to identify them). Note that a similar formula can be derived in the setting of the Sharp Change approach in Section (D.7).<sup>73</sup>

<sup>73</sup>In particular, the same logic in this section can be applied to the approximation for bunching at the end of Section (D.7) to show that:

$$\tilde{b}_2 = \varepsilon \left( \bar{z}_2 \frac{d\tau_2 - 2\phi^L}{1 - \tau_0 - \phi^L} \right) - (\bar{z}_2 - z^*) - \frac{\phi^F}{d\tau_1 - \phi^L}.$$

## D.7 Derivation of Formula for Bunching with a Pre-Existing Kink (Sharp Change Approach)

The Comparative Static approach abstracts from a key feature of our empirical setting. In particular, the Comparative Static approach models the transition from a budget set with no kink to one with a kink. This approach facilitates our basic intuition and provides a transparent bridge between our approach and existing bunching methods in the presence of a kink. However, in our context, we conduct analysis using data just before and just after the benefit reduction rate was decreased (in 1990 from 50 percent to 33.33 percent for 66-69 year olds; or from 33.33 percent to zero for this group in 2000; or from 33.33 percent to zero for 70-year-olds in 1990-1999). These changes involve moving from an initial state with a kink to a new state with a smaller kink. In a frictionless model, the distinction is immaterial. However, as we will show, this matters in the presence of a fixed adjustment cost. In particular, when the kink becomes more muted, the change in bunching will be attenuated due to the fixed adjustment cost.

We will assume that in the initial state, bunching is characterized as in Remark (2). Let the initial kink,  $K_1$ , be characterized by a lower marginal tax rate,  $\tau_0$ , to the left of  $z^*$ , and a higher marginal tax rate,  $\tau_1$ , to the right of  $z^*$ . The initial level of bunching is:

$$B_1 = \int_{z_1}^{z^* + \Delta z_1^*} h(\zeta) d\zeta$$

Now, consider a change in the kink to  $K_2$ , which retains the lower marginal tax rate  $\tau_0$  to the left of  $z^*$  but reduces the marginal tax rate to the right of  $z^*$  to  $\tau_2 < \tau_1$ . Had we begun with no kink and introduced  $K_2$ , bunching would have been:

$$B_2 = \int_{z_2}^{z^* + \Delta z_2^*} h(\zeta) d\zeta$$

Note that relative to  $K_1$ ,  $K_2$  provides a weaker incentive to bunch, *when starting from a baseline tax schedule with no kink*. Formally, we have  $z_2 \geq z_1$ ,  $\Delta z_2^* < \Delta z_1^*$  and  $B_2 \leq B_1$ .

However, in characterizing bunching when moving from  $K_1$  to  $K_2$ , individuals may be separated into several groups based on their optimal level of earnings  $z_0$  in the absence of a kink. First, there are individuals with  $z_0 < z^*$ . They will locate to the left of the kink under both  $K_1$  and  $K_2$ .

Second, we have individuals with  $z^* < z_0 \leq z_1$  (area *i* in Figure 8). These individuals would optimize in the presence of  $K_1$  by moving to  $z^*$ , were it not for the adjustment cost. Now, with a smaller kink  $K_2$ , these individuals continue to remain at the initial earnings level  $z_0 > z^*$ , as the utility gain to reoptimizing to  $z^*$  is even smaller than it was under  $K_1$ .

Third, we have those with  $z_1 < z_0 \leq z_2$  (area *ii* in Figure 8). When moving from no kink to  $K_1$ , these individuals locate at the kink,  $z^*$ . If the budget set had hypothetically transitioned from no kink to  $K_2$ , these individuals would have chosen to remain at  $z_0$ , due to the fixed adjustment cost. However, when moving from  $K_1$  to  $K_2$ , these agents remain at the kink  $z^*$ , because  $z^*$  is the optimal frictionless earnings level under  $K_2$ . Since these individuals already find themselves there and face a cost of adjusting their earnings, they continue to bunch at  $z^*$  when moving from  $K_1$  to  $K_2$ , even though they would not have bunched if we had started with no kink and moved directly to  $K_2$ .



Fourth, we have agents with  $\bar{z}_2 < z_0 \leq z^* + \Delta z_2$  (area *iii* in Figure 8). These individuals bunch at  $z^*$  when moving from no kink to either  $K_1$  or  $K_2$ . Thus, they remain bunching at  $z^*$  when moving from  $K_1$  to  $K_2$ .

Fifth, we have agents with  $z^* + \Delta z_2 < z_0 \leq z^* + \Delta z_1$  (areas *iv* and *v* in Figure 8). When starting from a budget set with no kink, these agents bunch under  $K_1$ , but not under  $K_2$ . Starting instead from  $K_1$ , they must choose between remaining at the kink  $z^*$  or moving to the frictionless optimum under  $K_2$ ,  $z_2 > z^*$ . We know that at least some of these individuals will remain bunching. To see this, consider an individual with earnings under no kink  $z_0 = z^* + \Delta z_2 + \delta_0$ . For small enough  $\delta_0$  optimal earnings under  $K_1$  is  $z^*$ , and optimal earnings under  $K_2$  tends to  $z^*$  as  $\delta_0$  tends to zero. Likewise, the net utility gain from relocating from  $z^*$  to  $z_2$  under  $K_2$  tends to zero as  $\delta_0$  tends to zero. Therefore, this individual will remain at  $z^*$  when moving from  $K_1$  to  $K_2$  for small enough  $\delta_0$ . In Figure 8, area *iv* shows those with initial earnings  $z^* + \Delta z_2 < z_0 < \bar{z}_0$  who remain bunching at the kink when transitioning from  $K_1$  to  $K_2$ . Area *v* shows those with initial earnings  $\bar{z}_0 < z_0 < z^* + \Delta z_1$ , who "debunch" from the kink when moving from  $K_1$  to  $K_2$ .

When reoptimizing is beneficial for at least some agents in this final group, we will have a reduction in bunching when transitioning from  $K_1$  to  $K_2$ .<sup>74</sup> Empirically, we observe such a reduction over time, so this is the case relevant to our setting. In this case, the marginal "de-buncher" will be defined by the following conditions:

$$\begin{aligned} -\frac{u_z(c_2, \bar{z}_2; \bar{n}_2)}{u_c(c_2, \bar{z}_2; \bar{n}_2)} &= (1 - \tau_2) \\ u((1 - \tau_2)\bar{z}_2 + R_2, \bar{z}_2; \bar{n}_2) - u((1 - \tau_2)z^* + R_2, z^*; \bar{n}_2) &\equiv \phi^* \\ -\frac{u_z(c_0, \bar{z}_0; \bar{n}_2)}{u_c(c_0, \bar{z}_0; \bar{n}_2)} &= (1 - \tau_0) \\ \bar{z}_0 &\leq z^* + \Delta z_1^* \end{aligned}$$

In words, the first line indicates that  $\bar{z}_2 > z^*$  is the optimal, frictionless level of earnings chosen by the top buncher in the presence of  $K_2$ . The second line requires that when facing  $K_2$ , this agent is indifferent between remaining at  $z^*$  and moving to  $\bar{z}_2$  through paying the adjustment cost. The third line defines  $\bar{z}_0$  as the initial level of earnings that this individual chooses when facing a constant marginal tax rate of  $\tau_0$  and no kink. The fourth line requires that this individual is initially bunching at  $z^*$  in response to  $K_1$ . If this last inequality is binding, then when moving from  $K_1$  to  $K_2$ , none of the bunchers "debunch" and the fraction bunching is unchanged. In that case, we have no variation available to identify both  $\varepsilon$  and  $\phi$ . As noted above, empirically we do observe that excess bunching falls around 1990 when the benefit reduction rate falls from 50 percent to 33.33 percent. Thus, we restrict attention to the case in which  $\bar{z}_0 < z^* + \Delta z_1$ .

Bunching at  $K_2$  following  $K_1$ , when  $\bar{z}_0 < z^* + \Delta z_1$ , can therefore be expressed as:

$$\tilde{B}_2 = \int_{\bar{z}_1}^{\bar{z}_0} h(\zeta) d\zeta$$

We can again solve this system of equations for  $\phi^*$  and  $\varepsilon$ . Note that  $\varepsilon$  is still identified by the

---

<sup>74</sup>In certain cases, it is possible that reoptimizing away from the kink is not optimal for anyone in the initial earnings range  $[z^* + \Delta z_2, z^* + \Delta z_1]$ . In that case, there is no change in bunching when moving from  $K_1$  to  $K_2$ .

potential adjustment of the top-most buncher:

$$\varepsilon = \frac{\bar{z}_0 - \bar{z}_2 (1 - \tau_0)}{\bar{z}_2 d\tau_2}$$

Note that when moving from  $K_1$  to  $K_2$ , the change in bunching is smaller than it would be if we had started with steady state bunching at  $K_1$  following no kink ( $B_1$ ) and then moved to steady state bunching  $K_2$  following no kink ( $B_2$ ). That is:

$$\begin{aligned} B_1 - \tilde{B}_2 &= \int_{\underline{z}_1}^{z^* + \Delta z_1^*} h(\zeta) d\zeta - \int_{\underline{z}_1}^{\bar{z}_0} h(\zeta) d\zeta \\ &\leq \int_{\underline{z}_1}^{z^* + \Delta z_1^*} h(\zeta) d\zeta - \int_{\underline{z}_2}^{z^* + \Delta z_2^*} h(\zeta) d\zeta \\ &= B_1 - B_2, \end{aligned}$$

where the second line follows from the fact that  $\bar{z}_0 \geq z^* + \Delta z_2^*$  and  $\underline{z}_1 \leq \underline{z}_2$ . We can again build intuition for this result by simplifying the formula for bunching in the second period. Assuming the density is constant over the range  $[z^*, z^* + \Delta z_1^*]$ , we have:

$$\begin{aligned} \tilde{b}_2 &= \bar{z}_0 - \underline{z}_1 \\ &= (\bar{z}_0 - \bar{z}_2) + \bar{z}_2 - \underline{z}_1 \\ &= \varepsilon \left( \bar{z}_2 \frac{d\tau_2}{1 - \tau_0} \right) + \bar{z}_2 - \underline{z}_1 \\ &= \varepsilon \left( \bar{z}_2 \frac{d\tau_2}{1 - \tau_0} \right) + (\bar{z}_2 - z^*) - \frac{\phi}{d\tau_1} \end{aligned}$$

where  $\tilde{b}_2 \equiv \tilde{B}_2/h(z^*)$ . On the third line, we have used the definition of the elasticity and on the fourth line, we used a first-order approximation to solve for  $\underline{z}_1$  as before in Section (D.6). Thus, we can see why bunching when moving from a larger to smaller kink is greater than would be predicted by the comparative static method. First, the term multiplying  $\varepsilon$  has a  $\bar{z}_2$  instead of a  $z^*$  and there is an additional term  $\bar{z}_2 - z^*$  – both of which increase bunching, since  $\bar{z}_2 > z^*$ . Both of these capture of the excess bunching from above, due to inertia. Finally, the third term has a  $d\tau_1$  in the denominator instead of a  $d\tau_2$ . The larger denominator increases bunching –  $d\tau_1 > d\tau_2$  – and captures the fact that there is less attenuation in bunching from below, also due to inertia.

