

Supplemental Appendix: Efficiency Criteria, Income Taxation, and Heterogeneous Elasticities

By JOHN STURM BECKO AND ANDRÉ SZTUTMAN

A. PROOFS

A1. Derivatives of household and aggregate quantities

Our theoretical results rely on expressions for the first two derivatives of household income and utility as well as aggregate tax revenue and welfare. We derive these below.

Lemma A.1 (Derivatives of household income). *Fix any reform $(\Delta, \Gamma) \in \Lambda$ and let $R(\iota) \equiv R_0 + \iota\Delta + (\iota^2/2)\Gamma$. There exists $\bar{\iota} > 0$ small enough that for all $h \in \mathcal{H}$, for all $\iota \in B_{\bar{\iota}}(0)$,*

$$\begin{aligned}
 \frac{d \log z^h(\iota)}{d\iota} &= \eta^h(\iota) \frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} + \varepsilon^h(\iota) \frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)} \\
 \frac{d^2 \log z^h(\iota)}{d\iota^2} &= \eta^h(\iota) \frac{R_{\iota\iota}(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} + \varepsilon^h(\iota) \frac{R''_{\iota\iota}(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)} \\
 &+ \left[-\eta^h(\iota) - 2 \frac{R'(z^h(\iota); \iota) z^h(\iota)}{R(z^h(\iota); \iota)} \eta^h(\iota)^2 + \eta_{+0}^h(\iota) \right] \left(\frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} \right)^2 \\
 &+ \left[-2 \left(\frac{R'(z^h(\iota); \iota) z^h(\iota)}{R(z^h(\iota); \iota)} + \frac{R''(z^h(\iota); \iota) z^h(\iota)}{R'(z^h(\iota); \iota)} \right) \eta^h(\iota) \varepsilon^h(\iota) \right. \\
 &\quad \left. + 2 \frac{R'(z^h(\iota); \iota) z^h(\iota)}{R(z^h(\iota); \iota)} \eta^h(\iota)^2 + \eta_{+1}^h(\iota) + \varepsilon_{+0}^h(\iota) \right] \frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} \frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)} \\
 &+ \left[-\varepsilon^h(\iota) + 2 \frac{R'(z^h(\iota); \iota) z^h(\iota)}{R(z^h(\iota); \iota)} \eta^h(\iota) \varepsilon^h(\iota) \right. \\
 &\quad \left. - 2 \frac{R''(z^h(\iota); \iota) z^h(\iota)}{R'(z^h(\iota); \iota)} \varepsilon^h(\iota)^2 + \varepsilon_{+1}^h(\iota) \right] \left(\frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)} \right)^2 \\
 &+ 2\eta^h(\iota) \varepsilon^h(\iota) \frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} \frac{R''_{\iota\iota}(z^h(\iota); \iota) z^h(\iota)}{R'(z^h(\iota); \iota)} + 2\varepsilon^h(\iota)^2 \frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)} \frac{R''_{\iota\iota}(z^h(\iota); \iota) z^h(\iota)}{R'(z^h(\iota); \iota)}
 \end{aligned} \tag{A1}$$

where

$$\begin{aligned}
 \eta_{+0}^h(\iota) &= \eta_c^h(\iota) \left(1 + \frac{R'(z^h(\iota); \iota) z^h(\iota)}{R(z^h(\iota); \iota)} \eta^h(\iota) \right) + \eta_z^h(\iota) \eta^h(\iota) \\
 \eta_{+1}^h(\iota) &= \frac{R'(z^h(\iota); \iota) z^h(\iota)}{R(z^h(\iota); \iota)} \eta_c^h(\iota) \varepsilon^h(\iota) + \eta_z^h(\iota) \varepsilon^h(\iota) \\
 \varepsilon_{+0}^h(\iota) &= \varepsilon_c^h(\iota) \left(1 + \frac{R'(z^h(\iota); \iota) z^h(\iota)}{R(z^h(\iota); \iota)} \eta^h(\iota) \right) + \varepsilon_z^h(\iota) \eta^h(\iota) \\
 \varepsilon_{+1}^h(\iota) &= \frac{R'(z^h(\iota); \iota) z^h(\iota)}{R(z^h(\iota); \iota)} \varepsilon_c^h(\iota) \varepsilon^h(\iota) + \varepsilon_z^h(\iota) \varepsilon^h(\iota)
 \end{aligned} \tag{A2}$$

Above, $\eta^h(\iota)$ and $\varepsilon^h(\iota)$ are income and compensated elasticities of labor supply and $\eta_c^h(\iota)$, $\eta_z^h(\iota)$,

$\varepsilon_c^h(\iota)$, and $\varepsilon_z^h(\iota)$ are super-elasticities of labor supply given by

$$(A3) \quad \begin{aligned} \eta^h(\iota) &\equiv N^h(c^h(\iota), z^h(\iota); \iota), & \eta_c^h(\iota) &\equiv \frac{\partial N^h(c^h(\iota), z^h(\iota); \iota)}{\partial \log c^h}, & \eta_z^h(\iota) &\equiv \frac{\partial N^h(c^h(\iota), z^h(\iota); \iota)}{\partial \log z^h}, \\ \varepsilon^h(\iota) &\equiv E^h(c^h(\iota), z^h(\iota); \iota), & \varepsilon_c^h(\iota) &\equiv \frac{\partial E^h(c^h(\iota), z^h(\iota); \iota)}{\partial \log c^h}, & \varepsilon_z^h(\iota) &\equiv \frac{\partial E^h(c^h(\iota), z^h(\iota); \iota)}{\partial \log z^h}, \end{aligned}$$

where

$$N^h(c, z; \iota) \equiv \frac{-\frac{\partial \log M^h(c, z)}{\partial \log c}}{\frac{d \log M^h(c, z)}{d \log c} \frac{d \log R(z; \iota)}{d \log z} + \frac{\partial \log M^h(c, z)}{\partial \log z} - \frac{d \log R'(z; \iota)}{d \log z}}$$

$$E^h(c, z; \iota) \equiv \frac{1}{\frac{\partial \log M^h(c, z)}{\partial \log c} \frac{d \log R(z; \iota)}{d \log z} + \frac{\partial \log M^h(c, z)}{\partial \log z} - \frac{d \log R'(z; \iota)}{d \log z}}$$

and where $M^h(c, z) \equiv -u_z^h(c, z)/u_c^h(c, z)$ is h 's marginal rate of substitution between labor and consumption.

PROOF:

We have already assumed that derivatives of all households' income exist for small enough ι . The proof shows that these derivatives satisfy (A1).

Since preferences and the retention schedule are differentiable, we have the first-order condition for all $\iota \in B_0(\bar{\iota})$

$$(A4) \quad R'(z^h(\iota), \iota) = M^h(R(z^h(\iota); \iota), z^h(\iota)) \equiv -\frac{u_z^h(R(z^h(\iota); \iota), z^h(\iota))}{u_c^h(R(z^h(\iota); \iota), z^h(\iota))}$$

Totally differentiating and solving for $d \log z^h(\iota)/d\iota$ implies

$$(A5) \quad \frac{d \log z^h(\iota)}{d\iota} = N^h(c^h(\iota), z^h(\iota); \iota) \frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} + E^h(c^h(\iota), z^h(\iota); \iota) \frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)}$$

where $c^h(\iota) \equiv R(z^h(\iota); \iota)$. Substituting the definitions of $\eta^h(\iota)$ and $\varepsilon^h(\iota)$ implies the first line of (A1).

Next, we differentiate:

$$(A6) \quad \begin{aligned} \frac{d^2 \log z^h(\iota)}{d\iota^2} &= N^h(c^h(\iota), z^h(\iota); \iota) \left[\frac{d}{d\iota} \left(\frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} \right) + \frac{d}{d \log z^h} \left(\frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} \right) \frac{d \log z^h(\iota)}{d\iota} \right] \\ &+ E^h(c^h(\iota), z^h(\iota); \iota) \left[\frac{d}{d\iota} \left(\frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)} \right) + \frac{d}{d \log z^h} \left(\frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)} \right) \frac{d \log z^h(\iota)}{d\iota} \right] \\ &+ \frac{\partial N^h(c^h(\iota), z^h(\iota); \iota)}{\partial \iota} \frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} + \frac{\partial E^h(c^h(\iota), z^h(\iota); \iota)}{\partial \iota} \frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)} \\ &+ \left(\frac{\partial N^h(c^h(\iota), z^h(\iota); \iota)}{\partial \log c^h} \frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} + \frac{\partial E^h(c^h(\iota), z^h(\iota); \iota)}{\partial \log c^h} \frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)} \right) \frac{d \log c^h(\iota)}{d\iota} \\ &+ \left(\frac{\partial N^h(c^h(\iota), z^h(\iota); \iota)}{\partial \log z^h} \frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} + \frac{\partial E^h(c^h(\iota), z^h(\iota); \iota)}{\partial \log z^h} \frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)} \right) \frac{d \log z^h(\iota)}{d\iota} \end{aligned}$$

Note that the definitions of $N^h(c, z; \iota)$, $E^h(c, z; \iota)$, $\eta^h(\iota)$, and $\varepsilon^h(\iota)$ imply

$$(A7) \quad \begin{aligned} \frac{\partial N^h(c^h(\iota), z^h(\iota); \iota)}{\partial \iota} &= \eta^h(\iota) \left[\eta^h(\iota) \left(\frac{R'_\iota(z^h(\iota); \iota) z^h(\iota)}{R(z^h(\iota); \iota)} - \frac{R'(z^h(\iota); \iota) z^h(\iota)}{R(z^h(\iota); \iota)} \frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} \right) \right. \\ &\quad \left. + \varepsilon^h(\iota) \left(\frac{R''_\iota(z^h(\iota); \iota) z^h(\iota)}{R'(z^h(\iota); \iota)} - \frac{R''(z^h(\iota); \iota) z^h(\iota)}{R'(z^h(\iota); \iota)} \frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)} \right) \right] \\ \frac{\partial E^h(c^h(\iota), z^h(\iota); \iota)}{\partial \iota} &= \varepsilon^h(\iota) \left[\eta^h(\iota) \left(\frac{R'_\iota(z^h(\iota); \iota) z^h(\iota)}{R(z^h(\iota); \iota)} - \frac{R'(z^h(\iota); \iota) z^h(\iota)}{R(z^h(\iota); \iota)} \frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} \right) \right. \\ &\quad \left. + \varepsilon^h(\iota) \left(\frac{R''_\iota(z^h(\iota); \iota) z^h(\iota)}{R'(z^h(\iota); \iota)} - \frac{R''(z^h(\iota); \iota) z^h(\iota)}{R'(z^h(\iota); \iota)} \frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)} \right) \right] \end{aligned}$$

Also, combining the definition of $c^h(\iota)$ and (A5) implies

$$(A8) \quad \frac{d \log c^h(\iota)}{d \iota} = \frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} + \frac{R'(z^h(\iota); \iota) z^h(\iota)}{R(z^h(\iota); \iota)} \left(\eta^h(\iota) \frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} + \varepsilon^h(\iota) \frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)} \right)$$

Substituting these expressions and the definitions of $\eta_c^h(\iota)$, $\eta_z^h(\iota)$, $\varepsilon_c^h(\iota)$, and $\varepsilon_z^h(\iota)$ into (A6) implies the second line of (A1).

Lemma A.2 (Derivatives of household utility). *Fix any reform $(\Delta, \Gamma) \in \Lambda$ and let $R(\iota) \equiv R_0 + \iota \Delta + (\iota^2/2)\Gamma$, and fix a set of cardinalizing functions $\{w^h\}$. There exists $\bar{\iota} > 0$ small enough that for all $h \in \mathcal{H}$, for all $\iota \in B_{\bar{\iota}}(0)$, if w^h is twice differentiable at $u^h(c^h(\iota), z^h(\iota))$, then*

$$(A9) \quad \begin{aligned} \frac{d(w^h \circ u^h)(\iota)}{d \iota} &= \lambda^h(\iota) R_\iota(z^h(\iota); \iota) \\ \frac{d^2(w^h \circ u^h)(\iota)}{d \iota^2} &= c^h(\iota) \left[c^h(\iota) \gamma^h(\iota) + \frac{R'(z^h(\iota); \iota) z^h(\iota)}{R(z^h(\iota); \iota)} \lambda^h(\iota) \frac{\eta^h(\iota)^2}{\varepsilon^h(\iota)} \right] \left(\frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} \right)^2 \\ &\quad + 2c^h(\iota) \frac{R'(z^h(\iota); \iota) z^h(\iota)}{R(z^h(\iota); \iota)} \lambda^h(\iota) \eta^h(\iota) \frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} \frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)} \\ &\quad + c^h(\iota) \frac{R'(z^h(\iota); \iota) z^h(\iota)}{R(z^h(\iota); \iota)} \lambda^h(\iota) \varepsilon^h(\iota) \left(\frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)} \right)^2 \\ &\quad + c^h(\iota) \lambda^h(\iota) \frac{R_{\iota\iota}(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} \end{aligned}$$

where $\lambda^h(\iota) \equiv \frac{\partial}{\partial c^h}(w^h \circ u^h)(c^h(\iota), z^h(\iota))$ and $\gamma^h(\iota) \equiv \frac{\partial^2}{\partial c^h^2}(w^h \circ u^h)(c^h(\iota), z^h(\iota))$.

PROOF:

We have already assumed that derivatives of all households' income—and therefore utility—exist for small enough ι . The proof shows that these derivatives satisfy (A9).

The first derivative in (A9) follows immediately from the envelope theorem (technically, we use the intensive-margin assumption that household income is differentiable in ι).

To compute the second derivative, we let $\hat{u}^h \equiv w^h \circ u^h$ and differentiate again:

$$(A10) \quad \begin{aligned} \frac{d^2(w^h \circ u^h)(\iota)}{d\iota^2} &= \hat{u}_{cc}^h(c^h(\iota), z^h(\iota))c^h(\iota) \frac{d \log c^h(\iota)}{d\iota} R_\iota(z^h(\iota); \iota) \\ &+ \hat{u}_{cz}^h(c^h(\iota), z^h(\iota))z^h(\iota) \frac{d \log z^h(\iota)}{d\iota} R_\iota(z^h(\iota); \iota) \\ &+ \hat{u}_c^h(c^h(\iota), z^h(\iota)) \left(R_{\iota\iota}(z^h(\iota); \iota) + R'_\iota(z^h(\iota); \iota)z^h(\iota) \frac{d \log z^h(\iota)}{d\iota} \right) \end{aligned}$$

Note that by the definition of $M^h(c, z) \equiv -u_z^h(c, z)/u_c^h(c, z)$ and h 's first-order condition,

$$(A11) \quad \begin{aligned} -\hat{u}_{cz}^h(c^h, z^h)c^h &= M^h(c^h, z^h)c^h \hat{u}_{cc}^h(c^h, z^h) + M_c^h(c^h, z^h)c^h \hat{u}_c^h(c^h, z^h) \\ &= R'(z^h; \iota) \left[\hat{u}_{cc}^h(c^h, z^h)c^h - \eta^h/\varepsilon^h \cdot \hat{u}_c^h(c^h, z^h) \right] \end{aligned}$$

for c^h , z^h , η^h , and ε^h evaluated at ι . Finally, we rewrite (A10), substituting for \hat{u}_{cz}^h using (A11), substituting for $d \log z^h(\iota)/d\iota$ and $d \log c^h(\iota)/d\iota$ using the expressions in the statement and the proof, respectively, of Lemma A.1, and substituting the definitions of $\lambda^h(\iota)$ and $\gamma^h(\iota)$. This results in the second line of (A9).

An immediate implication of Lemmas A.1 and A.2 is that for any reform $(\Delta, \Gamma) \in \Lambda$ there exists $\bar{\iota} > 0$ such that for $\iota \in (-\bar{\iota}, \bar{\iota})$, each household's contribution $z^h(\iota) - R(z^h(\iota); \iota)$ to aggregate tax revenue is twice continuously differentiable. Moreover, for any cardinalizing function $\{w^h\}_{h \in \mathcal{H}}$, if w^h is continuously differentiable (or twice differentiable, or twice-continuously differentiable), then so is h 's contribution $(w^h \circ V^h)(R(\iota))$ to aggregate welfare.

For the remainder of the paper, we abstract away from technical details associated with the order of integration and differentiation and assume that, for $\iota \in (-\bar{\iota}, \bar{\iota})$, (a) aggregate tax revenue and aggregate welfare inherit the smoothness properties of household-level contributions to them, respectively, and moreover (b) their derivatives in ι are equal to the integrals of the derivatives in ι of households' contributions to them. For a more technical treatment, see the working paper version, Sturm and Sztutman (2023).

Lemma A.3 (Derivatives of aggregate tax revenue). *Fix any reform $(\Delta, \Gamma) \in \Lambda$ and let $R(\iota) \equiv R_0 + \iota\Delta + (\iota^2/2)\Gamma$. If either Δ and Γ have bounded support or Assumption 1 holds, then the derivatives of the tax revenue function $Rev(R(\iota))$ satisfy*

$$(A12) \quad \begin{aligned} \left. \frac{dRev(R(\iota))}{d\iota} \right|_{\iota=0} &= \int \psi(z) \frac{\Delta(z)}{R_0(z)} g(z) dz \\ \left. \frac{d^2 Rev(R(\iota))}{d\iota^2} \right|_{\iota=0} &= \int \left[\psi(z) \frac{\Gamma(z)}{R_0(z)} + \Psi_0(z) \left(\frac{\Delta(z)}{R_0(z)} \right)^2 + \Psi_1(z) \left(\frac{\Delta'(z)}{R_0'(z)} \right)^2 \right] g(z) dz \end{aligned}$$

where $\psi(z)$, $\Psi_0(z)$, and $\Psi_1(z)$ are smooth functions and³²

$$(A13) \quad \begin{aligned} \psi(z) &= -R_0(z) + (1 - R'_0(z))z\eta(z) \\ &+ R_0(z) \left((1 - R'_0(z)) \left(\alpha(z) - \frac{d \log \varepsilon(z)}{d \log z} \right) + \frac{R''_0(z)z}{R'_0(z)} \right) \frac{\varepsilon(z)}{R'_0(z)} \\ \Psi_1(z) &= -z(1 + R'_0(z))\varepsilon(z) + z(1 - R'_0(z)) \left[\left(\alpha(z) - \frac{d \log \varepsilon^2(z)}{d \log z} \right) \varepsilon^2(z) + (\varepsilon\varepsilon^+) \right] \end{aligned}$$

³²Notice that we do not provide an explicit expression for $\Psi_0(z)$ in the lemma statement since our applications of this lemma only use the fact that $\Psi_0(z)$ is smooth. For an explicit expression, see the proof.

