

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

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

July 2013

## Abstract

Recent literature has suggested that individuals face frictions in adjusting their earnings in response to policy, such as the cost of gathering information about policy or of finding a job that pays their desired earnings level. Using a panel of administrative data from the Social Security Administration on 1% of the U.S. population from 1961 to 2006, we document descriptive evidence about such barriers and use this evidence in estimating the parameters of the adjustment process. We analyze "bunching" at the convex kink in individuals' budget sets created by the Social Security Annual Earnings Test (AET). When individuals transition out of being subject to the AET, the vast majority of earnings adjustment takes place within three years, suggesting rapid adjustment. We specify a model consistent with the descriptive patterns and estimate that the mean of the elasticity of earnings with respect to the implicit net-of-tax share is XX and the mean fixed cost of adjustment is \$XX. Using data from 1986-1988 and 1992-1994, during which period the AET became less stringent for those at Normal Retirement Age and above. We estimate a smaller elasticity of only 0.XX when we constrain adjustment costs to be zero, though the difference is not statistically significant.

---

\*Gelber: Wharton, agelber@wharton.upenn.edu; Jones: Harris and NBER, damonjones@uchicago.edu; Sacks: Wharton, dansacks@wharton.upenn.edu. We thank Raj Chetty, Jim Cole, Jim Davis, Mark Duggan, Jonathan Fisher, Richard Freeman, John Friedman, Adam Isen, Henrik Kleven, David Pattison, Emmanuel Saez, and numerous seminar participants. We acknowledge financial support from the Wharton Center for Human Resources and the Wharton Risk and Decision Processes Center, NIH grant #1R03 AG043039-01, and from a National Science Foundation Graduate Research Fellowship. This research was also supported by 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. This 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, individuals optimize their behavior frictionlessly in response to policies that affect the incentive to earn. However, literature has suggested that individuals may face important frictions in adjusting to such policies (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).

This paper develops evidence on the nature and size of frictions in adjusting earnings in response to policy. The 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 benefits once an individual begins to earn 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 marginal tax rates (MTRs) 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 (Friedberg, 1998, 2000).<sup>1</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 explain why individuals' earnings may still be expected 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

---

<sup>1</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 ET reduces Social Security benefits and is not administered through the tax system. Other papers on the ET include Burtless and Moffitt (1985), Gruber and Orszag (2003), and Song and Manchester (2007).

nonparametric documentation of behavioral responses.<sup>2</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 of bunching among the non-self-employed at any kink in the U.S.<sup>3</sup> Third, the AET is interesting to policy-makers in its own right, as it is an important factor that affects the earnings of the elderly in the U.S.

We make three main contributions to understanding adjustment frictions. First, we investigate the speed at which individuals adjust their earnings in response to AET policy changes, developing evidence consistent with adjustment frictions by documenting that in some contexts individuals do not adjust immediately to changes in AET. The speed of adjustment is of interest in part because 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 were 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 individuals to respond to policy changes.

Interestingly, we find that the speed of adjustment is relatively fast. Across several contexts—including both anticipated and unanticipated changes in policy—the vast majority of individuals’ adjustment occurs within at most three years. Adjustment is even faster in certain contexts, occurring within one year. Despite this fast adjustment, we do observe clear evidence of delays in some contexts, consistent with barriers to adjustment in these instances. 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. In the absence

---

<sup>2</sup>Other papers have examined bunching in the earnings schedule, including Blundell and Hoynes (2004), and Saez (2010), who shows that the amount of bunching can be related to the elasticity of earnings with respect to the net-of-tax share, defined as one minus the MTR. Our project is the first to estimate earnings elasticities through a similar method in the context of the ET.

<sup>3</sup>The lack of bunching at other kinks is consistent with the existence of adjustment costs, although this finding could also be explained by a low elasticity of earnings or hours worked with respect to the net-of-tax rate. As we discuss in greater detail in Section 6, Chetty et al. (2012b) do find evidence of more diffuse earnings responses to the Earned Income Tax Credit among the non-self-employed.

of barriers to adjustment, removal of a convex kink in the effective tax schedule should immediately lead bunching to vanish at the earnings level associated with the former kink.<sup>4</sup> Thus, any observed delay in responding to policy—such as the delays up to three years that we find—should reflect barriers to adjustment.

Second, we assess the mechanisms that underlie the patterns of adjustment we observe. We 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 responsive. Others subject to the AET appear to be unresponsive, raising the possibility that these individuals face substantial adjustment costs and suggesting heterogeneity in adjustment costs or elasticities in the population. We also 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>5</sup> We find little evidence that individuals not subject to the AET, i.e. too young to claim benefits, bunch at the kink. We do find, however, that those initially locating at the kink are particularly unlikely—relative to those initially some distance from the kink—to change employers while subject to the AET. This suggests that the AET affects earnings primarily when certain particularly responsive individuals choose to change their earnings while disproportionately staying at the same job.

Third, we specify a tractable model of earnings adjustment consistent with the above descriptive evidence that allows us to estimate a mean fixed adjustment cost and the mean elasticity of earnings with respect to the effective net-of-tax share. Recent work demonstrating the importance and existence of adjustment costs has raised the question of how to estimate the elasticity and adjustment cost in such a context. Our model leads to tractable methods that allow estimation of elasticities and adjustment costs with kinked budget sets. This complements Kleven and Waseem (forthcoming), who develop a method to estimate

---

<sup>4</sup>In our context, the only "appearance" of a new kink that we observe is the appearance of a kink at age 62. 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. Thus, we focus in this context on adjustment to the disappearance—rather than appearance—of kinks, while acknowledging that in other contexts it could be more illuminating to examine adjustment to the appearance of a kink.

<sup>5</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, and our evidence suggests such a role in the U.S.

related parameters in the presence of a notch in the budget set. The intuition that underlies our method is straightforward. In a single cross-section of individuals, the amount of bunching at a kink increases with the elasticity but decreases with the adjustment cost. With two or more cross-sections of individuals facing different tax rates in the region of the kink, we can recover the parameters of interest—the elasticity and the adjustment cost. Furthermore, when there is an existing kink and the size of the kink is reduced, there will be excess mass leftover at the kink, due to adjustment costs. We apply both insights to provide alternative estimates of the parameters of interest.

We apply our method to data spanning the decrease in the AET explicit marginal tax rate from 50 percent to 33 percent in 1990 for those aged 66 to 69. We estimate that the mean adjustment cost is between \$113 (in 2010 dollars) and that the mean earnings elasticity with respect to the net-of-tax share is 0.35. We estimate that the mean earnings elasticity is slightly smaller, 0.32, when we constrain the mean adjustment cost to be zero. Our results are similar when applying either of our two key approaches using 66 to 69 year olds before and after the 1990 change in policy. We can apply our methods to the disappearance of the kink when one ages from 69 to 70. In this case, we estimate slightly smaller adjustment cost, \$70, and a slightly larger elasticity, 0.45 – the elasticity reduces to 0.423 when we constrain the adjustment cost to be zero. Our result suggest a small adjustment cost in this setting, which corroborates the generally rapid adjustment to changes in policy we observe in the data. Intuitively, our results suggest that we may underestimate the elasticity if we were to ignore the adjustment cost, albeit only by a tiny bit. It may be the case that in the sample we study partially retired workers are in fairly flexible working arrangements. Nevertheless, the setting demonstrates convincingly how the method we develop may be applied to more general tax settings where adjustment costs may play a significant role.

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 ET, including how quickly individuals respond to changes in AET parameters across age groups and over time. Section 6 explores mechanisms underlying this behavioral response,

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 mean fixed adjustment cost and mean elasticity simultaneously. Section 8 concludes with discussion and avenues for future work.

## 2 Policy Environment

The two major components of the AET are the exempt amount and the benefit reduction rate for earnings above this amount. Figure 1 shows key features of the AET rules over the period we study, 1961 and 2006. The AET has become less stringent over this period. 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 rose faster than inflation. Starting in 1978, the AET had different rules for beneficiaries younger than Normal Retirement Age (NRA) and those of at least NRA but younger than the maximum age subject to the AET.<sup>6</sup> The exempt amount began to rise much faster for beneficiaries NRA and older than for younger beneficiaries. 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>7</sup> In 1990 and after, the benefit reduction rate fell to 33 percent for beneficiaries above the NRA. 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.<sup>8</sup>

---

<sup>6</sup>The NRA, the age at which workers can claim their full OASI benefits, was 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>7</sup>From 1961 to 1972, in addition to the threshold above which the benefit reduction rate was 50 percent that we examine in Figure 8, 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>8</sup>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 reductions we have described. This benefit is also reduced when the spouse works over 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; Coile, et al.

When current benefits are lost to the AET, future scheduled benefits are increased in some circumstances. This is sometimes called "benefit enhancement." The rules are different depending on if an individual is above or below NRA. Prior to 1972, the AET represented a pure loss in benefits for those NRA and above, as there was no benefit enhancement for these individuals. As a result, the AET implicit MTR during this period is also equal to the AET explicit MTR. For beneficiaries subject to the AET aged NRA and above, a 1 percent Delayed Retirement Credit (DRC) was introduced in 1972, meaning that each year's worth of benefits foregone led to a 1 percent increase in future yearly benefits. The DRC was raised to 3 percent in 1982 and gradually rose to 8 percent from 1990-2008 (though the AET was eliminated in 2000 for those above NRA). A 7-8 percent credit 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 receiving benefits today or foregoing benefits today and receiving higher benefits once she begins to collect OASI (Friedberg, 1998). Importantly, future benefits are not raised due to the DRC unless annual earnings are high enough above the exempt amount that an individual misses at least one month's worth of benefits due to the AET (Friedberg 1998; Social Security Administration 2012).<sup>9</sup> So long as one does not expect to earn significantly higher than the exempt amount, the implicit MTR is equal to the explicit MTR for an individual considering earning an extra dollar above the AET earnings threshold.

The AET could also affect an individual aged NRA and above who is considering whether to earn an amount sufficiently high enough to trigger benefit enhancement. Prior to the late 1990s, for those individuals who experienced benefit enhancement, the DRC was less than actuarially fair—on average individuals lost money in present value through the AET—implying that such decisions could still respond to changes in the AET explicit MTR. Even after the DRC became actuarially fair on average, an individual considering earning enough to trigger benefit enhancement might yet respond to the AET for several reasons. Those

---

2002), we model the ET as creating the MTRs associated with the benefit reduction rates described, because the ET reduces family benefits by these amounts (all else equal). Our data do not contain the information necessary to link spouses.

<sup>9</sup>For example, if an individual receives the average monthly benefit of \$1,200, is above NRA prior to 2000 and is therefore facing a benefit reduction rate of 33 percent, then she must earn at least \$3,600 beyond the exempt amount before receiving a DRC.

whose expected lifespan is shorter than average experience a net loss in benefits due to the AET. Liquidity-constrained individuals or those who discount faster than average could also reduce work in response to the AET. Finally, some individuals may mis-perceive the AET, as we discuss later.

For beneficiaries under NRA, an actuarial adjustment raises future benefits by 0.55 percent for each month’s worth of benefits loss due to the AET. Importantly, earning any amount over the AET threshold results in at least a month’s worth of benefits withheld for this group and therefore the full actuarial adjustment.<sup>10</sup> For those below NRA who are considering earning an extra dollar after already having crossed the AET earnings threshold, the AET explicit MTR is therefore equal to the AET implicit MTR, since the same enhancement of future benefits occurs no matter how much one earns above the AET exempt amount. It is important to note that this creates a notch in future benefits—as opposed to a 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 features a kink rather than a notch.

We 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 presumption we have described, and previous literature—and focus on these individuals in the framework we describe. In the empirical section, we explore evidence relating to the mechanisms that explain this response.<sup>11</sup>

<sup>10</sup>Social Security Administration (2012), Section 728.2; Gruber and Orszag (2003).

<sup>11</sup>In this study, we focus on the marginal incentives created by the AET and intensive margin responses, although the AET could also affect extensive margin decisions. In a companion paper, Gelber, Jones, and Sacks (2013) examine the extensive margin response in another context. In this paper, we follow previous literature based on the technique of Saez (2010) in focusing on intensive margin but not extensive margin responses to kinks.



### 3 Bunching Framework

Figure 2 shows the budget constraint and incentives created by the AET for those NRA and above in a model with no frictions.<sup>12</sup> 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,  $z - T(z)$ . Individuals derive utility  $u(c, z; n)$  from consumption,  $c$ , and gross-of-tax earnings,  $z$ , where increasing  $z$  requires costly effort and therefore reduces utility, and  $n$  is an index of heterogeneity governing this tradeoff, *i.e.* "ability". Consider first a world with 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, 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  $z^*$ .<sup>13</sup>

Panels C and D of Figure 2 depict densities of earnings we would expect to observe in the absence and presence of the ET, 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  $z^*$ ,

---

<sup>12</sup>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). Robert Moffitt (1986, 1990) surveys these methods and their applications.

<sup>13</sup>The AET potentially creates other distortions that (1) in the case of those post-NRA, are not likely to apply for bunchers and (2) in general do not appear to be empirically relevant for all ages. For these reasons, we abstract from these additional features in this section. These include: a non-convex kink in the budget constraint at the point at which OASI benefits are fully phased out, 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, and additional notch for those below the NRA earning more than the exempt due to the actuarial adjustment described above. We return to these incentives in our discussion of the empirical evidence and our estimation exercise.

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).

To measure the amount of bunching, we use a technique similar to Chetty, Friedman, Olsen, and Pistaferri (2011) and Kleven and Waseem (forthcoming), which we illustrate in Figure 3. The x-axis measures before-tax income,  $z$ , while the y-axis again measures the density of earnings. In Panel A, we reiterate the result from above, that the ex-post density of earnings in the presence of a kink is comprised of two groups: (A) those for whom optimal earnings are to the left of  $z^*$  under a lower marginal tax rate of  $\tau$  or are to the right of  $z^*$  under the higher marginal tax rate of  $\tau + d\tau$ , and (B) the bunchers who have optimal earnings to the right of  $z^*$  under the lower rate of  $\tau$  and to the left of  $z^*$  under the higher rate of  $\tau + d\tau$ . To estimate region B, we must subtract out Group A from the ex-post density, as is demonstrated in Panel B.

In Panel B, we divide the data into \$800 bins and estimate a seven-degree polynomial through the average density height within the bin. We exclude data near the kink by controlling for dummies for being in the seven bins nearest to the kink.<sup>14</sup> Our estimate of bunching,  $B$ , is the difference between the area under the empirical density in these seven bins and the area under the polynomial in this \$5600-wide region.<sup>15</sup> We estimate standard errors through a bootstrap procedure that we describe further in Appendix C.8 (and the results are similar under the delta method). Since we estimate bunching across different contexts, we report our bunching amount,  $B$ , normalized by the share of individuals in the

---

<sup>14</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). As we show in the Appendix, we have also experimented with other bandwidths and have estimated similar results.

<sup>15</sup>Our method is similar in spirit to that of Saez (2010), who subtracts the area under the empirical density to the left and right of the kink to estimate the area of B, but cannot easily accommodate curvature in the density. Our method is also similar to that of Chetty et al. (2010), who subtract a counterfactual density from the empirical density using a fixed-point method. Finally, Kleven and Waseem (2012) use similar techniques, but additionally allow for round number bunching in the counterfactual earnings distribution, above and beyond any kink or notch effects.

neighborhood  $[z^* - \delta, z^* + \delta]$  who belong to Group A (which we approximate as area under our polynomial over this range).

Some apparent limitations of our approach are worth discussion. First, 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 create earnings incentives near the bottom of the earnings distribution, such as Medicaid, Supplemental Security Income, Disability Insurance, or taxes such as unemployment insurance payroll taxes. While we acknowledge that other incentives represent a concern in principle—shared by 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>16</sup> The results show very clear evidence of bunching at the AET kink and no visible, systematic evidence of bunching in other regions.

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 ET, 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. 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. Nonetheless, as we discuss in greater detail later, the procedure we introduce can be interpreted as estimating the average elasticity and adjustment cost across a wide range of heterogeneous parameter values.

## 4 Data

We primarily rely on the restricted-access Social Security Administration Master Earnings File (MEF) and Master Beneficiary Record, described more fully in the Appendix. The data

---

<sup>16</sup>We have found that many other incentives, including income tax rates, are smooth on average around the ET convex kink.

contain a complete longitudinal earnings history with yearly information on earnings since 1951, the type and amount of Social Security benefits an individual receives in each year, year of birth, the year that claiming began (if any), and sex (among other variables). Starting in 1978, the measure of earnings in the MEF reflects total wage compensation, as measured on Internal Revenue Service (IRS) forms. Prior to 1978, the data were based on annual FICA earnings. For the years we examine in our results, separate information is available on self-employment earnings and non-self-employment earnings. In the main sample we pool data on men and women, since we have found similar patterns of adjustment in each sample separately. Our dataset contains a 1 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 have large sample size and are subject to little measurement error. Second, earnings (as measured in the dataset) are the Social Security tax base and are not subject to manipulation through deductions such as 401(k) contributions, credits, or other items. Third, 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>17</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 (capped) 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.

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). The data are based on unemployment insurance earnings records and longitudinally follow workers' earnings over time. The data have information on 90 percent of workers in covered states and their employers, though we are only able to use data on a 20 percent random subsample.

---

<sup>17</sup>We focus on a fixed sample to hold the sample constant across ages or years; as we describe later, we also investigate a number of other samples for robustness.

In covered states, the only workers not in the data are the self-employed and some state and local government workers. 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 sample size we obtain in the SSA data, and because the LEHD has quarterly data (in contrast to yearly SSA data).<sup>18</sup>

## 5 Earnings Response to Policy

We begin our empirical analysis with descriptive evidence on the nature of adjustment to policy changes in the AET. Our primary focus will be on the speed of adjustment following relatively sharp changes in policy with varying degrees of anticipation. We compliment this analysis with other evidence on more slowly evolving policy changes in the AET. Finally, we consider various robustness checks, across alternative samples and definitions of policy parameters.