## D.8 Estimating the Elasticity and Adjustment Cost

In this section, we describe in more detail how we use data on the amount of bunching to estimate the elasticity and adjustment cost. Let  $b = (b_1, b_2, \dots, b_K)$  be a vector of (estimated) bunching amounts normalized by the density at the kink, using the method described in Section (3). Let  $\tau = (\tau_1, \dots, \tau_K)$  be the tax schedule at each kink. The triplet  $\tau_k = (\tau_0^k, \tau_1^k, \tau_2^k)$  denotes the tax rate below ( $\tau_0^k$ ) and above ( $\tau_1^k$ ) the kink  $k$ ; when using the Sharp Change method,  $\tau_2^k$  denotes the *ex post* marginal tax rate above the kink after it has been reduced, as in Section (D.7). Let  $z^* = (z_1^*, \dots, z_K^*)$  be the earnings levels associated with each kink. To estimate  $(\varepsilon, \phi)$ ,

we seek the values of the parameters that make predicted bunching  $\hat{b}$  and actual (estimated) bunching  $b$  as close as possible on average.

Letting  $\hat{b}(\varepsilon, \phi) = (\hat{b}(\tau_1, z_1^*, \varepsilon, \phi), \dots, \hat{b}(\tau_K, z_K^*, \varepsilon, \phi))$ , our estimator is:

$$\left(\hat{\varepsilon}, \hat{\phi}\right) = \operatorname{argmin}_{(\varepsilon, \phi)} \left(\hat{b}(\varepsilon, \phi) - b\right)' W \left(\hat{b}(\varepsilon, \phi) - b\right), \quad (\text{D.6})$$

where  $W$  is a  $K \times K$  diagonal matrix whose diagonal entries are the inverse of the variances of the estimates of the  $b_k$ .

We obtain our estimates by minimizing equation (D.6) numerically. Solving this problem requires evaluating  $\hat{b}$  at each trial guess of  $(\varepsilon, \phi)$ .<sup>75</sup> Recall that in general bunching takes the form:

$$B_k(\tau_k, z_k^*; \varepsilon, \phi^*) = \int_{z_k^{lb}}^{z_k^{ub}} h(\zeta) d\zeta,$$

where  $(z_k^{lb}, z_k^{ub})$  are the ex-ante earnings levels of the lowest and highest earning bunchers, in the presence of linear tax at the lower tax rate,  $\tau_k$ . Define  $z_k^* + \Delta z_1^k$  as the *ex ante* earnings level for the highest earning buncher – in the absence of frictions – when the size of the kink is  $d\tau_1^k = \tau_1^k - \tau_0^k$ . As in the main text, we continue to assume that  $h(\cdot)$  is uniform in  $[z_k^*, z_k^* + \Delta z_1^k]$ , so that

$$b_k(\tau_k, z_k^*; \varepsilon, \phi^*) = z_k^{ub} - z_k^{lb},$$

where  $b = B/h(z_k^*)$ . The definitions of  $(z_k^{lb}, z_k^{ub})$  vary depending on the setting and are defined as follows. In the frictionless case (Saez 2010), we have:

$$\begin{aligned} z_k^{lb} &= z_k^* \\ z_k^{ub} &= z_k^* + \Delta z_1^k. \end{aligned}$$

In the presence of a fixed adjustment cost, we have under the Comparative Static approach:

$$\begin{aligned} z_k^{lb} &= \underline{z}_1^k \\ z_k^{ub} &= z_k^* + \Delta z_1^k, \end{aligned}$$

where  $\underline{z}_k$  is the ex-ante earnings of the marginal buncher from below given a fixed cost of adjustment. This is defined in the indifference condition above in equation (D.2). Finally, under the Sharp Change method, we have:

$$\begin{aligned} z_k^{lb} &= \underline{z}_1^k \\ z_k^{ub} &= \bar{z}_0^k, \end{aligned}$$

where  $\underline{z}_1^k$  is similarly the *ex ante* earnings of the marginal buncher from below (calculated using a kink with  $d\tau_1^k = \tau_1^k - \tau_0^k$ ). The *ex ante* earnings of the marginal buncher from above,  $\bar{z}_0^k$ , is defined in Section (D.7) where  $d\tau_1^k = \tau_1^k - \tau_0^k$  and  $d\tau_2^k = \tau_2^k - \tau_0^k$ .

Our estimator assumes a quasilinear utility function,  $u(c, z; n) = c - \frac{n}{1+1/\varepsilon} \left(\frac{z}{n}\right)^{1+1/\varepsilon}$ , following

<sup>75</sup>In solving problem (D.6), we impose that  $\phi \geq 0$ . When  $\phi < 0$ , every individual adjusts her earnings by at least some arbitrarily small amount, regardless of the size of  $\phi$ . This implies that  $\phi$  is not identified if it is less than zero.

Saez (2010), Chetty *et al.* (2011) and Kleven and Waseem (forthcoming). Note that because we have assumed quasilinearity,  $\phi^* = \phi$ ,  $\Delta z_1^k = z_k^* \left( \left( \frac{1-\tau_1^k}{1-\tau_0^k} \right)^\varepsilon - 1 \right)$  and  $n = z(\tau) / (1-\tau)^\varepsilon$ , where  $z(\tau)$  are the optimal, interior earnings under a linear tax of  $\tau$ . However, there typically is not a closed form solution for the  $(z_k^{lb}, z_k^{ub})$  in other cases. Instead, given  $\varepsilon$  and  $\phi$ , we find  $(z_k^{lb}, z_k^{ub})$  numerically as the solution to relevant indifference condition. For example,  $\underline{z}_1^k$  is defined implicitly by:

$$\underbrace{u((1-\tau_1^k)z_k^* + R_1^k, z_k^*; \underline{z}_1^k / (1-\tau_0^k)^\varepsilon)}_{\text{utility from adjusting to kink}} - \underbrace{u((1-\tau_1^k)\underline{z}_1^k + R_1^k, \underline{z}_1^k; \underline{z}_1^k / (1-\tau_0^k)^\varepsilon)}_{\text{utility from not adjusting}} = \phi,$$

The equation is continuously differentiable and has a unique solution. As such, Newton-type solvers are able to find  $\underline{z}_1^k$  accurately. Note that some combinations of  $\tau_k, z_k^*, \varepsilon$ , and  $\phi$  imply  $z_k^{lb} > z_k^{ub}$ . In this case, the lowest-earning adjuster does not adjust to the kink, and whenever this happens we set  $\hat{b}_k = 0$ . The predicted amount of bunching is therefore:

$$\hat{b}_k(\tau_k, z_k^*; \varepsilon, \phi) = \max(z_k^{ub} - z_k^{lb}, 0).$$

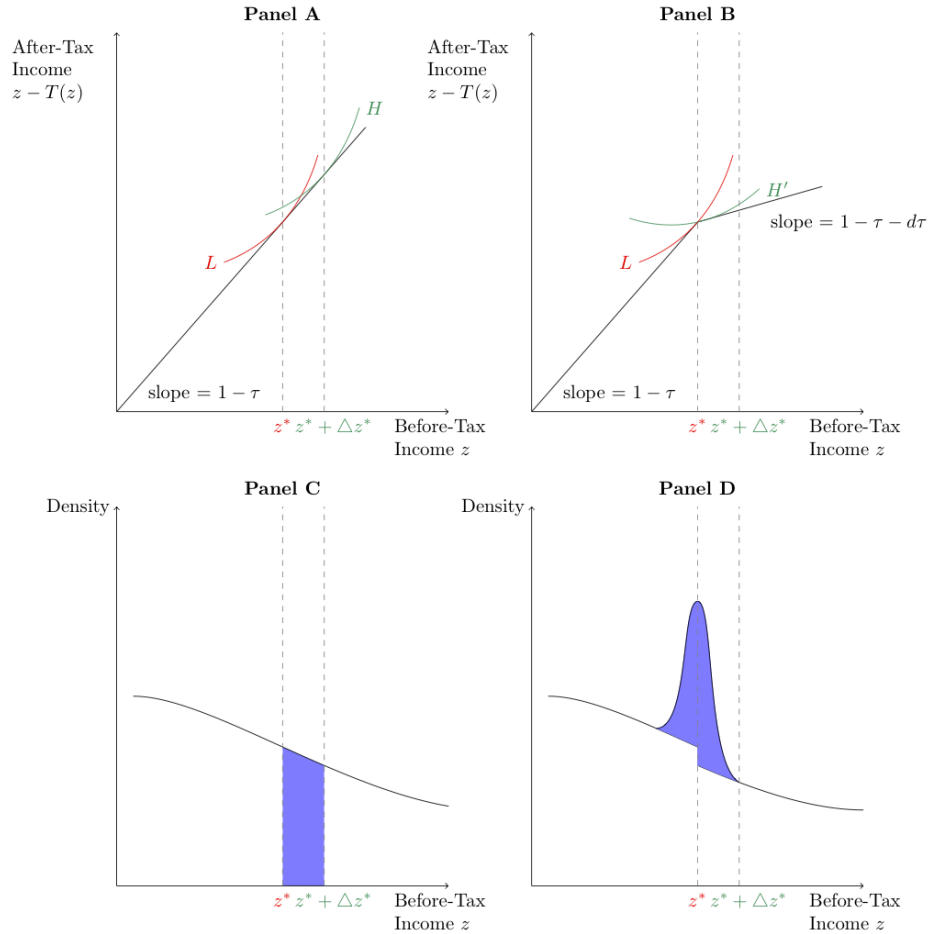
We have also shown a robustness check in our Tables in which we assume that the earnings distribution is lognormal, rather than assuming that that  $h(\cdot)$  is uniform in  $[z_k^*, z_k^* + \Delta z_1^k]$ . Specifically, we use the distribution of earnings at age 61 over 1986-1988 and 1992-1994 to estimate the parameters of a lognormal earnings distribution using maximum likelihood. (Individuals age 61 are not subject to the AET but are not far removed in age from those at retirement age, making this a reasonable counterfactual earnings density for those subject to the AET). We then solve for bunching using:

$$B_k(\tau_k, z_k^*; \varepsilon, \phi^*) = F(\log z_k^{ub}) - F(\log z_k^{lb})$$

We estimate bootstrapped standard errors. Observe that the estimated vector of parameters  $(\hat{\varepsilon}, \hat{\phi})$  is a function of the estimated amount of bunching; call this function  $\theta(b)$ . To compute bootstrapped standard errors, we use the bootstrap procedure of Chetty *et al.* (2011) to obtain 200 bootstrap samples of  $b$ . For each bootstrap sample, we compute  $\hat{\varepsilon}$  and  $\hat{\phi}$  as the solution to (D.6). The standard deviation of  $\hat{\varepsilon}$  and  $\hat{\phi}$  across bootstrap samples is the bootstrap standard error, and we compute confidence intervals analogously. We estimate whether an estimate is significantly different from zero by assessing how frequently the constraint  $\phi \geq 0$  binds in our estimation. Given this constraint, p-values are from a one-sided test of equality with zero. We have also estimated the standard errors using the delta method and obtained similar results.

