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

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

Online Appendix

A Proofs of Propositions and Other Claims

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

$$\frac{d\Pr\left(\left. z_{n1}>0\right|\tilde{z}_{n0}\right)}{d\tilde{z}_{n0}}=g\left(\left.\overline{q}_{n1}\right|n\right)\frac{d\overline{q}_{n,1}}{d\tilde{z}_{n0}}+\frac{\partial G\left(\left.\overline{q}_{n1}\right|n\right)}{\partial n}\frac{dn}{d\tilde{z}_{n0}} \tag{A.1}$$

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

$$\begin{aligned}\frac{d\overline{q}_{n,1}}{d\tilde{z}_{n0}} &= \frac{\partial v\left(\tilde{z}_{n1}-T\left(\tilde{z}_{n1}\right)+B_1\left(\tilde{z}_{n1}\right),\tilde{z}_{n1};n\right)}{\partial\tilde{z}_{n1}}\frac{d\tilde{z}_{n1}}{d\tilde{z}_{n0}}+\frac{\partial v\left(\tilde{z}_{n1}-T\left(\tilde{z}_{n1}\right)+B_1\left(\tilde{z}_{n1}\right),\tilde{z}_{n1};n\right)}{\partial n}\frac{dn}{d\tilde{z}_{n0}}\\ &= \frac{\partial v\left(\tilde{z}_{n1}-T\left(\tilde{z}_{n1}\right)+B_1\left(\tilde{z}_{n1}\right),\tilde{z}_{n1};n\right)}{\partial n}\frac{dn}{d\tilde{z}_{n0}}\end{aligned} \tag{A.2}$$

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

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

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

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

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

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

$$\lim_{\tilde{z}_{n0}\to z^{AET+}}\frac{d\Pr\left(\left. z_{n1}>0\right|\tilde{z}_{n0}\right)}{d\tilde{z}_{n0}}=\lim_{\tilde{z}_{n0}\to z^{AET-}}\frac{d\Pr\left(\left. z_{n1}>0\right|\tilde{z}_{n0}\right)}{d\tilde{z}_{n0}} \tag{A.4}$$

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

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

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

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

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

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

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

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

Thus, we now have:

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

Note also that:

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

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

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

where $\overline{q}_{n^{AET}1}$, n^{AET}, and $\lambda_{n^{AET}}$ all refer the individual for whom $\tilde{z}_{n0} = z^{AET}$.

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

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

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

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

where

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

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

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

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

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

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

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

B Model Extensions

B.1 Extension to Two or More Discrete Job Choices

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Note the following:

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

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

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

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

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

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

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

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

$$\begin{aligned}
\lim_{\tilde{z}_{n0}\to z^{AET+}} \frac{d\Pr\left(z_{n1}>0|\tilde{z}_{n0}\right)}{d\tilde{z}_{n0}} &= \lim_{\tilde{z}_{n0}\to z^{AET+}} \frac{dG\left(\overline{q}_{n1}|\, n\right)}{d\tilde{z}_{n0}} \\
&= g\left(\overline{q}_{n^{AET}1}|\, n^{*}\right)\Bigg[-db\cdot\lambda_{n^{*}} \\
&\qquad + \frac{\partial v\left(\tilde{z}_{n^{AET}0}-T\left(\tilde{z}_{n^{AET}0}\right)+B_1\left(\tilde{z}_{n^{AET}0}\right),\tilde{z}_{n^{AET}0};n^{AET}\right)}{\partial n}\frac{dn}{d\tilde{z}_{n0}}\Bigg] \\
&\quad + \frac{dG\left(\overline{q}_{n^{AET}1}|\, n^{AET}\right)}{dn}\frac{dn}{d\tilde{z}_{n0}}
\end{aligned} \tag{B.20}$$