### 5.1 Speed of Adjustment: Descriptive Evidence from Policy Variation Across Ages

We first examine the pattern of excess bunching across ages. Subsequent to 1982, ages 62-69 are subject to the AET. Thus, the treatment is activated for those reaching age 62 and shut down upon reaching age 70. The policy changes at ages 62 and 70 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>19</sup> Figure 4 plots earnings histograms

---

<sup>18</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 the age of 65. In our SSA data, 89 percent of people claim Social Security by age 65 during 1990 to 1999, and 97 percent claim by age 69, so it is a safe assumption that the great majority of the individuals observed in the LEHD data over age 65 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 are likely somewhat dulled by the inclusion of non-claimants in the sample. However, our primary interest concerns the patterns of responses to the Earnings Test across ages and over time, which prove to be visually and statistically clear in the LEHD.

<sup>19</sup>We examine other periods in Appendix E and obtain qualitatively similar results. The 1990-1999 period is relatively recent and affords us a relatively long, contiguous period over which the AET parameters were

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 red line.<sup>20</sup>

Figure 4 shows clear visual evidence of substantial excess bunching from ages 62-71 (inclusive).<sup>21</sup> Figure 4 also 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 ( $p < 0.01$  except at 65 and 71, where  $p = XX$  and  $XX$ , respectively). We find no robust evidence of adjustment in anticipation of future changes in policy, as those younger than 62 do not bunch.<sup>22</sup> We do find some evidence that unbunching may take more than one year, however, as those at ages 70 and 71 show some evidence of bunching.

At age 65, there is a small spike in the histogram to the left of the kink, whereas such a spike does not occur at other ages. One explanation is that this relates to the fact that the location of the kink changes 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. Earnings may spike to the left of the age 65 kink due to the fact that individuals have trouble adjusting their earnings to the new, higher kink within one year. This provides suggestive evidence that individuals are able to adjust to this policy change within two years, from age 64 to age 66; earnings appear smooth to the left of the kink at age 66 and older.

In Figure 4, bunching is positive and significant at age 70 ( $p < 0.01$ ) (and significantly less than at age 69 ( $p < 0.01$ )) and 71 ( $p < X$ ), but falls to levels statistically indistinguishable from zero by age 72 and above—indicating that complete adjustment to the expiration of AET at age 70 occurs by age 72.<sup>23</sup> The amount of bunching slowly rises from age 62 to 64,

stable. At the same time, we note that the Delayed Retirement Credit changed over this period (though this should not have affected the AET MTR for those near the kink as explained above).

<sup>20</sup>For ages younger than 62, we define the 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 kink in a given year as the kink that applies to post-NRA individuals in that year.

<sup>21</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) show, individuals tend to locate just below the kink much more frequently than just above the kink.

<sup>22</sup>If the cost of adjustment rose with the size of adjustment and this relationship were convex, we would expect anticipatory adjustment.

<sup>23</sup>However, 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 through nearly the entire year when she is classified as age 70 in our data. As a result, her yearly

which could suggest gradual adjustment. However, this could 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 substantially from age 62 to 64. To address this issue, in Appendix Figure E.1 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 64 (though less steep than in Figure 4).<sup>24</sup>

Similar patterns of adjustment occur when looking at the periods 1972-1982, 1983-1989 and 2000-2006 (Appendix Figures E.2 to E.4). In all cases, adjustment appears to take at most three years. However, we do find evidence of adjustment delays, as individuals continue to bunch at the kink at ages older than 69. While we interpret these patterns as responses to the onset and conclusion of the AET, these effects are not separately identified from general heterogeneity across age groups. Cutting against this is the evidence from Appendix Figure E.4. After 2000, individuals age out of the AET at age 65 rather than 70, the previous link between age and the timing of the AET is dramatically different, and yet we find similar patterns.

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

We next examine the speed of individuals' adjustment to a legislated change over time 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 applied to workers' earnings in October 2000 after having been passed in April 2000, and discussions prior to 2000 did not widely anticipate these changes.<sup>25</sup>

---

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. Despite this issue, it is nonetheless the case that significant bunching occurs at age 71, which cannot be due to this coarse measure of birth dates. Thus, while we must be careful to determine whether measured bunching at age 70 reflects a lack of adjustment or our coarse measure of age in the basic results, the overall results do suggest some delay in complete adjustment.

<sup>24</sup>We reiterate that these individuals face a notch in the budget set as opposed to a kink.

<sup>25</sup>The AET was also eliminated in 1983 for individuals aged 70 and 71, generating a similar test. 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.

Figure 5 shows the results starkly 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 no bunching visible, and this lack of bunching persists after 2000. Figure 5 also shows the amount of excess bunching estimated by year, along with 95 percent confidence intervals.<sup>26</sup> It is clear that the amount of bunching is significantly greater than zero in all years prior to 2000, and that estimates for 2000 and subsequent years show no significant bunching, with point estimates close to zero. Appendix Figure ?? shows bunching in the LEHD in 1999, 2000, and 2001. A spike at the kink is easily visible in 1999, and a small amount of bunching may be visible in 2000. By 2001, there is no visual evidence of bunching at the kink.<sup>27</sup>

These findings again suggest rapid adjustment—in this case, within one or two years. 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 at most three years. However, 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. 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 about their pay.<sup>28</sup> We return to these issues later, when we document substantial heterogeneity in the population in responsiveness to the AET.

In Appendix D, we discuss evidence on the speed of adjustment from quarterly earnings

---

<sup>26</sup>We have estimated this amount of excess bunching relative to four 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 setting the placebo kink equal to the exempt amount that had been scheduled prior to the 2000 legislation to take effect in each year; and 4) by using the exempt amount in a given year applicable to pre-NRA individuals. Given the lack of bunching visible in the histograms in 2000 and after, it is unsurprising that all of these methods show no significant bunching in these years.

<sup>27</sup>It should not be surprising that a small amount of bunching is visible in the LEHD data even though it is not visible in the SSA data, as the sample size is much larger in the LEHD. The standard errors on the 2001 bunching estimate in the LEHD are much smaller and allow us to rule out substantial amounts of bunching in this year.

<sup>28</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. This may have changed the counterfactual amount of bunching in 2000, even absent the elimination of the ET in this year. Nonetheless, it is worth emphasizing that this amount of bunching is unlikely to have been zero in the absence of the ET elimination, as the quarterly LEHD data show substantial evidence of bunching in quarterly earnings data prior to the fourth quarter of 2000, when the ET was eliminated.



data in the LEHD. This evidence has strengths and weaknesses relative to yearly SSA data. The LEHD data are quarterly and thus show a finer level of temporal disaggregation than the yearly SSA data. However, we show the LEHD quarterly data in the Appendix rather than the main text 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. In Appendix Figure ?? there is a small but significant amount of bunching in the quarter of the policy change (i.e. the fourth quarter of 2000) and in the following three quarters ( $p < 0.01$ ) but no significant bunching in subsequent quarters.<sup>29</sup> 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.

### 5.3 Other Evidence on Changes in Bunching

An important set of policy changes over time concerns changes in the Delayed Retirement Credit, which was introduced in 1972 at 1 percent, expanded to 3 percent in 1982, and expanded from 3 percent to 8 percent over the course of 1990-2008. However, Figure 6 shows that there is no sharp change in the amount of bunching around either 1972 or 1982, suggesting 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 steadily over this period.<sup>30</sup> A final piece of evidence concerns adjustment to the decrease in the AET

---

<sup>29</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 even though we do not estimate statistically significant bunching in the MEF in 2000.

<sup>30</sup>For example, the AET threshold amount rose much faster in the 1996-1999 period than in the previous period, and individuals may find it difficult to adjust earnings to a rapidly-increasing kink (or elasticities and adjustment costs may differ in this segment of the earnings distribution). It is possible that this helps to explain the decrease in the amount of bunching observed in these years.

marginal tax rate from 50 percent to 33 percent in 1990. As shown in Figure 6, the amount of bunching fell slightly from the years before 1990 to the years after 1990. We return to this evidence later when we use this policy change to estimate elasticities and adjustment costs.

We also conduct a variety of robustness tests. Previous work has demonstrated very different patterns of bunching among the self-employed and non-self-employed (Chetty et al. 2011) and also that sharp bunching in response to the US tax code is primarily driven by the self-employed (e.g. Chetty et al. 2012b). In Appendix Figure E.5 we show histograms for 1983-1989 of bunching among those with self-employment income. Perhaps surprisingly, there is much less bunching among these individuals. It follows that our results present the first evidence of significant bunching in the US in response to kinks among the non-self-employed. Given third-party verification of wage and salary income, these patterns are not likely to be driven by reporting responses. In Appendix Figure E.6, we also show that the histogram of bunching is similar for men and for women. Finally, it is possible to hypothesize that those with short expected lifespan should disproportionately bunch near the kink. For example, 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 and therefore lead to greater bunching at the kink among these individuals (though as we note, the DRC only takes effect at high earnings relative to the exempt amount). In Appendix Figure E.7, we show graphs illustrating that life expectancy is smooth near the kink, suggesting little evidence for such a mechanism.

Appendix Figure E.9 uses a bandwidth of \$500 instead of \$800, generating little change to our estimates (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 variations the order of polynomial we have tried have also shown similar results. In Appendix Figure E.11, we vary the region around the kink we exclude when estimating the amount of excess bunching (from \$2000 in the baseline to \$4000) and again estimate similar results. Limiting the sample to those who have substantial benefits (such as those with \$1000 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>31</sup>

---

<sup>31</sup>Since 1978, the earnings test has been assessed on yearly earnings, implying that we analyze the appro-

## 6 Mechanisms

This section probes the mechanisms that underlie patterns of adjustment to the AET, revealing patterns that suggest heterogeneity in responsiveness to the AET. Specifically, we examine which parts of the earnings distribution adjust to AET changes 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 claimants' earnings change from age 69 to age 70, when the AET is removed. We examine the period 1990-1999, during which the AET applied to individuals aged 62-69.<sup>32</sup> Figure 7 shows the mean percentage change in earnings 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 increase their earnings substantially from age 69 to age 70.<sup>33</sup>

This finding suggests that individuals locating near the kink at age 69 are different than other individuals in some way. The AET applies not only to individuals locating at the kink, but also to individuals in the neighborhood to the right of 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 among these individuals, as well. The fact that we do not suggests a degree of heterogeneity in adjustment costs or elasticities.<sup>34</sup> This finding raises the possibility that even though we

---

priate 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.

<sup>32</sup>1990-1999 represent natural years to investigate because the ET had a constant MTR and applied to the same age group throughout this period, and 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. Addressing this, for example by holding the sample constant in a subset of these years, yields very similar results.

<sup>33</sup>This spike in earnings growth in Figure 7 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. Meyer and Rosenbaum, 2001).

<sup>34</sup>The income effect of the AET also rises with income, which would also lead the mean percentage earnings

have observed relatively fast adjustment to the ET, this does not necessarily imply that mean adjustment costs in the population are small. For example, those initially locating at the kink may have low adjustment costs and react to the AET removal quickly, but those who never bunch at the kink to begin with may have higher adjustment costs.<sup>35</sup>

In Appendix Figure E.8, 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>36</sup> This again suggests that certain individuals are particularly responsive to the incentives created by the kink, in the sense that they serially bunch at the kink.

To get a better sense of the initial earnings of the bunchers we study, we examine more closely how the distribution of earnings differs across ages close in proximity but dramatically different in terms of AET incentives. Appendix Figure E.12 stacks the estimated distribution of earnings at ages 60, 61, and 62, as well as 69, 70, and 71. While the distribution of earnings changes modestly from year to year (such as from age 60 to 61, shown in the Figure) due to factors unrelated to the ET, the age 62 distribution shows a sharply different pattern than the age 60 or 61 distribution, with a sharp spike at the kink (particularly to the left of the kink), higher density immediately to the right of the kink, and generally lower density at earnings levels starting around \$5000 greater than the kink until meeting the age 61 distribution well above this level—consistent with the hypothesis that those bunching in the vicinity of the kink tend to come from a broad region well to the right of the kink.<sup>37</sup> Similarly, the age 69 distribution of earnings shows a sharply higher earnings density than the age 70 distribution

---

increase to fall as income rises. 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 at this lower level as earnings rises—consistent with the hypothesis that those initially locating at the kink are more responsive to the AET.

<sup>35</sup>The pattern is also consistent with such heterogeneity in elasticities, in addition or as opposed to adjustment costs.

<sup>36</sup>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 ET 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.

<sup>37</sup>A graph of the age-61 distribution of earnings conditional on locating in the vicinity of the kink at 62, and the graph of the age-70 distribution of earnings conditional on locating in the vicinity of the kink at age 69, show results consistent with this.

in the immediate region of the kink (particularly to its left) but shows a lower density than at age 70 at higher earnings levels, and eventually meeting a similar earnings density starting around \$6000 above the kink.<sup>38</sup> We return to this pattern of adjustment when motivating our model of adjustment costs below.

While our evidence suggests that the bunchers are in many ways much more responsive than others earning near the kink, we may still fail to pick up on more diffuse earnings responses among the latter group. 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. As a first step of addressing the possibility of diffuse bunching we vary the bandwidth that we choose for estimating excess bunching (in order to capture bunching over a wider range), as in Appendix Figure E.9.

## 6.2 Employers and the AET

We use the LEHD data to investigate a possible role for the employer in the responses to the AET. In Figure 9, 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 at which people face the ET, 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. This result is suggestive—though not definitive—evidence that excess bunching is accounted for in part by those with flexible work arrangements at their job, such as being paid hourly.<sup>39</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

---

<sup>38</sup>Note that some adjustment takes place after age 70; comparing the age 69 and age 71 or 72 distributions shows qualitatively similar patterns.

<sup>39</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, which in combination with the evidence that they are in industries that tend to be paid hourly provides suggestive evidence that those at the kink may have flexible work arrangements at their jobs. We have also 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.

earnings from age 69 to age 70. Thus, it appears that those who adjust by bunching are in part able to do so within a flexible working arrangement at their current employer.

Chetty et al. (2011) argue that employers drive a significant share of the bunching at kink points observed in Denmark through the equilibrium menu of earnings choices offered to workers. They observe that some individuals bunch at kinks even though they are not directly subject to the policy that creates the kink, concluding that these individuals bunch at the kink because employers create jobs that have those earnings levels. Phrased differently, some individuals bunch at kinks because their employer presents them a limited menu of earnings levels (including the kink earnings level) and would face costs of adjusting their earnings to a different level. We explore this possibility by testing for bunching among workers too young to claim OASI benefits, and therefore unaffected by the ET.

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. We extend this analysis by estimating bunching over the entire age distribution in the pre-1972 period, as Figure 8 shows. In this era, there was no Delayed Retirement Credit, and therefore, the incentives to bunch are arguable the strongest in our data set. This figure shows no significant excess bunching for ages younger than those subject to the ET. We interpret this as evidence that even in a period when employers may have faced the greatest incentive to coordinate workers at the kink, employers do not appear to drive the response to the AET in this way.

## 7 Estimating Elasticities and Adjustment Costs

The results thus far suggest a role for adjustment frictions in individuals' responses to policy changes, such as in generating slow adjustment to the disappearance of the kink at age 70, and in explaining the lack of response to the AET among those initially locating in the region above the kink area and still subject to the AET. As a first step in incorporating these frictions into an estimable model of earnings supply, we build upon the Saez (2010) model. That model uses bunching,  $B$ , to identify the elasticity of taxable earnings with respect to

the net-of-tax rate, and we extend it to allow for a cost of adjusting to tax changes.<sup>40</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 changes in the net-of-tax rate, we can jointly identify elasticities and adjustment costs. As discussed in Chetty (2012), Chetty et al. (2011) and Chetty et al. (2010), these parameters—adjustment costs and the elasticity that would characterize individuals’ behavior in the absence of adjustment costs—are relevant to the welfare cost of taxation.<sup>41</sup>

Our model relies on features of the empirical results that we have documented in the previous two sections. First, the descriptive results we have documented show evidence of adjustment frictions, which we model through a cost of adjusting earnings. Second, the empirical results suggest that employees’ choices are primarily responsible for patterns of bunching; this motivates a model in which employees choose their earnings. Third, we allow for heterogeneity in elasticities and adjustment costs, consistent with the results suggesting that individuals at the kink are particularly responsive to changes in the AET. Agents maximize utility over consumption and earnings (earning causes disutility)  $u(c, z; n)$  subject to a budget constraint  $c = (1 - \tau)z + R$ , where  $R$  is virtual income and the parameter  $n$  captures heterogeneity in the tradeoff between consumption and earnings supply.<sup>42</sup>

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 existing 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).<sup>43</sup> Our finding in Appendix Figure E.12

---

<sup>40</sup>Formally, the elasticity of taxable earnings with respect to the net-of-tax rate is defined as  $\epsilon \equiv -(\partial z / z) / (\partial \tau / (1 - \tau))$ .

<sup>41</sup>See Cogan (1981) on fixed costs of adjustment on the extensive margin of labor supply choices.

<sup>42</sup>In Appendix C we describe the model in more detail.

<sup>43</sup>Appendix Figure X shows that individuals do increase their earnings substantially from age 69 to age 70, as we would predict from such a model. However, 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

that the distribution of earnings at ages 62 and 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, is consistent with a simple model with fixed costs of adjustment that leads to a region of inaction and a region of adjustment. In Appendix C.9, 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 quantifying deviations from a frictionless model of earnings. Our first approach, a comparative static approach, relates attenuation in the *level* of bunching to the magnitude of adjustment costs. Our second approach, a sharp tax change approach, relates attenuation in the *change* in bunching to the magnitude of adjustment costs.