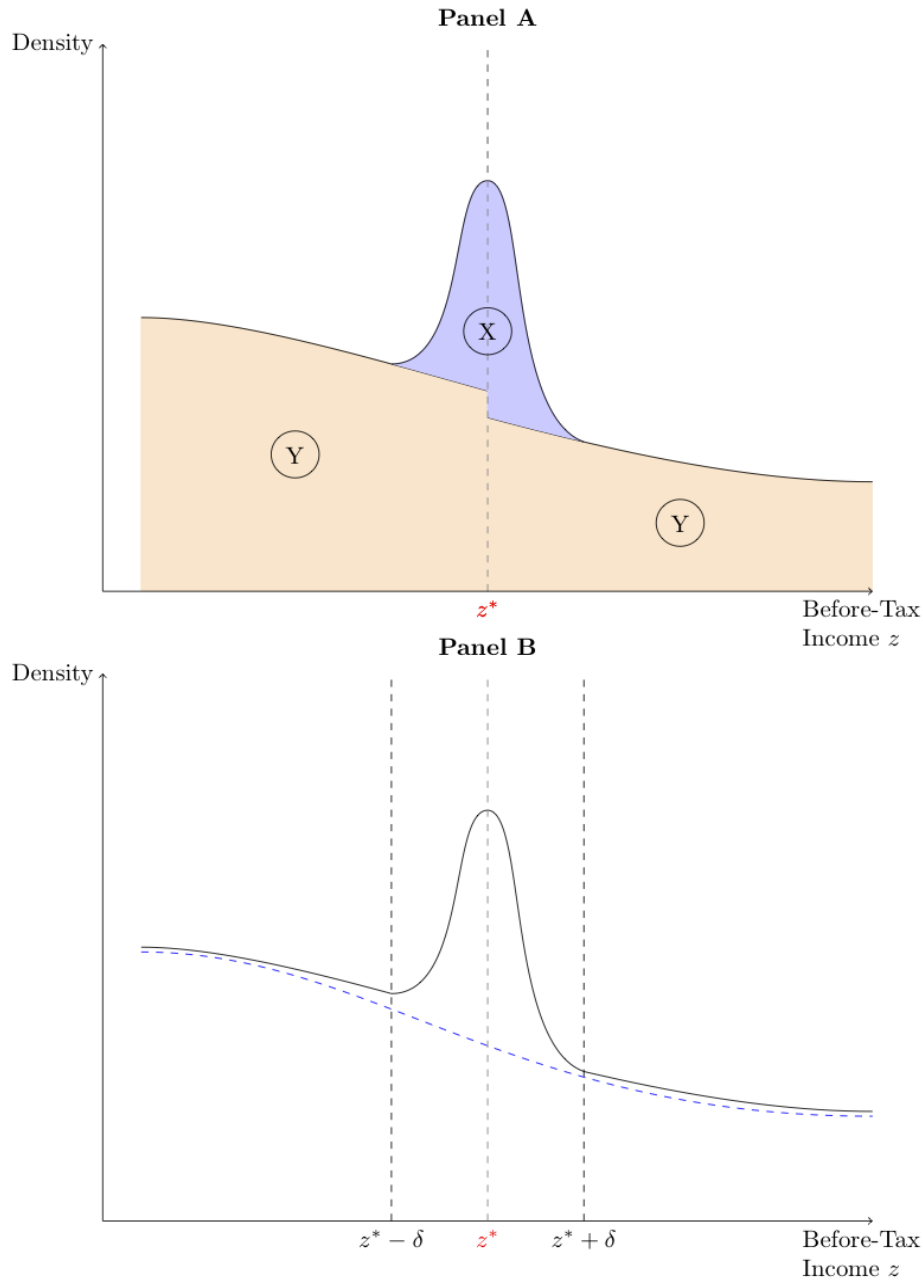
# E Appendix: Additional Figures

Figure E.1: Bunching Response to a Convex Kink (Frictionless Case)



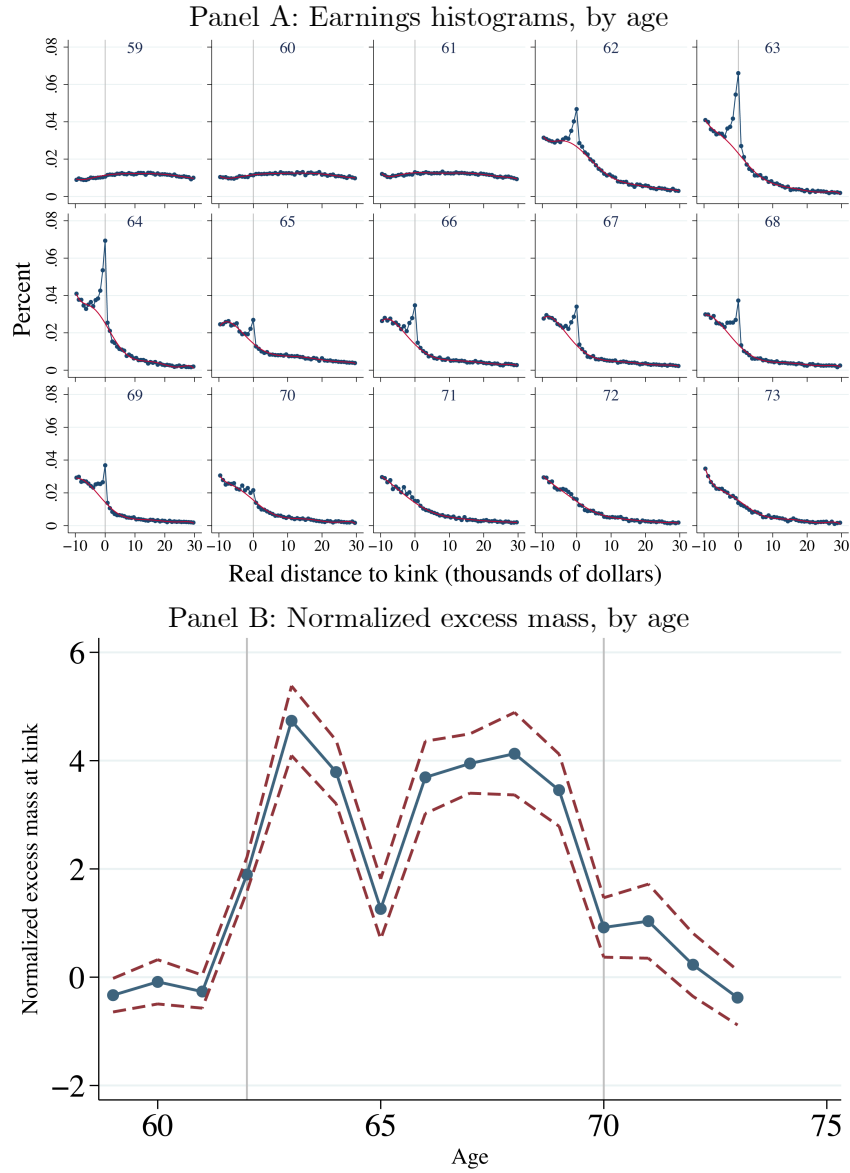
Note: When we move from a linear budget constraint (Panel A) to a convex kink (B), individuals with initial earnings between  $z^*$  and  $z^* + \Delta z^*$  relocate to the kink. As we move from a linear budget constraint (Panel C) to a convex kink (Panel D), a spike in the earnings density appears at the kink, corresponding to the density that was initially located between  $z^*$  and  $z^* + \Delta z^*$ . The spike is spread out in the vicinity of the kink in Panel D; this may result from several factors discussed in Saez (2010), such as inability to control earnings precisely.