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

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

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

B.2.1 Model with Mixture of Types

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

$$\lim_{\tilde{z}_{n0}\to z^{AET+}} \frac{d\Pr\left(z_{n1}>0|\,\tilde{z}_{n0}\right)}{d\tilde{z}_{n0}} - \lim_{\tilde{z}_{n0}\to z^{AET-}} \frac{d\Pr\left(z_{n1}>0|\,\tilde{z}_{n0}\right)}{d\tilde{z}_{n0}} = -\pi_{NA}^{AET}\cdot db\cdot g_{NA}\left(\overline{q}_{n^{AET}1}|\, n^{AET}\right) \tag{B.21}$$

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

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

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

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

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

B.3 Jointly Estimating the Structural Elasticity among Constrained Types and π_B^{AET}

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

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

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

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

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

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

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

We can show the following:

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

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

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

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

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

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

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

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

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

Now, consider estimating the following:

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

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

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

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

$$\begin{aligned}
\beta^{ATR} &= \frac{N_{RA,1}}{N_{R,1}}\cdot 0+\frac{N_{RNA,1}}{N_{R,1}}\cdot\left(-\frac{db}{z^{AET}}\right) \\
&= \frac{N_{R,0}}{N_{R,1}}\frac{N_{RNA,0}}{N_{R,0}}\frac{N_{RNA,1}}{N_{RNA,0}}\cdot\left(-\frac{db}{z^{AET}}\right) \\
&= \frac{\Pr\left(z_{n,1}>0\right|n\in R,NA)\,\pi_{NA}^{AET}}{\Pr\left(z_{n,1}>0\right|z^{AET}<\tilde{z}_{n0}<z^{AET}+\triangle z)}\left(-\frac{db}{z^{AET}}\right)
\end{aligned}$$

We thus have four equations:

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

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

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

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

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

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

B.4 Fully Dynamic Extension of the Model

In this section we briefly demonstrate under what conditions our results continue to hold once our model is extended to a multi-period setting with forward-looking agents. We again have two states of the world, state 0 and state 1. In this multi-period model, the tax and benefit schedule is the same across the two states for periods $1,\ldots,t-1$. However, in period t, there is no benefit reduction rate in state 0 and a kink at z^{AET} in state 1 created by a benefit reduction rate on earnings above z^{AET}. The tax and benefit schedules are once again the same across the two states during periods $t+1,\ldots,T$. (For simplicity we assume here that the tax and benefit schedules across the two states are once again the same during periods $t+1,\ldots,T$, but the model can be easily extended to assume that in each of these periods there is no benefit reduction rate in state 0 and a kink at z^{AET} in state 1.)

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

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

subject to a dynamic budget constraint:

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

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

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

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

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

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

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

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

$$\Pr\left(z_{n1t} > 0 \middle| \tilde{z}_{n0t}\right) = G\left(\overline{q}_{n1t} \middle| n, t\right) \tag{B.27}$$

where now:

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

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

$$\frac{d\Pr\left(z_{n1t} > 0 \middle| \tilde{z}_{n0t}\right)}{d\tilde{z}_{n0t}} = g\left(\overline{q}_{n1t} | n, t\right) \frac{d\overline{q}_{n1t}}{d\tilde{z}_{n0t}} + \frac{dG\left(\overline{q}_{n1t} \middle| n, t\right)}{dn} \frac{dn}{d\tilde{z}_{n0t}} \tag{B.29}$$

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

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

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

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

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

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

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

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

We now turn to the case in which we make the same assumptions, except that individuals are not able to adjust on the intensive margin, as in Section 6.3. We now have:

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

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

$$\left(1-T'\left(z\right)+B_1'\left(z\right)\right)v_c+v_z+V_z=\lambda\left[\left(1-\tau_0-db\right)+\frac{v_z+V_z}{v_c}\right] \tag{B.34}$$