Figure 10 illustrates how a fixed adjustment costs affects responses to a tax change in each case. In Panel A, we demonstrate the comparative static approach, where the marginal tax rate above  $z^*$  increases from  $\tau_0$  to  $\tau_1$ . Recall that our frictionless model in Section 3 predicts that the set of bunchers have initial earnings, i.e. in the absence of a kink, in the range  $[z^*, z^* + \Delta z_1]$ . Adjustment costs attenuate the level of bunching from below. We have drawn the indifference curves for an individual with initial earnings  $\underline{z}_1 \in [z^*, z^* + \Delta z_1]$  at point 0 when there is no kink. After the kink is introduced, this individual finds herself at point 1. Because this individual now faces a higher marginal tax rate, in the absence of an adjustment cost she would like to reduce her earnings to the kink at  $z^*$  marked by point 2. In order to do so, however, she will have to pay the adjustment cost. 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>44</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  $\underline{z}_1$  and moving to the kink earnings

---

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 bulk of the earnings distribution at age 70 should be closer to the kink than if individuals with low fixed costs of adjustment tend to have high elasticities.

<sup>44</sup>This is true, for example, if utility is quasilinear. 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. We explore the implications of this assumption in Appendix C.4.



level  $z^*$  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. - v.), whereas in the absence of an adjustment cost, individuals with initial earnings in the range  $[z^*, z^* + \Delta z_1]$  bunch (areas i. - 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,  $\underline{z}_1$ , is implicitly defined by the indifference condition drawn in Figure 10, 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. In Appendix C.8, we describe in more detail our method of solving this system of equations using a minimum distance estimator, as we will implement using data from different kinks created by AET policy change. In particular, we need at least two kinks to identify our parameters. In Panel B, we show the predicted bunching amount under a counterfactual kink where the marginal tax rate above  $z^*$  is instead  $\tau_2$ , where  $\tau_0 < \tau_2 < \tau_1$ . Bunching in this case is lower and is comprised 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. - iii.).

We can use a series of approximations to derive a tractable expression relating the elasticity and adjustment cost to the level of bunching, and to build intuition regarding our minimum distance estimation procedure. 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 C.5 that scaled bunching is approximately:

$$b_1(\boldsymbol{\tau}_1, z^*; \varepsilon, \phi) = \varepsilon \left( z^* \frac{d\tau_1}{1 - \tau_0} \right) - \phi \left( \frac{1}{d\tau_1} \right), \quad (3)$$

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 amount of the tax change. This generalizes and nests the formula developed in Saez (2010), which is equivalent in the case where there is no adjustment cost and we therefore drop the second term (i.e.  $-\phi \left(\frac{1}{d\tau}\right)$ ).

Equation (3) also shows the features of the data that allow us to identify  $\varepsilon$  and  $\phi$ . Again, focusing on the first term, the Saez (2010) formula describes how bunching should vary between two different kinks in a frictionless model. The extent to which observed bunching deviates from this pattern is attributed to frictions via the second term. We need to observe bunching at two or more kinks, with variation in the change in tax rate  $d\tau$ .<sup>45</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, consider a regression of  $b$  on  $z^*(d\tau/(1-\tau))$  and  $1/d\tau$ , with the constant omitted. The coefficient on the first term is  $\varepsilon$ , and the coefficient on the second term is  $\phi$ .

We now turn to our second approach to quantifying adjustment costs. Let  $K_1$  and  $K_2$  be kinks with 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 direct transition from  $K_1$  to  $K_2$  and assume  $d\tau_2 < d\tau_1$ . Our previous comparative static approach would predict that the range of bunchers would decrease from  $[\underline{z}_1, z^* + \Delta z_1]$  (areas ii.-v. in Figure 10 Panel B) to  $[\underline{z}_2, z^* + \Delta z_2]$  (area iii.). However, since we are moving from an initial kink to a smaller kink, we must take into account the pre-existing bunching. It turns out that in this case, the fixed adjustment cost not only attenuates bunching in a cross-section, but also attenuates the change in bunching between two cross-sections in response to a change in the size of the kink. The first source of attenuation 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 frictionless optimum under  $K_2$  is  $z^*$  for everyone initially earning in the range  $[z^*, z^* + \Delta z_2]$ .

---

<sup>45</sup>Variation in  $z^*$  alone would let us identify  $\varepsilon$  and  $\phi$ , but this identification would come solely from functional form.

The second source of attenuation is drive by individuals in area iv. of Panel B. Panel C of Figure 10 demonstrates this. At point 0, we have drawn an individual's initial earnings  $\bar{z}_0 \in [z^*, z^* + \Delta z_1]$  under no kink and 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. Now, we mute the kink by transitioning to  $K_2$ . Note that since initial earnings,  $\bar{z}_0$ , are greater than  $z^* + \Delta z_2$ , this individual would not have bunched had we gone directly from no kink to  $K_2$ . Her optimal earnings under  $\tau_2$  is at  $\bar{z}_2 > z^*$  (marked as point 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 likewise remain at the kink. Thus, bunching under  $K_2$ , now higher due to this inertia, is:

$$B_2(\tau_2, z^*; \varepsilon, \phi^*) = \int_{\bar{z}_1}^{\bar{z}_0} h(\zeta) d\zeta. \quad (4)$$

It follows that the change in bunching from  $K_1$  to  $K_2$  will be smaller (arear v. versus areas ii., iv. and v. in Panel B). In Appendix C.7 we show that scaled bunching under  $K_1$  will be as before in equation (3)<sup>46</sup>, whereas scaled bunching under  $K_2$  is approximately:

$$b_2(\tau_2, z^*; \varepsilon, \phi) = \varepsilon \left( \frac{d\tau_2}{1 - \tau_2} + 1 \right) \bar{z}_2 - \phi \left( \frac{1}{d\tau_1} \right), \quad (5)$$

and the critical earnings level  $\bar{z}_2$  is defined implicitly by the indifference condition in Panel C:

$$\phi^* \equiv u((1 - \tau_2)\bar{z}_2 + R_2, \bar{z}_2; \bar{n}) - u((1 - \tau_0)z^* + R_0, z^*; \bar{n}). \quad (6)$$

It is helpful to provide intuition for the procedures we describe. 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. Under our comparative static approach, suppose that instead we have two cross-sections of data featuring different changes in marginal tax rates at the kink. The change in the amount of bunching from

---

<sup>46</sup>In practice, we vary the assumption that pre-period bunching is attenuated, alternatively setting the pre-period bunching level to the frictionless amount described in Section 3.

one cross-section to the other will depend on the adjustment cost.<sup>47</sup> Thus, both the level of bunching (in a single cross-section) and the change in bunching from one cross-section to another give us information about the adjustment cost and elasticity, helping us to estimate both of these parameters. Turning to the sharp tax change approach, 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 post period. In the extreme case where a kink has been eliminated, we can attribute any residual bunching to adjustment costs. The former approach is perhaps best suited for comparing bunching after two different kinks have been introduced and adjustment has settled into its steady-state, while the latter is better suited to analyze within kink variation in the marginal tax rate jump or the elimination of a kink, preferably just before and after the policy change.

A few additional points regarding the methodology above are worth noting. First, although the closed-form solutions in (3) and (5) help illuminate the intuition behind our methodology, we do not use these simplified approximations in our actual estimation. Instead, we make use of a minimum distance estimator to solve the nonlinear system of equations laid out in Appendix C.8. The key assumptions underlying that method are (1) we assume that the density of initial earnings  $h(z)$  is uniform over the range  $[z^*, z^* + \Delta z^*]$  and (2) we use a quasi-linear, isoelastic utility function. The first assumption is commonly made in the bunching literature (see Chetty et al. (2011) or Kleven and Waseem (Forthcoming), for example) as is the functional form for utility (see Saez (2010), Chetty et al. (2011) and Kleven and Waseem (Forthcoming), for example).<sup>48</sup> Second, the expository derivations in (3) and (5) do not impose quasilinearity, but similarly make use of the uniform density assumption and a first-order approximation for utility in the neighborhood of the kink.

## 7.1 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  $(\varepsilon_i, \phi_i, n_i)$

---

<sup>47</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.

<sup>48</sup>Note, we also relax the first assumption, and instead estimate a log-normal distribution of earnings,  $z$ , as explained in Appendix C.8.

be jointly distributed according to a smooth CDF  $G(\cdot)$ , which translates into a smooth, joint distribution of elasticities, fixed costs and earnings  $\tilde{H}(z, \varepsilon, \phi)$ . We again assume that the density of earnings,  $\tilde{h}(z, \varepsilon, \phi)$ , is constant over the interval  $[\underline{z}, z^* + \Delta z^*]$ . We now have a generalized expression for observed bunching:

$$\begin{aligned} B &= \iiint_{\underline{z}}^{z^* + \Delta z^*} \tilde{h}(\zeta, \varepsilon, \varphi) d\zeta d\varepsilon d\varphi \\ &\approx h(z^*) \cdot \left[ \bar{\varepsilon} \left( z^* \frac{dt}{1 - \tau} \right) - \bar{\phi} \left( \frac{1}{d\tau} \right) \right], \end{aligned} \quad (7)$$

where we have used approximations for  $\Delta z^*$  and  $\underline{z}$  (as described in greater detail in Appendix C.6), and  $h(z^*) = \iint \tilde{h}(z^*, \varepsilon, \varphi) d\varepsilon d\varphi$ .<sup>49</sup> Rearranging terms, we derive a generalized formula analogous to equation (3):

$$b = \bar{\varepsilon} \left( z^* \frac{dt}{1 - \tau} \right) - \bar{\phi} \left( \frac{1}{d\tau} \right).$$

The parameters  $\bar{\varepsilon}$  and  $\bar{\phi}$  are the average elasticity and adjustment cost for those who bunch at the kink.

## 7.2 Estimates of Elasticity and Adjustment Cost

To estimate  $\varepsilon$  and  $\phi$ , we first rely on a change in the AET benefit reduction rate in 1990. We begin with a graphical depiction of the patterns driving the results. Figure 6 shows excess bunching among 66-69 year-olds, for whom the benefit reduction fell from 50 percent to 33 percent in 1990. It is clear that in the 1986-1994 period, excess bunching generally fell, corresponding to the fall in the marginal tax rate.<sup>50</sup> When apply our "comparative static" approach, we use data from 1989 and 1991, avoiding the year of change in order to allow individuals sufficient time to adjust to the policy change. When we apply our "sharp tax change" method, we use data immediately before and after the tax change to make use of change in bunching for estimation purposes. Finally, we also apply the "sharp tax change" method to the case where individuals age out of the AET at age 69.

<sup>49</sup>We are thankful to Henrik Kleven for suggesting this interpretation in the presence of heterogeneity.

<sup>50</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.

As we describe in more detail in Appendix C.8, our main estimates using the comparative static approach are based on a solution for (1) and (2) in the case of quasilinear utility.<sup>51</sup> When we apply the sharp tax change approach, we similar solve equations (4) and (6). 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—and estimates of the implicit marginal tax rate associated with the AET. We begin by ignoring the baseline tax rate and – following Friedberg (1998, 2000) and the theoretical presumptions we have described – we assume that people treat the AET as a true tax, and we set  $d\tau_{pre} = 0.5$  and  $d\tau_{post} = 0.33$ . We then vary these assumptions in various dimensions – we substitute baseline tax rate of 0.15<sup>52</sup> or, following Colie, Gruber and Diamond (2002), we translate the 0.50 and 0.33 explicit tax rates to 0.36 and 0.24 respectively, after taking into account actuarial adjustments to future benefits.

Table 2 presents our results from the comparative static approach in Panel A. In the baseline case, in which  $\tau = 0.15$ ,  $d\tau_{pre} = 0.50$ , and  $d\tau_{post} = 0.33$ , we estimate an elasticity of 0.350 (standard error 0.XX) in Column (1) and an adjustment cost of \$114 (standard error \$XX) in Column (2), both significantly different from zero at the 1 percent level.<sup>53</sup> When we make the same assumptions about tax rates but constrain the adjustment cost to zero, as most previous literature has implicitly done, we estimate a slightly smaller elasticity of only 0.325 (standard error 0.XX) in Column (5). It makes sense that the estimated elasticity is lower when we do not allow for adjustment costs than when we do, as adjustment costs dull individuals' responses just as low elasticities do. However, in our case the adjustment cost do appear to matter significantly for the elasticity estimate. In the next three rows, we consider alternative specifications. When we use an earlier pre-period, we estimate slightly larger estimate, but similar qualitative patterns. When we dampen the tax changes to account for

---

<sup>51</sup>Numerical solutions in this literature typically impose quasilinearity (e.g. Saez 2010; Kleven, Landais, Saez, and Schultz 2012; Kleven and Waseem 2012; Chetty, Friedman, Olson, and Pistaferri 2011). If we were to relax the assumption of quasilinearity, we would need to specify outside wealth, which we do not observe in the data.

<sup>52</sup> $\tau = 0.15$  is the average effective marginal tax rate for 65 year-olds in 1985-1994, with earnings in the neighborhood of the kink, calculated using TAXSIM and the Statistics of Income tax return files.

<sup>53</sup>As we describe in Appendix C.8, we have estimated the standard errors using both the delta method and the bootstrap, which give similar results. The standard errors reported above are based on the bootstrap.

the DRC, we estimate larger elasticities. This makes intuitive sense – for the same behavioral response, if we assume a less pronounced tax change, we will infer a larger elasticity. Finally, raising the baseline tax rate slightly reduces the elasticity, but leaves the qualitative patterns unaltered. This also makes intuitive sense – for the same behavioral response and size of kink, raising the baseline tax rate increases the percentage change in the net-of-tax rate at the kink, and we therefore infer a smaller elasticity.

In Panel B, we apply the sharp tax change method. Across similar specifications as in Panel A, we arrive at remarkably similar quantitative and qualitative patterns, looking at Columns (1), (2) and (5). This is not a surprise, since the estimates rely on essentially the same variation in tax rates. In this case, we are able to demonstrate exactly where our identification is coming from. In Columns (4)–(6), we estimate an elasticity from bunching, assuming zero adjustment cost. Comparing Columns (5) and (6), we see that following the reduction in the size of the kink, the naive elasticity estimate actually *increases* from 0.332 to 0.535 in the baseline case. This corresponds nicely with our theory, which predicts that following a reduction in the kink, there will be excess bunching due to inertia (see area .iv in Figure 10, Panel B). Once we allow for an adjustment cost, this excess bunching is attributed to an adjustment cost.

In Panel C, we demonstrate an additional application of the sharp tax change method – it is able to use the disappearance of a kink to inform us about the adjustment cost. In this Panel, we focus on those workers that age out of the ET at age 70. In this case, the residual bunching at age 70 is informative about frictions. We focus on workers aged 69 and 70 in 1990–1999, where the relevant explicit tax rate is 0.33. We estimate a somewhat larger elasticity for this population, but observe similar patterns for the size and implications of the adjustment costs.<sup>54</sup>

---

<sup>54</sup>Additional robustness checks are provided in our appendix. We show the estimates under alternative assumptions about the tax rates; using a Tobit functional form for the distribution of earnings rather than a uniform distribution and changing the bandwidth from \$800 to \$500 (all of which are described in greater detail in the Appendix). Our results are relatively unaltered by these variations. We also consider estimates over alternative time spans and age ranges.

## 8 Conclusion

In the context of the Social Security 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 introduction and disappearance of convex kinks in the effective tax schedule. 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 to elasticities estimated in a medium-run time frame of a few years. However, the evidence showing delayed adjustment in some cases suggests that regressing earnings or taxable income on contemporaneous tax rates may not be the correct specification in some contexts.

Second, we investigate the mechanisms that underlie the patterns of adjustment. Adjustment to removal of kinks occurs primarily through higher than average 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. Furthermore, when this earnings growth occurs, those at the kink who adjust the most are disproportionately likely to remain within the same employer. We additionally investigate the extent to which firms may help coordinate bunching, but looking for bunching among those too young to be directly affected by the AET. The responses appear to be driven mainly by employees, as only those subject to the AET bunch at the AET kink. The combination of these pieces of evidence suggests that the bunching primarily results from the choices of certain particularly responsive employees who are able to flexibly vary their earnings while disproportionately staying at their job.

Third, motivated by these empirical findings, we specify a model of employees' earnings adjustment to policy in which we relate the mean earnings elasticity and the mean fixed adjustment cost to empirical estimates of bunching. When we consider different static levels of bunching using data from 1989 and 1991, we estimate that the mean elasticity is 0.35 and the mean adjustment cost is \$114. When we constrain adjustment costs to be zero, the elasticity we estimate (0.323) is only slightly smaller. When we use the change in bunching levels following a sharp tax change, we arrive at very similar results. The small



adjustment costs we estimate are consistent with our empirical observation that bunching responds rapidly to changes in policy. Nonetheless, our estimates demonstrate the potential importance of allowing for adjustment costs when estimating elasticities in general.

The analysis leaves open a number of avenues of further inquiry. First, we find heterogeneity in the speed of adjustment across contexts. As we have emphasized, many factors may differ across these contexts, including the degree to which policy changes are anticipated, the degree of publicity surrounding the policy changes, the ages affected, the calendar year, and the distribution of individuals' earnings relative to their desired earnings in the absence of adjustment costs. Interestingly, one might expect that unanticipated changes would be associated with the slowest adjustment (as individuals take time to incorporate new information in their behavior), yet individuals aged 66-69 appear to adjust to the 2000 elimination of the AET within one year at most—faster than the adjustment the anticipated change of aging out of the AET. Given that 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.