PROOF:

Since $c^h(\iota) = R(z^h(\iota); \iota)$, we have

$$\begin{aligned}
\left. \frac{d(z^h(\iota) - c^h(\iota))}{d\iota} \right|_{\iota=0} &= -\Delta(z_0^h) + (1 - R'_0(z_0^h))z_0^h \left. \frac{d \log z^h(\iota)}{d\iota} \right|_{\iota=0} \\
\left. \frac{d^2(z^h(\iota) - c^h(\iota))}{d\iota^2} \right|_{\iota=0} &= -\Gamma(z_0^h) - 2\Delta'(z_0^h)z_0^h \left. \frac{d \log z^h(\iota)}{d\iota} \right|_{\iota=0} \\
&\quad - R''_0(z_0^h)(z_0^h)^2 \left(\left. \frac{d \log z^h(\iota)}{d\iota} \right|_{\iota=0} \right)^2 \\
&\quad + (1 - R'_0(z_0^h))z_0^h \left(\left(\left. \frac{d \log z^h(\iota)}{d\iota} \right|_{\iota=0} \right)^2 + \left. \frac{d^2 \log z^h(\iota)}{d\iota^2} \right|_{\iota=0} \right)
\end{aligned} \tag{A14}$$

Combining this observation with the expressions in Lemma A.1 for the derivatives of household income, we obtain

$$\left. \frac{d\text{Rev}(R(\iota))}{d\iota} \right|_{\iota=0} = \int [A_1(z)\Delta(z) + B_1(z)\Delta'(z)] dz \tag{A15}$$

where

$$A_1(z) = \left(-1 + \frac{(1 - R'_0(z))z}{R_0(z)} \eta(z) \right) g(z) \quad \text{and} \quad B_1(z) = \frac{(1 - R'_0(z))z}{R'_0(z)} \varepsilon(z) g(z). \tag{A16}$$

and

$$\begin{aligned}
\left. \frac{d^2\text{Rev}(R(\iota))}{d\iota^2} \right|_{\iota=0} &= \int \left[A_2(z)\Delta(z)^2 + 2B_2(z)\Delta(z)\Delta'(z) + C_2(z)\Delta'(z)^2 \right. \\
&\quad \left. + D_2(z)\Delta(z)\Delta''(z) + 2E_2(z)\Delta'(z)\Delta''(z) \right. \\
&\quad \left. + F_2(z)\Gamma(z) + G_2(z)\Gamma'(z) \right] dz \tag{A17}
\end{aligned}$$

where

$$\begin{aligned}
A_2(z) &= \frac{(1 - R'_0(z))z}{R_0(z)^2} \left(-\eta(z) + \left(1 - 2\frac{R'_0(z)z}{R_0(z)} - \frac{R''_0(z)z}{1 - R'_0(z)} \right) \eta(z)^2 + \eta_{+0}(z) \right) g(z) \\
B_2(z) &= \left[-\frac{z}{R_0(z)} \eta(z) \right. \\
&\quad \left. + \frac{(1 - R'_0(z))z}{R_0(z)R'_0(z)} \left(\left(1 - \frac{R'_0(z)z}{R_0(z)} - \frac{R''_0(z)z}{R'_0(z)} - \frac{R''_0(z)z}{1 - R'_0(z)} \right) (\eta\varepsilon)(z) \right. \right. \\
&\quad \left. \left. + \frac{R'_0(z)z}{R_0(z)} \eta^2(z) + \frac{\eta_{+1}(z) + \varepsilon_{+0}(z)}{2} \right) \right] g(z) \\
C_2(z) &= \left[-\frac{(1 + R'_0(z))z}{R'_0(z)^2} \varepsilon(z) + \frac{(1 - R'_0(z))z}{R'_0(z)^2} \left(1 - 2\frac{R''_0(z)z}{R'_0(z)} - \frac{R''_0(z)z}{1 - R'_0(z)} \right) \varepsilon^2(z) \right. \\
&\quad \left. + 2\frac{(1 - R'_0(z))z^2}{R_0(z)R'_0(z)} (\eta\varepsilon)(z) + \frac{(1 - R'_0(z))z}{R'_0(z)^2} \varepsilon_{+1}(z) \right] g(z) \tag{A18}
\end{aligned}$$

and where

$$\begin{aligned}
D_2(z) &= 2 \frac{(1 - R'_0(z))z^2}{R_0(z)R'_0(z)} (\eta\varepsilon)(z)g(z) \\
E_2(z) &= \frac{(1 - R'_0(z))z^2}{R'_0(z)^2} \varepsilon^2(z)g(z) \\
F_2(z) &= \left(-1 + \frac{1 - R'_0(z)}{R_0(z)} z\eta(z) \right) g(z) \\
G_2(z) &= \frac{1 - R'_0(z)}{R'_0(z)} z\varepsilon(z)g(z)
\end{aligned}
\tag{A19}$$

Next, we integrate (A15) and (A17) by parts, leveraging the assumptions that R_0 , Δ , Γ , and all income-conditional elasticity moments are sufficiently smooth.

$$\begin{aligned}
\frac{d\text{Rev}(R(\iota))}{d\iota} \Big|_{\iota=0} &= \int \left[A_1(z) - B'_1(z) \right] \Delta(z) dz + \left[B_1(z) \Delta(z) \right] \Big|_0^\infty \\
\frac{d^2\text{Rev}(R(\iota))}{d\iota^2} \Big|_{\iota=0} &= \int \left[\left(A_2(z) - B'_2(z) + \frac{D''_2(z)}{2} \right) \Delta(z)^2 \right. \\
&\quad \left. + (C_2(z) - D_2(z) - E'_2(z)) \Delta'(z)^2 \right. \\
&\quad \left. + (F_2(z) - G'_2(z)) \Gamma(z) \right] dz \\
&\quad + \left[\left(B_2(z) - \frac{D'_2(z)}{2} \right) \Delta(z)^2 + D_2(z) \Delta(z) \Delta'(z) \right. \\
&\quad \left. + E_2(z) \Delta'(z)^2 + G_2(z) \Gamma(z) \right] \Big|_0^\infty
\end{aligned}
\tag{A20}$$

Computing the derivatives above reveals that

$$\begin{aligned}
A_1(z) - B'_1(z) &= F_2(z) - G'_2(z) = \psi(z)g(z)/R_0(z) \\
C_2(z) - D_2(z) - E'_2(z) &= \Psi_1(z)g(z)/R'_0(z)^2
\end{aligned}
\tag{A21}$$

where $\psi(z)$ and $\Psi_1(z)$ are as in the statement of the lemma,³³ and we define

$$\Psi_0(z) \equiv \left(A_2(z) - B'_2(z) + \frac{D''_2(z)}{2} \right) R_0(z)^2 / g(z)
\tag{A23}$$

The smoothness of these functions follows from our assumptions that taxes, the income density, as well as income-conditional elasticities and super-elasticities are smooth functions of income.

To complete the proof, it remains to show that the boundary terms in (A20) are zero under each of the two conditions provided in the lemma statement. This is immediate in the case where both Δ and Γ have bounded support. Alternatively, suppose instead that Δ and Γ may have unbounded support but Assumption 1 holds. Toward evaluating the boundary terms in the $z \rightarrow \infty$ limit, we

³³We also use the observation that

$$\varepsilon^{+h} \equiv d\mathcal{E}^h(z_0^h) / d\log z_0^h = dE^h(R_0(z_0^h), z_0^h; 0) / d\log z_0^h = d\log R'_0(z_0^h) / d\log z_0^h \varepsilon_c^h(0) + \varepsilon_z^h(0)
\tag{A22}$$

where \mathcal{E}^h is as defined in Section III.A and E^h , ε_c^h , and ε_z^h are as defined in Lemma A.1.

first explicitly compute $D'_2(z)$:

$$(A24) \quad \begin{aligned} D'_2(z) &= 2 \frac{(1 - R'_0(z))z}{R_0(z)R'_0(z)} (\eta\varepsilon)(z)g(z) \left[1 - \frac{d \log R_0(z)}{d \log z} - \frac{d \log R'_0(z)}{d \log z} - \alpha(z) \right] \\ &\quad + 2 \frac{(1 - R'_0(z))z}{R_0(z)R'_0(z)} z(\eta\varepsilon)'(z)g(z) - 2 \frac{zR''_0(z)z}{R_0(z)R'_0(z)} (\eta\varepsilon)(z)g(z) \end{aligned}$$

Note that under the assumptions in (3) and (4), there are constants k_i for which

$$(A25) \quad \begin{aligned} |D'_2(z)\Delta(z)\Delta'(z)| &\leq (1 + R'_0(z))z|(\eta\varepsilon)(z)|g(z) (k_1 + k_2|\alpha(z)|) \\ &\quad + (k_3z + k_4R_0(z)) \left| \frac{d(\eta\varepsilon)(z)}{d \log z} \right| g(z) + k_5R_0(z)|(\eta\varepsilon)(z)|g(z) \end{aligned}$$

The same assumptions imply there are constants k_i for which

$$(A26) \quad \begin{aligned} |B_1(z)\Delta(z)| &\leq k_6(1 + R'_0(z)) \frac{R_0(z)z}{R'_0(z)} \varepsilon(z)g(z) \\ |B_2(z)\Delta(z)^2| &\leq k_7zR_0(z)|\eta(z)|g(z) + k_8 \frac{R_0(z)z}{R'_0(z)} |(\eta\varepsilon)(z)|g(z) + k_9zR_0(z)|(\eta\varepsilon)(z)|g(z) \\ &\quad + \left(\frac{R_0(z)z}{R'_0(z)} + R_0(z) \right) (k_{10}|\eta^2(z)| + k_{11}|\eta_{+1}(z) + \varepsilon_{+0}(z)|) g(z) \\ |D_2(z)\Delta(z)\Delta'(z)| &\leq k_{12}(1 + R'_0(z))z^2|(\eta\varepsilon)(z)|g(z) \\ |E_2(z)\Delta'(z)^2| &\leq k_{13}(1 + R'_0(z))z^2|\varepsilon^2(z)|g(z) \\ |G_2(z)\Gamma(z)| &\leq k_{14}(1 + R'_0(z)) \frac{R_0(z)z}{R'_0(z)} |\varepsilon(z)|g(z) \end{aligned}$$

Since we have assumed all elasticities and super-elasticities have upper bounds in the population at R_0 , as does $d(\eta\varepsilon)(z)/d \log z$, Assumption 1 implies all of these terms converge in the $z \rightarrow 0$ and $z \rightarrow \infty$ limits, as desired.

Lemma A.4 (Derivatives of welfare). *Fix any reform $(\Delta, \Gamma) \in \Lambda$ and let $R(\iota) \equiv R_0 + \iota\Delta + (\iota^2/2)\Gamma$. Further, fix any twice differentiable cardinalizing functions $\{w^h\}_{h \in \mathcal{H}}$, let $W(\iota) \equiv E_h[w^h \circ V^h(R(\iota))]$, and let $\lambda^h \equiv \partial(w^h \circ u^h)(c_0^h, z_0^h)/\partial c_0^h$ and $\gamma^h \equiv \partial^2(w^h \circ u^h)(c_0^h, z_0^h)/\partial (c_0^h)^2$.*

If Assumption 1 holds and $(\lambda\eta)(z)$ is continuously differentiable as well as bounded in the $z \rightarrow 0$ and $z \rightarrow \infty$ limits, then

$$(A27) \quad \begin{aligned} \left. \frac{dW(\iota)}{d\iota} \right|_{\iota=0} &= \int \phi(z) \frac{\Delta(z)}{R_0(z)} g(z) dz \\ \left. \frac{d^2W(\iota)}{d\iota^2} \right|_{\iota=0} &= \int \left[\phi(z) \frac{\Gamma(z)}{R_0(z)} + \Phi_0(z) \left(\frac{\Delta(z)}{R_0(z)} \right)^2 + \Phi_1(z) \left(\frac{\Delta'(z)}{R'_0(z)} \right)^2 \right] g(z) dz \end{aligned}$$

where

$$(A28) \quad \begin{aligned} \phi(z) &= \lambda(z)R_0(z) & \Phi_1(z) &= (\lambda\varepsilon)(z)R'_0(z)z \\ \Phi_0(z) &= \left[R_0(z)\gamma(z) + \frac{R'_0(z)z}{R_0(z)} \left(\lambda \frac{\eta^2}{\varepsilon} \right) (z) + (\lambda\eta)(z) \left(\alpha(z) + \frac{R'_0(z)z}{R_0(z)} \right) - z(\lambda\eta)'(z) \right] R_0(z) \end{aligned}$$

If moreover $\lambda(z)$, $\gamma(z)$, $(\lambda\varepsilon)(z)$, and $(\lambda\eta^2/\varepsilon)(z)$ are continuous functions, then so are $\phi(z)$, $\Phi_0(z)$,

and $\Phi_1(z)$.

PROOF:

From the expressions in Lemma A.2 for the derivatives of each household's contribution to welfare, we obtain

$$(A29) \quad \left. \frac{dW(\iota)}{d\iota} \right|_{\iota=0} = \int \lambda(z) \Delta(z) g(z) dz$$

and

$$(A30) \quad \left. \frac{d^2W(\iota)}{d\iota^2} \right|_{\iota=0} = \int \left[A(z) \Delta(z)^2 + 2B(z) \Delta(z) \Delta'(z) + C(z) \Delta'(z)^2 + D(z) \Gamma(z) \right] dz$$

where

$$(A31) \quad \begin{aligned} A(z) &= \frac{1}{R_0(z)} \left[R_0(z) \gamma(z) + \frac{R'_0(z) z}{R_0(z)} \left(\lambda \frac{\eta^2}{\varepsilon} \right) (z) \right] g(z) \\ B(z) &= \frac{z(\lambda\eta)(z)}{R_0(z)} g(z) & C(z) &= \frac{z(\lambda\varepsilon)(z)}{R'_0(z)} g(z) & D(z) &= \lambda(z) g(z) \end{aligned}$$

Next, we integrate (A30) by parts, leveraging the assumptions that R_0 , Δ , and Γ are smooth and our assumption in the lemma statement that $(\lambda\eta)(z)$ is continuously differentiable.

$$(A32) \quad \left. \frac{d^2W(\iota)}{d\iota^2} \right|_{\iota=0} = \int \left[(A(z) - B'(z)) \Delta(z)^2 + C(z) \Delta'(z)^2 + D(z) \Gamma(z) \right] dz + \left[B(z) \Delta(z)^2 \right] \Big|_0^\infty$$

Note that $D(z) = \phi(z)g(z)/R_0(z)$ and $C(z) = \Phi_1(z)g(z)/R'_0(z)^2$, and $A(z) - B'(z) = \Phi_0(z)g(z)/R_0(z)^2$. It is clear from these expressions—combined with our assumptions that taxes and income density are smooth functions of income—that $\phi(z)$, $\Phi_0(z)$, and $\Phi_1(z)$ are continuous functions of income if $\lambda(z)$, $\gamma(z)$, $(\lambda\varepsilon)(z)$, and $(\lambda\eta^2/\varepsilon)(z)$ are.

To complete the proof, it remains to show that the boundary term in (A20) is zero under the conditions provided in the lemma statement. To see this, note that, under the assumptions in (4), there is a constant k for which

$$(A33) \quad \left| B(z) \Delta(z)^2 \right| \leq k(\lambda\eta)(z) z R_0(z) g(z).$$

Under Assumption 1 and the assumption on $(\lambda\eta)(z)$ in the lemma statement, the RHS converges to zero in the $z \rightarrow 0$ and $z \rightarrow \infty$ limits, as desired.