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

$$\frac{d\Pr\left(\left.z_{n1t}>0\right|\tilde{z}_{n0t}\right)}{d\tilde{z}_{n0t}}=\begin{cases}
\begin{array}{l} g\left(\overline{q}_{n1t}|n,t\right)\Big(\left[v_n+V_n\right]\frac{dn}{d\tilde{z}_{n0t}} \\ \quad+\left[\left(1+r_{t-1}\right)v_c-v_A^0\right]\frac{dA_{n1,t-1}}{d\tilde{z}_{n0t}}\Big) \\ \quad+\frac{dG\left(\overline{q}_{n1t}|n,t\right)}{dn}\frac{dn}{d\tilde{z}_{n0t}}\end{array} & \text{if } \tilde{z}_{n0t}<z^{AET} \\[2ex]
\begin{array}{l} g\left(\overline{q}_{n1t}|n,t\right)\Big(\lambda_n\left[\left(1-\tau_0-db\right)+\frac{v_z+V_z}{v_c}\right]+\left[v_n+V_n\right]\frac{dn}{d\tilde{z}_{n0t}} \\ \quad+\left[\left(1+r_{t-1}\right)v_c-v_A^0\right]\frac{dA_{n1,t-1}}{d\tilde{z}_{n0t}}\Big)+\frac{dG\left(\overline{q}_{n1t}|n,t\right)}{dn}\frac{dn}{d\tilde{z}_{n0t}}\end{array} & \text{if } \tilde{z}_{n0t}\geq z^{AET}
\end{cases} \tag{B.35}$$

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

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

Maintaining our smoothness assumptions for this section, we thus have our original result when earnings in state 1 are constrained ($\tilde{z}_{n1t} = \tilde{z}_{n0t}$):

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

B.5 Model Extension: Joint Claiming and Employment Decision

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

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

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

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

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

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

$$\frac{\partial \overline{q}_n^0}{\partial \tilde{z}_{n0}} = \frac{\partial v_n^0}{\partial \tilde{z}_{n0}} = 0 \tag{B.44}$$

Additionally, we have:

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

B.5.1 Claiming Decision

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

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

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

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

$$\frac{d \Pr\left(claim | \, \tilde{z}_{n0} \right)}{d\tilde{z}_{n0}} = \frac{\partial \Pr\left(claim | \, \tilde{z}_{n0} \right)}{\partial \tilde{z}_{n0}} + \frac{\partial \Pr\left(claim | \, \tilde{z}_{n0} \right)}{\partial n} \frac{dn}{d\tilde{z}_{n0}} \tag{B.47}$$

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

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

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

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

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

$$\lim_{\tilde{z}_{n0}\to z^{*}+}\frac{\partial \Pr\left(claim|\,\tilde{z}_{n0}\right)}{\partial \tilde{z}_{n0}}-\lim_{\tilde{z}_{n0}\to z^{*}-}\frac{\partial \Pr\left(claim|\,\tilde{z}_{n0}\right)}{\partial \tilde{z}_{n0}}=\delta \tag{B.50}$$

where

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

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

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

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

B.5.2 Extensive Margin Response with Endogenous Claiming

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

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

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

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

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

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

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

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

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

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

where

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

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

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

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

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

B.5.3 Bounding the Behavioral Response

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

If our assumption holds, we can show the following:

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

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

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

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

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

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

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

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

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

C Procedure for estimating excess mass

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

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

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

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

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

D Method of Adjusting for Measurement Error

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

D.1 Estimation

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

$$\hat{\eta} = \frac{\hat{\beta}_{PE}}{\hat{\alpha}_{ANTR}} \cdot \frac{\hat{\mu}_{ANTR}}{\hat{\mu}_{PE}}, \tag{D.62}$$

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

D.1.1 Calibrating and drawing from the earnings growth distribution

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

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

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

D.1.2 Simulating the first stage

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

1. Resample from the distribution of age 60 earnings, shown in Figure 2 (assuming an exempt amount of \$10,000). This yields a data set on age 60 earnings, z_i, corresponding to the running variable used in our analysis.