Second, , we present little evidence distinguishing among the possible reasons for reaction to the AET—such as liquidity constraints or misperceptions—and this 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 adjustment simply as a "black box" cost, without modeling the process that leads to this cost, such as information acquisition or job search. Future research could fruitfully model such processes—which could include explicitly modeling dynamic considerations relating to adjustment costs—and continue to 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. Studying the nature of 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*, ed. David Card, Richard Blundell, and Richard B. Freeman, 411–59. Chicago: 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." *The Quarterly Journal of Economics*, 126.
- Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber (2012a), "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.), volume 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 (2012b), "Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings." Working Paper, 1–75.
- Coile, Courtney, Peter Diamond, Jonathan Gruber, and Alain Joustien (2002), "Delays in

- claiming social security benefits.” *Journal of Public Economics*, 84, 357–385.
- 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.” *The 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*, ed. Henry J. Aaron and Joseph A. Pechman, 27–71. Washington, DC: Brookings Institution.
- 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,” London School of Economics Working Paper.
- Kleven, Henrik and Mazhar Waseem (Forthcoming), ”Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan”, *The Quarterly Journal of Economics*.
- Kline, Patrick, and Melissa Tartari (2013), ”What Distributional Impacts Mean: Welfare Reform Experiments and Competing Margins of Adjustment,” UC Berkeley Working Paper.
- 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.” *The 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(3): 317–28.
- Moffitt, Robert. 1990. ”The Econometrics of Kinked Budget Constraints.” *Journal of Economic Perspectives*, 4(2): 119–39.
- Saez, Emmanuel (2010), ”Do Taxpayers Bunch at Kink Points?” *American Economic Jour-*

nal: 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.

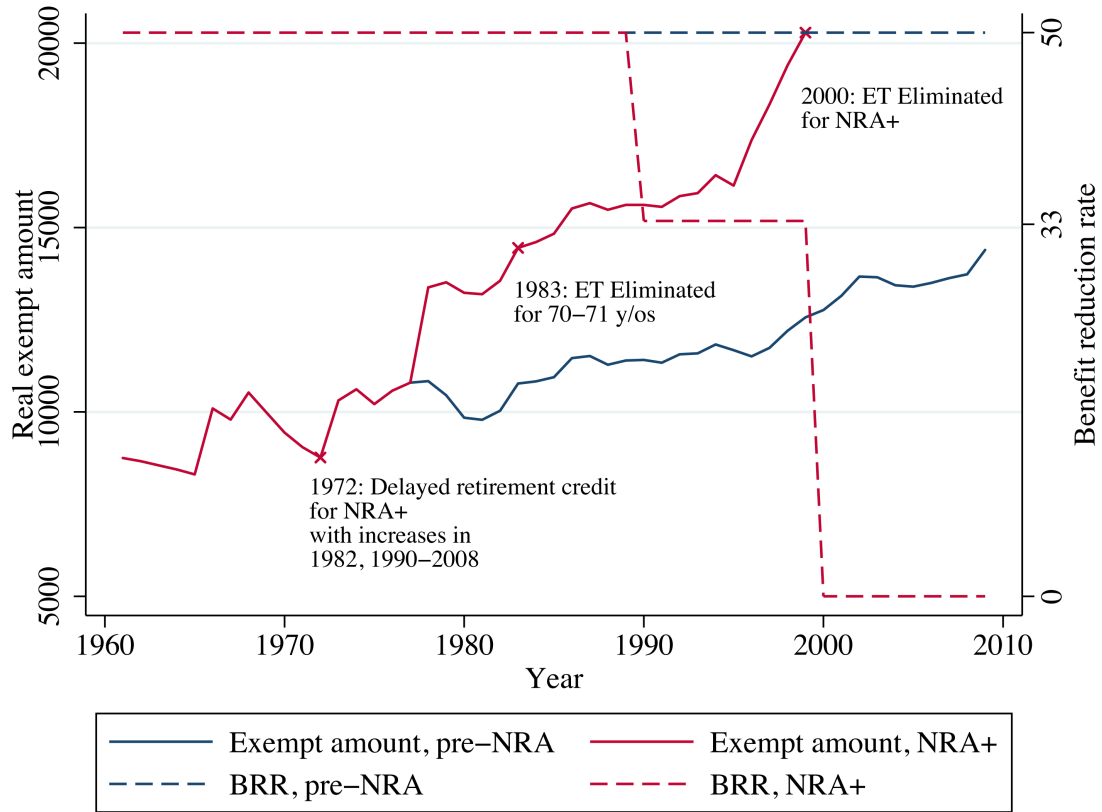
Simon, Herbert A (1955), "A Behavioral Model of Rational Choice." *The Quarterly Journal of Economics*, 69, 99–118.

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

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

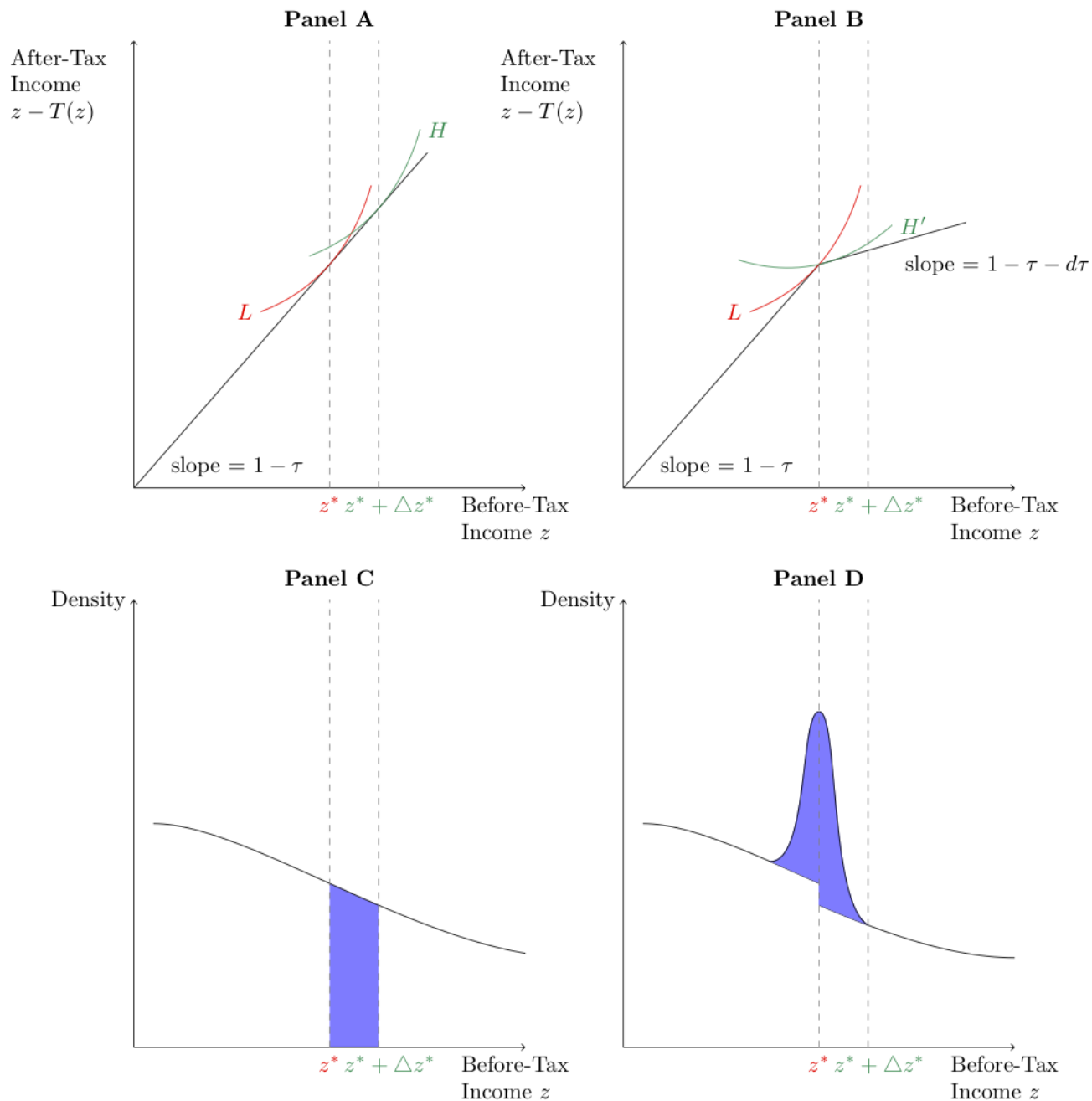
Song, Jae G and Joyce Manchester (2007), "New evidence on earnings and benefit claims following changes in the retirement earnings test in 2000." *Journal of Public Economics*, 91, 669–700.

Figure 1: Policy Variation: Earnings Test Rules over Time



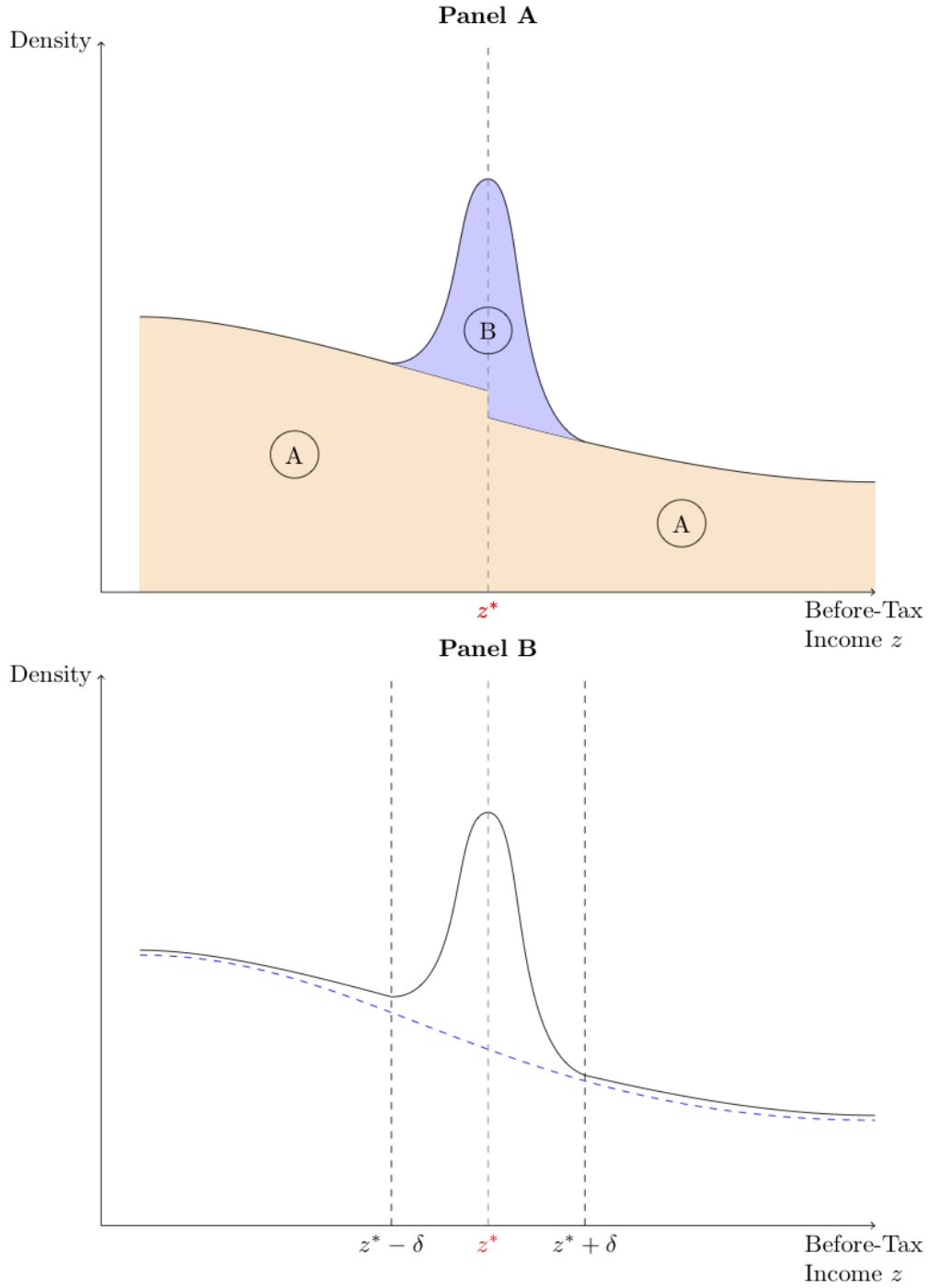
Note: The figure illustrates the primary relevant variation in Earnings Test parameters from 1961 to 2009. The left vertical axis measures the real value of the exempt amount over time, which is plotted by a solid line. The right vertical axis measures the benefit reduction rate in Social Security Payments for every dollar earned beyond the exempt amount. NRA refers to the Normal Retirement Age.

Figure 2: Bunching Response to a Convex Kink



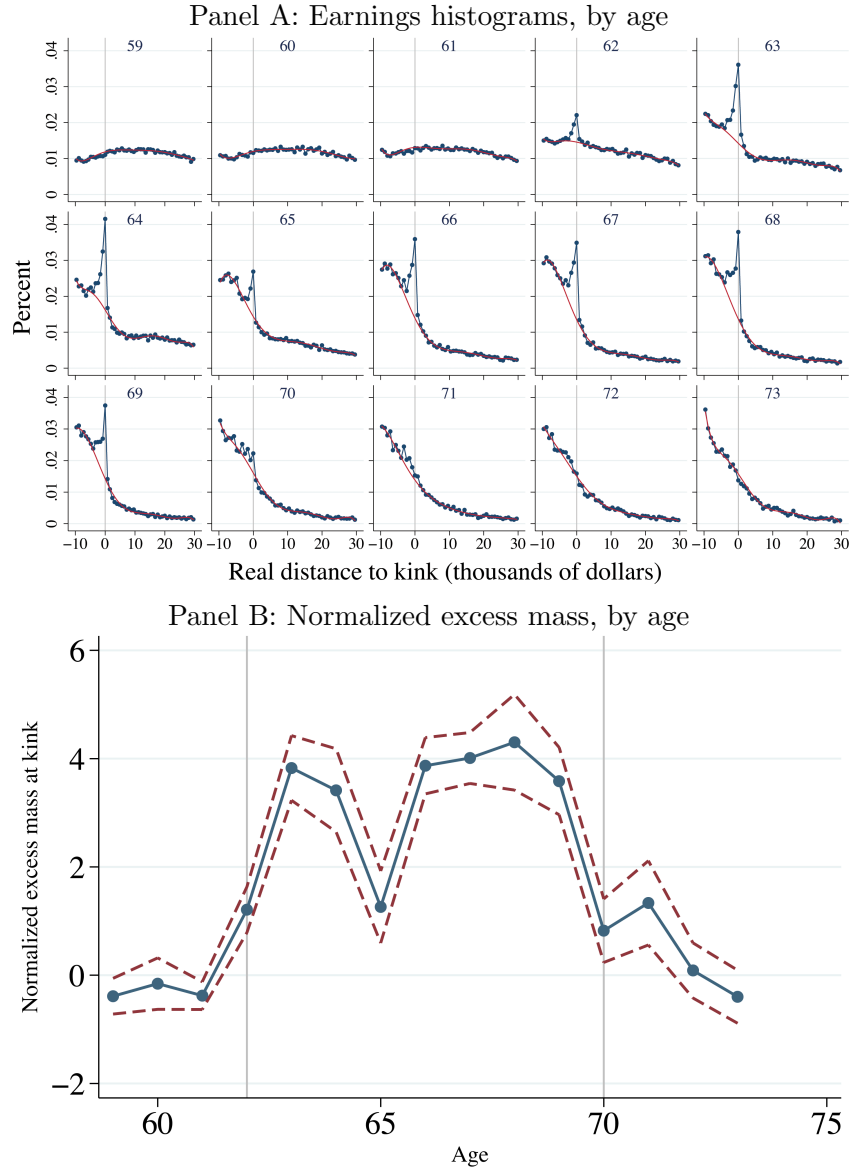
Note: Panels A and B show that when we move from a linear budget constraint to a budget constraint with a convex kink, a range of individuals choose to locate at the kink. Panels C and D show that as we move from a linear budget constraint to a constraint with a convex kink, a spike in the earnings density appears at the kink. Panel D shows that this spike may be smoother in the presence of a variety of factors, such as inability to control earnings precisely, as discussed in Saez (2010).

Figure 3: Bunching Estimation Methodology



Note: Panel A decomposes the ex-post distribution into three groups. Group A is comprised of those who locate to the left of the kink under the initial lower marginal tax rate  $\tau$ . The bunchers, group B, are those who bunch at the kink in the presence of the higher marginal tax rate  $\tau + d\tau$ . Group C is comprised of those who locate to the right of the kink under the higher marginal tax rate  $\tau + d\tau$ . Panel B demonstrates how, excluding data in a neighborhood of  $z^*$ , the distribution of earnings in the absence of the kink is estimated to recover the share in Group C.

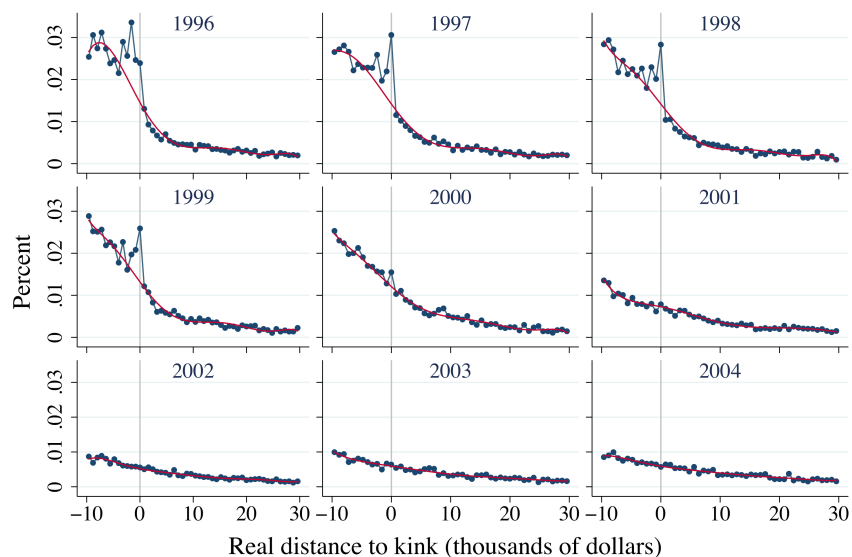
Figure 4: Adjustment Across Ages: Histograms of Earnings and Normalized Excess Mass, 59-73-year-old Social Security Claimants by Age 65, 1990-1999



Note: The figure shows earnings histograms (Panel A) and normalized excess bunching (Panel B) from a 1 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. In Panel A, the bin width is \$800, and each earnings distribution is centered on the level of earnings associated with the kink in each year; the earnings level 0, shown by the vertical red line, denotes the kink. Earnings are measured in CPI-U-adjusted year 2010 dollars in each year. "Claimants" in this and all other figures in the main text refers to those 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).