A2. Proof of Theorem 4

By Lemma A.3, since $\Delta^*(z, \ell)$ has bounded support, we have

$$(A34) \quad \begin{aligned} \delta \text{Rev}(z) &= \lim_{\ell \rightarrow +0} \frac{1}{\ell} \int_{z-\ell}^{z+\ell} \psi(\tilde{z}) \frac{\Delta^*(\tilde{z}; z, \ell)}{R_0(\tilde{z})} g(\tilde{z}) d\tilde{z} \\ \delta^2 \text{Rev}(z) &= \lim_{\ell \rightarrow +0} \ell \int_{z-\ell}^{z+\ell} \left[\Psi_0(\tilde{z}) \left(\frac{\Delta^*(\tilde{z}; z, \ell)}{R_0(\tilde{z})} \right)^2 + \Psi_1(\tilde{z}) \left(\frac{\Delta^{*'}(\tilde{z}; z, \ell)}{R'_0(\tilde{z})} \right)^2 \right] g(\tilde{z}) d\tilde{z} \end{aligned}$$

where $\psi(z)$, $\Psi_0(z)$, and $\Psi_1(z)$ are continuous functions. By the continuity of these terms, R_0 , R'_0 , and $g(z)$; by the strict positivity of $R_0(z)$ and $R'_0(z)$ across all $z > 0$; and noting that $\Delta^*(\tilde{z}; z, \ell) =$

$\Delta^*((\tilde{z} - z)/\ell; 0, 1)$; we have

$$\begin{aligned}
\delta \text{Rev}(z) &= \lim_{\ell \rightarrow +0} \left(\frac{\psi(z)g(z)}{R_0(z)} + o(\ell) \right) \ell^{-1} \int_{z-\ell}^{z+\ell} \Delta^*((\tilde{z} - z)/\ell; 0, 1) d\tilde{z} \\
&= \frac{\psi(z)g(z)}{R_0(z)} \int_{-1}^1 \Delta^*(x; 0, 1) dx \\
\text{(A35)} \quad \delta^2 \text{Rev}(z) &= \lim_{\ell \rightarrow +0} \left(\frac{\Psi_0(z)g(z)}{R_0(z)^2} + o(\ell) \right) \ell \int_{z-\ell}^{z+\ell} \Delta^*((\tilde{z} - z)/\ell; 0, 1)^2 d\tilde{z} \\
&\quad + \lim_{\ell \rightarrow +0} \left(\frac{\Psi_1(z)g(z)}{R'_0(z)^2} + o(\ell) \right) \ell \int_{z-\ell}^{z+\ell} \ell^{-2} \Delta^{*'}((\tilde{z} - z)/\ell; 0, 1)^2 d\tilde{z} \\
&= 0 + \frac{\Psi_1(z)g(z)}{R'_0(z)^2} \int_{-1}^1 \Delta^{*'}(x; 0, 1)^2 dx
\end{aligned}$$

Define $\kappa_1(z) \equiv g(z) \int_{-1}^1 \Delta^*(x; 0, 1) dx$ and $\kappa_2(z) \equiv zg(z)/R'_0(z)^2 \cdot \int_{-1}^1 \Delta^{*'}(x; 0, 1)^2 dx$, and note that $\int_{-1}^1 \Delta^*(x; 0, 1) dx, \int_{-1}^1 \Delta^{*'}(x; 0, 1)^2 dx > 0$. Note that $\kappa_1(z)$ and $\kappa_2(z)$ are smooth, by the smoothness of $g(z)$ and $R_0(z)$, and strictly positive for $z > 0$, by the assumptions that R_0 is strictly positively valued and g has full support, and by the construction of Δ^* . The result then follows from (A35) and the expressions for $\psi(z)$ and $\Psi_1(z)$ in Lemma A.3.

A3. Proof of Theorem 1

Since rationalizability implies Pareto efficiency, it suffices to show that (a) if for some $z > 0$, $\delta \text{Rev}(z) > 0$, then R_0 is locally Pareto inefficient and (b) if for all $z > 0$, $\delta \text{Rev}(z) < 0$, then R_0 is locally rationalizable by some continuously differentiable cardinalizing functions.

We begin with (a), supposing that for some $z > 0$, $\delta \text{Rev}(z) > 0$. By the definition of $\delta \text{Rev}(z)$, there exists $\iota, \ell > 0$ small enough that

$$\text{(A36)} \quad \text{Rev}(R_0 + \iota \Delta^*(z, \ell)) > \text{Rev}(R_0).$$

Since there is weakly higher retention at all incomes under $R_0 + \iota \Delta^*(z, \ell)$ than under R_0 , the reform also improves welfare for all households, strictly so for those with initial incomes close enough to z , who constitute a positive mass since $g(\cdot)$ has full support. So it is a Pareto improvement.

Now turning to (b), suppose $\delta \text{Rev}(z) < 0$ for all z . We will construct a set of continuously differentiable cardinalizing functions $\{w^h\}_{h \in \mathcal{H}}$ that rationalize R_0 . To this end, we first define $\{\hat{\lambda}^h\}_{h \in \mathcal{H}}$ by

$$\text{(A37)} \quad \hat{\lambda}^h = -\delta \text{Rev}(z_0^h) / \kappa_1(z_0^h)$$

where $\kappa_1(z_0^h)$ is as in the statement of Theorem 4. We now define

$$\text{(A38)} \quad w^h(u) \equiv \left(\frac{\hat{\lambda}^h}{u^h(c_0^h, z_0^h)} - 1 \right) u + \int_0^{u - V^h(R_0)} \exp[-\text{sgn}(x)|x|^{1/2}] dx$$

Note that w^h is strictly increasing and continuously differentiable, as desired. Simple computations reveal that, for all $h \in \mathcal{H}$, $\lambda^h \equiv d(w^h \circ u^h)(c_0^h, z_0^h) / dc_0^h = \hat{\lambda}^h > 0$ and, for all $(c, z) \in (\mathbb{R}_{>0})^2$ with $u^h(c, z) \neq V^h(R_0)$, $\gamma^h(c, z) \equiv d^2(w^h \circ u^h)(c, z) / dc^2$ is defined and continuously approaches ∞ in any limit with $u^h(c, z) \rightarrow V^h(R_0)$.

Now consider any reform $(\Delta, \Gamma) \in \Lambda$. Define $R(\iota) \equiv R_0 + \iota \Delta + (\iota^2/2)\Gamma$ and $W(\iota) \equiv E_h[w^h \circ V^h(R(\iota))]$. We will argue that for all sufficiently small ι , either $\text{Rev}(R(\iota)) < \text{Rev}(R_0)$, $W(\iota) < W(0)$,

or $\text{Rev}(R(\iota)) = \text{Rev}(R_0)$ and $W(\iota) - W(0)$. Note that it suffices to consider the case where $\Delta \neq 0$, since if $\Delta = \Gamma = 0$ the result holds trivially, and if $\Delta = 0$ and $\Gamma \neq 0$ the result is implied by the case of the alternative reform ($\tilde{\Delta} = \Gamma, \tilde{\Gamma} = 0$).

To do this, define $\{\tilde{w}^h\}_{h \in \mathcal{H}}$ by

$$(A39) \quad \tilde{w}^h(u) \equiv \left(\frac{\hat{\lambda}^h}{u^h(c_0^h, z_0^h)} - 1 \right) u + \int_0^{u - V^h(R_0)} \exp[\hat{\gamma}(z_0^h)x] dx$$

for some function $\hat{\gamma}(z)$ specified later in the proof. Note that for all h , $\tilde{w}^h(V^h(R_0)) = w^h(V^h(R_0))$, $\partial(\tilde{w}^h \circ u^h)(c_0^h, z_0^h)/\partial c_0^h = \partial(w^h \circ u^h)(c_0^h, z_0^h)/\partial c_0^h$, and there exists $d^h > 0$ such that for all $u \in B_{d^h}(V^h(R_0))$, we have $w^h(u) \leq \tilde{w}^h(u)$. By the continuity of w^h , \tilde{w}^h , u^h , R_0 , Δ , and Γ , and by the intensive margin assumption, this implies there similarly exists $\iota^h > 0$ such that $w^h(V^h(R(\iota))) \leq \tilde{w}^h(V^h(R(\iota)))$ for all $\iota \in B_{\iota^h}(0)$. This in turn implies there exist smooth functions $\{\hat{w}^h\}_{h \in \mathcal{H}}$ so that (a) for each h , $\hat{w}^h(V^h(R(\iota))) = \tilde{w}^h(V^h(R(\iota)))$ for $\iota \in B_{\iota^h/2}(0)$, (b) $w^h(u) \leq \hat{w}^h(u)$ for all u , and (c) $\hat{w}^h(V^h(R_0)) = w^h(V^h(R_0))$. Letting $\hat{W}(\iota) \equiv E_h[\hat{w}^h \circ V^h(R(\iota))]$, it therefore suffices to show that there exists $\bar{\iota} > 0$ such that

$$(A40) \quad 0 \in \arg \max_{\iota \in (-\bar{\iota}, \bar{\iota})} \hat{W}(\iota) \quad \text{s.t.} \quad \text{Rev}(R(\iota)) \geq \text{Rev}(R(0)).$$

By Lagrangian sufficiency it suffices in turn to show that within some non-empty interval $(-\bar{\iota}, \bar{\iota})$, $\iota = 0$ maximizes $\hat{W}(\iota) + \text{Rev}(R(\iota))$. This in turn follows from the definition of first and second derivatives, provided that

$$(A41) \quad \left. \frac{d}{d\iota} \right|_{\iota=0} (\hat{W}(\iota) + \text{Rev}(R(\iota))) = 0 \quad \text{and} \quad \left. \frac{d^2}{d\iota^2} \right|_{\iota=0} (\hat{W}(\iota) + \text{Rev}(R(\iota))) < 0$$

To this end, we compute the derivatives above using Lemmas A.3 and A.4. These lemmas apply because we have imposed Assumption 1 and—letting $\check{\lambda}^h \equiv \partial(\hat{w}^h \circ u^h)(c_0^h, z_0^h)/\partial c_0^h$ and $\check{\lambda}(z) \equiv E_h[\check{\lambda}^h | z_0^h = z]$ —the construction of w^h and \hat{w}^h implies $\check{\lambda}^h = \lambda^h = -\delta \text{Rev}(z_0^h)/\kappa_1(z_0^h)$. This ensures that (a) $\check{\lambda}(z) = -\delta \text{Rev}(z)/\kappa_1(z)$ is smooth by Theorem 4 and the smoothness of primitives, (b) $(\check{\lambda}\eta)(z) = \check{\lambda}(z)\eta(z)$, where recall $\eta(z)$ is bounded by assumption, and (c) by Theorem 4, $\check{\lambda}(z) \propto -\delta \text{Rev}(z)/g(z)$ which is bounded in the $z \rightarrow 0$ and $z \rightarrow \infty$ limits by Assumption 1, the assumptions in (3), the assumption elasticities and $\alpha(z)$ are bounded, and the assumption $\varepsilon(z)$'s derivative in log income is bounded.

We begin with the first derivative of the Lagrangian. Lemmas A.3 and A.4 and Theorem 4 imply

$$(A42) \quad \left. \frac{d}{d\iota} \right|_{\iota=0} (\hat{W}(\iota) + \text{Rev}(R(\iota))) = \int \left[\check{\lambda}(z) + \frac{\delta \text{Rev}(z)}{\kappa_1(z)} \right] \Delta(z) g(z) dz = 0$$

We now turn to the second derivative of the Lagrangian. Let $\hat{\gamma}^h \equiv \partial^2(\hat{w}^h \circ u^h)(c_0^h, z_0^h)/\partial(c_0^h)^2$ and $\hat{\gamma}(z) \equiv E_h[\hat{\gamma}^h | z_0^h = z]$, and note that $\hat{\gamma}^h = \hat{\gamma}(z_0^h)$ by construction. By Lemmas A.3 and A.4, we

have

$$(A43) \quad \frac{d^2}{d\iota^2} \Big|_{\iota=0} (\hat{W}(\iota) + \text{Rev}(R(\iota))) = \int \left[(\Phi_0(z) + \Psi_0(z)) \left(\frac{\Delta(z)}{R_0(z)} \right)^2 + (\Phi_1(z) + \Psi_1(z)) \left(\frac{\Delta'(z)}{R_0'(z)} \right)^2 + \cancel{(\phi(z) + \psi(z))} \frac{\Gamma(z)}{R_0(z)} \right] g(z) dz$$

where the cancellation uses that $\phi(z) = \check{\lambda}(z)R_0(z)$ and $\psi(z) = \delta \text{Rev}(z)R_0(z)/\kappa_1(z)$. By the expressions for $\Phi_0(z)$, $\Psi_0(z)$, $\Phi_1(z)$, and $\Psi_1(z)$ in Lemmas A.3 and A.4, and by the assumption that R_0 is strictly positive, there exists—for every $z > 0$ with $\Delta(z) \neq 0$ —a value of $\hat{\gamma}(z)$ negative enough to guarantee that the integrand in (A43) is strictly negative at that z . Take the functions $\{\hat{w}^h\}_{h \in \mathcal{H}}$ to have been constructed using such a $\hat{\gamma}$ function. Since (a) by the smoothness of Δ and the assumption Δ is nonzero, there exists positive measure of z s at which $\Delta(z) \neq 0$ and (b) there are at most countably many z with $\Delta'(z) \neq 0$ and $\Delta(z) = 0$, we conclude that the expression in (A43) is strictly negative. This completes the proof.

A4. Proof of Theorems 2.a and 2.b

Theorems 2.a and 2.b both rely on the existence of a tax reform that has positive second-order effects on revenue and all households' welfare. We begin by providing such a reform in the lemma below and then apply it to both theorems. Part of the lemma relies on several technical results implied by Assumption 2; we relegate these lemmas to Supplemental Appendix A7

Lemma A.5. *Suppose $\delta^2 \text{Rev}(z^*) > 0$ at some income $z^* > 0$. Fix any set of cardinalizing functions $\{w^h\}_{h \in \mathcal{H}}$ with locally bounded curvature.*

Then for all sufficiently small $\ell > 0$, there exists a reform $(\Delta, \Gamma) \in \Lambda$ with $\text{supp } \Delta \cup \text{supp } \Gamma = (z^ - \ell, z^* + \ell)$,*

$$(A44) \quad \frac{d^2}{d\iota^2} \Big|_{\iota=0} \text{Rev}(R(\iota)) > 0, \quad \text{and} \quad \frac{d^2}{d\iota^2} \Big|_{\iota=0} (w^h \circ V^h)(R(\iota)) \geq 0$$

for all households $h \in \mathcal{H}$, with strict inequality if $z_0^h \in (z^ - \ell, z^* + \ell)$ and $w^{h'}(V^h(R_0)) > 0$, where $R(\iota) \equiv R_0 + \iota\Delta + (\iota^2/2)\Gamma$.*

If the assumptions above hold with cardinalizing functions equal to the identity and, in addition, Assumption 2 holds, then there exists $\bar{\iota} > 0$ such that for all $\iota \in (-\bar{\iota}, \bar{\iota})$,

$$(A45) \quad \frac{d^2}{d\iota^2} V^h(R(\iota)) \geq 0 \quad \text{for all } h \in \mathcal{H}$$

PROOF:

Define $\Delta = \Delta^*(z^*, \ell/2)$ and $\Gamma = \ell^{-3/2} \Delta^*(z^*, \ell)$. Note that $(\Delta, \Gamma) \in \Lambda$ and $\Delta(z_0^h) = \Gamma(z_0^h) = 0$ if $z_0^h \notin (z^* - \ell, z^* + \ell)$. We will argue that for small enough $\ell > 0$, this reform delivers the desired second derivatives.

First, consider revenue effects. By Lemma A.3, and using the observation that $\Delta(z; z^*, \ell/2) =$

$\Delta^*((z - z^*)/(\ell/2); 0, 1)$,

$$(A46) \quad \frac{d^2}{dt^2} \Big|_{t=0} \text{Rev}(R(t)) = \int \left[\psi(z) \frac{\ell^{-3/2} \Delta^*(z; z^*, \ell)}{R_0(z)} + \Psi_0(z) \left(\frac{\Delta^*(z; z^*, \ell/2)}{R_0(z)} \right)^2 + \Psi_1(z) \left(\frac{2\ell^{-1} \Delta^{*'}((z - z^*)/(\ell/2); 0, 1)}{R'_0(z)} \right)^2 \right] g(z) dz$$

where $\psi(z)$, $\Psi_0(z)$, and $\Psi_1(z)$ are continuous functions defined in Lemma A.3 and where $\delta^2 \text{Rev}(z^*) > 0$ implies $\Psi_1(z^*) > 0$ by Theorem 4. By this continuity and the continuity and positivity of R_0 and R'_0 , $\psi(z) \Delta^*(z; z^*, \ell) g(z)/R_0(z)$ and $\Psi_0(\Delta^*(z; z^*, \ell/2) g(z)/R_0(z))^2$ are bounded in absolute value within $[z^* - \ell, z^* + \ell]$ by some functions $\zeta_1(\ell)$ and $\zeta_2(\ell)$ that may be taken to converge to $|\psi(z^*) g(z^*)/R_0(z^*)|$ and $|\Psi_0(z^*) g(z^*)/R_0(z^*)^2|$, respectively, in the $\ell \rightarrow 0$ limit. Similarly, $\max_{z \in [z^* - \ell, z^* + \ell]} |\Psi_1(z) g(z)/R'_0(z)^2 - \Psi_1(z^*) g(z^*)/R'_0(z^*)^2| \rightarrow 0$ as $\ell \rightarrow 0$. Since $|\Delta^*(z; z^*, \ell/2)|, |\Delta^*(z; z^*, \ell)| \leq 1$, we therefore have

$$(A47) \quad \begin{aligned} \frac{d^2}{dt^2} \Big|_{t=0} \text{Rev}(R(t)) &\geq \int_{z^* - \ell}^{z^* + \ell} \left[\ell^{-3/2} \zeta_1(\ell) + \zeta_2(\ell) + 2\ell^{-2} \left(\frac{\Psi_1(z^*) g(z^*)}{R_0(z^*)^2} - o(\ell) \right) \Delta^{*'} \left(\frac{z - z^*}{\ell/2}; 0, 1 \right)^2 \right] dz \\ &= 2\ell^{-1/2} \zeta_1(\ell) + 2\ell \zeta_2(\ell) \\ &\quad + 2\ell^{-1} \left(\frac{\Psi_1(z^*) g(z^*)}{R_0(z^*)^2} - o(\ell) \right) \int_{-1}^1 \Delta^{*'}(u; 0, 1)^2 du \end{aligned}$$

Since the integral in the final line above is strictly positive, as are $\Psi_1(z^*)$, $g(z^*)$, and $R_0(z^*)$, the RHS is strictly positive for small enough $\ell > 0$.