Figure E.2: Bunching Estimation Methodology



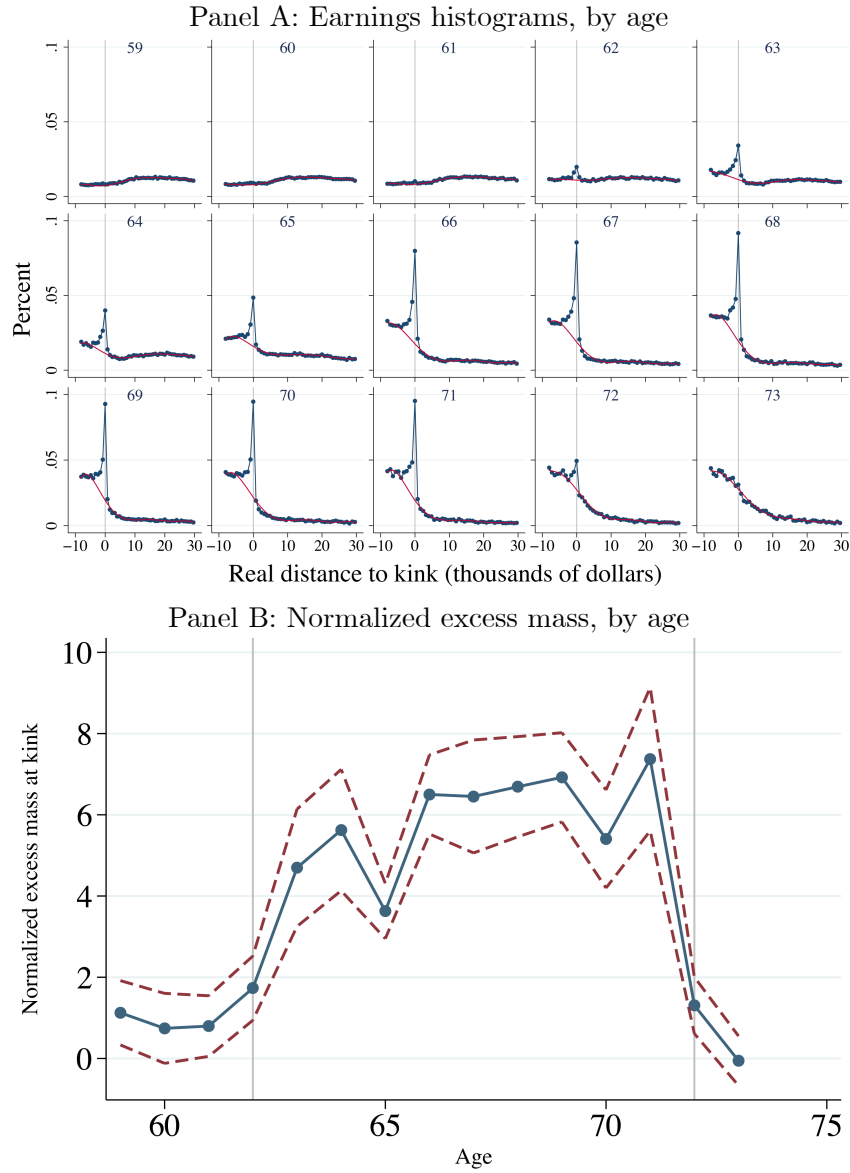
Note: Panel A decomposes the ex-post earnings distribution shown in Appendix Figure E.1 Panel D into two groups. The bunchers, group  $X$ , are those who bunch at the kink in the presence of the higher marginal tax rate  $\tau + d\tau$  but not at the lower marginal tax rate  $\tau$ . The non-bunchers, group  $Y$ , are comprised of those who locate to the left of the kink under the initial lower marginal tax rate  $\tau$ , and those who locate to the right of the kink under the higher marginal tax rate  $\tau + d\tau$ . Panel B demonstrates how the distribution of earnings in the absence of the kink is estimated to recover the share of bunchers, by excluding data in a neighborhood of  $z^*$ .

Figure E.3: Adjustment Across Ages: Histograms of Earnings and Normalized Excess Mass, 59-73-year-old OASI Claimants, 1990-1999



See notes to Figure 2. The figure differs from Figure 2 only because the sample in year  $t$  consists only of people who have claimed OASI in year  $t$  or before (whereas in Figure 2 it consists of all those who claimed by age 65).

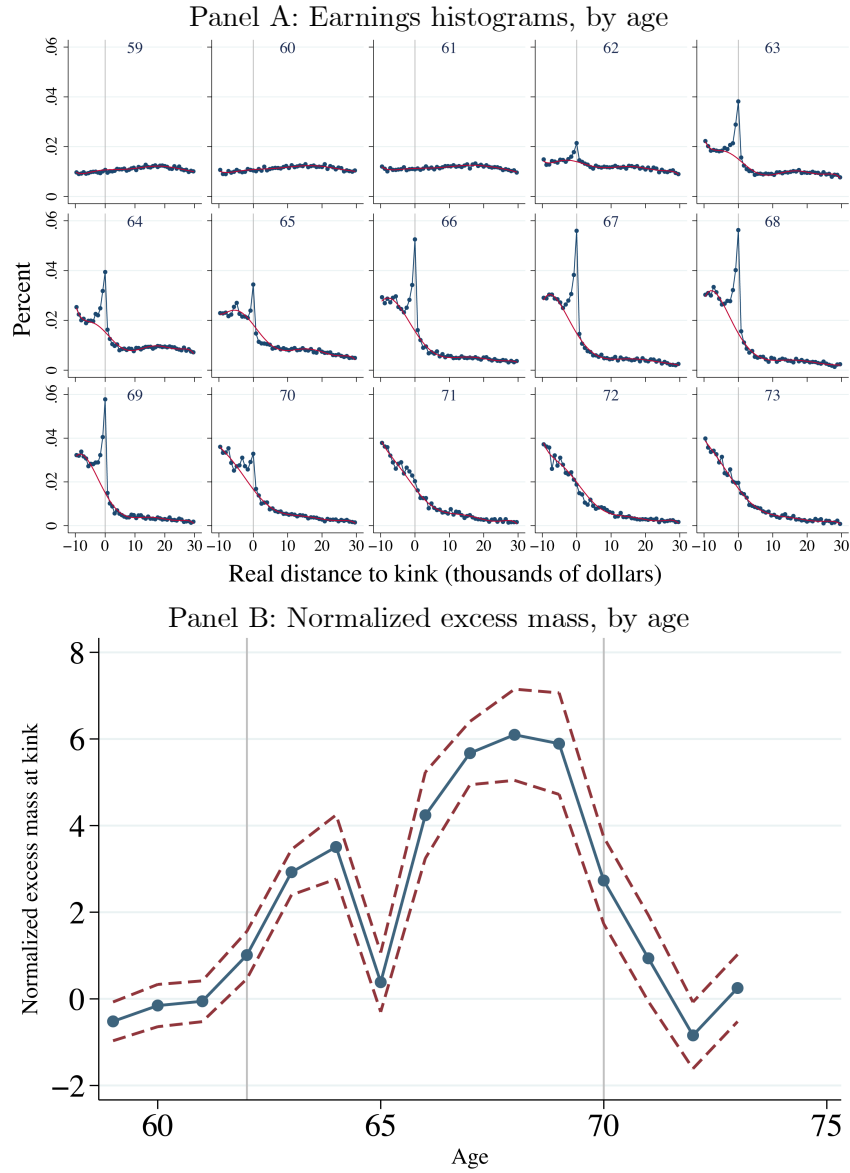
Figure E.4: Adjustment Across Ages: Histograms of Earnings and Normalized Excess Mass, 59-73-year-olds Claiming OASI by Age 65, 1972-1982



See notes to Figure 2. The figure differs from Figure 2 only because the years examined are 1972-1982 (whereas in Figure 2 the years examined are 1990-1999).

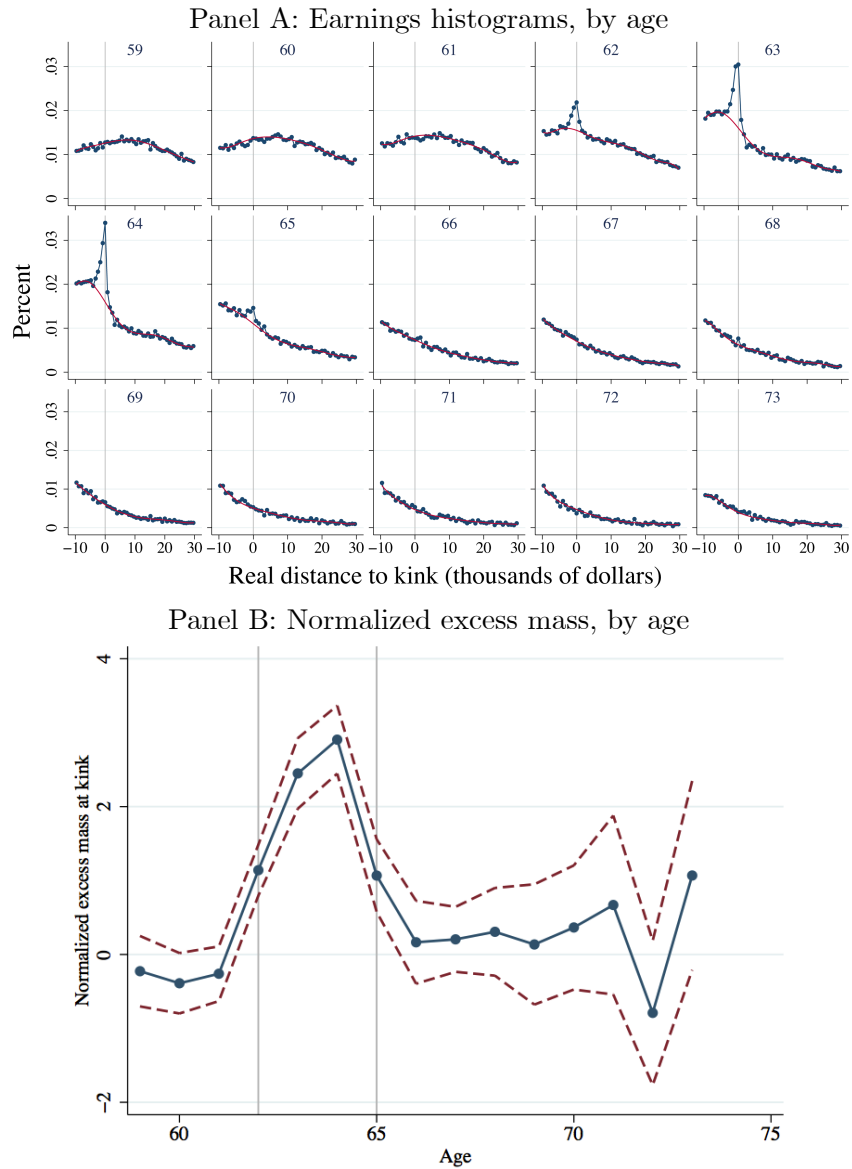


Figure E.5: Adjustment Across Ages: Histograms of Earnings and Normalized Excess Mass, 59-73-year-olds Claiming OASI by Age 65, 1983-1989



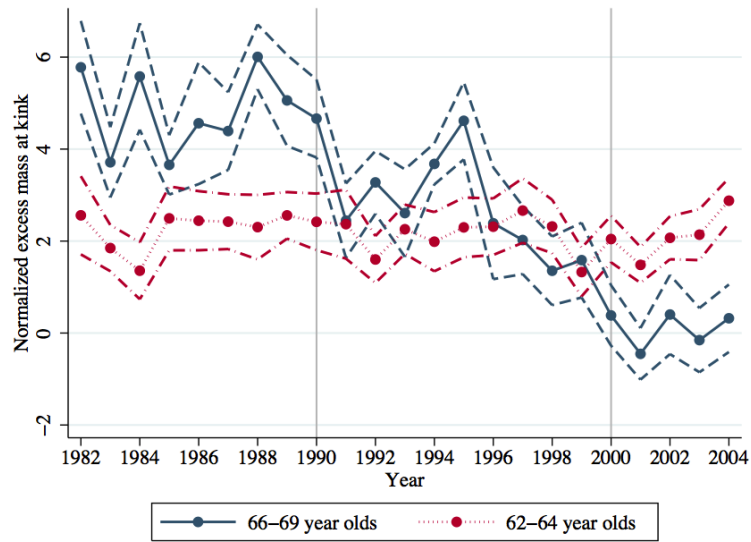
See notes to Figure 2. The figure differs from Figure 2 only because the years examined are 1983-1989 (whereas in Figure 2 the years examined are 1990-1999).