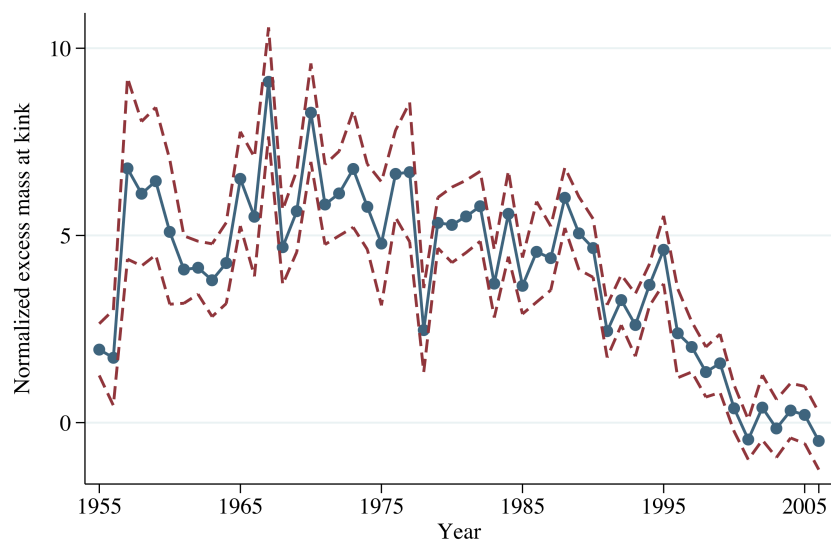


Figure 5: Adjustment Across Years: Histograms of Earnings and Normalized Excess Mass, 66-69 year old Social Security Claimants, 1996-2004



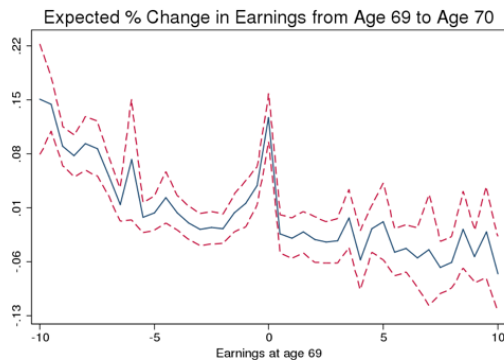
Note: The figure shows histograms of earnings from a 1 percent random sample of SSA administrative data on Social Security claimants aged 66-69 in each year from 1996 to 2004 (inclusive). See other notes to Figure 4 Panel A.

Figure 6: Excess Bunching by Year, 1955-2006



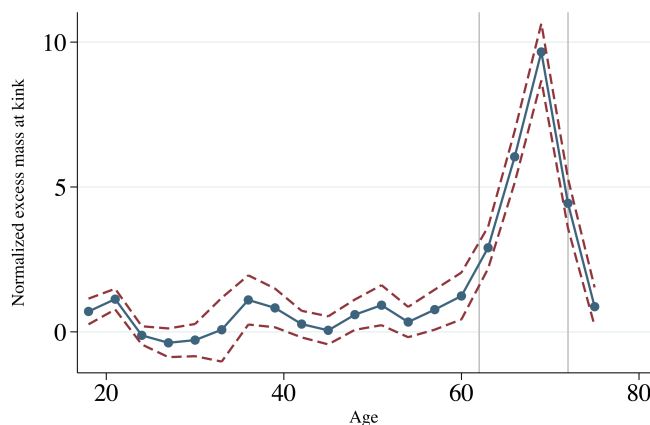
Note: The figure shows normalized excess bunching from a 1 percent random sample of SSA administrative data on Social Security claimants aged 66-69 in each year between 1955 and 2006 (inclusive). See other notes to Figure 4 Panel B.

Figure 7: Mean Percent Change in Earnings from Age 69 to Age 70, by Age 69 Earnings, 1990-1999



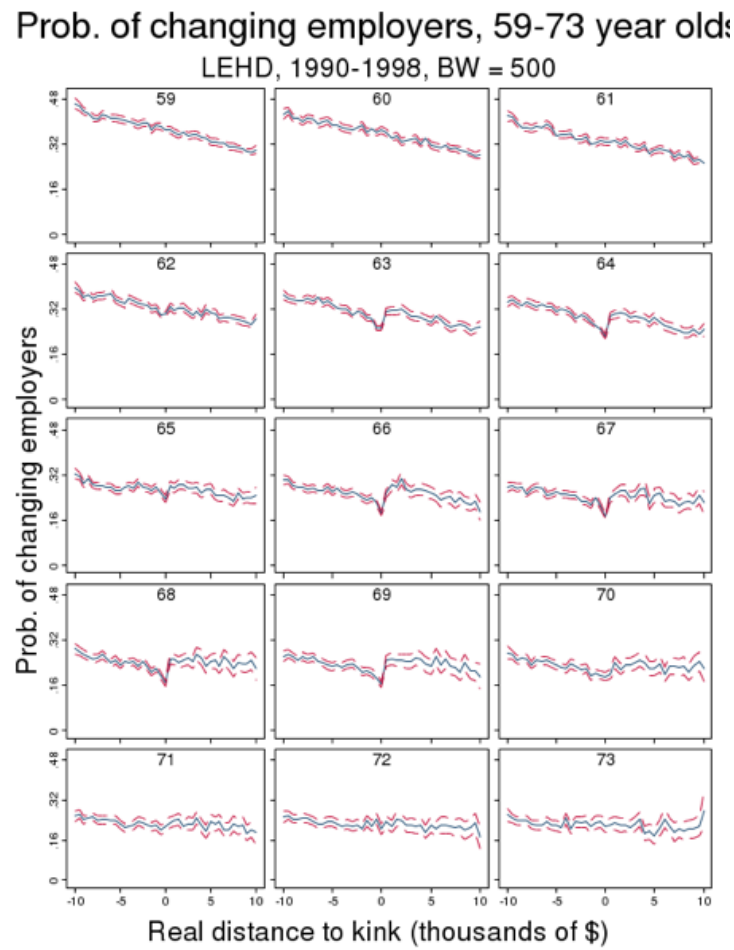
Note: The figure shows the mean percentage change in earnings from age 69 to age 70 in a 20 percent sample of the LEHD in 1990-1998 (y-axis), plotted against earnings at age 69 (x-axis). Earnings are measured relative to the kink, shown at zero with the vertical line. The bin width is \$800. See other notes to Figure 4 Panel A.

Figure 8: Normalized Excess Mass at Kink by Age among Social Security Claimants, 1966-1971



Note: The figure shows normalized excess bunching from a 1 percent random sample of SSA administrative data on Social Security claimants (i.e. who claimed by age 65) for each age between 18 and 80 (inclusive) for 1966-1971. The vertical lines show the ages at which the AET first applies (62) and the age at which the AET ceases to apply (71). Normalized excess bunching is calculated by grouping bins of three adjacent years of ages together (e.g. 18 to 21); doing so in the case of time periods post-1971 does not show statistically significant positive normalized bunching at ages younger than those subject to the Earnings Test. See other notes to Figure 4 Panel B.

Figure 9: Probability of Changing Employers from Age  $t$  to Age  $t+1$ , by Age  $t$  Earnings, 1990-1999



Note: The figure shows the percentage of workers who change employers from age  $t$  to age  $t+1$  in a 20 percent sample of the LEHD in 1990-1998 (y-axis), plotted against the earnings at age 69 (x-axis). See notes to Figure 7.

Figure 10: Bunching Responses to a Convex Kink, with Fixed Adjustment Costs

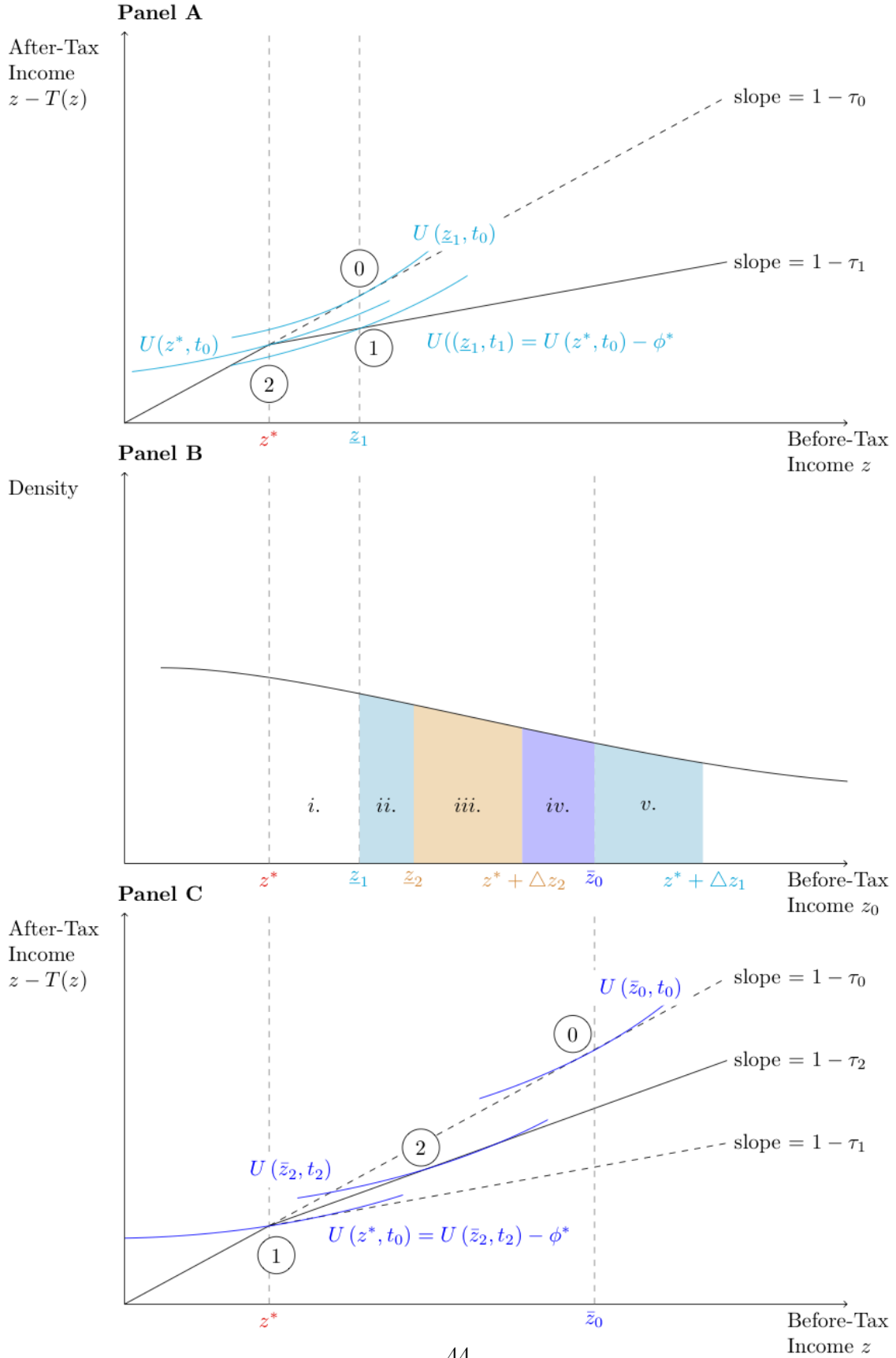


Table 1: Summary Statistics

	(1)	(2)
	Ages 18-75	Ages 62-69
Mean Earnings	35,352 (262,388)	28,843 (762,917)
10th Percentile	4,909	1,966
25th Percentile	14,510	7,276
50th Percentile	31,423	17,543
75th Percentile	44,441	36,682
90th Percentile	62,281	57,421
Share with Positive Earnings	0.45 (0.27)	0.41 (0.37)
Observations	15,840,014	1,684,122
Individuals	651,938	577,973

Note: The data in Panel A are taken from a 1 percent random sample of earnings records from SSA administrative data. The data for ages 18-75 covers all those in the sample in years 1955-2005 (inclusive). The data for ages 62-69 cover the main sample we use most often, the group of individuals who have claimed Social Security by age 65. Earnings are expressed in 2010 dollars.

Table 2: Estimates of Elasticity and Adjustment Cost

	(1)	(2)	(3)	(4)	(5)
<b>Panel A: Comp. Static Method</b>	$\varepsilon$	$\phi$	$\varepsilon \phi = 0$		
Baseline	0.350	\$113.78	0.325		
1988 vs. 1991	0.428	\$252.86	0.370		
$\tau_1 = 0.36; \tau_2 = 0.24$	0.559	\$128.75	0.497		
$\tau_0 = 0.1$	0.296	\$89.01	0.279		
<b>Panel B: 1990 Sharp Change Method</b>	$\varepsilon$	$\phi$	$\varepsilon \phi = 0$		
			Pooled	Pre-1990	1990
Baseline	0.348	\$105.40	0.377	0.332	0.535
	(0.026)	(55.30)	(0.027)	(0.30)	(0.044)
1988 to 1990	0.372	\$71.80	0.383	0.361	0.535
	(0.017)	(40.80)	(0.017)	(0.018)	(0.044)
$\tau_1 = 0.36; \tau_2 = 0.24$	0.537	\$62.90	0.581	0.516	0.781
	(0.040)	(38.30)	(0.041)	(0.047)	(0.065)
$\tau_0 = 0.15$	0.274	\$117.10	0.296	0.260	0.436
	(0.020)	(55.90)	(0.021)	(0.024)	(0.036)
<b>Panel C: Age 70 Sharp Change Method</b>	$\varepsilon$	$\phi$	$\varepsilon \phi = 0$		
Baseline	0.450	\$70.30	0.423		
	(0.029)	(42.30)	(0.028)		
69 vs. 70 yr olds	0.404	\$42.30	0.387		
	(0.033)	(32.80)	(0.035)		
$\tau_2 = 0.24$	0.654	\$47.60	0.617		
	(0.042)	(27.90)	(0.041)		
$\tau_0 = 0.15$	0.368	\$73.90	0.345		
	(0.024)	(45.70)	(0.023)		

Note: The table shows estimates of the elasticity and adjustment cost using the procedures described in the text, with bootstrap standard errors shown in parentheses. Panel A uses the "comparative static" method, with a baseline specification estimated using the sample of 66-69 year olds in 1989 and 1991. Panel B uses the "sharp tax change" method about the reduction in the benefit reduction rate in 1990, with a baseline specification estimated using the sample of 66-69 year olds in 1989-1990. Panel C uses the "sharp tax change" method at the phase out of the AET at age 70, with a baseline specification of 68-70 year olds in 1990-1999. All baseline specifications use a bin width of \$800, a uniform density and assume  $\tau_0 = 0$ ,  $\tau_1 = 0.33$  and  $\tau_2 = 0.50$ . Alternative specifications deviate from the baseline as noted. Columns (1) and (2) report joint estimates with an unrestricted  $\phi$ , while Columns (4)-(6) impose the restriction  $\phi = 0$ .

## A Data Appendix: Social Security Data and Earnings Test

Our data are a 1 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. Our data consist of a 1 percent random sample of Social Security Numbers (SSNs); we randomly extract SSNs from the database and follow each of these individuals over the full time period. The Earnings Test is based on earnings as measured in this dataset. Prior to 1977, the data have information on annual FICA earnings; since 1977, the data have information on uncapped wage compensation. Before 1977, 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 post-1977 data. 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. 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) persons 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 included in Medicare coverage and are therefore present in the MBR.

We construct our main sample as follows. Because the AET only affects people who claim Social Security, we do not want to include too many people who have not claimed. Nonetheless, the claiming decision is endogenous and may depend on people's awareness of the AET and their anticipated reaction to it. Our main sample therefore consists of individuals who claim in the year they turn 65 or earlier. Because we focus on the intensive margin response, we further limit the sample to observations with positive earnings in our main analysis.

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 through 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. Despite this issue, it is nonetheless the case that significant bunching occurs at age 71, which cannot be due to this coarse measure of birth dates. Thus, while we must be careful to determine whether measured bunching at age 70 reflects a lack of adjustment or our coarse measure of age in the basic results, the overall results do suggest some delay in complete adjustment.<sup>55</sup>

---

<sup>55</sup>In order to investigate whether our coarse measure of age could be driving the results at age 70, we investigated the pattern of bunching by age among those born in January, as well as those born in December. Those born in December show more bunching at the kink, but those born in January still show significant ( $p < 0.01$ ) bunching at the kink, indicating that some of the bunching at age 70 is not due to our coarse

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.

## B Data 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 contain 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 generally do not cover federal employees, most agricultural workers, many churches and nonprofits, and the self-employed.

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, and information about this firm’s NAICS code is available. We consider an employee to have changed their employer 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.

No earnings data are available before 1989 in any state. We select data from 1990-1999 in order to examine a time period in which the AET explicit marginal tax rate was constant. 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. We examine a 20 percent random sample of the original LEHD file, as this was the largest amount of data that our programs could handle on the limited server space available.

[Figures Forthcoming]

## C Model of Earnings Response

### C.1 Baseline Model

We start with a baseline frictionless model of earnings, following Saez (2010). We briefly sketch the key highlights of this model for comparison to our model with a fixed cost of adjustment. Individuals maximize utility over consumption,  $c$ , and costly earnings,  $z$ , with the tradeoff depending on an “ability” parameter  $n$ :

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

---

measure of age. Those born in other months show levels of bunching intermediate between those in January and December.