Second, consider welfare effects. By Lemma A.2,

$$(A48) \quad \begin{aligned} \frac{d^2(w^h \circ V^h)(R(t))}{dt^2} \Big|_{t=0} &\geq c_0^h \left[c_0^h \gamma^h \left(\frac{\Delta^*(z_0^h; z^*, \ell/2)}{R_0(z_0^h)} \right)^2 + \lambda^h \frac{\ell^{-3/2} \Delta^*(z_0^h; z^*, \ell)}{R_0(z_0^h)} + 2 \frac{R'_0(z_0^h) z_0^h}{R_0(z_0^h)} \lambda^h \eta^h \frac{\Delta^*(z_0^h; z^*, \ell/2)}{R_0(z_0^h)} \frac{\Delta^{*'}(z_0^h; z^*, \ell/2)}{R'_0(z_0^h)} \right] \end{aligned}$$

where $\lambda^h \equiv \partial(w^h \circ u^h)(c_0^h, z_0^h)/\partial c_0^h$ and $\gamma^h \equiv \partial^2(w^h \circ u^h)(c_0^h, z_0^h)/\partial (c_0^h)^2$, and where the inequality comes from omitting terms that can be signed as weakly positive. Using the observations that (a) $\Delta^*(z; z^*, \ell) \geq \Delta^*(z; z^*, \ell/2)$ and $\Delta^*(z; z^*, \ell) \in [0, 1]$, (b) $\Delta^{*'}(z; z^*, \ell/2) = 2\ell^{-1} \Delta^{*'}((z - z^*)/\ell; 0, 1)$, and (c) $d \log R_0(z)/d \log z \leq B^R$ by the assumptions in (3), we have

$$(A49) \quad \begin{aligned} \frac{d^2(w^h \circ V^h)(R(t))}{dt^2} \Big|_{t=0} &\geq \lambda^h \Delta^*(z_0^h; z^*, \ell) \left[- \frac{\hat{\rho}(z^*, \ell)}{R_0(z_0^h)} + \ell^{-3/2} - 4\ell^{-1} B^R \bar{\eta} \frac{|\Delta^{*'}((z_0^h - z^*)/(\ell/2); 0, 1)|}{R'_0(z_0^h)} \right] \end{aligned}$$

where $\hat{\rho}(z^*, \ell) = \max_{z \in [z^* - \ell, z^* + \ell]} \bar{\rho}(z)$ and where $\bar{\rho}$ is the bounding function in the definition of locally bounded curvature. Since R'_0 and $\Delta^{*'}(z^*, 1)$ are continuous functions of income and R'_0 is strictly positive at all positive incomes, $1/R_0(z)$ and $|\Delta^{*'}((z - z^*)/(\ell/2); z^*, 1)|/R'_0(z)$ are bounded across $z \in [z^* - \ell, z^* + \ell]$ for all small enough ℓ . So is $\hat{\rho}(z^*, \ell)$ since, as $\bar{\rho}$ is continuous by assumption,

$\hat{\rho}(z^*, \ell) \rightarrow \bar{\rho}(z^*)$ as $\ell \rightarrow 0$. We conclude that for small enough ℓ , the term in brackets on the RHS of (A49) is strictly positive for all $h \in \mathcal{H}$. So $d^2(w^h \circ V^h)(R(\iota))/d\iota^2|_{\iota=0} \geq 0$ for all $h \in \mathcal{H}$, and this inequality is strict if and only if $\lambda^h = w^{h'}(V^h(R_0))u_c^h(c_0^h, z_0^h) > 0$ and $z_0^h \in (z^* - \ell, z^* + \ell)$, so that $\Delta^*(z_0^h; z^*, \ell) > 0$.

Third and finally, consider the case where Assumption 2 holds. Let $R(\iota, \ell)$ and $z^h(\iota, \ell)$ denote taxes and income as a function of both the length of the reform ℓ and its magnitude ι (and similarly for other household variables). By Lemma A.2 as well as our assumption that households act on the intensive margin for small enough reforms, for any ℓ , there exists $\bar{\iota}^\ell$ such that for $\iota \leq \bar{\iota}^\ell$, for all $h \in \mathcal{H}$,

$$(A50) \quad \frac{d^2 V^h(R(\iota, \ell))}{d\iota^2} \geq \lambda^h(\iota, \ell) \left[\frac{c^h(\iota, \ell)\gamma^h(\iota, \ell)}{\lambda^h(\iota, \ell)} \frac{R_\iota(z^h(\iota, \ell); \iota, \ell)^2}{R(z^h(\iota, \ell); \iota, \ell)} + R_{\iota\iota}(z^h(\iota, \ell); \iota, \ell) \right. \\ \left. + 2 \frac{R'(z^h(\iota, \ell); \iota, \ell)z^h(\iota, \ell)}{R(z^h(\iota, \ell); \iota, \ell)} \eta^h(\iota, \ell) \frac{R_\iota(z^h(\iota, \ell); \iota, \ell)R'_\iota(z^h(\iota, \ell); \iota, \ell)}{R'(z^h(\iota, \ell); \iota, \ell)} \right]$$

where $\lambda^h(\iota, \ell) \equiv u_c^h(c^h(\iota, \ell), z^h(\iota, \ell))$ and $\gamma^h(\iota, \ell) \equiv u_{cc}^h(c^h(\iota, \ell), z^h(\iota, \ell))$. We let $X^h(\iota, \ell)$ denote the term in brackets above. The remainder of the proof shows there exists $\ell > 0$ such that for all small enough ι , $X^h(\iota, \ell) \geq 0$ for all $h \in \mathcal{H}$. Note that it suffices to show this for h with $z_0^h \in [z^* - \ell, z^* + \ell]$ because, by Lemma A.1, other households' incomes do not respond to the reform, so—as the reform does not affect retention at their initial incomes—they are unaffected by it for any ι .

To this end, we now make several observations. First, under Assumption 2, Lemma A.11 implies there exist bounding functions $f_1(\iota, \ell)$, $f_2(\iota, \ell)$, $f_3(\iota, \ell)$ such that for all ℓ , (a) each $f_i(\iota, \ell) \rightarrow 0$ as $\iota \rightarrow 0$ and (b) there exists small enough $\bar{\iota}^\ell > 0$ such that for all $|\iota| < \bar{\iota}^\ell$, for all $h \in \mathcal{H}$,

$$(A51) \quad \left| \frac{c^h(\iota, \ell)\gamma^h(\iota, \ell)}{\lambda^h(\iota, \ell)} - \frac{c^h(0, \ell)\gamma^h(0, \ell)}{\lambda^h(0, \ell)} \right| \leq f_1(\iota, \ell) \\ \left| \frac{R'(z^h(\iota, \ell); \iota, \ell)z^h(\iota, \ell)}{R(z^h(\iota, \ell); \iota, \ell)} - \frac{R'(z^h(0, \ell); 0, \ell)z^h(0, \ell)}{R(z^h(0, \ell); 0, \ell)} \right| \leq f_2(\iota, \ell) \\ \left| \eta^h(\iota, \ell) - \eta^h(0, \ell) \right| \leq f_3(\iota, \ell)$$

Second, the assumption in the lemma statement about locally bounded curvature, the assumption that income elasticities at R_0 are bounded in magnitude by some $\bar{\eta}$, and the assumptions in (3) imply that the quantities above are, at $\iota = 0$, bounded across all h with $z_0^h \in [z^* - \ell, z^* + \ell]$ by $\hat{\rho}(z^*, \ell)$, B^R , and $\bar{\eta}$, respectively. Third, we note that for the particular reform constructed above, $|R_\iota(z; \iota, \ell)| \leq (1 + |\iota|\ell^{-3/2})\mathbb{1}_{z \in [z^* - \ell/2, z^* + \ell/2]}$, $|R'_\iota(z; \iota, \ell)| \leq \ell^{-1}\bar{s}(2 + |\iota|\ell^{-3/2})\mathbb{1}_{z \in [z^* - \ell/2, z^* + \ell/2]}$ for $\bar{s} \equiv \max_{z \in [-1, 1]} |\Delta^{*'}(z; 0, 1)|$, and $R_{\iota\iota}(z; \iota, \ell) \geq \underline{\Delta}\ell^{-3/2}\mathbb{1}_{z \in [z^* - \ell/2, z^* + \ell/2]}$ for $\underline{\Delta} \equiv \min_{z \in [-1/2, 1/2]} \Delta^*(z; 0, 1) > 0$. Fourth, let $\underline{R} \equiv \min_{z \in [z^*/2, 2z^*]} R_0(z)$ and $\underline{r} \equiv \min_{z \in [z^*/2, 2z^*]} R'_0(z)$, and note that $\underline{R}, \underline{r} > 0$ since we have assumed the level and slope of retention are positive at all positive incomes and R_0 is smooth. For the particular reform constructed above, this implies $R(z^h(\iota, \ell); \iota, \ell) \geq \underline{R} - |\iota| - \frac{|\iota|^2}{2}\ell^{-3/2}$ and $R'(z^h(\iota, \ell); \iota, \ell) \geq \underline{r} - \bar{s}\ell^{-1} \left(2|\iota| + \frac{|\iota|^2}{2}\ell^{-3/2} \right)$.

Putting these observations together, (A50) implies that for any $\ell < z^*/2$, for any ι small enough

that $|\iota| < \hat{\iota}^\ell, \bar{\iota}^\ell$ and $|\iota| + \frac{|\iota|^2}{2} < \underline{R}$, and $\bar{s}\ell^{-1} \left(2|\iota| + \frac{|\iota|^2}{2}\ell^{-3/2} \right) < \underline{r}$, we have

$$(A52) \quad X^h(\iota, \ell) \geq \mathbf{1}_{z^h(\iota, \ell) \in [z^* - \ell/2, z^* + \ell/2]} \left[-(\hat{\rho}(z^*, \ell) + f_1(\iota, \ell)) \frac{(1 + |\iota|\ell^{-3/2})^2}{\underline{R} - |\iota| - \frac{|\iota|^2}{2}\ell^{-3/2}} + \underline{\Delta}\ell^{-3/2} \right. \\ \left. + 2(B_R + f_2(\iota, \ell))(\bar{\eta} + f_3(\iota, \ell)) \frac{(1 + |\iota|\ell^{-3/2})\bar{s} \left(2\ell^{-1} + |\iota|\ell^{-1/2} \right)}{\underline{r} - \bar{s}\ell^{-1} \left(2|\iota| + \frac{|\iota|^2}{2}\ell^{-3/2} \right)} \right]$$

for all h with $z_0^h \in [z^* - \ell, z^* + \ell]$. Recalling that $\hat{\rho}(z^*, \ell) \rightarrow \bar{\rho}(z^*)$ as $\ell \rightarrow 0$, note that there is ℓ small enough that the term in brackets above is strictly positive when $\iota = 0$. Fixing such a ℓ , note there is $\check{\iota}^\ell$ small enough that the same term is strictly positive for $\iota \leq \check{\iota}^\ell$. So we have ℓ for which $X^h(\iota, \ell) \geq 0$ for all h , for all small enough ι , as desired.

We now apply this lemma to prove Theorems 2.a and 2.b.

PROOF OF THEOREM 2.A.

Take (Δ, Γ) as in Lemma A.5 and construct a random tax reform $(\mathbf{\Delta}, \mathbf{\Gamma})$ defined by $(\mathbf{\Delta}(\omega), \mathbf{\Gamma}(\omega)) = (\Delta, \Gamma)$ with probability $\frac{1}{2}$ and $(\mathbf{\Delta}(\omega), \mathbf{\Gamma}(\omega)) = (-\Delta, \Gamma)$ with probability $\frac{1}{2}$. Since $(\Delta, \Gamma) \in \Lambda$, $(-\Delta, \Gamma) \in \Lambda$ as well, so $(\mathbf{\Delta}, \mathbf{\Gamma}) \in \mathbf{\Lambda}$. It suffices—in order for the theorem to follow from Taylor's theorem—to show that, letting $\mathbf{R}(\iota) = R_0 + \iota\mathbf{\Delta} + (\iota^2/2)\mathbf{\Gamma}$,

$$(A53) \quad \left. \frac{d}{d\iota} \right|_{\iota=0} \mathbb{E}_\omega[\text{Rev}(\mathbf{R}(\iota))] = 0 \quad \text{and} \quad \left. \frac{d}{d\iota} \right|_{\iota=0} \mathbb{E}_\omega[V^h(\mathbf{R}(\iota))] = 0 \quad \text{for all } h \in \mathcal{H}$$

and there exists $\bar{\iota}$ such that for all $\iota \in B_{\bar{\iota}}(0)$,

$$(A54) \quad \frac{d^2}{d\iota^2} \mathbb{E}_\omega[\text{Rev}(\mathbf{R}(\iota))] > 0 \quad \text{and} \quad \frac{d^2}{d\iota^2} \mathbb{E}_\omega[V^h(\mathbf{R}(\iota))] \geq 0 \quad \text{for all } h \in \mathcal{H}$$

and there exists $h \in \mathcal{H}$ for whom the final inequality is strict for all $\iota \in B_{\bar{\iota}}(0)$.

The first claim—concerning the reform's first-order effects at $\iota = 0$ —is immediate from its construction and Lemmas A.2 and A.3.

Next, consider the claim concerning the convexity of revenues. Recall from Lemma A.5 that, at $\iota = 0$, revenues have a strictly positive second derivative in ι for the reform (Δ, Γ) . It follows from Lemma (A.3) that the same is true for $(-\Delta, \Gamma)$. The twice continuous differentiability of revenues (see the discussion following Lemma A.2) then implies there exists $\hat{\iota}$ so that for all $\iota \in B_{\hat{\iota}}(0)$, $d^2\text{Rev}(R(\iota))/d\iota^2 > 0$ for both deterministic reforms, and therefore also $d^2\mathbb{E}_\omega[\text{Rev}(\mathbf{R}(\iota))]/d\iota^2 > 0$ as desired.

Finally, consider the claims concerning the convexity of household utility. Recall from Lemma A.5 that there exists $\check{\iota}$ such that for all $\iota \in B_{\check{\iota}}(0)$, $d^2\mathbb{E}_\omega[V^h(R(\iota))]/d\iota^2 \geq 0$ for all $h \in \mathcal{H}$, where $R(\iota) = R_0 + \iota\Delta + (\iota^2/2)\Gamma$. By Lemma A.2, the same is true for the reform $(-\Delta, \Gamma)$, and so the same is also true for the random reform. It remains to show that there is some household h for which utility has a strictly positive second derivative in the random reform for all $\iota \in B_{\check{\iota}}(0)$ for some $\check{\iota} > 0$. Since we have already shown the second derivative of utility weakly positive in both deterministic reforms, it suffices to show there is some household h for which the second derivative of utility is strictly positive in the one of the deterministic reforms for all $\iota \in B_{\check{\iota}}(0)$. By Lemma A.5, for households h with z_0^h close enough to z , this derivative is strictly positive for (Δ, Γ) at $\iota = 0$. Since taxes, tax reforms, and utility are smooth and since behavior is twice-continuously differentiable in ι , h 's utility is twice-continuously differentiable in ι , so h 's second derivative of utility is strictly positive on some positive interval as desired.

PROOF OF THEOREM 2.B

Fix any cardinalizing functions $\{w^h\}$ with locally bounded curvature. We will construct a reform that strictly increases revenues and welfare locally.

Toward constructing this reform, note that since $\delta^2 \text{Rev}(z) > 0$, Theorem 4 and the assumptions that taxes, the income distribution, and income-conditional elasticity moments are sufficiently smooth in income imply there exists z_1, z_2 close enough to z that $\delta^2 \text{Rev}(z_1), \delta^2 \text{Rev}(z_2) > 0$. By Lemma A.5, there exist reforms $(\Delta_1, \Gamma_1), (\Delta_2, \Gamma_2) \in \Lambda$ that have bounded and non-overlapping support and—letting $R_1(\iota) = R_0 + \iota\Delta_1 + (\iota^2/2)\Gamma_1$ and $R_2(\iota) = R_0 + \iota\Delta_2 + (\iota^2/2)\Gamma_2$ —that satisfy

$$(A55) \quad \left. \frac{d^2 \text{Rev}(R_1(\iota))}{d\iota^2} \right|_{\iota=0} > 0 \quad \text{and} \quad \left. \frac{d^2 w^h \circ V^h(R_1(\iota))}{d\iota^2} \right|_{\iota=0} \geq 0$$

for all households h , with strict inequality if $w^h(V^h(R_0)) = 0$ and either $\Delta_1(z_0^h) \neq 0$ or $\Gamma_1(z_0^h) \neq 0$, and

$$(A56) \quad \left. \frac{d^2 \text{Rev}(R_2(\iota))}{d\iota^2} \right|_{\iota=0} > 0 \quad \text{and} \quad \left. \frac{d^2 w^h \circ V^h(R_2(\iota))}{d\iota^2} \right|_{\iota=0} \geq 0$$

for all households h , with strict inequality if $w^h(V^h(R_0)) = 0$ and either $\Delta_2(z_0^h) \neq 0$ or $\Gamma_2(z_0^h) \neq 0$.

We construct a new reform as follows: First, we note that by the expression for the first derivative of tax revenue in Lemma A.3, there exist k_1, k_2 , such that any reform given to first-order by $k_1\Delta_1 + k_2\Delta_2$ has zero first-order revenue effects.³⁴ Second, recall we have assumed $\{w^h\}_{h \in \mathcal{H}}$ is such that $w^h(V^h(R_0)) > 0$ for a positive measure of households. This implies that either (a) $w^h(V^h(R_0)) > 0$ for a positive measure of households with $\Delta_1(z_0^h)$ or $\Gamma_1(z_0^h)$ non-zero or $\Delta_2(z_0^h)$ or $\Gamma_2(z_0^h)$ or (b) $w^h(V^h(R_0)) > 0$ for a positive measure of households within some interval $(\hat{z} - \ell, \hat{z} + \ell)$ with $\hat{z} > \ell > 0$ that does not overlap with the support of $(\Delta_1, \Gamma_1), (\Delta_2, \Gamma_2)$. Third, we construct a new reform (Δ, Γ) : In case (b) above we take $\Delta = k_1\Delta_1 + k_2\Delta_2$ and $\Gamma = (k_1)^2\Gamma_1 + (k_2)^2\Gamma_2 + k^2\Delta^*(\hat{z}, \ell)$ for some small $k > 0$; in case (a) we set $k = 0$. Note that since $(\Delta_1, \Gamma_1), (\Delta_2, \Gamma_2), (0, \Delta^*(\hat{z}, \ell)) \in \Lambda$ with non-overlapping support, $(\Delta, \Gamma) \in \Lambda$. Importantly, by the expressions in Lemma A.3 for the second derivative of revenue and the fact that $(\Delta_1, \Gamma_1), (\Delta_2, \Gamma_2)$, and $\Delta^*(z, \ell)$ do not overlap, we may choose k small enough that

$$(A57) \quad \left. \frac{d^2 \text{Rev}(R(\iota))}{d\iota^2} \right|_{\iota=0} = (k_1)^2 \left. \frac{d^2 \text{Rev}(R_1(\iota))}{d\iota^2} \right|_{\iota=0} + (k_2)^2 \left. \frac{d^2 \text{Rev}(R_2(\iota))}{d\iota^2} \right|_{\iota=0} + k^2 \left. \frac{d^2 \text{Rev}(\hat{R}(\iota))}{d\iota^2} \right|_{\iota=0}$$

is strictly positive, where $\hat{R}(\iota) \equiv R(\iota) = R_0 + \frac{1}{2}k^2\Delta^*(\hat{z}, \ell)$. So (Δ, Γ) 's revenue impact is zero to first order and strictly positive to second order.