Figure E.6: Adjustment Across Ages: Histograms of Earnings and Normalized Excess Mass, 59-73-year-olds Claiming OASI by Age 65, 2000-2006



See notes to Figure 2. The figure differs from Figure 2 only because the years examined are 2000-2006 (whereas in Figure 2 the years examined are 1990-1999). As explained in the main text, the NRA slowly rose from 65 for cohorts that reached age 62 during this period; the results are extremely similar when the sample is restricted to those who claimed by 66, instead of 65. In the year of attaining NRA, the AET applies for months prior to such attainment.

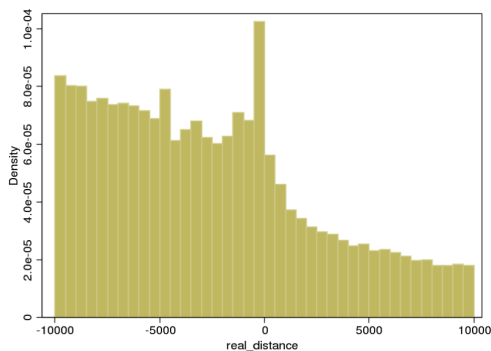
Figure E.7: Comparison of Normalized Excess Bunching Among 62-64 Year-Olds and 66-69 Year-Olds, 1982-2004



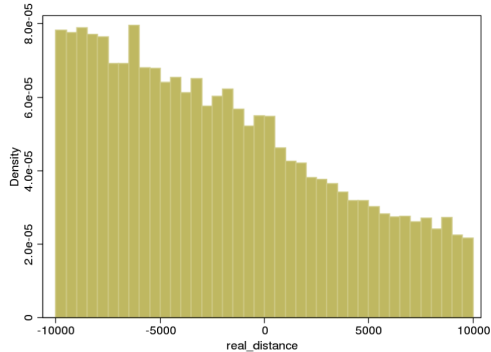
Note: the figure shows excess normalized bunching among 62-64 year-olds and 66-69 year-olds in each year from 1982 to 2004. Note the caveat that the 62-64 year-old group faces a notch at the exempt amount, as opposed to the kink faced by those 66-69.

Figure E.8: Adjustment Across Ages: Histograms of Earnings, 66-69 Year-Olds, 1999-2001

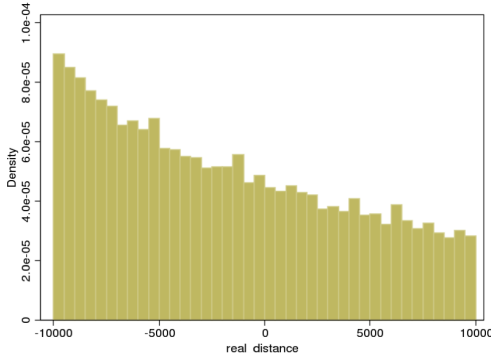
Panel A: Earnings histogram, 66-69 year-olds, 1999



Panel B: Earnings histogram, 66-69 year-olds, 2000

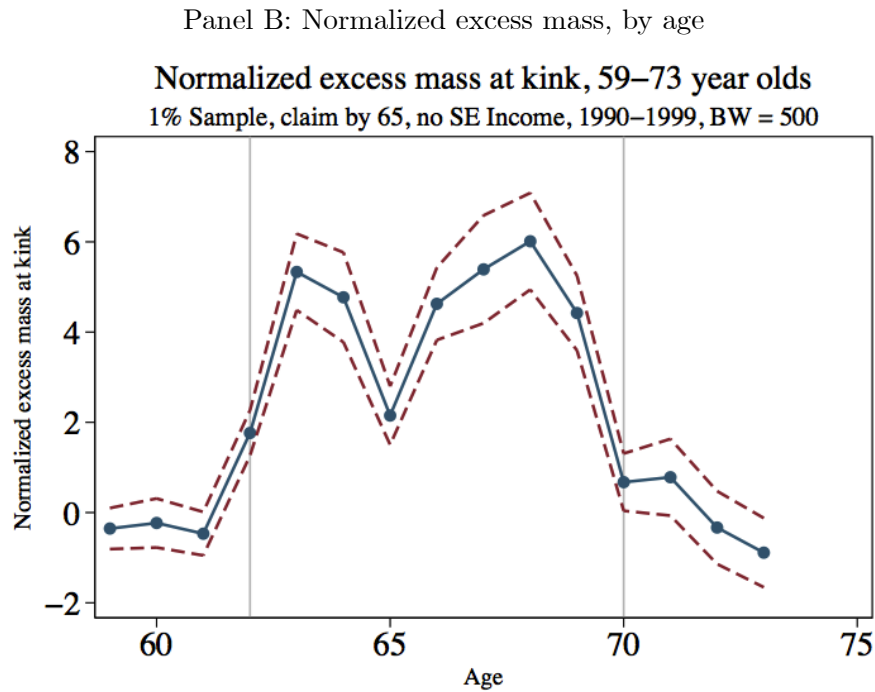
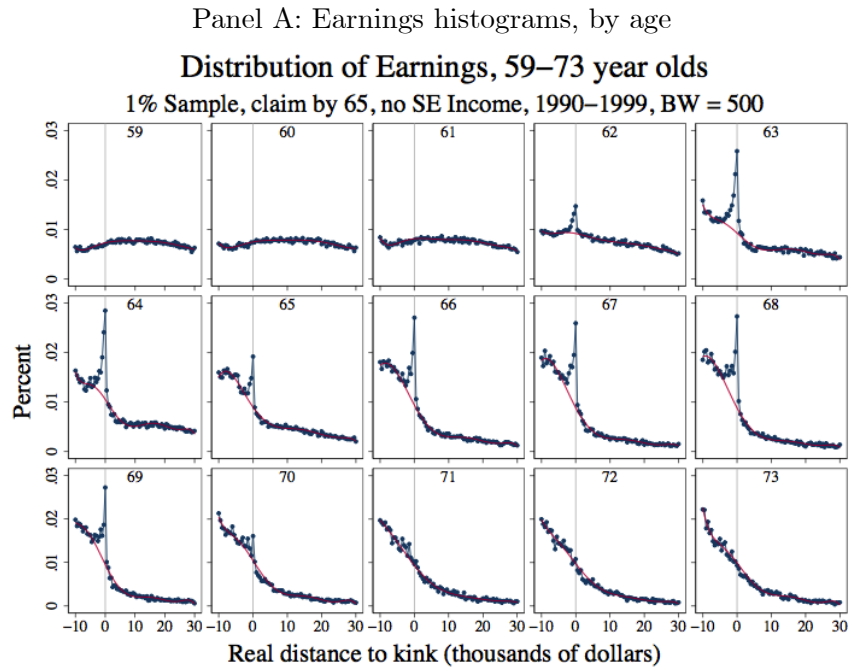


Panel C: Earnings histogram, 66-69 year-olds, 2001



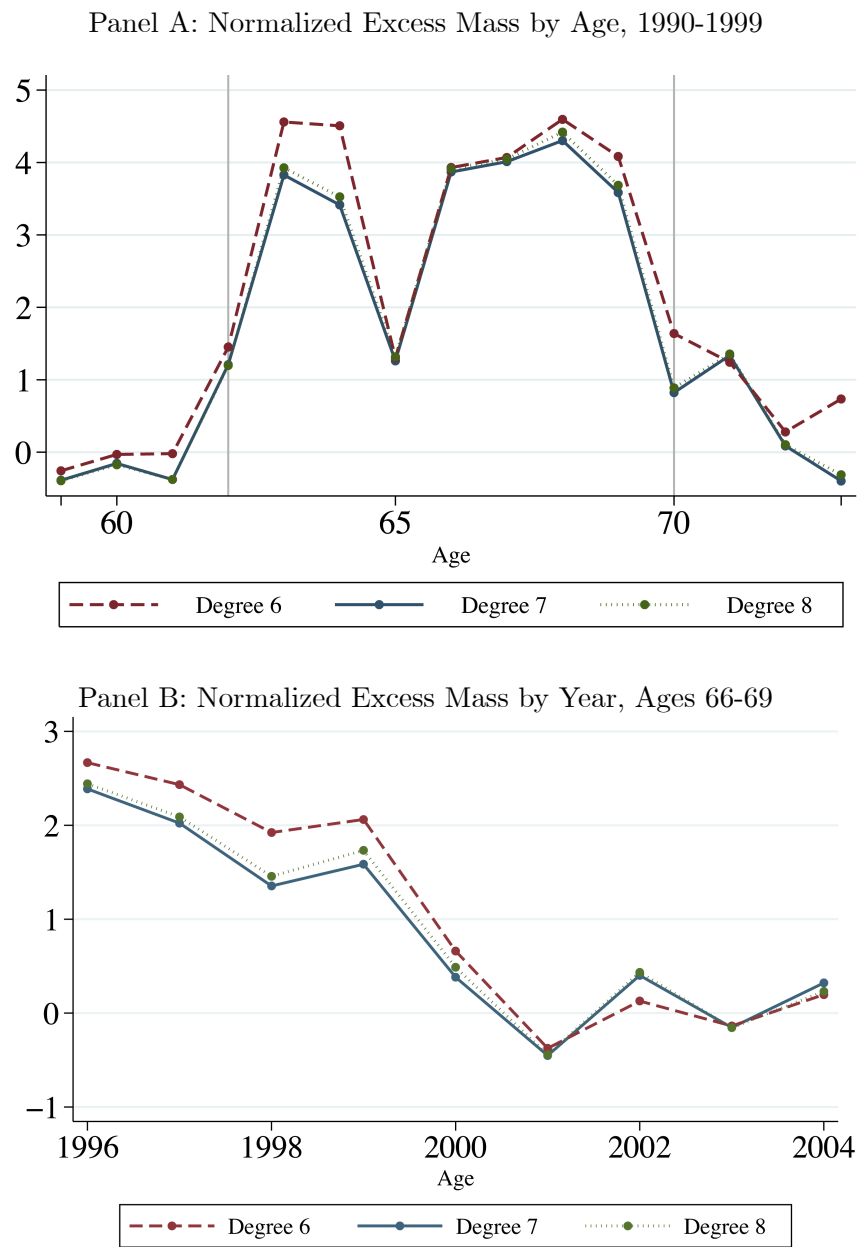
Note: the figure shows a histogram of earnings in 1999, 2000, and 2001, using LEHD data on 66-69 year-olds (which have a larger sample size than the SSA data). Earnings are measured relative to the kink, shown at zero.