Heterogeneity is parameterized by  $n$ , 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:

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

## C.2 Linear Tax Schedule

We 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-Mirlees 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. Define  $H_1(\cdot)$  and  $h_1(\cdot)$  as the smooth CDF and density of earnings under a higher, constant marginal tax rate  $\tau_1$ . Under quasilinear utility, we have:

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

$H_1$  is defined similarly as a function of  $\tau_1$ .

## C.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 will have 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  $h(\cdot) = H'(\cdot)$ . 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  $[z^*, z^* + \Delta z^*]$ . 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*:

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

where  $d\tau = \tau_1 - \tau_0$ .<sup>56</sup>

## C.4 Fixed Adjustment Costs

We now extend the model to include a fixed cost of adjusting earnings. We assume that the adjustment cost reflects an additional disutility of  $\phi^*$  of changing 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  $n$ . This is true, for example, if utility is quasilinear. 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>57</sup> If the gain in utility is monotonically increasing in initial earnings, and

<sup>56</sup>This formula holds if there is a single elasticity  $\varepsilon$  in the population. We investigate cases with heterogeneity below.

<sup>57</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.

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.

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  $\phi^*$ , 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 change in the MTR to  $\tau_1 > \tau_0$  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, so there exists a unique level of initial earnings at which the agent is indifferent between adjusting and staying at the initial earnings level—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 constant adjustment of  $\phi^*$ , and  $z^* \geq z_{1,\phi}$ , then:*

1. *Individuals with initial earnings below a unique threshold  $\underline{z}$  do not adjust their earnings.*<sup>58</sup>

2. *The threshold level of earnings is implicitly defined by the following:*

$$u((1 - \tau_1) z^* + R_1, z^*) - u((1 - \tau_1) \underline{z} + R_1, \underline{z}) \equiv \phi^* \\ z^* \leq \underline{z} \leq z^* + \Delta z^*.$$

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:*

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

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 initial earnings are 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  $\tau_1$ .*

---

<sup>58</sup>  $z_{1,\phi}$  is again the *ex post* level of earnings for the individual who initially locates at  $\underline{z}$ , the initial earnings level over which individuals adjust their earnings. Note that  $\underline{z}$  denotes this the threshold in the non-linear budget set case, whereas  $z_{0,\phi}$  denotes this threshold in the linear budget set case.

## C.5 Derivation of Formula for Bunching

As we discuss in the Section 7, the amount of bunching is equal to the integral of the initial earnings density in the range  $[\underline{z}, z^* + \Delta z^*]$ :

$$B(\tau, z^*; \varepsilon, \phi^*) = \int_{\underline{z}}^{z^* + \Delta z^*} h(\zeta) d\zeta \quad (\text{C.2})$$

If the density is locally uniform, the integral in (1) is:

$$B(\tau, z^*; \varepsilon, \phi^*) \approx h(\underline{z})(z^* + \Delta z - \underline{z}) \quad (\text{C.3})$$

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

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

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

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

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  $\Delta z^* = z^*(d\tau/(1 - \tau))\varepsilon$ . The excess mass at the kink can now be expressed as a linear function of the parameters:

$$b(\tau, z^*; \varepsilon, \phi) = \varepsilon \left( z^* \frac{d\tau}{1 - \tau} \right) - \phi \left( \frac{1}{d\tau} \right). \quad (\text{C.4})$$

## C.6 Derivation of Formula for Bunching with Heterogeneity

In 7.1, we derive the formula for bunching  $B$  in the presence of heterogeneity as follows:

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

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

## C.7 Derivation of Formula for Bunching with a Pre-Existing Kink

Our demonstration of the method for estimating the elasticity of earnings and adjustment costs is not quite tailored to our empirical setting. We have thus far modeled 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 will conduct analysis using data just before and just after the benefit reduction rate was decreased in 1990 from 50 percent to 33 percent for 65 to 69 year olds. This change involves 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, the change in bunching when the kink is changed will be attenuated by the fixed adjustment cost.

We will assume that in the initial state, bunching is characterized as before in Remark (2). Let the initial kink,  $K_1$ , be characterized by a lower marginal tax rate,  $\tau_1$ , to the left of  $z^*$ , and a higher marginal tax rate,  $\tau'_1$ , to the right of  $z^*$ . The initial level of bunching will be:

$$B_1 = \int_{\underline{z}_1}^{z^* + \Delta z_1^*} h(\zeta) d\zeta$$

Now, consider a change in the kink to  $K_2$ , which retains the lower marginal tax rate,  $\tau_2 = \tau_1$  to the left of  $z^*$  but reduces the marginal tax rate to the right of  $z^*$  to  $\tau'_2 < \tau'_1$ . If we would have begun with no kink and introduced  $K_2$ , bunching would be:

$$B_2 = \int_{\underline{z}_2}^{z^* + \Delta z_2^*} h(\zeta) d\zeta$$

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  $\underline{z}_2 \geq \underline{z}_1$ ,  $\Delta z_2^* < \Delta z_1^*$  and  $B_2 \leq B_1$ .

In characterizing bunching when moving from  $K_1$  to  $K_2$ , individuals may be separated into 5 distinct groups based on their optimal level of earnings in the absence of any kink  $z_0$ . First, there are individuals with  $z_0 < z^*$ . They will locate to the left of  $K_1$  and remain there in response to  $K_2$ . Second, we have individuals with  $z^* < z_0 \leq \underline{z}_1$ . 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, these individuals continue to remain with the initial earnings level  $z_0 > z^*$ , as the utility gain to reoptimizing is even smaller now. For these first two groups, there is no behavioral response to either  $K_1$  or  $K_2$ .

Third, there are those for whom  $\underline{z}_1 < z_0 \leq \underline{z}_2$ . When moving from no kink to  $K_1$ , these individuals locate at the kink,  $z^*$ . However, when moving from no kink to  $K_2$  these individuals would choose to remain at  $z_0$ , due to the fixed adjustment cost. In both cases, however,  $z^*$  would be the optimal level of earnings in the absence of the friction. Therefore, when moving from  $K_1$  to  $K_2$ , these agents remain at the kink  $z^*$ . It follows that these individuals bunch in the presence of  $K_2$  when moving from  $K_1$ , but would not have bunched had we started with no kink and moved directly to  $K_2$ . Next, we have agents with  $\underline{z}_2 < z_0 \leq z^* + \Delta z_2^*$ . 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  $K_2$ .

Finally, we have agents with  $z^* + \Delta z_2^* < z_0 \leq z^* + \Delta z_1^*$ . These agents bunch under  $K_1$ , but not under  $K_2$ , when starting from a budget set with no kink. Starting 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 initial earnings, under no kink, at  $z_0 = z^* + \Delta z_2^* + \delta_0$ . For small enough  $\delta_0$  the optimal earnings under  $K_1$  is  $z^*$ , and the 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 with  $\delta_0$ . However, the fixed adjustment cost does not tend to zero; therefore, this individual will remain at  $z^*$  when moving from  $K_1$  to  $K_2$  for small enough  $\delta_0$ . In certain cases, it is even possible that reoptimizing away from the kink is not worth it for anyone in this final range. In that case, there is no change in bunching 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. In that case, the marginal "de-buncher" will be defined by the following conditions:

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

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 this agent is indifferent between remaining at  $z^*$  and moving to  $\bar{z}_2$  and paying the adjustment cost,

when facing  $K_2$ . 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  $\tau_2$  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  $\varepsilon$  and  $\phi$ . Thus, we restrict attention to the case where  $\bar{z}_0 < z^* + \Delta z_1^*$ .

Bunching at  $K_2$  following  $K_1$ , when  $\bar{z}_0 < z^* + \Delta z_1^*$ , can be summarized 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 adjustment of the top-most buncher. That is:

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

We can further simplify the estimation provided we assume that the density of baseline earnings is uniform over the range  $[z^*, z^* + \Delta z_1^*]$ . In that case, bunching under  $K_1$ , i.e. before the change in the kink, is:

$$\begin{aligned} B_1 &= \int_{\bar{z}_1}^{z^* + \Delta z_1^*} h(\zeta) d\zeta \\ &= h(z^*)(z^* + \Delta z_1^* - \bar{z}_1) \end{aligned}$$

and normalized bunching,  $b \equiv B/h$ , can be written as in Appendix C.5:

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

Bunching at the new, smaller kink,  $K_2$ , will now be:

$$\begin{aligned} \tilde{B}_2 &= \int_{\bar{z}_1}^{\bar{z}_0} h(\zeta) d\zeta \\ &= h(z^*)(\bar{z}_0 - \bar{z}_1) \end{aligned}$$



Normalizing bunching, we have:

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

We therefore have two equations  $(b_1, b_2)$  and three unknowns:  $(\varepsilon, \phi, \bar{z}_2)$ . We close the system by using the following identity for  $\bar{z}_2$ :

$$u((1 - \tau'_2) \bar{z}_2 + R'_2, \bar{z}_2; \bar{n}_2) - u((1 - \tau'_2) z^* + R'_2, z^*; \bar{n}_2) \equiv \phi^*$$

Note that when moving from  $K_1$  to  $K_2$ , the change in bunching is smaller than would be recorded had we started with steady state bunching at  $K_1$  following no kink ( $B_1$ ) and then move 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  $\underline{z}_1 \leq \underline{z}_2$ .

## C.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, let  $\boldsymbol{\tau} = (\boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_K)$  be the tax schedule at each kink (with  $\boldsymbol{\tau}_k = (\tau_k, \tau'_k)$  denoting the tax rate below and above kink  $k$ ), and let  $z^* = (z_1^*, \dots, z_K^*)$  be the exempt amount, *i.e.* the earnings level 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  $\hat{b}(\varepsilon, \phi) = (\hat{b}(\boldsymbol{\tau}_1, z_1^*, \varepsilon, \phi), \dots, \hat{b}(\boldsymbol{\tau}_K, z_K^*, \varepsilon, \phi))$ , our estimator is:

$$(\hat{\varepsilon}, \hat{\phi}) = \operatorname{argmin}_{(\varepsilon, \phi)} \left( \hat{b}(\varepsilon, \phi) - b \right)' W \left( \hat{b}(\varepsilon, \phi) - b \right), \quad (\text{C.5})$$

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

We obtain our estimates by minimizing equation (C.5) numerically. Solving this problem

requires evaluating  $\hat{b}$  at each trial guess of  $(\varepsilon, \phi)$ .<sup>59</sup> Recall that:

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

As in the main text, we continue to assume that  $h(\cdot)$  is uniform in  $[z^*, z^* + \Delta z]$ , so that

$$b(\tau, z^*; \varepsilon, \phi^*) = z^* + \Delta z - \underline{z},$$

where  $b = B/h(z^*)$ . Our estimator assumes a quasilinear utility function,  $u(c, z; n) = c - \frac{n}{1+1/\varepsilon} \left(\frac{z}{n}\right)^{1+1/\varepsilon}$ . Note that because we have assumed quasilinearity,  $\phi^* = \phi$  and  $\Delta z = z^* \left( \left(\frac{1-\tau'}{1-\tau}\right)^\varepsilon - 1 \right)$ . Even with quasilinear utility, however, there is no closed form for  $\underline{z}$ . Instead, given  $\varepsilon$  and  $\phi$ , we find  $\underline{z}$  numerically as the solution to:

$$\underbrace{u((1-\tau')z^* + R', z^*; \underline{z}/(1-\tau)^\varepsilon)}_{\text{utility from adjusting to kink}} - \underbrace{u((1-\tau')\underline{z} + R', \underline{z}; \underline{z}/(1-\tau)^\varepsilon)}_{\text{utility from not adjusting}} = \phi,$$

where  $\underline{n} = \underline{z}/(1-\tau)^\varepsilon$  is derived earlier in the Appendix. The equation is continuously differentiable and has a unique solution. As such, Newton-type solvers, such as the simplex method we use, are able to find  $\underline{z}$  accurately. Note that some combinations of  $\tau, z^*, \varepsilon$ , and  $\phi$  imply  $\underline{z} > z^* + \Delta z$ . 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:

$$\hat{b}(\tau, z^*; \varepsilon, \phi) = \max(z^* + \Delta z - \underline{z}, 0).$$

We have also shown a robustness check in Table 2 in which we assume that the earnings distribution is lognormal, rather than assuming that  $h(\cdot)$  is uniform in  $[z^*, z^* + \Delta z]$ . 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, as individuals age 61 are not subject to the AET but are not far removed in age from those at retirement age.

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 100 bootstrap samples of  $b$ . For each bootstrap sample, we compute  $\hat{\varepsilon}$  and  $\hat{\phi}$  as the solution to (C.5). The standard deviation of  $\hat{\varepsilon}$  and  $\hat{\phi}$  across bootstrap samples is the bootstrap standard error. We have also estimated the standard errors using the delta method and obtained similar results.

<sup>59</sup>In solving problem (C.5), 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.

## C.9 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 level of earnings 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 faced an effective marginal tax rate of  $\tilde{\tau} = \tau + \phi \cdot \text{sgn}(z - z_0)$ . It follows that our predictions about earnings adjustments are similar to our previous predictions, except that the effective marginal tax rate  $\tilde{\tau}$  appears, instead of  $\tau$ . 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  $d\tau$ :

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

Since  $(-\phi)$  appears in the denominator, the estimate of the elasticity increases as attenuation in bunching due to the linear adjustment cost increases.

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

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

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

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

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

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

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 tax rates, we are not able to estimate these parameters using data: in the time period our data covers, we are able to examine only one change in the MTR, from 50 percent to 33 percent for 65-69 year-olds in 1990.

## D Appendix: Analysis of Bunching Using Quarterly LEHD Data

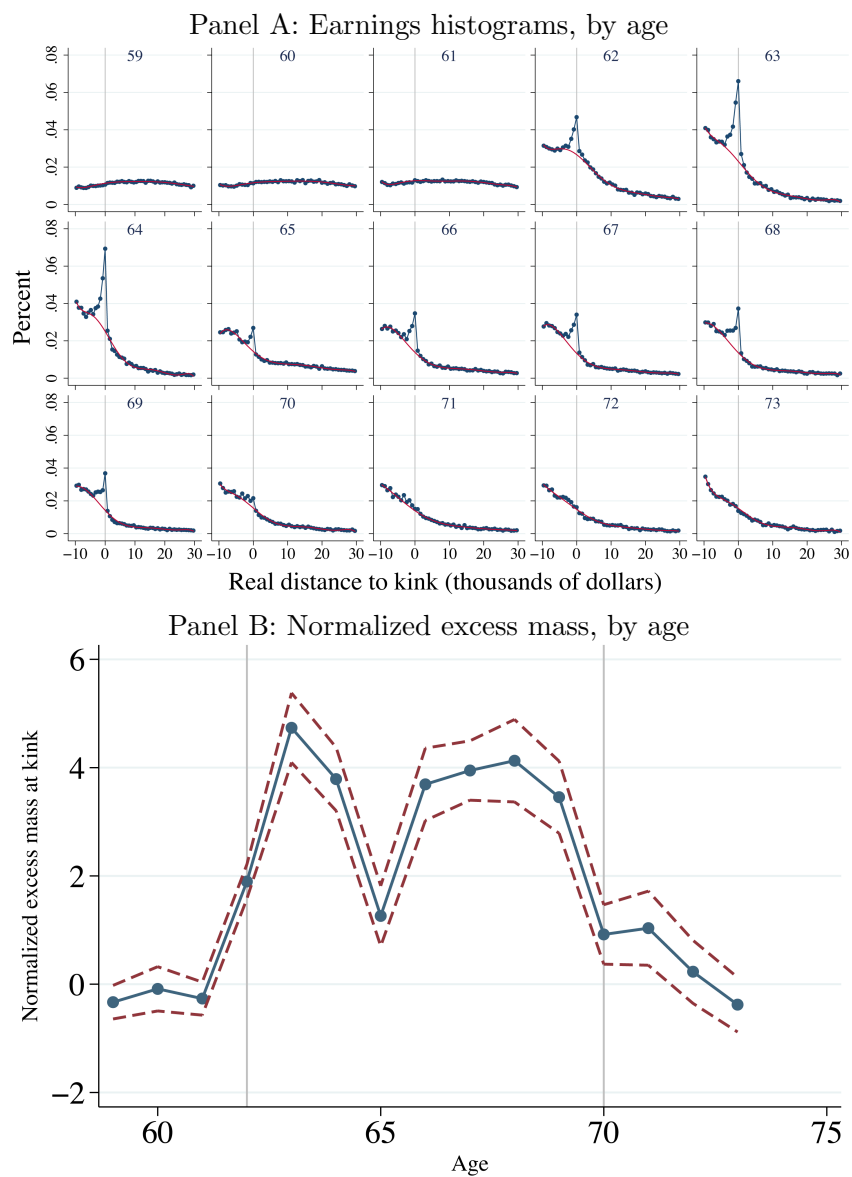
We present a supplementary analysis of bunching in the quarterly earnings data in the LEHD. We define a "quarterly kink earnings level" in a given calendar year as one-quarter of the earnings level associated with the kink in that year. When interpreting these graphs, it is important to bear in mind an important limitation: the AET is assessed yearly, and thus individuals can appear at the quarterly kink in a given quarter even though their yearly earnings does not ultimately put them at the applicable kink (though there is no reason that this should produce bunching at the quarterly kink). Conversely, individuals could have a yearly earnings level associated with the kink even though their quarterly earnings in each quarter does not put them at the kink (though as we note in the text, there is no incentive for individuals to do so if they are targeting the kink earnings level for the year).

We investigate adjustment to the 2000 elimination of the AET for ages 66-69. Appendix Figure Y shows clear evidence of excess bunching in 1999 and the first quarters of 2000; by the fourth quarter of 2000, excess bunching is at least greatly reduced. By the final quarter of 2001, it is clear that no excess bunching is visible. Excess bunching is significantly greater than zero in the first three quarters of 2000 ( $p < 0.01$ ) but becomes insignificant in the subsequent five quarters. Thus, the evidence suggests that the vast majority of excess bunching is eliminated within a few quarters of the 2000 disappearance of the AET.