We now turn to (Δ, Γ) 's welfare impact. First, letting $W(\iota) \equiv E_h[w^h \circ V^h(R(\iota))]$, suppose $dW(\iota)/d\iota|_{\iota=0} \neq 0$. Then, by the definition of first and second derivatives, there exists $\bar{\iota} > 0$ such that for all $\iota \in B_{\bar{\iota}}(0)$, $R(\iota) > R(0)$ and $W(\iota) > W(0)$ or $R(-\iota) > R(0)$ and $W(-\iota) > W(0)$, completing the proof. Alternatively, suppose $dW(\iota)/d\iota|_{\iota=0} = 0$. Using the expression for the second derivative of welfare in Lemma A.4 and the fact that $(\Delta_1, \Gamma_1), (\Delta_2, \Gamma_2)$, and $\Delta^*(\hat{z}, \ell)$ are non-overlapping, we have

$$(A58) \quad \left. \frac{d^2 W(\iota)}{d\iota^2} \right|_{\iota=0} = (k_1)^2 \left. \frac{d^2 W_1(\iota)}{d\iota^2} \right|_{\iota=0} + (k_2)^2 \left. \frac{d^2 W_2(\iota)}{d\iota^2} \right|_{\iota=0} + k^2 \left. \frac{d^2 \hat{W}(\iota)}{d\iota^2} \right|_{\iota=0}$$

³⁴In order to apply Lemma A.3, we use that since $(\Delta_1, \Gamma_1), (\Delta_2, \Gamma_2) \in \Lambda$ and these reforms have non-overlapping support, any linear combination of them is also contained in Λ .

where $W_1(\iota) \equiv E_h[w^h \circ V^h(R_0 + \iota\Delta_1 + (\iota^2/2)\Gamma_1)]$, $W_2(\iota) \equiv E_h[w^h \circ V^h(R_0 + \iota\Delta_2 + (\iota^2/2)\Gamma_2)]$, and $\hat{W}(\iota) \equiv E_h[w^h \circ V^h(\hat{R}(\iota))]$. In the case where a positive measure of households with $\Delta_1(z_0^h)$, $\Gamma_1(z_0^h)$, $\Delta_2(z_0^h)$, or $\Gamma_2(z_0^h)$ non-zero have $w^{h'}(V^h(R_0)) > 0$, (A55) and (A56) imply $d^2W_1(\iota)/d\iota^2|_{\iota=0} > 0$, $d^2W_2(\iota)/d\iota^2|_{\iota=0} \geq 0$ with strict inequality in at least one case, and $d^2\hat{W}(\iota)/d\iota^2|_{\iota=0} = 0$. In the complementary case, (A55) and (A56) imply $d^2W_1(\iota)/d\iota^2|_{\iota=0} > 0$, $d^2W_2(\iota)/d\iota^2|_{\iota=0} \geq 0$, and integrating the expression for welfare's second derivative in Lemma A.2 implies $d^2\hat{W}(\iota)/d\iota^2|_{\iota=0} > 0$. We conclude that if $dW(\iota)/d\iota|_{\iota=0} = 0$, then $d^2W(\iota)/d\iota^2|_{\iota=0} > 0$. The definition of the second derivative then implies that for all sufficiently small ι , $R(\iota) > R(0)$ and $W(\iota) > W(0)$, completing the proof.

A5. Proof of Theorem 3

We will define a set of cardinalizing functions with locally bounded curvature and then show that they locally rationalize R_0 .

To this end, let $\psi(z)$, $\Psi_0(z)$, $\Psi_1(z)$ be as in Lemma A.3 and, for all $z > 0$, define

$$(A59) \quad \hat{\lambda}(z) \equiv -\psi(z)/R_0(z) \quad \text{and} \quad \hat{\varepsilon}(z) \equiv -\frac{\Psi_1(z)/2}{\hat{\lambda}(z)R_0'(z)z}$$

Note that $\hat{\lambda}(z)$ and $\hat{\varepsilon}(z)$ are continuously differentiable by the assumed smoothness of primitives, which implies $\psi(z)$, $\Psi_1(z)$ are continuously differentiable by the expressions for $\psi(z)$, $\Psi_1(z)$ in Lemma A.3. They are also strictly positive since $\psi(z)$, $\Psi_1(z) < 0$ by the assumption that $\delta\text{Rev}(z)$, $\delta^2\text{Rev}(z) < 0$, using Lemma A.3 and Theorem 4. Moreover, $\hat{\lambda}(z)$ is bounded in the limits as $z \rightarrow 0$ and $z \rightarrow \infty$ by the expression for $\psi(z)$ in Lemma A.3, Assumption 1, the assumptions in (3), the assumption that elasticities and $\alpha(z)$ are bounded, and the assumption that $\varepsilon(z)$ has bounded derivatives in log-income. Similarly, $\Psi_1(z)/z$ is bounded above zero in the limits as $z \rightarrow 0$ and $z \rightarrow \infty$, by the assumption that $\delta^2\text{Rev}(z) > 0$, the assumption that $\delta^2\text{Rev}(z)/(zg(z))$ is bounded away from zero in the limits as $z \rightarrow 0$ and $z \rightarrow \infty$, the expressions for $\Psi_1(z)$ and $\delta^2\text{Rev}(z)$ in Lemma A.3 and Theorem 4, and the bounds on marginal retention from Assumption 1. These observations and Assumption 1 imply that $\hat{\varepsilon}(z)$, too, is bounded above zero in the limits as $z \rightarrow 0$ and $z \rightarrow \infty$. The positivity and continuity of $\hat{\varepsilon}(z)$ therefore implies $\hat{\varepsilon}(z)$ is bounded above some value $\hat{\underline{\varepsilon}}$ across all z . By assumption, $P_h[\varepsilon^h \leq \hat{\varepsilon}(z) \mid z_0^h = z] \geq P_h[\varepsilon^h \leq \hat{\underline{\varepsilon}} \mid z_0^h = z] > 0$ for all $z > 0$.

Defining

$$(A60) \quad \bar{\eta}(z) \equiv E_h[\eta^h \mid z_0^h = z, \varepsilon^h \leq \hat{\varepsilon}(z)] \quad \text{and} \quad \left(\frac{\bar{\eta}^2}{\varepsilon}\right)(z) \equiv E_h[(\eta^h)^2/\varepsilon^h \mid z_0^h = z, \varepsilon^h \leq \hat{\varepsilon}(z)],$$

Assumption 3 implies $(\bar{\eta}^2/\varepsilon)(z)$ and $\bar{\eta}(z)$ are continuously differentiable. Next, define

$$(A61) \quad \hat{\rho}(z) \equiv \frac{-R_0(z) - \Psi_0(z)}{R_0(z)\hat{\lambda}(z)} - \frac{R_0'(z)z}{R_0(z)} \left(\frac{\bar{\eta}^2}{\varepsilon}\right)(z) - \bar{\eta}(z) \left(\alpha(z) + \frac{R_0'(z)z}{R_0(z)}\right) + z\bar{\eta}'(z) + \frac{d \log \hat{\lambda}(z)}{d \log z}$$

The smoothness of primitives and the observations above imply $\hat{\rho}(z)$ is continuous in z . Finally, define

$$(A62) \quad \hat{\lambda}^h \equiv \frac{\mathbb{1}_{\varepsilon^h \leq \hat{\varepsilon}(z_0^h)}}{P_h[\varepsilon^h \leq \hat{\varepsilon}(z) \mid z_0^h = z]} \hat{\lambda}(z_0^h) \quad \hat{\rho}^h \equiv \mathbb{1}_{\varepsilon^h \leq \hat{\varepsilon}(z_0^h)} \hat{\rho}(z_0^h)$$

$$k_1^h \equiv \hat{\lambda}^h/u_c^h(c_0^h, z_0^h) \quad k_2^h \equiv \left(\hat{\rho}^h - \frac{c_0^h u_{cc}^h(c_0^h, z_0^h)}{u_c^h(c_0^h, z_0^h)}\right) / \left(c_0^h u_c^h(c_0^h, z_0^h)\right)$$

We can now define our set of cardinalizing functions $\{w^h\}_{h \in \mathcal{H}}$.

$$(A63) \quad w^h(u) \equiv \begin{cases} \frac{k_1^h}{k_2^h} \exp \left[k_2^h \left(u - u^h(c_0^h, z_0^h) \right) \right] & \text{if } \varepsilon^h \leq \hat{\varepsilon}(z_0^h) \text{ and } k_2^h \neq 0 \\ k_1^h u & \text{if } \varepsilon^h \leq \hat{\varepsilon}(z_0^h) \text{ and } k_2^h = 0 \\ \left(u - u^h(c_0^h, z_0^h) \right)^3 & \text{if } \varepsilon^h > \hat{\varepsilon}(z_0^h) \end{cases}$$

Note that for all households h , w^h is strictly increasing and smooth. For the finite measure of households with $\varepsilon^{h'} \leq \hat{\varepsilon}(z_0^h)$, we moreover have $w^h(V^h(R_0)) > 0$. Simple computations reveal that

$$(A64) \quad \lambda^h \equiv \frac{\partial(w^h \circ u^h)(c_0^h, z_0^h)}{\partial c_0^h} = \hat{\lambda}^h \quad \text{and} \quad \gamma^h \equiv \frac{\partial^2(w^h \circ u^h)(c_0^h, z_0^h)}{\partial (c_0^h)^2} = \frac{\hat{\lambda}^h \hat{\rho}^h}{c_0^h}$$

We next argue the regularity conditions required for Lemma A.4 apply, given these cardinalizing functions. First, our earlier observation that $\hat{\lambda}(z)$ is continuously differentiable implies $\lambda(z) \equiv E_h[\lambda^h | z_0^h = z] = \hat{\lambda}(z)$ is continuously differentiable as well. Second, $\gamma(z) \equiv E_h[\gamma^h | z_0^h = z] = \hat{\lambda}(z) \hat{\rho}(z) / R_0(z)$ is continuous by our earlier observation that $\hat{\lambda}(z)$ and $\hat{\rho}(z)$ are continuous and by the smoothness of R_0 . Third—letting $\bar{\varepsilon}(z) \equiv E_h[\varepsilon^h | z_0^h = z, \varepsilon^h \leq \hat{\varepsilon}(z)]$ —we have $(\lambda \varepsilon)(z) = \hat{\lambda}(z) \bar{\varepsilon}(z)$ and $(\lambda \eta^2 / \varepsilon)(z) = \hat{\lambda}(z) (\bar{\eta}^2 / \bar{\varepsilon})(z)$ continuous since $\hat{\lambda}(z)$ is continuous and since $\bar{\varepsilon}(z)$ and $(\bar{\eta}^2 / \bar{\varepsilon})(z)$ are continuous by Assumption 3 and the continuous differentiability of $\hat{\varepsilon}(z)$. Fourth, $(\lambda \eta)(z) = \hat{\lambda}(z) \bar{\eta}(z)$ is continuously differentiable since $\hat{\lambda}(z)$ is, as observed above, and since $\bar{\eta}(z)$ is, by Assumption 3 and the continuous differentiability of $\hat{\varepsilon}(z)$. Fifth and finally, $(\lambda \eta)(z) = \hat{\lambda}(z) \bar{\eta}(z)$ is bounded as $z \rightarrow 0$ and $z \rightarrow \infty$, since $\hat{\lambda}(z)$ is, as shown above, and $\bar{\eta}(z)$ is by the assumption that elasticities are bounded in the population.

We now argue R_0 is locally utilitarian-optimal given the cardinalizing functions $\{w^h\}_{h \in \mathcal{H}}$. It suffices to show that for arbitrary $(\Delta, \Gamma) \in \Lambda$, there exists $\bar{\iota} > 0$ such that

$$(A65) \quad 0 \in \arg \max_{\iota \in (-\bar{\iota}, \bar{\iota})} W(\iota) \quad \text{s.t.} \quad \text{Rev}(R(\iota)) \geq \text{Rev}(R(0))$$

where $R(\iota) \equiv R_0 + \iota \Delta + (\iota^2 / 2) \Gamma$ and $W(\iota) \equiv E_h[w^h \circ V^h(R(\iota))]$. Note that it suffices to consider the case where $\Delta \neq 0$, since if $\Delta = \Gamma = 0$ the desired conclusion holds trivially, and if $\Delta = 0$ and $\Gamma \neq 0$, then the desired conclusion follows from considering the reform $(\Delta = \Gamma, \tilde{\Gamma} = 0)$. By Lagrangian sufficiency, it suffices to show that within some non-empty interval $(-\bar{\iota}, \bar{\iota})$, $\iota = 0$ maximizes $W(\iota) + \text{Rev}(R(\iota))$. This in turn follows from the definition of first and second derivatives, provided that

$$(A66) \quad \left. \frac{d}{d\iota} \right|_{\iota=0} \left(W(\iota) + \text{Rev}(R(\iota)) \right) = 0 \quad \text{and} \quad \left. \frac{d^2}{d\iota^2} \right|_{\iota=0} \left(W(\iota) + \text{Rev}(R(\iota)) \right) < 0$$

To this end, we compute the derivatives above using Lemmas A.3 and A.4.

These derivatives are determined by the terms $\phi(z), \Phi_0(z), \Phi_1(z), \psi(z), \Psi_0(z), \Psi_1(z)$ in the lemmas, about which we accordingly make several observations: First,

$$(A67) \quad \psi(z) + \phi(z) = \psi(z) + R_0(z) \lambda(z) = \psi(z) + R_0(z) \hat{\lambda}(z) = 0$$

by the construction of $\hat{\lambda}(z)$. Second,

$$(A68) \quad \begin{aligned} \Psi_1(z) + \Phi_1(z) &= \Psi_1(z) + (\lambda \varepsilon)(z) R_0'(z) z = \Psi_1(z) + \hat{\lambda}(z) \bar{\varepsilon}(z) R_0'(z) z \\ &\leq \Psi_1(z) + \hat{\lambda}(z) \hat{\varepsilon}(z) R_0'(z) z = \Psi_1(z) / 2 < 0 \end{aligned}$$

by the definition of $\hat{\varepsilon}(z)$. Third, $\Psi_0(z) + \Phi_0(z)$ is equal to

$$\begin{aligned}
(A69) \quad &= \Psi_0(z) + \left[R_0(z)\gamma(z) + \frac{R'_0(z)z}{R_0(z)} \left(\lambda \frac{\eta^2}{\varepsilon} \right) (z) + (\lambda\eta)(z) \left(\alpha(z) + \frac{R'_0(z)z}{R_0(z)} \right) - z(\lambda\eta)'(z) \right] R_0(z) \\
&= \Psi_0(z) + \hat{\lambda}(z) \left[\hat{\rho}(z) + \frac{R'_0(z)z}{R_0(z)} (\overline{\eta^2/\varepsilon})(z) + \bar{\eta}(z) \left(\alpha(z) + \frac{R'_0(z)z}{R_0(z)} \right) - \frac{1}{\hat{\lambda}(z)} \frac{d(\hat{\lambda}(z)\bar{\eta}(z))}{d \log z} \right] R_0(z)
\end{aligned}$$

which is equal to $-R_0(z)$ and hence strictly negative by the definition of $\hat{\rho}(z)$. By Lemmas A.3 and A.4, these observations imply (A66) provided $\Delta \neq 0$.

A6. Proof of claim in Section III.C

For each household h , let $\text{Rev}^h(R) \equiv z^h(R) - c^h(R)$ denote the revenue generated by household h under the schedule R . For each group $i \in I$, we define

$$(A70) \quad \text{Rev}^i(R) \equiv \text{E}_h \left[\text{Rev}^h(R) \mid i(h) = i \right]$$

and similarly

$$\begin{aligned}
(A71) \quad \delta \text{Rev}^i(z) &= \lim_{\ell \rightarrow +0} \frac{1}{\ell} \frac{d}{d\ell} \Big|_{\ell=0} \text{Rev}^i(R_0 + \iota \Delta^*(z, \ell)) \\
\delta^2 \text{Rev}^i(z) &= \lim_{\ell \rightarrow +0} \ell \frac{d^2}{d\ell^2} \Big|_{\ell=0} \text{Rev}^i(R_0 + \iota \Delta^*(z, \ell)).
\end{aligned}$$

We now claim that for all $i \in I$ and $z > 0$, if $\delta \text{Rev}^i(z) < 0$, then $\delta^2 \text{Rev}^i(z) < 0$. It suffices to show that, in an alternative economy in which all households are replaced with copies of the households in any particular group i , $\delta \text{Rev}(z) < 0$ implies $\delta^2 \text{Rev}(z) < 0$. We will argue this follows from the expressions for $\delta \text{Rev}(z)$ and $\delta^2 \text{Rev}(z)$ given by Theorem 4.