Figure E.9: Adjustment Across Ages: Histograms of Earnings and Normalized Excess Mass, 59-73-year-olds Claiming OASI by Age 65, 1990-1999



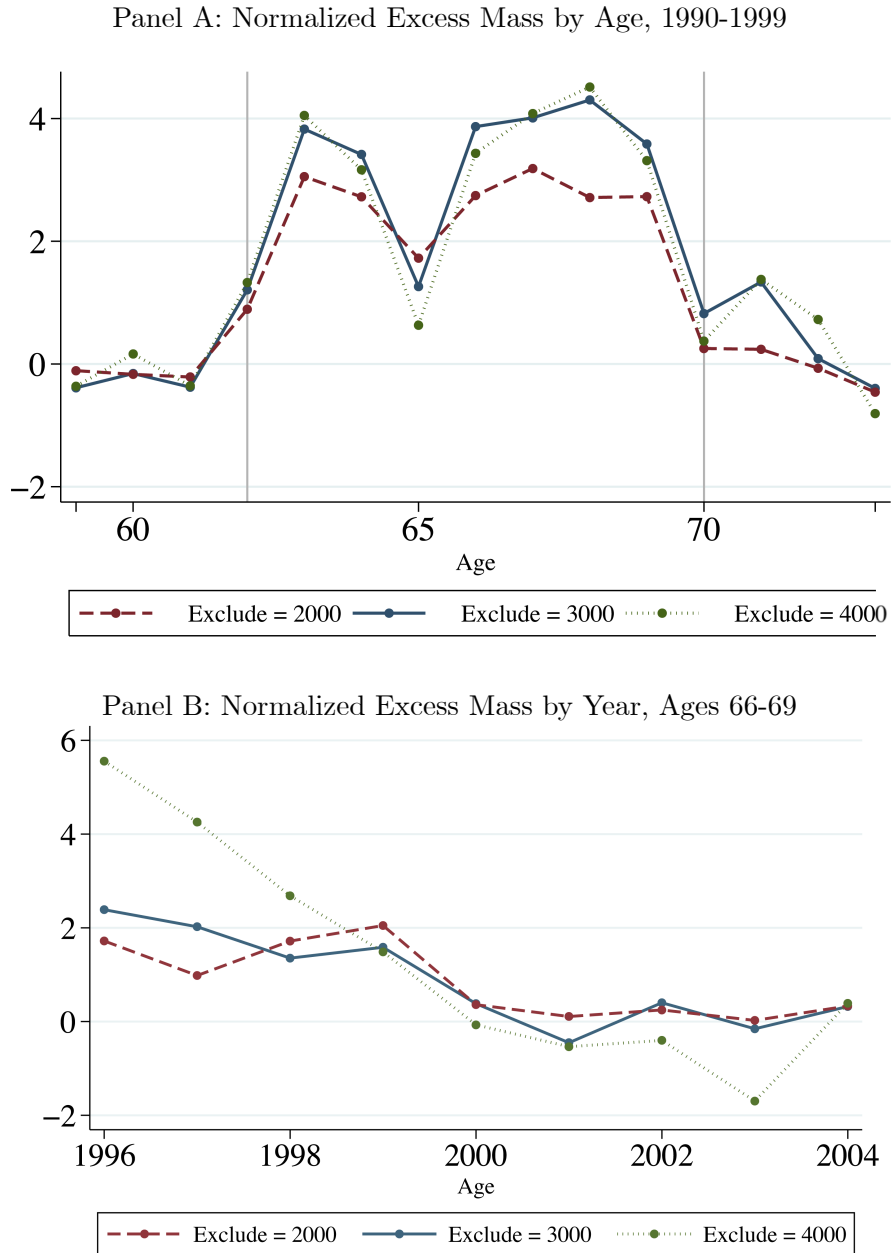
See notes to Figure 2. The figure differs from Figure 2 only because the bandwidth is \$500 (whereas in Figure 2 it is \$800).

Figure E.10: Robustness to Polynomial Degree: Normalized Excess Mass by Age and Year, OASI Claimants by 65



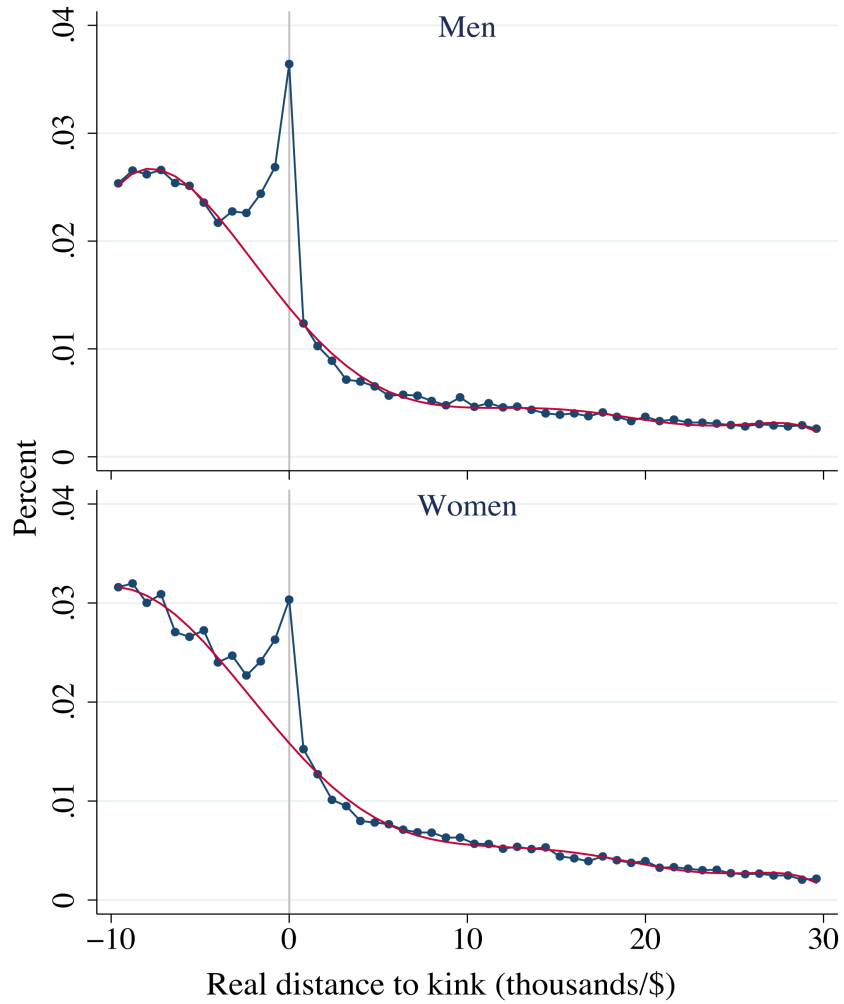
Notes: The figure shows the difference in estimates of normalized excess bunching as we vary the degree of the polynomial used. For additional notes on the samples see Figure 2 for Panel A and Figure 3 for Panel B.

Figure E.11: Robustness to the Excluded Region: Normalized Excess Mass by Age and Year, OASI Claimants by 65



Notes: The figure shows the difference in estimates of normalized excess bunching as we vary the region about the kink that is "dummied out" in the polynomial estimation. For additional notes on the samples see Figure 2 for Panel A and Figure 3 for Panel B.

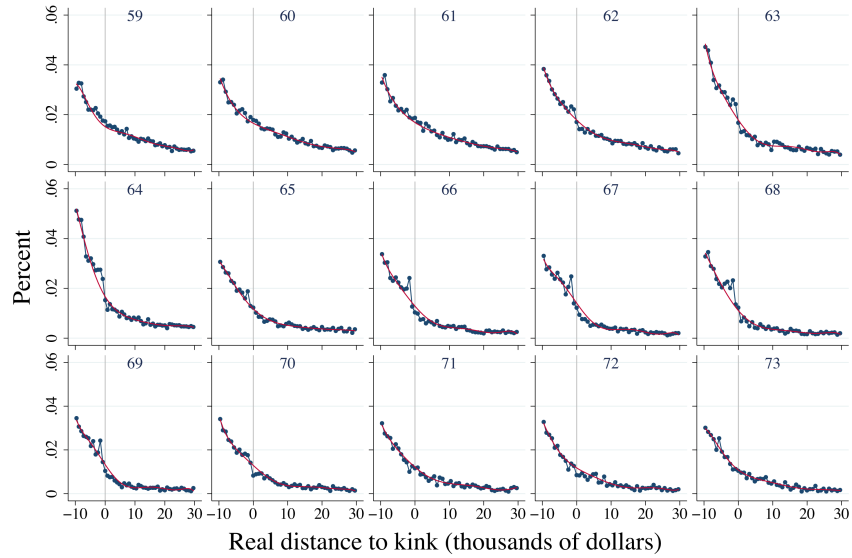
Figure E.12: Adjustment by Sex: Histograms of Earnings, 59-73-year-olds Claiming OASI by Age 65, 1990-1999



See notes to Figure 2. The sample examined is the same as in Figure 2 but examines men and women separately.