2. For each observation, draw from the calibrated earnings growth distribution and simulate earnings forward three years (as explained above). This yields a value of age 63 earnings and distance to the exempt amount, z_i^*, for each observation.

3. Given z_i^*, find the ANTR at age 63, $ANTR_i^*$, using the statutory formula.

4. Estimate an RK of age 63 $ANTR_i^*$ on z_i, using the main bandwidth in estimation (\$2,800). Record the estimated kink and the mean $ANTR_i^*$ at $Z_i = 0$.

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

D.2 Inference

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

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

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

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

D.3 Validation

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

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

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

$$E\left[PE_i|z_i^*, ANTR_i^*\right] = f\left(z_i^*\right) + \beta \cdot ANTR_i^*, \tag{D.64}$$

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

Simulation details for validation exercise The simulation is the same as the process for simulating the first stage, except we also simulate the outcome, employment, and we let the first stage bandwidth differ from iteration to iteration. The simulation approach is similar to that used in Card et al. (2017). The simulation works as follows.

1. Fix the structural parameter β (explained above) and fix the measurement error process.

2. Draw a data set on the observed running variable, Z_i; the desired earnings at age 63 relative to the threshold, Z_i^*; the ANTR at age 63 given desired earnings, $ANTR_i^*$, and the outcome PE_i, as follows.

   (a) Resample from the distribution of age 60 earnings, shown in Figure 2 (and assuming a kink point of \$10,000). This yields a data set on age 60 earnings, z_i, corresponding to the running variable used in our analysis.

   (b) For each observation, simulate earnings forward three years according to the calibrated earnings growth distribution (as explained above). This yields a value of age 63 earnings and distance to the exempt amount, z_i^*, for each observation.

   (c) Given z_i^*, find the ANTR at age 63, $ANTR_i^*$, using the statutory formula.

   (d) Given z_i^* and $ANTR_i^*$, find $E[PE_i|z_i^*, ANTR_i^*]$ as

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

   Where f and β are calibrated as described above.

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

   $$PE_i = \mathbf{1}\left\{e_i \leq E\left[PE_i\left|z_i^*, ANTR_i^*\right.\right]\right\}, \tag{D.66}$$

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

3. Estimate an RK of PE_i on z_i, using the CCT procedure to find the bandwidth. Record the $\hat{\beta}_{PE}$ and $\hat{\mu}_{PE}$, as well as their standard errors.

4. Estimate an RK of $ANTR_i^*$ on z_i using the same bandwidth as in step (3). Record the estimated kink and the mean $ANTR_i^*$ at $Z_i = 0$.

5. Repeat steps 2-4 1,000 times, yielding 1,000 sets of estimates.

6. Find the average estimated kink in ANTR and mean ANTR. Obtain the elasticity in each iteration by dividing using the iteration-specific reduced form and the *average* first stage kink and mean. Calculate delta method standard errors (using the asymptotic standard errors of the reduced form and the standard deviation of estimates in the first stage).

7. Repeat steps 1-6 for each measurement error process (the three described above, as well as a no measurement error benchmark), and for two assumed structural parameters, $\beta = 0$ and $\beta = 1.85$. These parmaeters correspond to an elasticity of zero and 2.41, the latter being five times our baseline estimate, yielding an elasticity which is roughly our estimate after adjusting for measuring error (as reported in Appendix Table E.4).