[Figures Forthcoming]

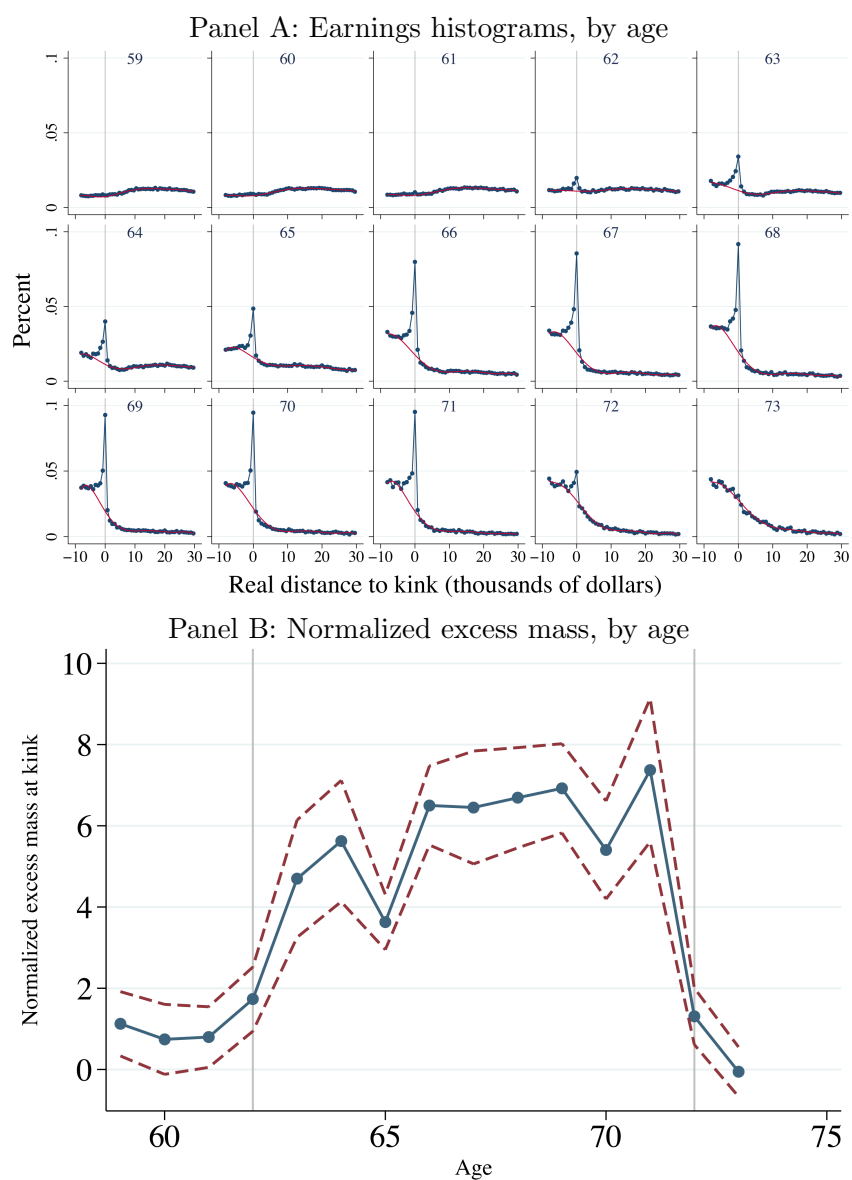
## E Appendix: Additional Figures

Figure E.1: Adjustment Across Ages: Histograms of Earnings and Normalized Excess Mass, 59-73-year-old Social Security Claimants, 1990-1999



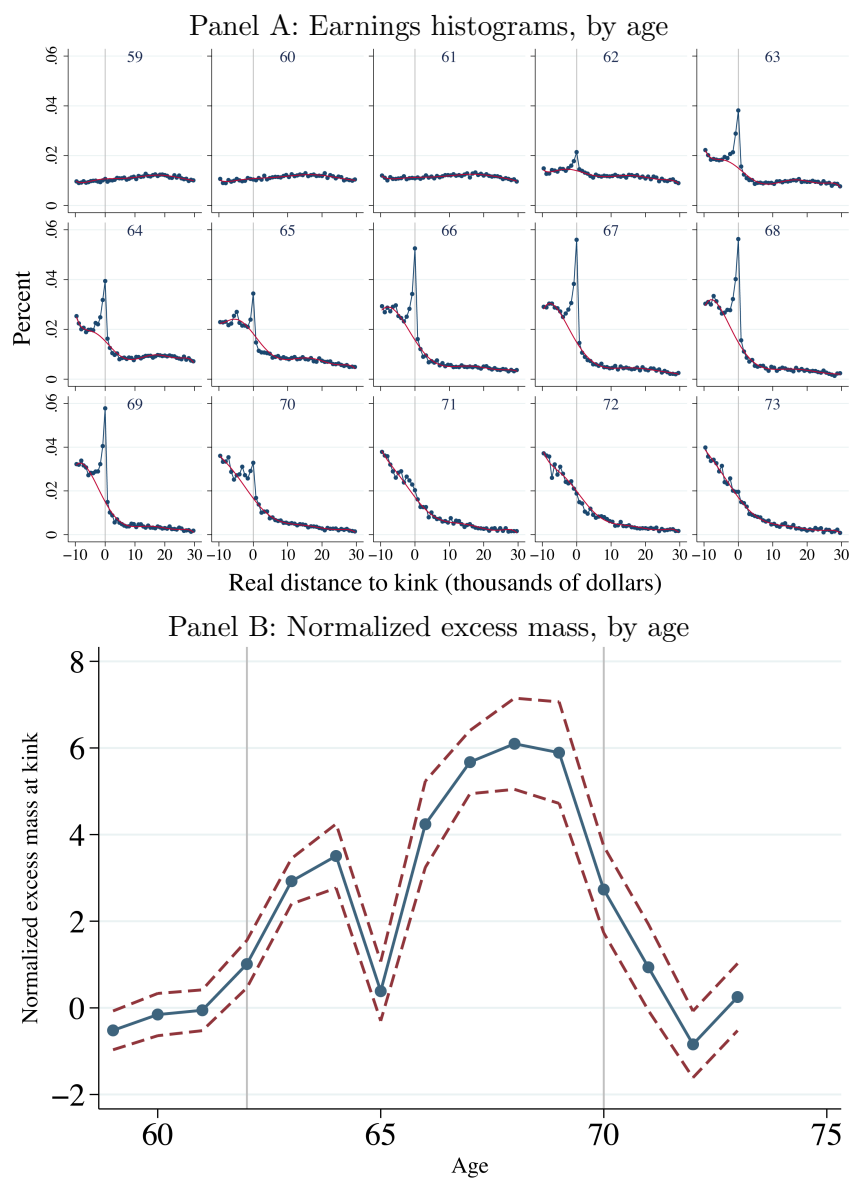
See notes to Figure 4. The figure differs from Figure 4 only because the sample in year  $t$  consists only of people who have claimed OASI in year  $t$  or before (whereas in Figure 4 it consisted of all those who claimed by age 65).

Figure E.2: Adjustment Across Ages: Histograms of Earnings and Normalized Excess Mass, 59-73-year-old Social Security Claimants, 1972-1982



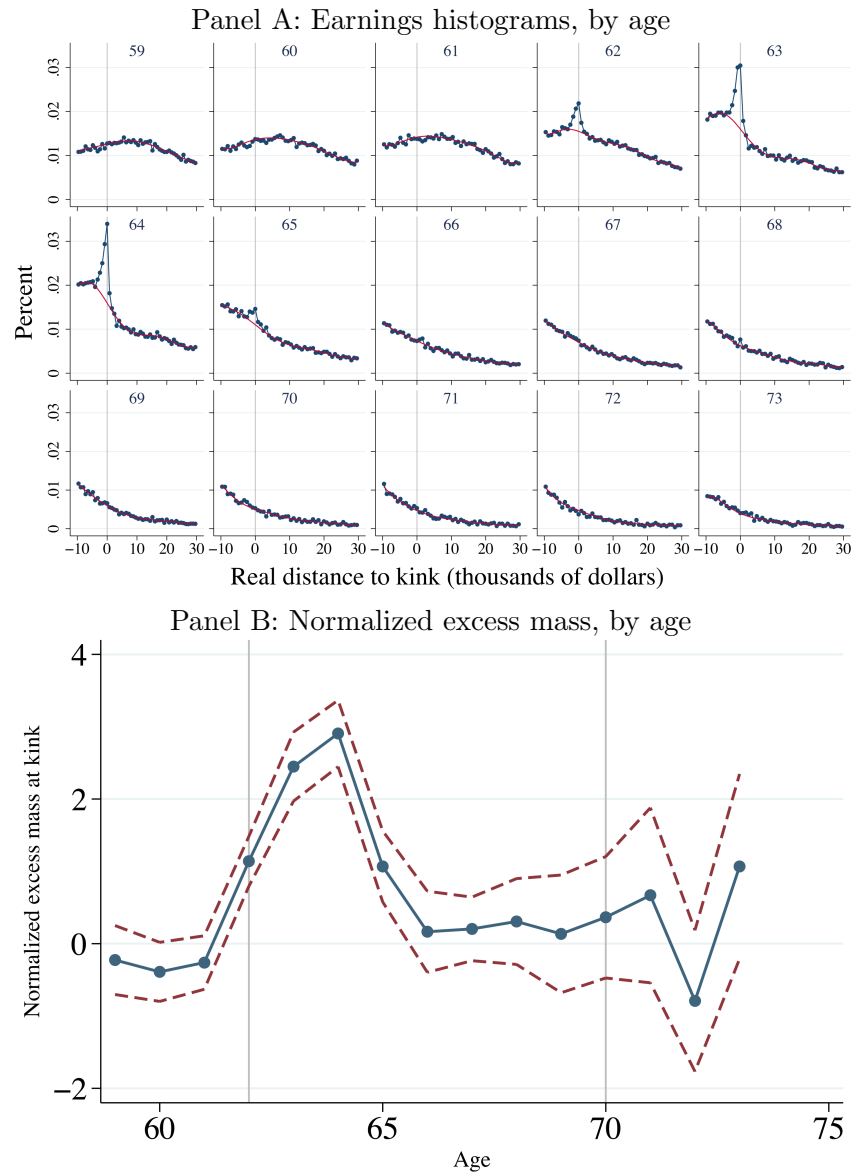
See notes to Figure 4. The figure differs from Figure 4 only because the years examined are 1972-1982 (whereas in Figure 4 the years examined were 1990-1999).

Figure E.3: Adjustment Across Ages: Histograms of Earnings and Normalized Excess Mass, 59-73-year-old Social Security Claimants, 1983-1989



See notes to Figure 4. The figure differs from Figure 4 only because the years examined are 1983-1989 (whereas in Figure 4 the years examined are 1990-1999).

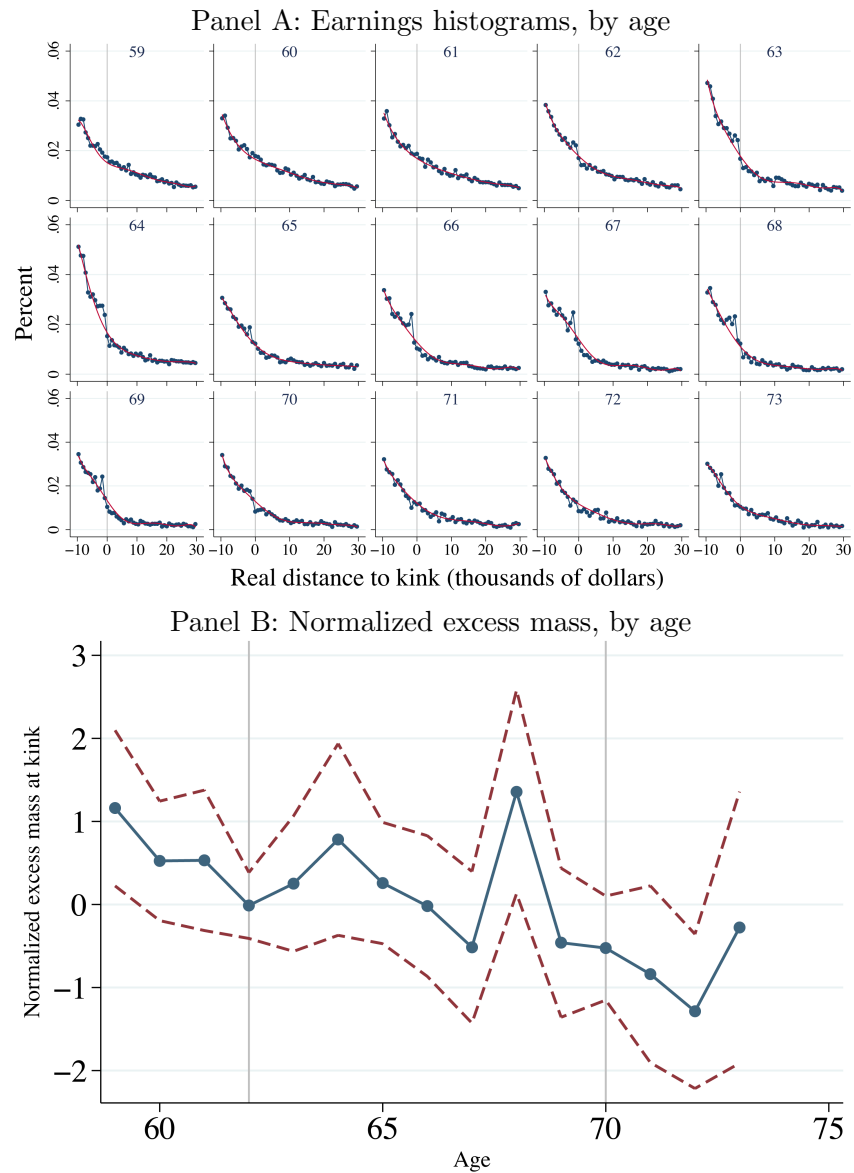
Figure E.4: Adjustment Across Ages: Histograms of Earnings and Normalized Excess Mass, 59-73-year-old Social Security Claimants, 2000-2006



See notes to Figure 4. The figure differs from Figure 4 only because the years examined are 2000-2006 (whereas in Figure 4 the years examined are 1990-1999). As explained in the main text, the Normal Retirement Age 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.

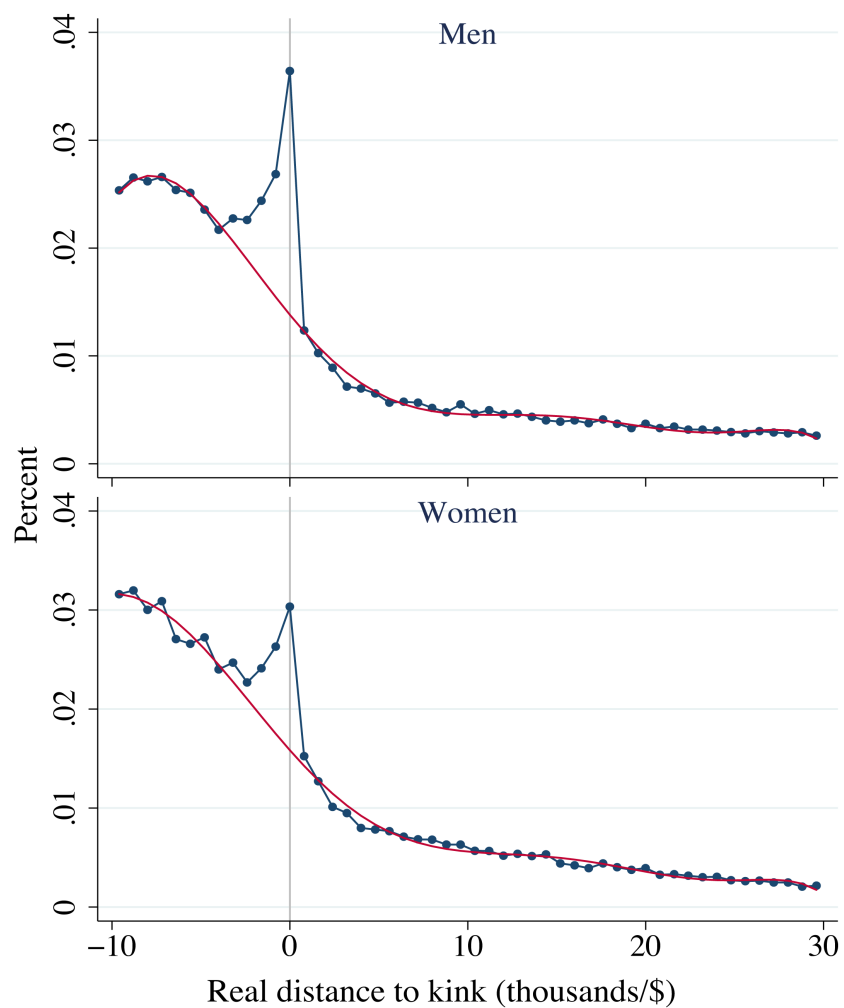


Figure E.5: Adjustment Across Ages: Histograms of Earnings and Normalized Excess Mass, 59-73-year-old Social Security Claimants by 65 with Self-Employment Income, 1990-1999



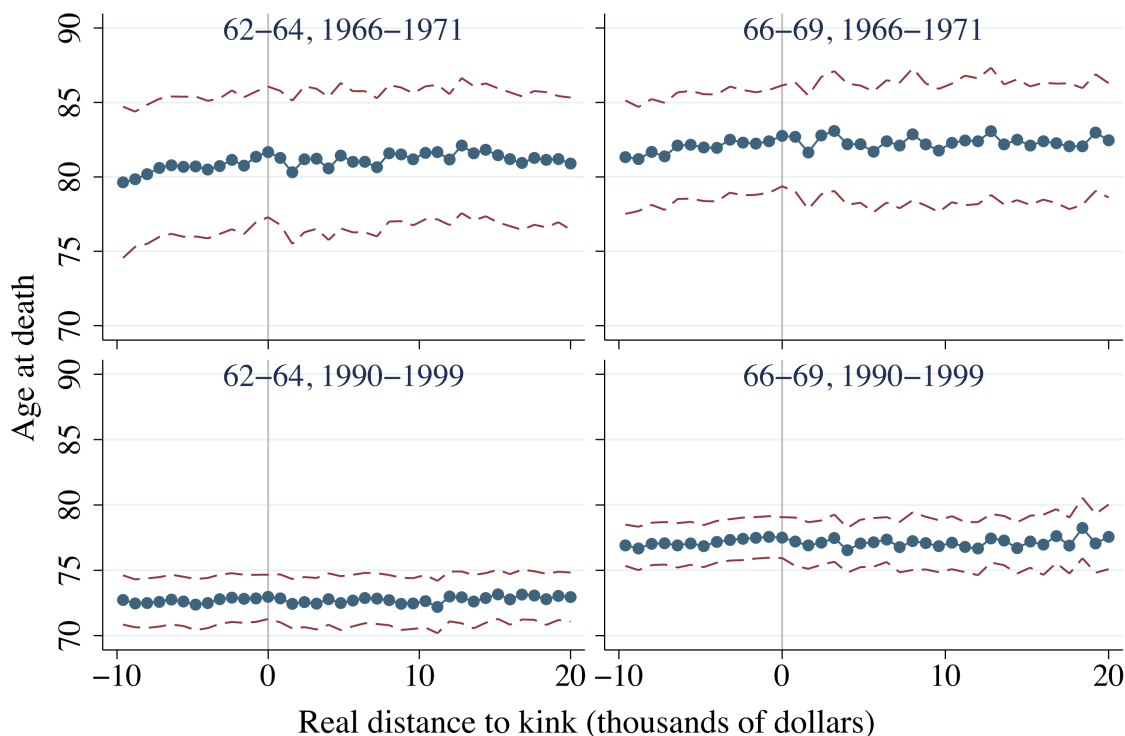
See notes to Figure 4. The figure differs from Figure 4 only because the sample consists of those with positive self-employment income (whereas in Figure 4 those with positive self-employment income are excluded).