Figure E.13: Adjustment Across Ages: Histograms of Earnings, 59-73-year-olds Claiming OASI by Age 65 with Self-Employment Income, 1990-1999



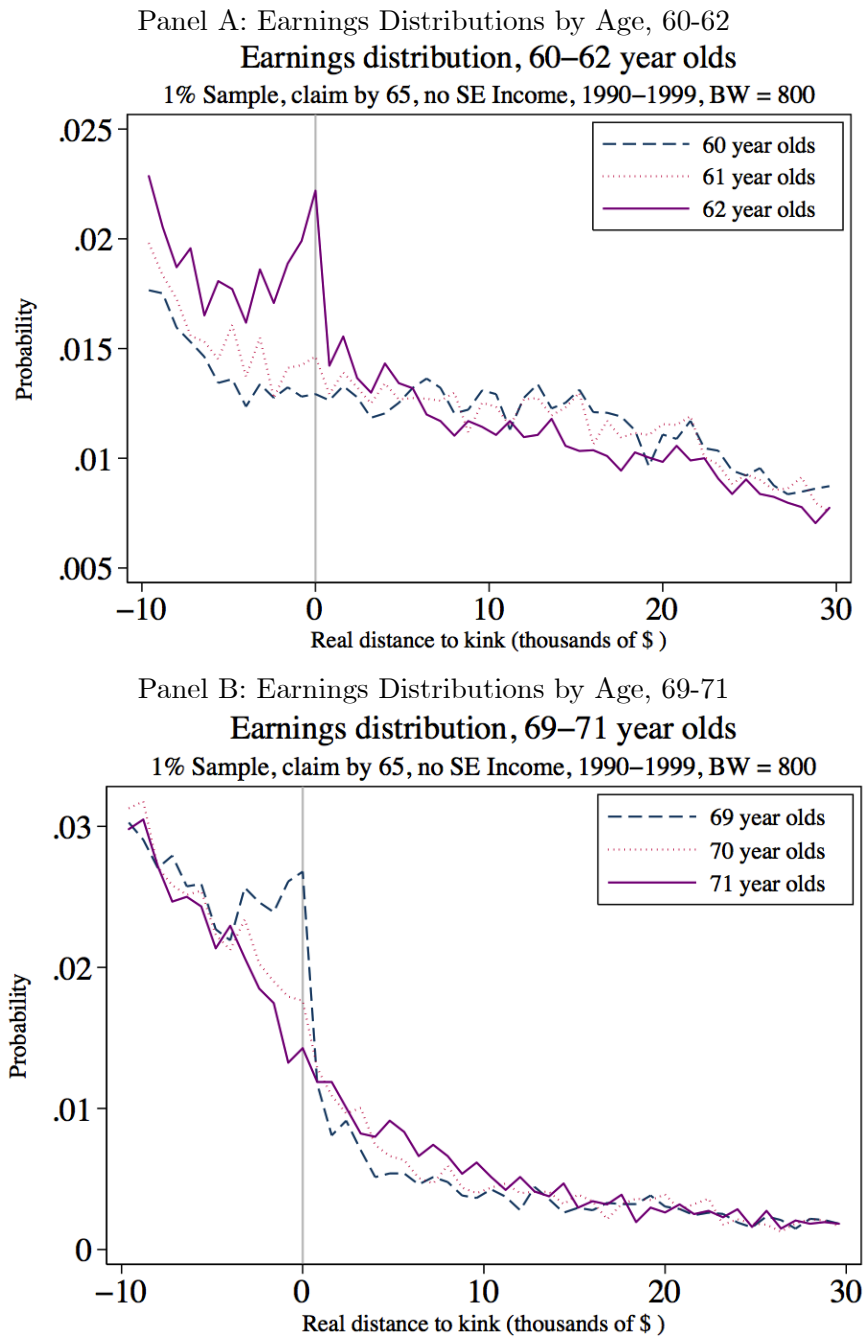
See notes to Figure 2. The figure differs from Figure 2 only because the sample consists of those with positive self-employment income (whereas in Figure 2 those with positive self-employment income are excluded).

Figure E.14: Probability that Earnings Move with Kink, 1990-1998



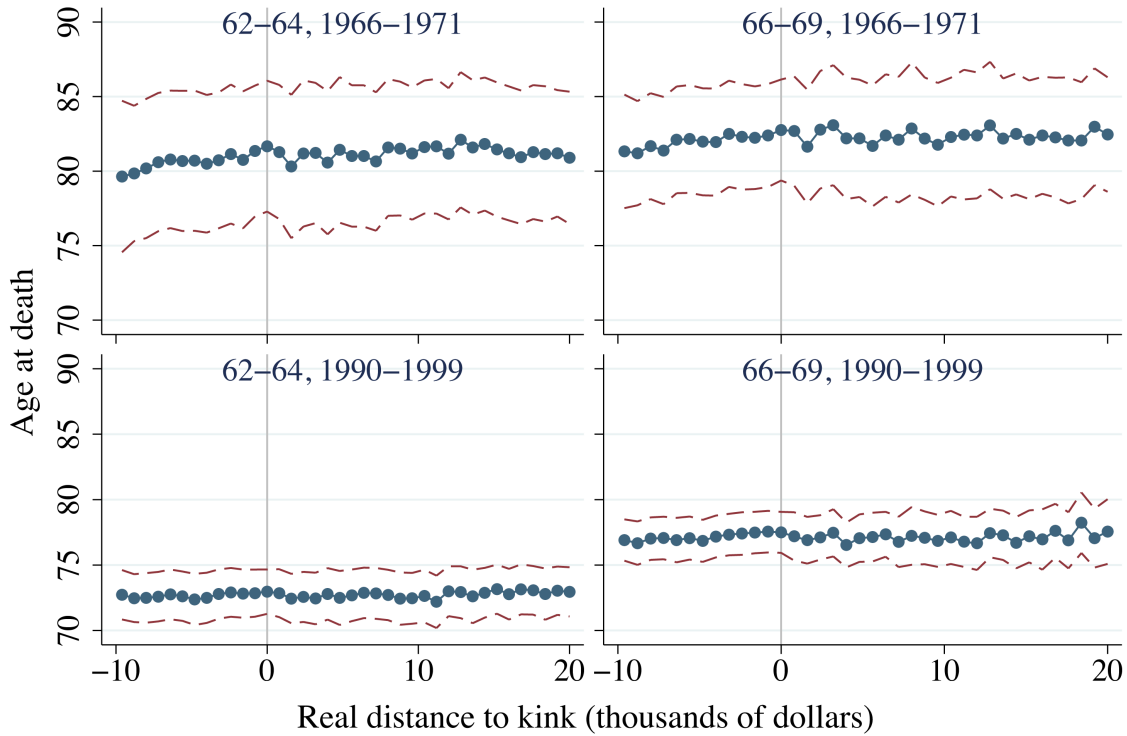
Note: The figure shows the probability that individual earnings move with the kink from year to year (i.e. the probability that an individual locates at the kink in year  $t + 1$ , conditional on locating at the kink in year  $t$ ), for age groups 58 to 60, 62 to 63, and 65 to 68. (Each of these ages refers to age in year  $t$ .) Results are similar when considering similar age bins. The kink is defined as the region within \$2800 of the exempt amount. We exclude the year 1999 because the AET is eliminated for those over NRA in 2000. See other notes to Figure 2.

Figure E.15: Earnings Distributions by Age, OASI Claimants by Age 65, 1990-1999



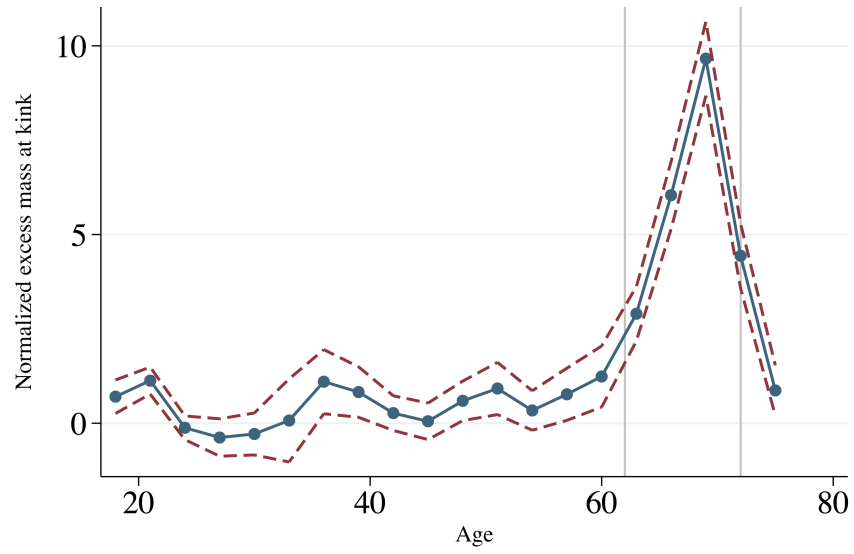
Notes: The figure shows earnings distributions at ages 60, 61, and 62 (Panel A) and at ages 69, 70, and 71 (Panel B).

Figure E.16: Mortality Analysis: Mean Age at Death, 62-69-year-olds Claiming OASI by Age 65, 1966-1971 and 1990-1999



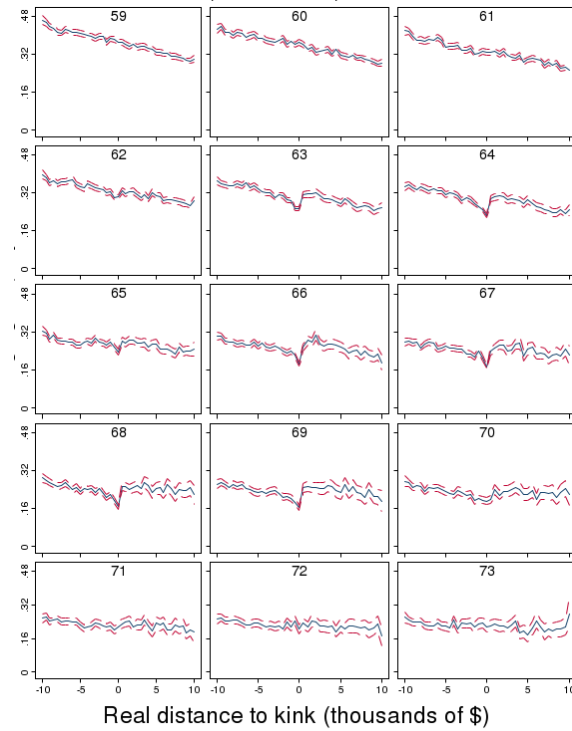
Note: The figure shows mean age at death from a one percent random sample of SSA administrative data on individuals aged 59-73, claiming OASI by age 65, between 1966 and 1971 (inclusive) in the top panel, and between 1990 and 1999 (inclusive) in the bottom panel. The figure shows no clearly noticeable patterns at the kink that are different from those away from the kink. This holds true for those 62-64 and 66-69, in a period prior to the introduction of the Delayed Retirement Credit (*i.e.* 1966-1971) and subsequent to its introduction (*i.e.* 1990-1999). Results are similar for other time periods.