To this end, we make a useful observation about super-elasticities when preferences fall into the class we have imposed in the main text (see (34)). Specifically, we have $M^h(c, z) = \theta^h v_z^{i'}(z)/v_c^{i'}(c)$ and therefore

$$(A72) \quad \mathcal{E}^h(z) = \left(\frac{d \log v_z^{i'}(z)}{d \log z} - \frac{d \log v_c^{i'}(R_0(z))}{d \log z} - \frac{d \log R'_0(z)}{d \log z} \right)^{-1}$$

is independent of h , given z . So for any household h with income z , we have $\varepsilon^{h+} \equiv d\mathcal{E}^h(z)/d \log z = d\varepsilon(z)/d \log z$. Using this observation, the fact that the population contains a single i group, and Theorem 4 the condition $\delta^2 \text{Rev}(z) < 0$ is equivalent to

$$(A73) \quad 1 + R'_0(z) > (1 - R'_0(z)) \left(\alpha(z)\varepsilon(z) - \frac{d\varepsilon(z)}{d \log z} \right).$$

We wish to show this follows from $\delta \text{Rev}(z) < 0$, which can be written as

$$(A74) \quad R'_0(z) - (1 - R'_0(z)) \frac{R'_0(z)z}{R_0(z)} \eta(z) - \frac{R''_0(z)z}{R'_0(z)} \varepsilon(z) > (1 - R'_0(z)) \left(\alpha(z)\varepsilon(z) - \frac{d\varepsilon(z)}{d \log z} \right).$$

As these expressions have the same right-hand sides, it suffices to show the left-hand side of (A74)

is weakly lower than that of (A73), i.e.,

$$(A75) \quad 1 + (1 - R'_0(z)) \frac{R'_0(z)z}{R_0(z)} \eta(z) + \frac{R''_0(z)z}{R'_0(z)} \varepsilon(z) \geq 0$$

Using the expressions for ε^h and η^h in Lemma A.1, we have

$$(A76) \quad \begin{aligned} & 1 + (1 - R'_0(z)) \frac{R'_0(z)z}{R_0(z)} \eta(z) + \frac{R''_0(z)z}{R'_0(z)} \varepsilon(z) \\ &= \frac{-\frac{d \log v_c^i(c)}{d \log c} \frac{d \log R_0(z)}{d \log z} + \frac{d \log v_z^i(z)}{d \log z} - \frac{d \log R'_0(z)}{d \log z} + (1 - R'_0(z)) \frac{d \log v_c^i(c)}{d \log c} \frac{d \log R_0(z)}{d \log z} + \frac{d \log R'_0(z)}{d \log z}}{-\frac{d \log v_c^i(c)}{d \log c} \frac{d \log R_0(z)}{d \log z} + \frac{d \log v_z^i(z)}{d \log z} - \frac{d \log R'_0(z)}{d \log z}} \\ &= \varepsilon(z) \left(\frac{d \log v_z^i(z)}{d \log z} - R'_0(z) \frac{d \log v_c^i(c)}{d \log c} \frac{d \log R_0(z)}{d \log z} \right) \end{aligned}$$

The last line is indeed weakly positive, by the positivity of compensated elasticities, the concavity of v_c^i and the convexity of v_z^i .

A7. Local bounding arguments

The proof of Theorem 2.a uses the boundedness—ensured by Assumption 2—of several quantities across households. We derive these bounds through a series of lemmas below.

Lemma A.6. *Fix any $(\Delta, \Gamma) \in \Lambda$ and let $R(\iota) \equiv R_0 + \iota\Delta + (\iota^2/2)\Gamma$. There exist $\bar{\iota}$ and smooth functions f_0 and f_1 satisfying $f_0(0) = f_1(0) = 0$ such that for all $\iota \in B_{\bar{\iota}}(0)$, for all $h \in \mathcal{H}$,*

$$(A77) \quad \left| \frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} - \frac{R_\iota(z^h(0); 0)}{R(z^h(0); 0)} \right|, \left| \frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)} - \frac{R'_\iota(z^h(0); 0)}{R'(z^h(0); 0)} \right| \leq f_0(\iota) + f_1(|\log z^h(\iota) - \log z_0^h|)$$

PROOF:

By the definition of $R(\iota)$ and the triangle inequality, we have

$$(A78) \quad \begin{aligned} & \left| \frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} - \frac{R_\iota(z^h(0); 0)}{R(z^h(0); 0)} \right| \\ & \leq \left| \frac{\iota \frac{\Gamma(z^h(\iota))}{R_0(z^h(\iota))}}{1 + \iota \frac{\Delta(z^h(\iota))}{R_0(z^h(\iota))} + \frac{\iota^2}{2} \frac{\Gamma(z^h(\iota))}{R_0(z^h(\iota))}} \right| + \left| \frac{\Delta(z^h(\iota))}{R_0(z^h(\iota))} \right| \left| \frac{1}{1 + \iota \frac{\Delta(z^h(\iota))}{R_0(z^h(\iota))} + \frac{\iota^2}{2} \frac{\Gamma(z^h(\iota))}{R_0(z^h(\iota))}} - 1 \right| \\ & + \left| \frac{\Delta(z^h(\iota))}{R_0(z^h(\iota))} \right| \left| 1 - \frac{R_0(z^h(\iota))}{R_0(z^h(0))} \right| + \left| \frac{\int_{\log z^h(0)}^{\log z^h(\iota)} \frac{d\Delta(z^h)}{d \log z^h} d \log z^h}{R_0(z^h(0))} \right| \end{aligned}$$

For $|\iota| + \frac{|\iota|^2}{2} \leq B_\Gamma^\Delta$, the assumptions in (3) and (4) imply the RHS above is

$$(A79) \quad \leq \frac{|\iota| B_\Gamma^\Delta}{1 - (|\iota| + \frac{|\iota|^2}{2}) B_\Gamma^\Delta} + B_\Gamma^\Delta \frac{(|\iota| + \frac{|\iota|^2}{2}) B_\Gamma^\Delta}{1 - (|\iota| + \frac{|\iota|^2}{2}) B_\Gamma^\Delta} + 2B_\Gamma^\Delta \left(e^{B^R |\log z^h(\iota) - \log z^h(0)|} - 1 \right)$$

Analogous arguments apply to the case of R' .

Lemma A.7. Fix any $(\Delta, \Gamma) \in \Lambda$ and let $R(\iota) \equiv R_0 + \iota\Delta + (\iota^2/2)\Gamma$. There exist $\bar{\iota}$ and smooth functions f_0 and f_1 satisfying $f_0(0) = f_1(0) = 0$ such that for all $\iota \in B_{\bar{\iota}}(0)$, for all $h \in \mathcal{H}$,

$$(A80) \quad \left| \frac{d \log R(z^h(\iota); \iota)}{d \log z} - \frac{d \log R(z^h(0); 0)}{d \log z} \right|, \left| \frac{d \log R'(z^h(\iota); \iota)}{d \log z} - \frac{d \log R'(z^h(0); 0)}{d \log z} \right| \leq f_0(\iota) + f_1 \left(\left| \log z^h(\iota) - \log z_0^h \right| \right)$$

PROOF:

By the definition of $R(\iota)$ and the triangle inequality, we have

$$(A81) \quad \begin{aligned} & \left| \frac{d \log R(z^h(\iota); \iota)}{d \log z} - \frac{d \log R(z^h(0); 0)}{d \log z} \right| \\ & \leq \left| \frac{\left(\iota \Delta'(z^h(\iota)) + \frac{\iota^2}{2} \Gamma'(z^h(\iota)) \right) z^h(\iota)}{R_0(z^h(\iota)) + \iota \Delta(z^h(\iota)) + \frac{\iota^2}{2} \Gamma(z^h(\iota))} \right| \\ & + \left| \frac{R'_0(z^h(\iota)) z^h(\iota)}{R_0(z^h(\iota))} \left| \frac{R_0(z^h(\iota))}{R_0(z^h(\iota)) + \iota \Delta(z^h(\iota)) + \frac{\iota^2}{2} \Gamma(z^h(\iota))} - 1 \right| \right| \\ & + \left| \frac{R'_0(z_0^h(0)) z_0^h}{R_0(z_0^h)} \left| e^{|\log R'_0(z^h(\iota)) - \log R'_0(z_0^h)| + |\log z^h(\iota) - \log z_0^h| + |\log R_0(z^h(\iota)) - \log R_0(z_0^h)|} - 1 \right| \right| \end{aligned}$$

For $|\iota| + \frac{|\iota|^2}{2} \leq B_\Gamma^\Delta$, the assumptions in (3) and (4) imply the RHS above is

$$(A82) \quad \leq 2 \frac{B_\Gamma^\Delta B^R \left(|\iota| + \frac{|\iota|^2}{2} \right)}{1 - B_\Gamma^\Delta \left(|\iota| + \frac{|\iota|^2}{2} \right)} + B^R \left| e^{(1+2B^R)|\log z^h(\iota) - \log z_0^h|} - 1 \right|$$

Analogous arguments apply to the case of R' .

Lemma A.8. Fix any $(\Delta, \Gamma) \in \Lambda$, let $R(\iota) \equiv R_0 + \iota\Delta + (\iota^2/2)\Gamma$, and suppose Assumption 2 holds. There exist $\bar{\iota} > 0$ and smooth functions $f_0(\iota)$ and $f_1(\iota)$ satisfying $f_0(0) = f_1(0) = 0$ such that for all $h \in \mathcal{H}$, if $\iota \in B_{\bar{\iota}}(0)$ and $|\log c^h(\hat{\iota}) - \log c_0^h| + |\log z^h(\hat{\iota}) - \log z_0^h| < \bar{\iota}$ for all $\hat{\iota} \in \bar{B}_\iota(0)$, then

$$(A83) \quad \left| \eta^h(\iota) - \eta^h(0) \right|, \left| \varepsilon^h(\iota) - \varepsilon^h(0) \right| \leq f_0(\iota) + f_1 \left(|\log c^h(\iota) - \log c_0^h| + |\log z^h(\iota) - \log z_0^h| \right)$$

PROOF:

We begin by establishing the claim in the lemma statement for $1/\varepsilon^h$. The definition of $\varepsilon^h(\iota)$ in Lemma A.1 implies

$$(A84) \quad \begin{aligned} \left| \frac{1}{\varepsilon^h(\iota)} - \frac{1}{\varepsilon^h(0)} \right| & \leq \left| \frac{d \log R(z^h(\iota); \iota)}{d \log z} \right| \left| \frac{\partial \log M^h(c^h(\iota), z^h(\iota))}{\partial \log c} - \frac{\partial \log M^h(c^h(0), z^h(0))}{\partial \log c} \right| \\ & + \left| \frac{\partial \log M^h(c^h(0), z^h(0))}{\partial \log c} \right| \left| \frac{d \log R(z^h(\iota); \iota)}{d \log z} - \frac{d \log R(z^h(0); 0)}{d \log z} \right| \\ & + \left| \frac{\partial \log M^h(c^h(\iota), z^h(\iota))}{\partial \log z} - \frac{\partial \log M^h(c^h(0), z^h(0))}{\partial \log z} \right| \\ & + \left| \frac{d \log R'(z^h(\iota); \iota)}{d \log z} - \frac{d \log R'(z^h(0); 0)}{d \log z} \right| \end{aligned}$$

Take $\bar{\iota} > 0$ small enough that Lemma A.7 and $\bar{\iota} < d$, where d is as in Assumption 2. Assumption 2, the assumptions in (3) and (4), and Lemma A.7 then imply that for some smooth f_0 and f_1 satisfying $f_0(0) = f_1(0) = 0$, the RHS of (A84) is

$$(A85) \quad \begin{aligned} &\leq (1 + B^R)b \left(|\log c^h(\iota) - \log c_0^h| + |\log z^h(\iota) - \log z_0^h| \right) \\ &+ \left(1 + \frac{\bar{\eta}}{\underline{\varepsilon}} \right) \left(f_0(\iota) + f_1(|\log z^h(\iota) - \log z_0^h|) \right) \end{aligned}$$

where $\bar{\eta}$ and $\underline{\varepsilon}$ are the upper and lower bounds on the magnitudes at R_0 of income and compensated elasticities, respectively, that exist by assumption. The desired statement for ε^h follows from the observation that, when $|\varepsilon^h(\iota) - \varepsilon(0)| < \underline{\varepsilon}$,

$$(A86) \quad \left| \varepsilon^h(\iota) - \varepsilon(0) \right| = \left| \frac{\varepsilon^h(0) \left(\varepsilon^h(\iota)^{-1} - \varepsilon^h(0)^{-1} \right)}{\varepsilon^h(0)^{-1} - \left(\varepsilon^h(\iota)^{-1} - \varepsilon^h(0)^{-1} \right)} \right| \leq \frac{\bar{\varepsilon} |\varepsilon^h(\iota)^{-1} - \varepsilon^h(0)^{-1}|}{\underline{\varepsilon} - |\varepsilon^h(\iota)^{-1} - \varepsilon^h(0)^{-1}|}$$

where $\bar{\varepsilon}$ is the assumed upper bound on the magnitude of compensated elasticities. We next establish the claim in the lemma statement for η^h/ε^h . This follows immediately from the definitions of these elasticities in Lemma A.1, which imply $\eta^h(\iota)/\varepsilon^h(\iota) = -\partial \log M(c^h(\iota), z^h(\iota))/\partial \log c^h$, under Assumption 2. The desired statement for η^h then follows from the bounds already established and the observation that

$$(A87) \quad \begin{aligned} \left| \eta^h(\iota) - \eta^h(0) \right| &\leq \varepsilon^h(\iota) \left| \frac{\eta^h(\iota)}{\varepsilon^h(\iota)} - \frac{\eta^h(0)}{\varepsilon^h(0)} \right| + \left| \varepsilon^h(\iota) \eta^h(0) \right| \left| \frac{1}{\varepsilon^h(0)} - \frac{1}{\varepsilon^h(\iota)} \right| \\ &\leq \left(\bar{\varepsilon} + |\varepsilon^h(\iota) - \varepsilon^h(0)| \right) \left(\left| \frac{\eta^h(\iota)}{\varepsilon^h(\iota)} - \frac{\eta^h(0)}{\varepsilon^h(0)} \right| + \bar{\eta} \left| \frac{1}{\varepsilon^h(0)} - \frac{1}{\varepsilon^h(\iota)} \right| \right) \end{aligned}$$

Lemma A.9. *Fix any $(\Delta, \Gamma) \in \Lambda$, let $R(\iota) \equiv R_0 + \iota\Delta + (\iota^2/2)\Gamma$, and suppose Assumption 2 holds. There exist $\bar{\iota} > 0$ and smooth functions $f_0(\iota)$ and $f_1(\iota)$ satisfying $f_0(0) = f_1(0) = 0$ such that for all $h \in \mathcal{H}$, if $\iota \in B_{\bar{\iota}}(0)$ and $|\log c^h(\tilde{\iota}) - \log c_0^h| + |\log z^h(\tilde{\iota}) - \log z_0^h| < \bar{\iota}$ for all $\tilde{\iota} \in \bar{B}_{\bar{\iota}}(0)$, then*

$$(A88) \quad \begin{aligned} &\left| \frac{d \log c^h(\iota)}{d \iota} - \frac{d \log c^h(0)}{d \iota} \right|, \left| \frac{d \log z^h(\iota)}{d \iota} - \frac{d \log z^h(0)}{d \iota} \right| \\ &\leq f_0(\iota) + f_1 \left(|\log c^h(\iota) - \log c_0^h| + |\log z^h(\iota) - \log z_0^h| \right) \end{aligned}$$

PROOF:

By Lemma A.1, there exists $\bar{\iota} > 0$ such that all $z^h(\iota)$ are differentiable at all $\iota \in B_{\bar{\iota}}(0)$ and

$$(A89) \quad \frac{d \log z^h(\iota)}{d \iota} = \eta^h(\iota) \frac{R_\iota(z^h(\iota); \iota)}{R(z^h(\iota); \iota)} + \varepsilon^h(\iota) \frac{R'_\iota(z^h(\iota); \iota)}{R'(z^h(\iota); \iota)}$$

By the assumptions in (4) and the assumption that elasticities are bounded at $\iota = 0$, the terms above are bounded at $\iota = 0$. By Lemmas A.6, and A.8, each of their changes in ι are bounded in the same manner as described in the lemma statement for some $\bar{\iota} < \bar{\iota}$. It follows that their products and sums are similarly bounded, as desired.

Next, consider $c^h(\iota)$. Here, we use the observation that, since $c^h(\iota) = R(z^h(\iota); \iota)$,

$$(A90) \quad \frac{d \log c^h(\iota)}{d \iota} = \frac{d \log R(z^h(\iota))}{d \log z^h} \frac{d \log z^h(\iota)}{d \iota} + \frac{R_\iota(z^h(\iota), \iota)}{R(z^h(\iota), \iota)}$$

for all $h \in \mathcal{H}$ if $\iota \in B_{\bar{\iota}}(0)$. By the assumptions in (3) and (4) and the observations in the previous paragraph, the terms above are bounded at $\iota = 0$. By Lemmas A.7 and A.6 and the first part of the lemma, shown above, each of their changes in ι are bounded in the same manner as described in the lemma statement for some $\bar{\iota} < \bar{\iota}$. It follows that their products and sums are similarly bounded, as desired.

Lemma A.10. *Fix any $(\Delta, \Gamma) \in \Lambda$, let $R(\iota) \equiv R_0 + \iota\Delta + (\iota^2/2)\Gamma$, and suppose Assumption 2 holds. There exists $\bar{\iota} > 0$ and a continuous function $b(\iota)$ satisfying $b(0) = 0$ such that for all $h \in \mathcal{H}$, for all $\iota \in B_{\bar{\iota}}(0)$,*

$$(A91) \quad |\log c^h(\iota) - \log c_0^h|, |\log z^h(\iota) - \log z_0^h| \leq b(\iota)$$

PROOF:

Define $x^h(\iota) \equiv |\log c^h(\iota) - \log c_0^h| + |\log z^h(\iota) - \log z_0^h|$. Note it suffices to show there exists $\bar{\iota}$ and a continuous function $b(\iota)$ satisfying $b(0) = 0$ such that for all $\iota \in B_{\bar{\iota}}(0)$ and $h \in \mathcal{H}$, $x^h(\iota) \leq b(\iota)$.