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

D.3.1 Results

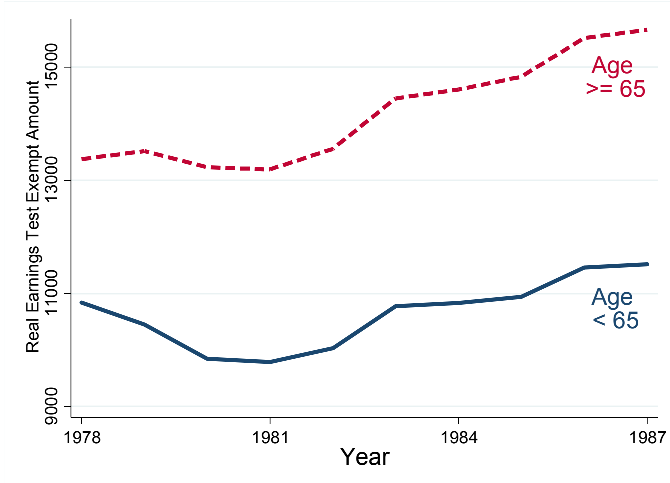
Our results from two sets of simulations are reported in Appendix Table E.5. The first table reports estimates using $\beta = 1.85$ and the second using $\beta = 0$, corresponding to elasticities of 2.41 and 0. The tables report the mean and SD of the estimated reduced form and first stage kinks, as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

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

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

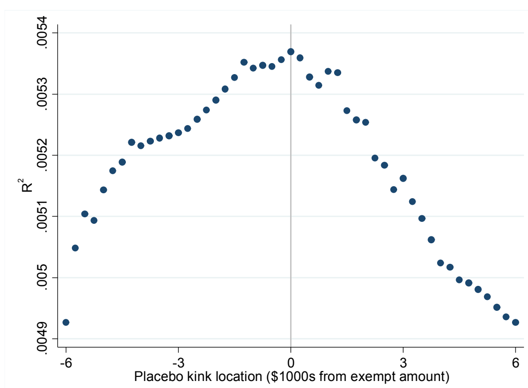
E Appendix Figures and Tables

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



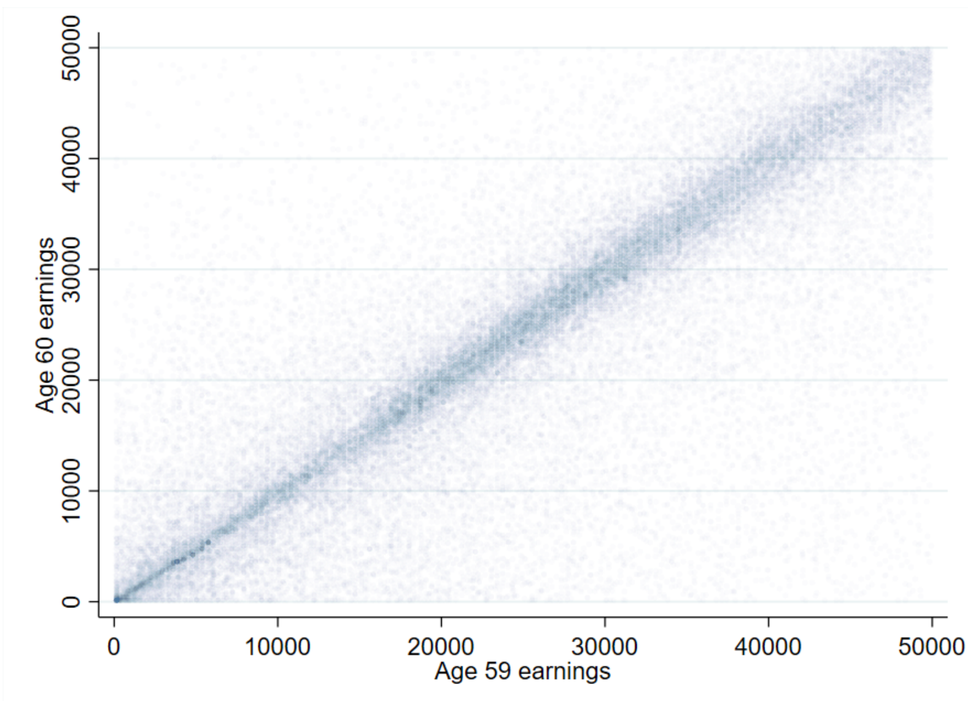
Notes: The figure shows the real value of the exempt amount over time among those 62-64 years old (labeled “Age<65” in the graph) and those above (labeled “Age>=65”). The AET applied to earnings of claimants from ages 62 to 71 from 1978 to 1982, but only to claimants aged 62 to 69 from 1983 to 1989.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