Figure E.17: Normalized Excess Mass at Kink by Age, 1966-1971



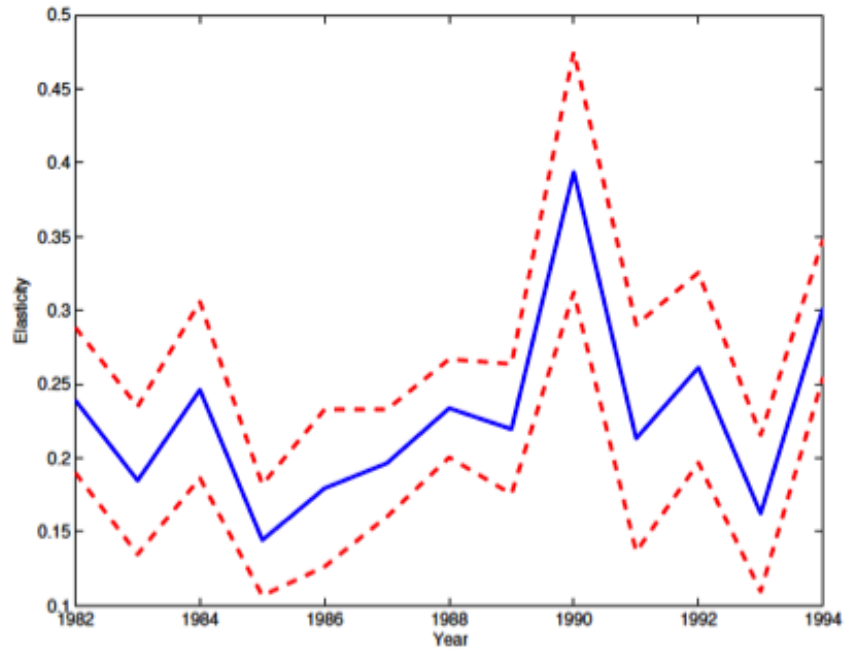
Note: The figure shows normalized excess bunching, ages 18 to 75, 1966-1971. We group ages into three-year bins. See other notes to Figure 2.

Figure E.18: Fraction of Workers Changing Employers from Age  $t$  to Age  $t+1$ , by Age  $t$  Earnings, 1990-1998



Note: The figure shows the fraction of workers who change employers from age  $t$  to age  $t + 1$  (y-axis), plotted against earnings at age 69 (x-axis). For example, 0.16 on the y-axis implies that 16 percent of workers change employers from age  $t$  to age  $t + 1$ . Earnings are measured relative to the kink, shown at zero on the x-axis. The data are a 20 percent random sample of the LEHD in 1990-1998. The bin width is \$800. Solid (dotted) lines show point estimates (95 percent confidence intervals).

Figure E.19: Elasticity Estimates by Year, Saez (2010) Method, 1982-1994



Note: The figure shows elasticities estimated using the Saez (2010) method, by year from 1982 to 1994, among 66-69 year-old OASI claimants. We use our methods for estimating normalized excess bunching but use Saez' (2010) formula to calculate elasticities.

## F Appendix: Additional Estimates of $\varepsilon$ and $\phi$

Table F.1: Estimates of Elasticity and Adjustment Cost Using Sharp Change Method and Disappearance of Kink at Age 70

	(1)	(2)	(3)
	$\varepsilon$	$\phi$	$\varepsilon \phi = 0, \text{ Age } 69$
Basic	0.28 [0.25, 0.34]***	89.93 [25.51, 240.00]***	0.25 [0.22, 0.30]***
Lognormal	0.30 [0.26, 0.36]***	\$87.19 [31.88, 256.39]***	0.27 [0.23, 0.32]***
Benefit Enhancement	0.42 [0.37, 0.50]***	\$58.01 [16.70, 146.86]***	0.38 [0.33, 0.45]***
Excluding FICA	0.36 [0.31, 0.43]***	\$83.01 [23.79, 213.17]***	0.33 [0.28, 0.38]***
Bandwidth = \$500	0.20 [0.22, 0.32]***	\$17.05 [0.69, 76.00]***	0.26 [0.16, 0.23]***
68-70 year-olds	0.30 [0.27, 0.35]***	\$79.10 [24.61, 189.34]***	0.28 [0.25, 0.32]***

Note: The table estimates elasticities and adjustment costs using the removal of the AET at age 70, using data on 69-71 year-olds. We cannot estimate the constrained elasticity using only data on age 70 because the benefit reduction rate is zero at that age. We show these results in the Appendix, rather than the main text, because the estimates of excess bunching at age 70 are potentially affected by the coarse measure of age that we use, as explained above. To address this issue, we use age 71 in estimating these results. Using only age 70 shows very similar results, which is unsurprising because Figure 2 shows that normalized excess bunching is similar at ages 70 and 71. The row labeled "68-70 year-olds" uses data from ages within this range. See also notes to Table 2.



Table F.2: Estimates of Elasticity and Adjustment Cost Using Sharp Change Method and 1990 Policy Change, Assuming no Pre-Period Bunching Attenuation

	(1)	(2)	(3)	(4)
	$\varepsilon$	$\phi$	$\varepsilon \phi = 0$	
			1990	1989
Baseline	0.22 [0.18, 0.26]***	\$139.53 [46.25, 359.40]***	0.39 [0.33, 0.48]***	0.22 [0.18, 0.26]***
Lognormal density	0.24 [0.19, 0.29]***	\$149.87 [42.55, 426.32]***	0.43 [0.35, 0.54]***	0.24 [0.19, 0.29]***
Benefit Enhancement	0.37 [0.30, 0.43]***	\$76.24 [17.28, 237.63]***	0.59 [0.49, 0.73]***	0.37 [0.30, 0.43]***
Excluding FICA	0.30 [0.25, 0.35]***	\$119.84 [32.33, 344.45]***	0.50 [0.42, 0.62]***	0.30 [0.25, 0.35]***
Bandwidth = \$500	0.24 [0.19, 0.30]***	\$83.80 [5.86, 296.54]***	0.38 [0.30, 0.50]***	0.24 [0.19, 0.30]***
68-70 year-olds	0.31 [0.27, 0.34]***	\$106.32 [25.21, 291.94]***	0.50 [0.42, 0.62]***	0.31 [0.27, 0.34]***

Note: The table applies the Sharp Change (SC) method to the 1990 policy change, using data from 1989 and 1990, but assumes that bunching in 1989 is not attenuated by adjustment frictions. The constrained estimate of bunching using data only from 1989 is mechanically the same as the unconstrained estimate, as both rely on the Saez (2010) formula for bunching (as the first-order conditions reveal). See also notes to Appendix Table F.1.

Table F.3: Estimates of Elasticity and Adjustment Cost Using Sharp Change Method and Elimination of Earnings Test in 2000 for 66-69 Year-Olds

	(1)	(2)	(3)
	$\varepsilon$	$\phi$	$\varepsilon \phi = 0, 1999$
Baseline	0.10 [0.03, 0.19]***	\$23.23 [0.00, 331.52]*	0.09 [0.04, 0.14]***
Lognormal density	0.10 [0.03, 0.21]***	\$20.17 [0.00, 346.90]*	0.09 [0.04, 0.14]***
Benefit Enhancement	0.15 [0.04, 0.36]***	\$14.93 [0.00, 350.46]*	0.14 [0.07, 0.21]***
Excluding FICA	0.13 [0.03, 0.28]***	\$21.31 [0.00, 402.69]*	0.12 [0.06, 0.18]***
Bandwidth = \$500	0.10 [0.07, 0.21]***	\$56.48 [0.03, 367.05]**	0.09 [0.05, 0.13]***
1998 to 2000	0.12 [0.08, 0.55]***	\$23.09 [0.00, 957.13]*	0.11 [0.07, 0]***

Note: The table applies the SC method to the 2000 policy change, using data on 1999 and 2000. It is not possible to perform the constrained estimate using a cross-section from the year 2000 because the marginal tax rate is zero in this year, implying that the elasticity is undefined. The final row uses data from 1998 and 2000, rather than 1999 and 2000. Note that the left endpoint of the confidence interval is sometimes at 0.00 even though the p-value indicates significance at the 5 percent level; this is because a few of the bootstrap replications show point estimates that are positive and very small so round to 0.00. See also notes to Appendix Table F.1.

Table F.4: Estimates of Elasticity and Adjustment Cost Using Sharp Change Method and Disappearance of Kink at Age 70

	(1)	(2)	(3)	(4)
	$\varepsilon$	$\phi$	$\varepsilon \phi = 0$	
			1993	1989
Baseline	0.21 [0.18, 0.34]***	\$0.00 [0.00, 902.56]	0.21 [0.14, 0.31]***	0.21 [0.17, 0.26]***
Lognormal density	0.24 [0.20, 0.36]***	\$33.35 [0.00, 577.01]	0.22 [0.15, 0.34]***	0.23 [0.19, 0.29]***
Benefit Enhancement	0.37 [0.30, 0.50]***	\$78.09 [0.00, 352.03]	0.32 [0.21, 0.47]***	0.36 [0.29, 0.43]***
Excluding FICA	0.30 [0.25, 0.42]***	\$58.54 [0.00, 492.47]	0.27 [0.18, 0.40]***	0.29 [0.24, 0.35]***
Bandwidth = \$500	0.25 [0.21, 0.80]***	\$0.00 [0.00, 4188.55]	0.25 [0.20, 0.35]***	0.25 [0.19, 0.32]***
1988 and 1993	0.26 [0.23, 0.33]***	\$146.33 [0.00, 408.86]	0.21 [0.14, 0.31]***	0.25 [0.22, 0.28]***
Linear Approximation	0.46 [0.58, 0.60]***	\$379.10 [351.27, 406.93]***	0.30 [0.19, 0.41]***	0.39 [0.31, 0.47]***

Note: The table applies the Comparative Static (CS) method to the 1990 policy change, using data from 1989 and 1993. "Linear approximation" refers to estimates based on the linearized formulas presented in the text. See also notes to Appendix Table F.1.