By Lemma A.9, there exists $\bar{\iota}, \bar{x} > 0$ and smooth functions f_0 and f_1 satisfying $f_0(0) = f_1(0) = 0$ with the following property: For all $h \in \mathcal{H}$, if $\iota \in \bar{B}_{\bar{\iota}}(0)$ and $x(\iota) \leq \bar{x}$ for all $\tilde{\iota} \in \bar{B}_{\bar{\iota}}(0)$, then

$$(A92) \quad x(\iota) \leq f_0(\iota) + \left| \int_0^\iota f_1(x(\tilde{\iota})) d\tilde{\iota} \right| \leq x(\iota) \leq \bar{f}_0(\iota) + \bar{f}_1 \text{sgn}(\iota) \int_0^\iota x(\tilde{\iota}) d\tilde{\iota}$$

where $\bar{f}_0(\iota) \equiv \max_{\tilde{\iota} \in \bar{B}_0(\iota)} f_0(\tilde{\iota})$ and $\bar{f}_1 \equiv \max_{x \in \bar{B}_{\bar{x}}(0)} |f_1'(x)|$. Note that \bar{f}_0 is continuous, weakly increasing in ι , and satisfies $\bar{f}_0(0) = 0$.

For $\iota \in \bar{B}_{\bar{\iota}}(0)$, define

$$(A93) \quad b(\iota) \equiv \bar{f}_0(\iota) \exp[\bar{f}_1 \iota]$$

Let $\hat{\iota} > 0$ be some value at which $b(\hat{\iota}) < \bar{x}$. For each $h \in \mathcal{H}$, let $\bar{\iota}^h$ be the infimum of the set of $|\iota|$ for which $x^h(\iota) > b(\iota)$, if such ι exist, or $\min[\bar{\iota}, \hat{\iota}]$ otherwise. The remainder of the proof argues that in fact $\bar{\iota}^h \geq \min[\bar{\iota}, \hat{\iota}]$ for all $h \in \mathcal{H}$; this establishes the claim in the lemma.

We now fix some $h \in \mathcal{H}$ and suppose toward a contradiction that $\bar{\iota}^h < \min[\bar{\iota}, \hat{\iota}]$. Since by the construction of $\bar{\iota}^h$ and $\hat{\iota}$ this implies $x^h(\iota) \leq b(\iota) < \bar{x}$ for all $\iota \in \bar{B}_{\bar{\iota}^h}(0)$, the continuous differentiability of $x^h(\iota)$ implies $x^h(\iota) < \bar{x}$ for all $\iota \in \bar{B}_{\bar{\iota}^h}(0)$ for some $\hat{\iota}^h > \bar{\iota}^h \geq 0$, implying that (A92) holds. But then Gronwall's inequality implies $x(\iota) \leq b(\iota)$ for all $\iota \in \bar{B}_{\bar{\iota}^h}(0)$, which contradicts the construction of $\bar{\iota}^h$.

As an immediate corollary of Lemmas A.7, A.8, A.9, and A.10, as well as the assumption on the curvature of marginal utilities in Assumption 2, we obtain the following.

Lemma A.11. *Fix any $(\Delta, \Gamma) \in \Lambda$, let $R(\iota) \equiv R_0 + \iota\Delta + (\iota^2/2)\Gamma$, and suppose Assumption 2 holds. There exists $\bar{\iota} > 0$ and a continuous function f satisfying $f(0) = 0$ such that for all $h \in \mathcal{H}$, for all $\iota \in B_{\bar{\iota}}(0)$, $f(\iota)$ is weakly larger than each of*

$$(A94) \quad \begin{aligned} & \left| \frac{d \log R(z^h(\iota); \iota)}{d \log z} - \frac{d \log R(z^h(0); 0)}{d \log z} \right|, & \left| \frac{d \log R'(z^h(\iota); \iota)}{d \log z} - \frac{d \log R'(z^h(0); 0)}{d \log z} \right|, \\ & \left| \eta^h(\iota) - \eta^h(0) \right|, & \left| \varepsilon^h(\iota) - \varepsilon^h(0) \right|, \\ & \left| \frac{d \log c^h(\iota)}{d \iota} - \frac{d \log c^h(0)}{d \iota} \right|, & \left| \frac{d \log z^h(\iota)}{d \iota} - \frac{d \log z^h(0)}{d \iota} \right|, \\ & \left| \frac{d \log u_c^h(c^h(\iota), z^h(\iota))}{d \log c} - \frac{d \log u_c^h(c_0^h, z_0^h)}{d \log c} \right|. \end{aligned}$$

B. EXAMPLES AND SIMULATION EXERCISE

B1. *Examples of one-dimensional preferences with convex revenues*

In Section III.C we have extensively discussed how heterogeneity in compensated elasticities conditional on income can lead revenues to be convex—i.e., $\delta\text{Rev}(z) > 0$ —at the same time as they are decreasing—i.e., $\delta\text{Rev}(z) < 0$. We have also shown that decreasing revenues imply concave revenues when households are drawn from a one-dimensional type space and their preferences fall in a particular class. In the main text, we alluded to the fact that, outside of this preference class, it is possible to have decreasing but concave revenues even when households are homogeneous conditional on their incomes. We now elaborate on this statement.

In both of the examples we will provide below, households have no income effects, there is no variation in elasticities conditional on income, and taxes are affine. Under these conditions, rearranging the expressions in Theorem 4 implies that $\delta\text{Rev}(z) < 0 < \delta^2\text{Rev}(z)$ —i.e., revenues are decreasing but concave—if and only if

$$(B1) \quad 0 < R'_0(z) - (1 - R'_0(z)) \left(\alpha(z)\varepsilon(z) - \frac{d\varepsilon(z)}{d\log z} \right) < -1 + (1 - R'_0(z)) \left(-\frac{d\varepsilon(z)}{d\log z} + \varepsilon^+(z) \right).$$

Note that given values of the statistics above except for $\alpha(z)$, one may—provided $R'_0(z) \neq 0$ —always find a value of $\alpha(z)$ for which the condition above holds, provided $1 < (1 - R'_0(z)) (-d\varepsilon(z)/d\log z + \varepsilon^+(z))$. We therefore provide examples of preference families for which the right-hand side of this inequality can be made arbitrarily large.

Example 1. *Taxes are linear, i.e., $R_0(z) = rz$ for $r \in (0, 1)$. A one-dimensional continuum of households h have additive-CES preferences with idiosyncratic elasticities $\beta(\theta^h)$:*

$$(B2) \quad u^h(c, z) = c - \frac{z^{1+1/\beta(\theta^h)}}{1 + 1/\beta(\theta^h)} \Big/ \theta^h.$$

Finally, $\beta(\theta)$ is very sharply decreasing around some type $\theta^{h(z^*)}$.

It is easy to verify that when preferences fall into this class, $\varepsilon^+(z) = 0$ for all z but $d\varepsilon(z^*)/d\log z$ can be made arbitrarily negative at some z^* , as desired.

Example 2. *Taxes are linear, i.e., $R_0(z) = rz$ for $r \in (0, 1)$. A one-dimensional continuum of households h have preferences:*

$$(B3) \quad u^h(c, z) = c - v\left(z/\theta^h\right)$$

for some increasing and strictly convex function v with variable elasticity. Finally, $d\log v'(n)/d\log n$ is very sharply decreasing locally to $n = v'^{-1}(r)$.

When preferences fall into this class, each household supplies labor $z^h = \theta^h v'^{-1}(r)$ and has a (common) compensated elasticity $\varepsilon^h = (d\log v'(n)/d\log n|_{n=v'^{-1}(r)})^{-1}$. It is easy to verify that, moreover, super-elasticities $\varepsilon^+(z)$ can be made arbitrarily large, as desired.

B2. *Simulation exercise*

We now describe in detail the simulation exercise referred to in Section V.

Concretely, each household h has constant compensated elasticities β^h and no income elasticity. Within every income level, 80% of households have $\beta^h = 0.01$ and 20% have $\beta^h = 2$. We roughly

approximate the distribution of income in our data by assuming that productivity is distributed according to a generalized Pareto distribution with location 0, scale e^9 , and shape 1/2 (this implies a Pareto tail of shape 2). Marginal taxes are initially 50% at every income level. The planner places a constant but not common welfare weight on each household, is indifferent between households with the same initial income, and ensures that her first-order condition holds at the initial tax schedule. Under this calibration, the planner’s first-order condition holds exactly but the second-order condition fails for incomes above \approx \$90,000, as shown in Figure B1.

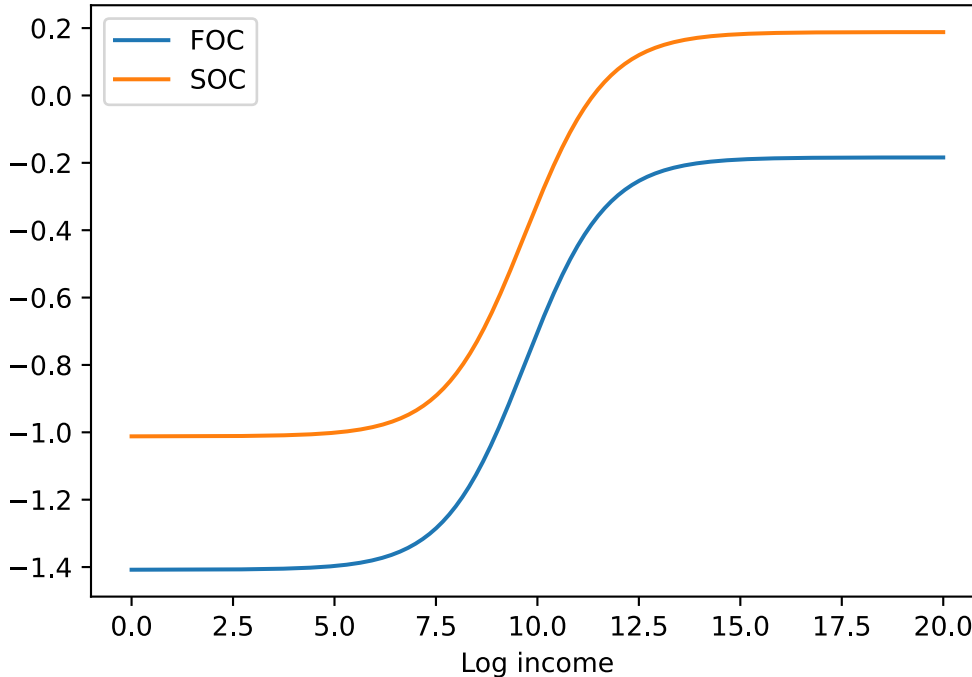


FIGURE B1. FIRST- AND SECOND-ORDER TESTS AT INITIAL TAX SCHEDULE

Note: The first-order test passes, but the second-order test fails (crosses above zero at some incomes).

Next, we perturb marginal taxes slightly, lowering them by 1% at every income level. After computing how households respond to this change, we recompute the planner’s first-order condition, which no longer holds. We then move taxes slightly in the direction in which the first-order condition fails. This does not simply push taxes back to where they began, as—at high incomes—taxes were initially at a local minimum. We iterate this procedure until it converges.

Figure B2 shows how marginal retention evolves between the initial tax schedule and the final schedule to which our procedure converges. Notably, taxes converge to a new, much lower marginal rate at high incomes—where the second-order condition initially failed—but are barely changed at lower incomes—where the second order condition initially held.³⁵ Intuitively, an increase in marginal retention sorts high-elasticity households into higher income levels, rationalizing further decreases in marginal taxes at the top and so on until eventually taxes are high enough.

³⁵We show the schedule after a maximum of 5000 iterations. We have verified that the schedule is visually indistinguishable after 50,000 iterations.

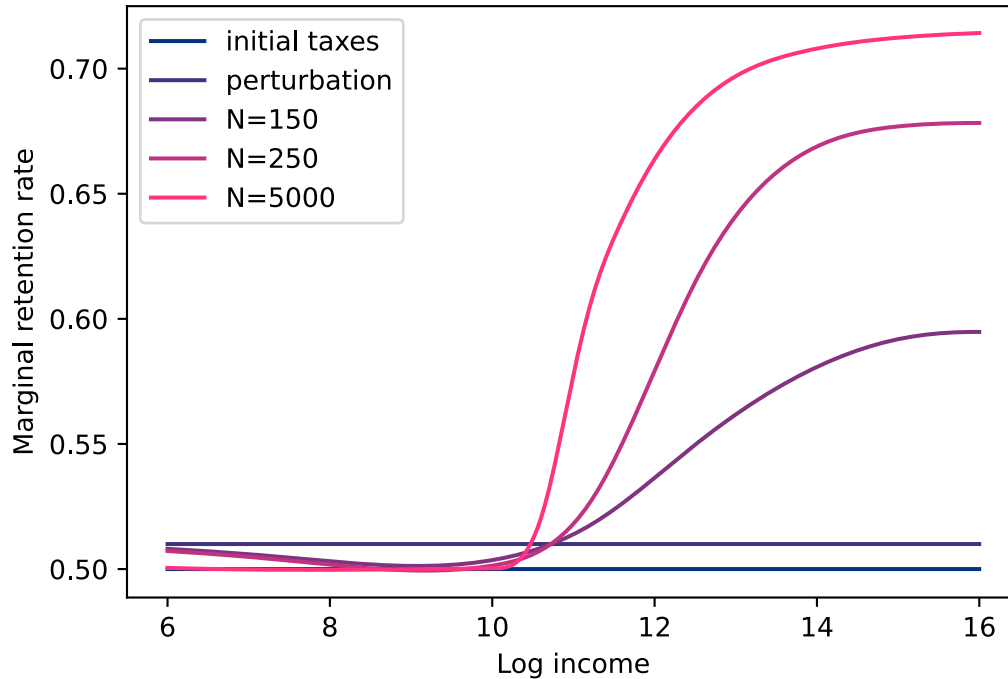


FIGURE B2. CONVERGENCE OF TAXES FOLLOWING A SMALL PERTURBATION

Note: Schedule of marginal retention rates following a small perturbation away from initial taxes under which the first-order revenue test passes but the second-order revenue test fails.

C. EMPIRICAL APPENDIX

C1. Data

We use the NBER panel and of tax returns from 1979 to 1990 (IRS-SOI, NBER, 1990). This sample period includes major tax reforms such as the Economic Recovery Tax Act 1981, which decreased marginal rates in 3 years from 1982 to 1984; the Tax Reform Act 1986, which decreased the number of brackets and reduced the top marginal rate to 28%; the 1987 EITC expansion; and some state level tax reforms.

The data include limited demographic information, individual-specific federal and state income tax schedules,³⁶ and various measures of income. To construct a consistent measure of taxable income for the entire period, we borrow from Gruber and Saez (2002). The measure we use includes wage, business and capital income, and subtracts exemptions, as well as standard and itemized deductions. The procedure of Gruber and Saez (2002) computes changes in income and marginal tax rates (at initial incomes) over 3-year windows. We drop individual-years with changes in marital status or initial (pre-deductions) income lower than \$10 thousand dollars (in 1990 terms).

The panel sample consists of a random selection of four digit endings of social security numbers. There is purely random attrition; in some years, a random subset of the sample social security numbers are excluded. The repeated cross-section sample is larger but not a panel; it over-represents

³⁶We use the tax liabilities as they are computed in Gruber and Saez (2002). These were computed using the NBER TAXSIM program, which calculates liabilities under US Federal and State income tax laws from individual data (Feenberg and Coutts, 1993).

higher-income individuals, which we account for with sample weights.

This data set is appropriate for our exercise for two reasons. First, it contains tax changes that affect people throughout the income distribution, allowing us to estimate ETIs at each income level. Second, the data includes substantial tax variation at both the state and federal level, and as such workers' behavioral responses to these changes can be thought of as typical for real tax reforms.

C2. Estimation details omitted from main text

ESTIMATION OF MEAN ETI

Methodologically, we implement our estimation of (43) by differencing out by year fixed-effects and then running local-polynomial regressions in year-demeaned-log-income space with a constant bandwidth and Gaussian kernel. We select this bandwidth—and those used in each of the local regressions described below—using the leave-one-out cross-validation procedure discussed under "Choice of kernels and selection of bandwidths for local regressions." We compute confidence intervals by using the basic bootstrap method described in Chapter 5 of Davison and Hinkley (1997).

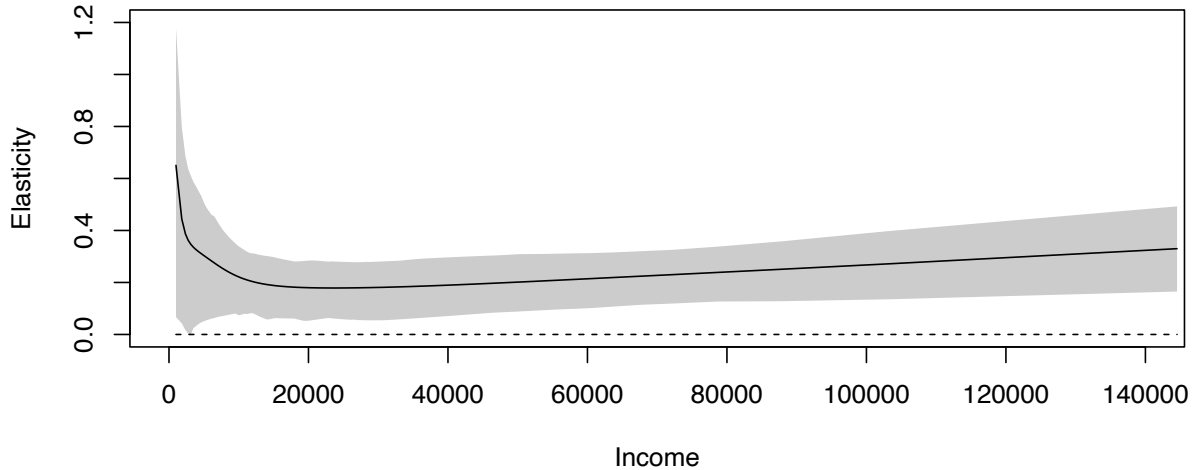


FIGURE C1. MEAN ETI BY INCOME LEVEL IN 1990 USD, AND 95% CONFIDENCE BANDS

Note: As discussed above, we estimate ETI locally in year-demeaned-log-income space. The figure converts these estimates to income space in 1990 by combining log income levels that are the same distance from average log incomes in their respective years.