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



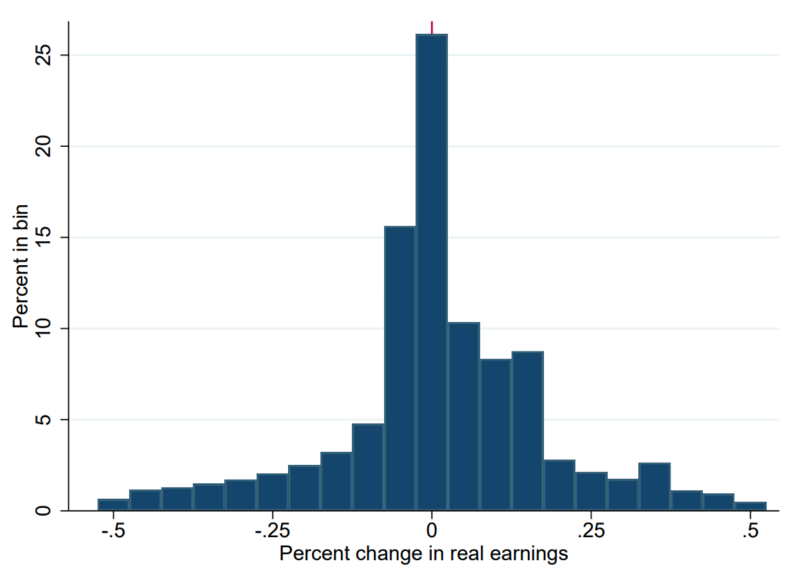
Notes: The figure plots the R-squared of our baseline specification against the “placebo” kink location relative to the exempt amount, following Landais (2015). The vertical line denotes the actual location of the exempt amount. As described in the text, we estimate a set of placebo changes in slope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

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



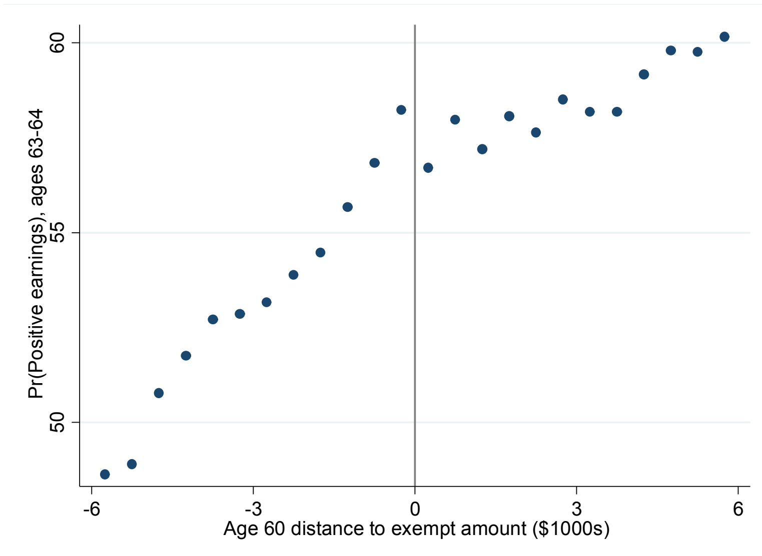
Notes: The figure reproduces Panel B of 8, but omits the 45 degree line. See Figure 8 for additional details.