Figure E.6: Adjustment Across Ages: Histograms of Earnings, 59-73-year-old Male and Female Social Security Claimants by 65, 1990-1999



See notes to Figure 4. The sample examined is the same as in Figure 4 but examines men and women separately.

Figure E.7: Life Expectancy Analysis: Mean Age at Death, 62-69-year-old Social Security Claimants, 1966-1971 and 1990-1999



Note: The figure shows mean age at death from a 1 percent random sample of SSA administrative data on Social Security claimants aged 59-73 between 1966 and 1971 (inclusive) in the top panel, and between 1990 and 1999 (inclusive) in the bottom panel. The bin width is \$800. The earnings level 0, shown by the vertical line, denotes the kink. Earnings are measured in CPI-U-adjusted year 2010 dollars in each year. The figure shows no clearly noticeable patterns at the kink, both for those 62-64 and 66-69, and 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).

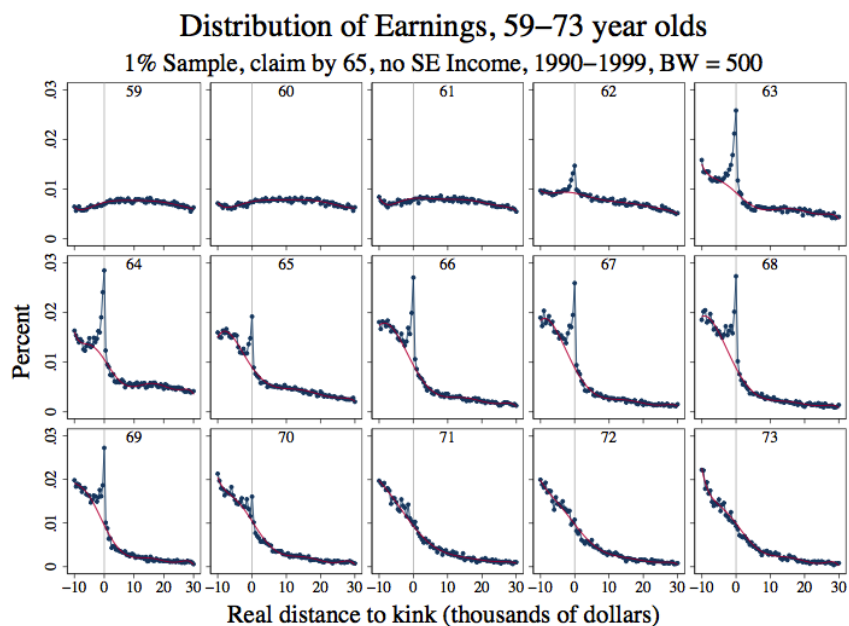
Figure E.8: Probability of Earnings Moving with Kink, 1990-1998



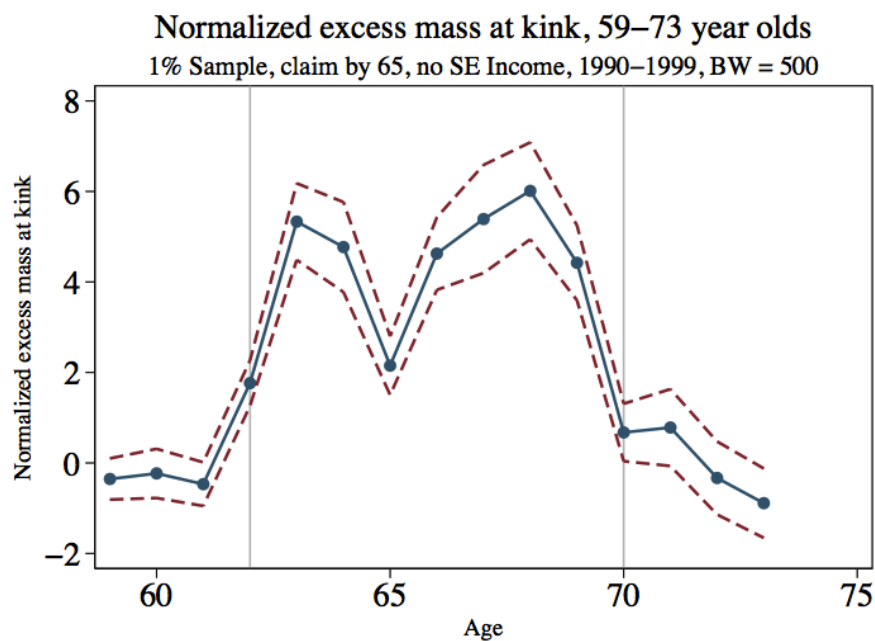
Note: The figure shows the probability that individual earnings move with the kink from year to year (i.e. the probability that an individual located at the kink in year  $t + 1$ , conditional on locating at the kink in year  $t$ ), for age groups 58 to 60, 62 to 63, and 65 to 68. (Each of these ages refers to age in year  $t$ .) The kink is defined as the region within \$2800 of the exempt amount. See other notes to Figure 4.

Figure E.9: Adjustment Across Ages: Histograms of Earnings and Normalized Excess Mass, 59-73-year-old Social Security Claimants, 1990-1999

Panel A: Earnings histograms, by age

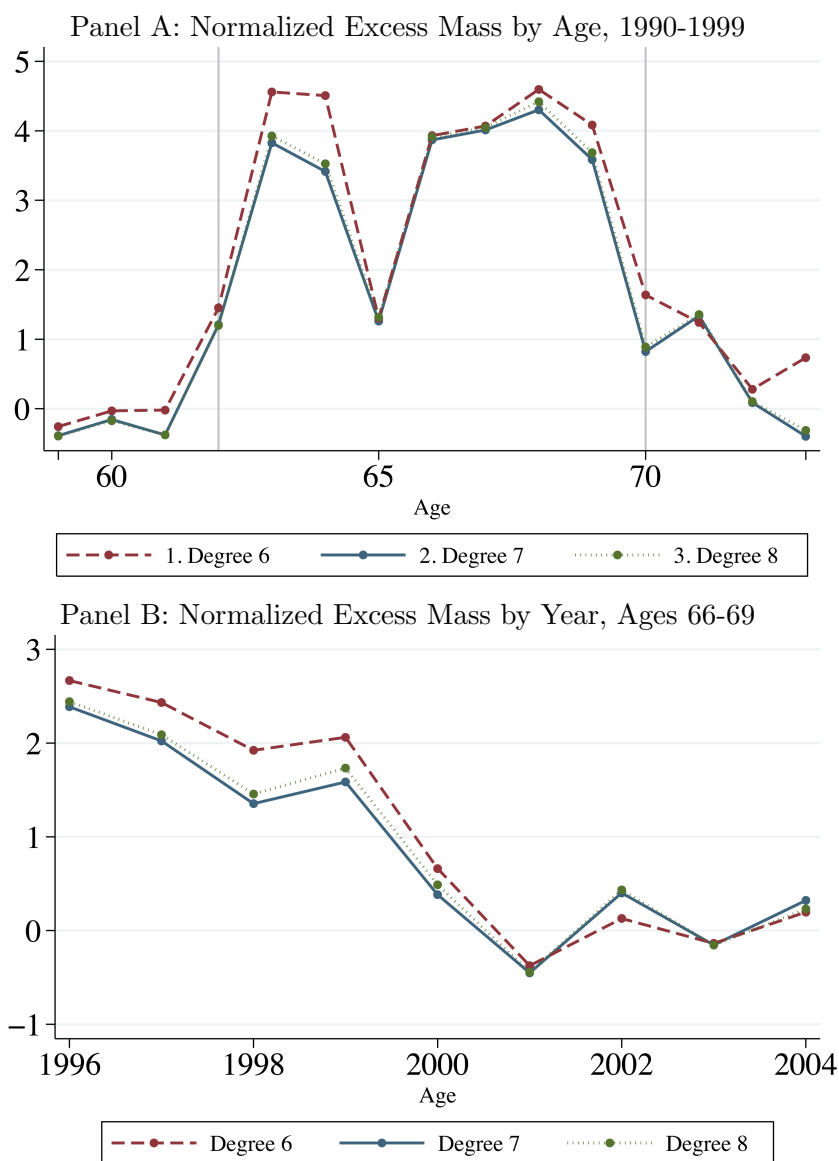


Panel B: Normalized excess mass, by age



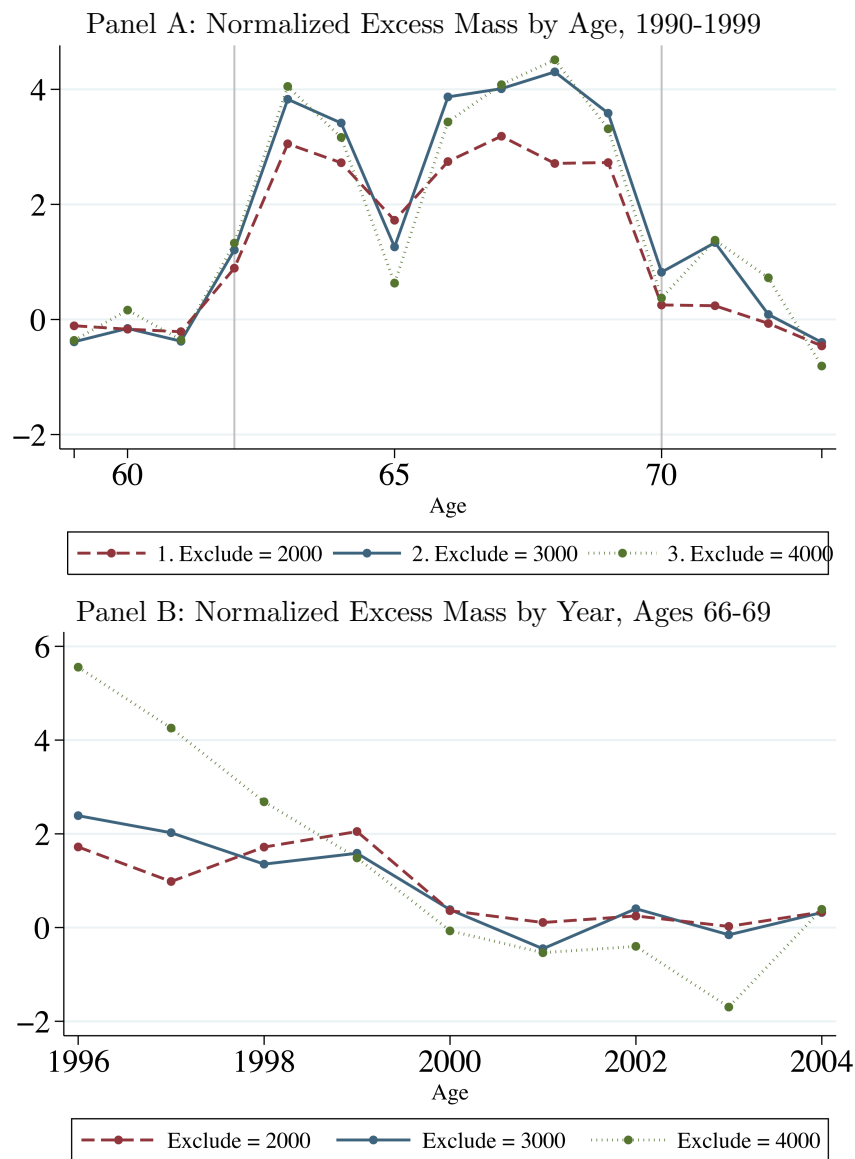
See notes to Figure 4. The figure differs from Figure 4 only because the bandwidth is \$500 (whereas in Figure 4 it is \$800).

Figure E.10: Robustness to Polynomial Degree: Normalized Excess Mass by Age and Year, Social Security 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 4 for Panel A and Figure 5 for Panel B.

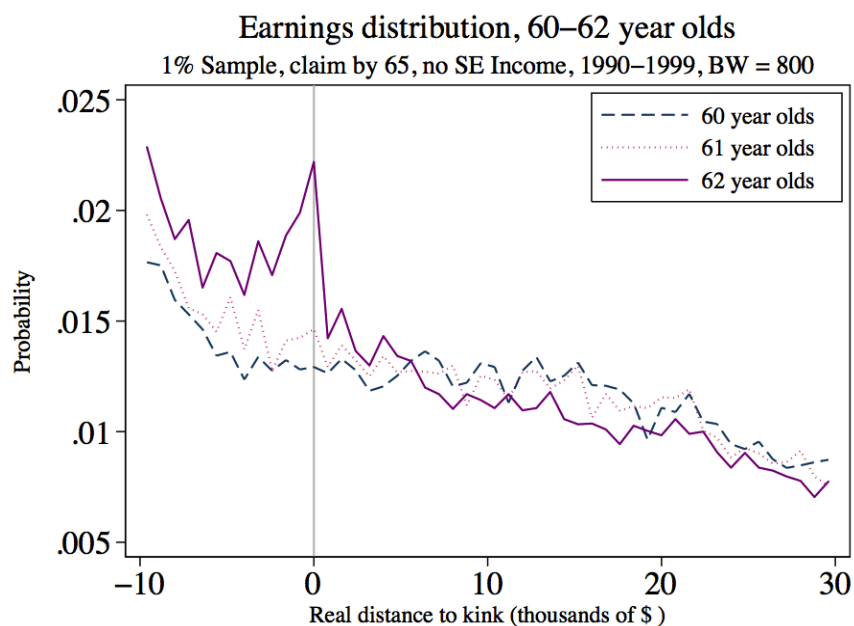
Figure E.11: Robustness to the Excluded Region: Normalized Excess Mass by Age and Year, Social Security Claimants by 65



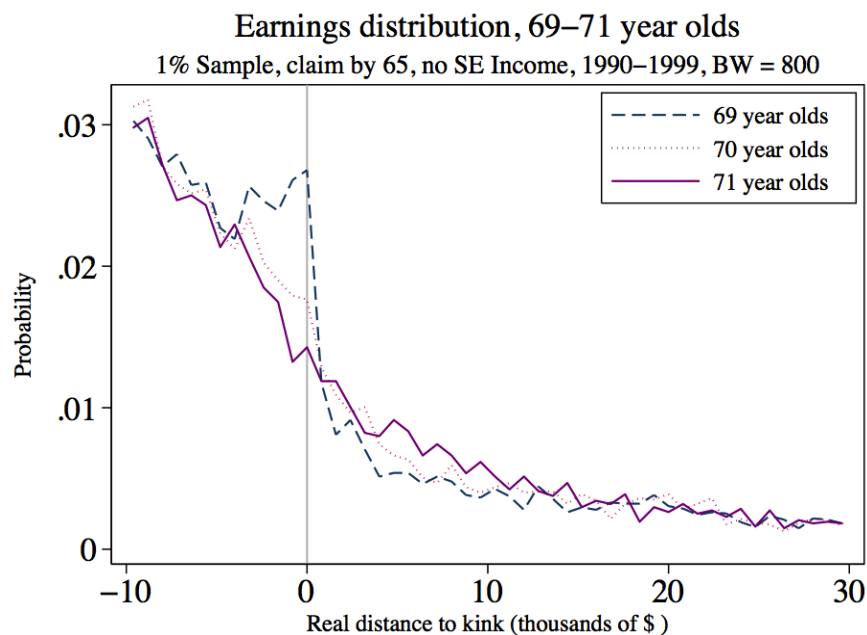
Notes: The figure shows the difference in estimates of normalized excess bunching as we vary the region about the kink excluded from the polynomial estimation. For additional notes on the samples see Figure 4 for Panel A and Figure 5 for Panel B.

Figure E.12: Adjustment Across Ages: Histograms of Earnings and Normalized Excess Mass, 59-73-year-old Social Security Claimants by 65 with Self-Employment Income, 1983-1989

Panel A: Earnings Distributions by Age, 60-62



Panel B: Earnings Distributions by Age, 69-71



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