ESTIMATION OF MEAN ETI BY ITEMIZATION STATUS

To estimate mean ETIs by itemization status, first, we compute the conditional mean value of itemizations by income, with locally linear regressions of the value of itemizations on year-demeaned log taxable income, as in (44). Then we additionally compute the conditional variance of the value of itemizations at each income level, using locally constant regressions of the form:

$$(C1) \quad \left(item_{t-1}^h - \overline{item}(\tilde{z}) \right)^2 = a(\tilde{z}) + \xi_{t-1}^h,$$

where $\overline{item}(\tilde{z})$ is the mean itemization level conditional on log taxable income \tilde{z} . For our estimates of variance based on two ETI groups, we estimate the regression in (45) locally in year-demeaned

log-income space with the same methodology used for income-conditional mean ETIs, labeling as heavy-itemizers those who have an average number of itemizations larger than the mean for that income level. For our estimates of variance that rely on dividing into three or more groups, we further subdivide the heavy-itemization group into two or more groups. To create cutoffs that subdivide the above-mean-itemizations group into roughly equally groups across each income level, we compute first the income-unconditional quantiles of the normalized number of itemizations \widetilde{item}_{t-1}^h , defined as:

$$(C2) \quad \widetilde{item}_{t-1}^h = \frac{item_{t-1}^h - \overline{item}(\tilde{z})}{sd(item(\tilde{z}))}$$

where $sd(item(\tilde{z}))$ is the estimated income conditional standard deviation of itemizations, computed from (C1). We find the cutoffs that subdivide the above mean \widetilde{item}_{t-1}^h into two, three, and four equally sized groups (here we use cutoffs that do not depend on income). We use these cutoffs to create indicator variables $HI(g; k)_{t-1}^h$ which are equal to 1 if, out of the $k \in \{3, 4, 5\}$ groups we are using in a given specification, household h 's value of \widetilde{item}_{t-1}^h is between the upper and lower cutoffs for group g .

We then estimate the local regression:

$$(C3) \quad \begin{aligned} \Delta \log z_t^h = & a(\tilde{z}) + \varepsilon_{L,t}(\tilde{z}) \cdot \Delta \log R_t'(z_{t-1}^h) + \sum_{g=2}^k \delta_{g,k}(\tilde{z}) \cdot HI(g; k)_{t-1}^h \cdot \Delta \log R_t'(z_{t-1}^h) \\ & + c(\tilde{z}) \cdot mrs_{t-1}^h + d_{t-1} + \xi_t^h \end{aligned}$$

We then estimate the shares $(s_{g,k})$ of each of groups by running locally linear regression of the indicator variables $HI(g; k)_{t-1}^h$ on \tilde{z} . We combine our estimates of shares and our estimates of the differences in elasticities $(\delta_{g,k}(\tilde{z}))$ to compute our estimates of the lower bound on the variance of elasticities at each income level \tilde{z} .

$$(C4) \quad \hat{V}_k(\tilde{z}) = \sum_{g=1}^k \hat{s}_{g,k}(\tilde{z}) \cdot (\hat{\delta}_{g,k}(\tilde{z}) - \bar{\delta}_k(\tilde{z}))^2$$

where $\bar{\delta}_k(\tilde{z}) = \sum_{g=1}^k \hat{s}_{g,k}(\tilde{z}) \cdot \hat{\delta}_{g,k}(\tilde{z})$, and where we set $\hat{s}_{1,k} = 1 - \sum_{g=1}^k \hat{s}_{g,k}$ and set, without loss for the purposes of computing variances, $\hat{\delta}_{1,k}(\tilde{z}) = 0$.

To compute confidence intervals, we bootstrap this entire procedure holding fixed the choices for the bandwidths, and use the basic bootstrap confidence interval.

ESTIMATION OF LOCAL SHAPE PARAMETER

To compute $\alpha(z)$ in each year, we estimate a smooth functional form for the CDF of taxable income $G_t(z)$ in each year t with local polynomial regressions to third degree in log taxable income. Here we follow Fan and Gijbels (1996), who recommend including terms up to one order above the derivative of interest. We then translate our estimated coefficients for the first and second derivatives of $G_t(z)$ in log income into an estimate of $\alpha_t(z)$. Namely, our estimate of $\alpha_t(z)$ is equal to the negative of our local estimate of $\frac{d^2 G_t(z)}{d \log z^2}$ divided by our local estimate of $\frac{d G_t(z)}{d \log z}$. We obtain confidence intervals by bootstrapping the whole procedure while holding the estimated optimal bandwidth constant.

Our estimates are broadly consistent across years and similar to other recent work such as those

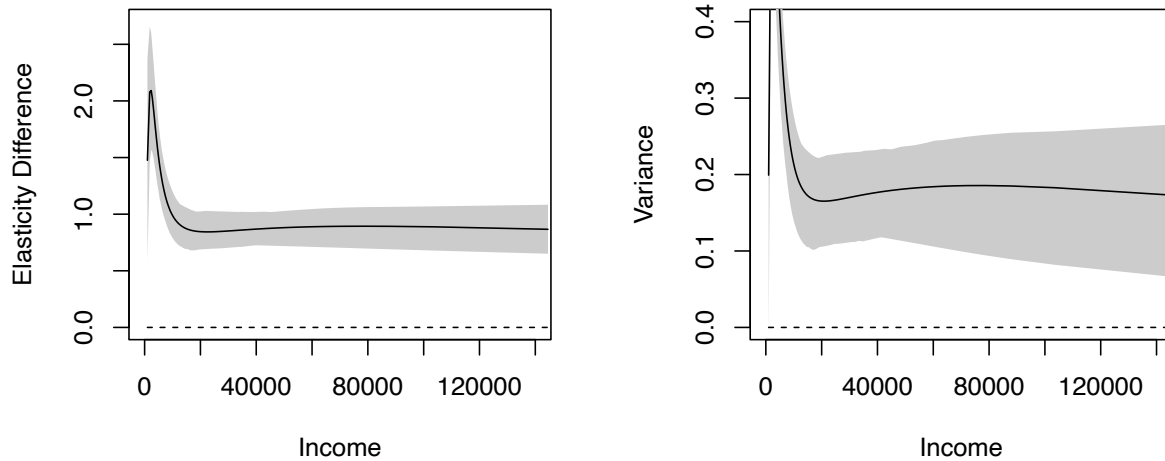


FIGURE C2. ETI DIFFERENCE BY ITEMIZER STATUS AND IMPLIED VARIANCE BOUND

Note: Left panel: Difference in elasticities between heavy and light itemizers, by income. Right panel: Implied lower bound on variance in elasticity, by income. 95% confidence bands. To pool households across years, we represent their (log) incomes in deviations from the annual mean; plots show incomes relative to the 1990 mean.

of Hendren (2020). Figure C4 shows our estimates in 1990.

CHOICE OF KERNELS AND SELECTION OF BANDWIDTHS AND FOR LOCAL REGRESSIONS

For the regressions that rely on the cross-sectional tax return data, we used Epanechnikov kernels (since those have bounded support, they save on computational time). For all other local regressions, we used Gaussian kernels.

To select bandwidths, we minimize the average of squared leave-one-out residuals, i.e.,

$$(C5) \quad LOOCV = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{g}_{-i}(x_i))^2,$$

where $\hat{g}_{-i}(x_i)$ is the predicted value for y_i using the estimated model \hat{g}_{-i} that does use the observation i , but evaluated at the covariate values x_i . The average of these residuals is an estimate for the asymptotic mean integrated square error (AMISE) of model with bandwidth h . By minimizing it, we pick the model with lowest estimated AMISE (Li and Racine, 2007).

For regressions with a large number of observations—that is, the regressions that we use to compute the shape parameter of the income distribution based on the cross-sectional tax return data—the procedure above is computationally demanding. To speed up computations we use the procedure described in Racine (1993), where the leave-one-out cross-validation is computed in subsamples and then is scaled down, using the fact that the optimal bandwidth should be proportional to $c\sigma_x n^{-1/5}$, where c is a constant that does not depend on the number of observations n , and σ_x is the standard deviation of x . In that procedure, we create forty subsamples of one thousand observations each for each set of local regressions.

D. ADDITIONAL FIGURES

This appendix presents the three figures discussed in Section V. Each figure shows a simulation of the revenue, household welfare, and welfare generated by earners with initial incomes

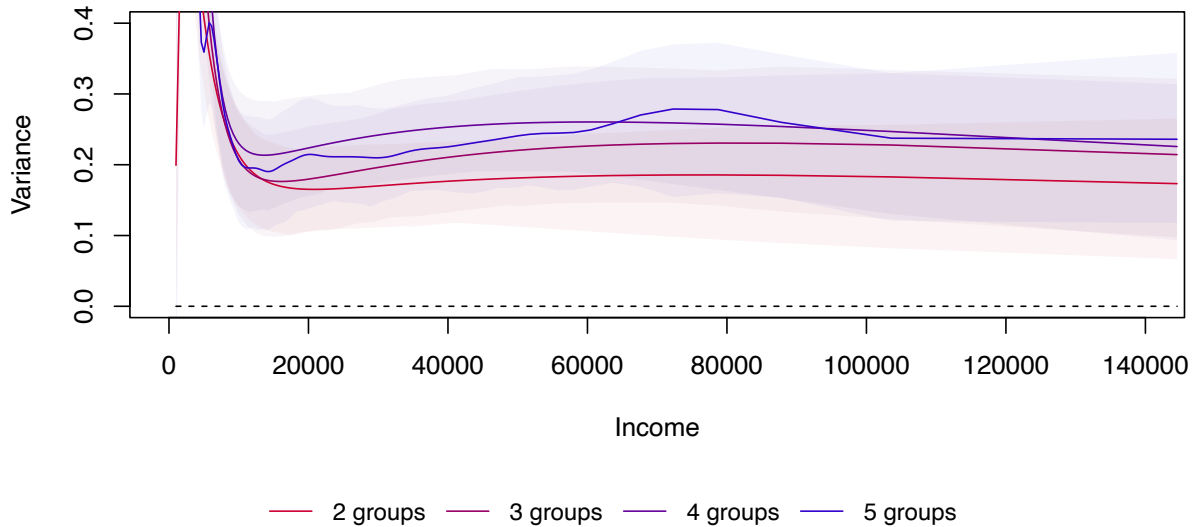


FIGURE C3. LOWER BOUNDS ON ETI VARIANCE FOR DIFFERENT NUMBERS OF ITEMIZER GROUPS

Note: For different numbers of groups, implied lower bound on variance in elasticity, by income. 95% confidence bands. To pool households across years, we represent their (log) incomes in deviations from the annual mean; plots show incomes relative to the 1990 mean.

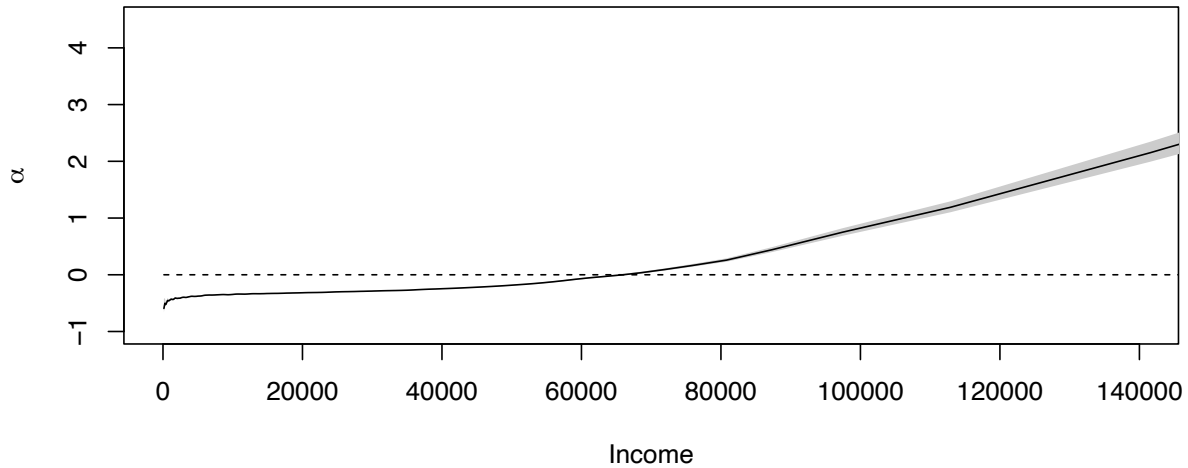


FIGURE C4. LOCAL SHAPE PARAMETER OF THE 1990 INCOME DISTRIBUTION

Note: Local shape parameter of the 1990 income distribution, by income level. 95% confidence bands.

above \$100,000, as a function of the top marginal retention rate that applies to income above this threshold. Simulations assume that each household h has quasilinear preferences with a constant compensated elasticity and a productivity distributed independently of its elasticity. Elasticities are distributed according to a truncated log-normal distribution with mean 0.3, maximum 12, and standard deviation σ_{top} set equal to either 0.625, 0.75, or 0.875. Productivities are distributed so that top incomes have a Pareto tail with Pareto parameter 2.5. We compute welfare generated by top earners by assuming that the planner places a value $\lambda_{top} = 0.25$ on transfers to top earners

relative to government revenue.

The three plots below simulate the impacts of the top tax rate at different values of σ_{top} .

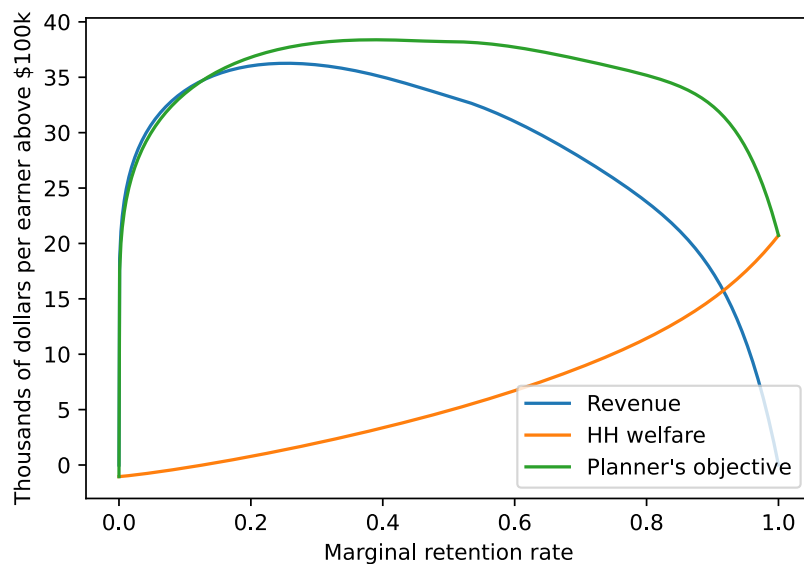


FIGURE D1. SIMULATED IMPACTS OF THE TOP TAX RATE WHEN $\sigma_{top} = 0.625$

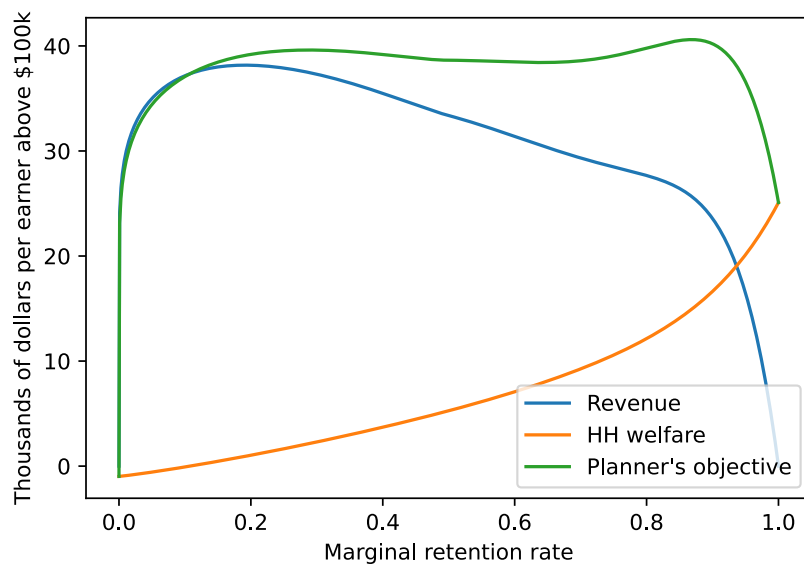
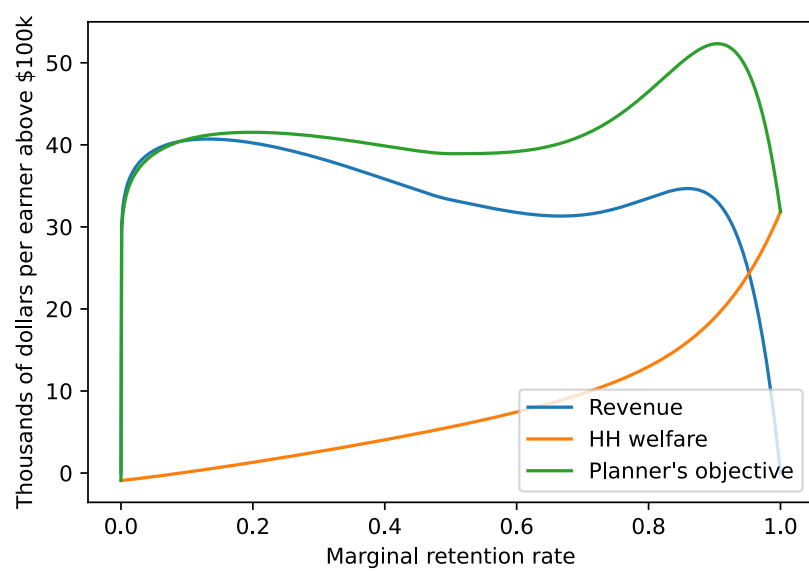


FIGURE D2. SIMULATED IMPACTS OF THE TOP TAX RATE WHEN $\sigma_{top} = 0.75$

FIGURE D3. SIMULATED IMPACTS OF THE TOP TAX RATE WHEN $\sigma_{top} = 0.875$