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



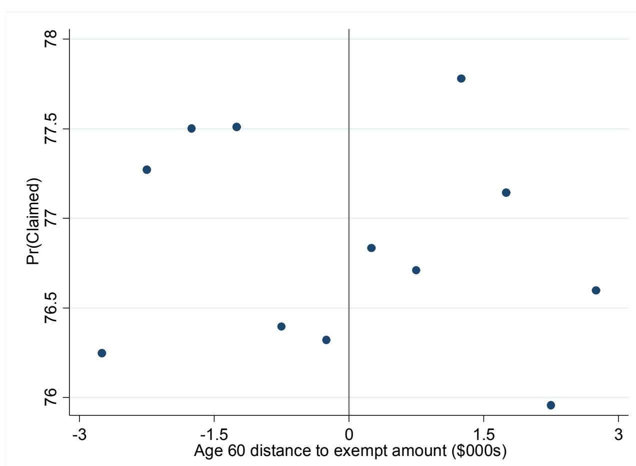
Notes: This histogram shows that there is a large mass near zero percent nominal earnings growth from one year to the next in a wider set of ages, 28 to 59, than we focus on in the main analysis. This indicates that among a broad set of ages, a substantial mass of individuals have no growth in desired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

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



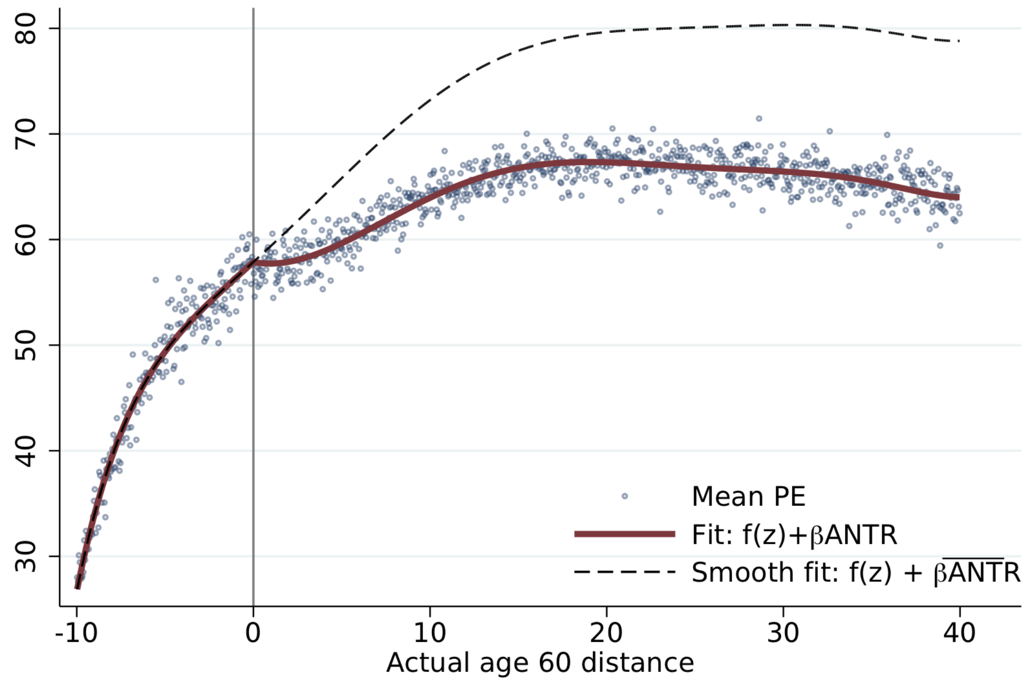
Notes: See the notes to Figure 4. This figure is identical to Figure 4, except that the range of the x-axis on this figure runs from -\$6,000 to \$6,000. Like Figure 4, this figure also shows a clear, discontinuous change in slope at the exempt amount.

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



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

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



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

Appendix Table E.1: Robustness of Elasticity Estimates

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

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

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

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

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

Appendix Table E.3: Elasticity Estimates Accounting for Claiming

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

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

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

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

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

Appendix Table E.5: Validating the Measurement Error Correction

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

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