# Online Appendix for "Generalized Social Marginal Welfare Weights Imply Inconsistent Comparisons of Tax Policies"

Itai Sher
University of Massachusetts Amherst

**Contents**

# A Definitions and proofs of results stated in main text

## A.1 Well-behaved families of tax policies

In this section, I spell out the requirements for a well-behaved families of tax policies introduced in Section II.A more formally and completely, both for individualized tax policies that can depend on $i$, and for non-individualized tax policies that do not depend on $i$, as in Section V.A.

### A.1.1 Individualized tax policies

A family of tax policies $\left(T^{\theta}\right)_{\theta\in\Theta}$ is **well-behaved** if

1. for each $i$ and $\theta$, $i$'s optimal income in response to $T^{\theta}$, $z_i(\theta)$ exists, is unique, and $z_i(\theta) > 0$, and the second order condition for $i$'s optimization problem, when facing $T^{\theta}$, holds with strict inequality at the optimum: $\left.\frac{\mathrm{d}^2}{\mathrm{d}z^2}\right|_{z=z_i(\theta)} U_i\left(z - T_i\left(z,\theta\right), z\right) < 0$, and
2. (a) for all $i$, the map $(z,\theta) \mapsto T_i(z,\theta)$ is smooth, and
   (b) there exists a finite set subset of $\{i_0, i_1, \ldots, i_n\}$ of $I$, with $n \geq 1$ and $i_0 = 0 < i_1 < i_2 < \ldots < i_n = 1$ such that the map $(i, z, \theta) \mapsto T_i(z,\theta)$ is smooth on $(i_{k-1}, i_k) \times Z \times \Theta$, for $k = 1, \ldots n$.

To eliminate any possible ambiguity, $\left.\frac{\mathrm{d}^2}{\mathrm{d}z^2}\right|_{z=z_i(\theta)} U_i\left(z - T_i\left(z,\theta\right), z\right)$ is the second derivative of the function $z \mapsto U_i\left(z - T_i\left(z,\theta\right), z\right)$. Assuming quasilinear utility, $\left.\frac{\mathrm{d}^2}{\mathrm{d}z^2}\right|_{z=z_i(\theta)} U_i\left(z - T_i\left(z,\theta\right), z\right) =$

$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\Big|_{z=z_i(T)} u\left(z - T_i\left(z,\theta\right) - v_i\left(z\right)\right)$. As mentioned in the main text, condition 2b allows for a finite number of discontinuities in $i$.

A tax policy is **regular** if there exists a well-behaved family $\left(T^{\theta}\right)_{\theta\in\Theta}$ and $\theta' \in \Theta$ such that $T^{\theta'} = T$. Given this definition, it is easy to see that a tax policy $T$ is regular if and only if

1. for each $i$, $z_i\left(T\right)$ exists and is unique, $z_i\left(T\right) > 0$, and $\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}\Big|_{z_i=z_i(T)} U_i\left(z_i - T_i\left(z_i\right), z_i\right) < 0$, and
2. (a) for all $i$, the map $z \mapsto T_i\left(z\right)$ is smooth, and,
   (b) there exists a finite set subset of $\{i_0, i_1, \ldots, i_n\}$ of $I$, with $n \geq 1$ and $i_0 = 0 < i_1 < i_2 < \ldots < i_n = 1$ such that the map $(i, z) \mapsto T_i\left(z\right)$ is smooth on $(i_{k-1}, i_k) \times Z$, for $k = 1, \ldots n$.

It follows immediately from the definitions of well-behaved families of tax policies and regular tax policies that any regular tax policy must satisfy the above conditions. Going in the other direction, if $T$ satisfies the above conditions then the family $\left(T^{\theta}\right)$, defined by $T^{\theta} = T, \forall\theta$ is well-behaved. So the above conditions are sufficient for a tax policy to be regular as well.

A doubly parameterized family $\left(T^{\theta,\epsilon}\right)_{\theta\in\Theta,\epsilon\in E}$ is **well-behaved** if

1. for each $i, \theta$, and $\epsilon$, $z_i\left(\theta, \epsilon\right)$ exists and is unique, $z_i\left(\theta, \epsilon\right) > 0$, and the second order condition holds with strict inequality: $\frac{\mathrm{d}^2}{\mathrm{d}z^2}\Big|_{z=z_i(\theta,\epsilon)} U_i\left(z - T_i\left(z, \theta, \epsilon\right), z\right) < 0$, and
2. (a) for all $i$, the map $(z, \theta, \epsilon) \mapsto T_i\left(z, \theta, \epsilon\right)$ is smooth, and
   (b) there exists a finite set subset of $\{i_0, i_1, \ldots, i_n\}$ of $I$, with $n \geq 1$ and $i_0 = 0 < i_1 < i_2 < \ldots < i_n = 1$ such that the map $(i, z, \theta, \epsilon) \mapsto T_i\left(z, \theta, \epsilon\right)$ is smooth on $(i_{k-1}, i_k) \times Z \times \Theta \times E$, for $k = 1, \ldots n$.

### A.1.2 Non-individualized tax policies

When taxes are not individualized, and hence are the same for all agents and do not depend on $i$, the requirements for well-behavedness simplify. In particular, in this case, a tax policy $T$ is regular if and only if

1. for each $i$, $z_i\left(T\right)$ exists and is unique, $z_i\left(T\right) > 0$, and $\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}\Big|_{z_i=z_i(T)} U_i\left(z_i - T\left(z_i\right), z_i\right) < 0$, and
2. the map $z \mapsto T\left(z\right)$ is smooth.

Likewise, when taxes are not individualized, a family $\left(T^{\theta,\epsilon}\right)$ is well behaved if

1. for each $i, \theta$, and $\epsilon$, $z_i\left(\theta, \epsilon\right)$ exists and is unique, $z_i\left(\theta, \epsilon\right) > 0$, and the second order condition holds with strict inequality: $\frac{\mathrm{d}^2}{\mathrm{d}z^2}\Big|_{z=z_i(\theta,\epsilon)} U_i\left(z - T\left(z, \theta, \epsilon\right), z\right) < 0$, and
2. the map $(z, \theta, \epsilon) \mapsto T\left(z, \theta, \epsilon\right)$ is smooth.

The following observation is useful

**Observation A.1** *A family of non-individualized tax policies $\left(T^{\theta,\epsilon}\right)$ is well behaved if and only if (i) for all $\theta$ and $\epsilon$, $T^{\theta,\epsilon}$ is regular and (ii) the map $(z, \theta, \epsilon) \mapsto T\left(z, \theta, \epsilon\right)$ is smooth.*

## A.2 Proof of Proposition 1

Assume that $g$, $\left(T^{\theta}\right)$ and $\theta_0$ are as in the hypothesis of the proposition. Now, first assume that $\int g_i\left(\theta_0\right) \frac{\partial}{\partial\theta}\big|_{\theta=\theta_0} T_i\left(z_i\left(\theta_0\right),\theta\right) \mathrm{d}i < 0$. It follows from the smoothness of welfare weights, utility functions and parameterized families of tax policies that if $\theta_1$ is such that $\theta_1 > \theta_0$ and $\theta_1$ is sufficiently close to $\theta_0$, then for all $\theta' \in [\theta_0, \theta_1]$, $\int g_i\left(T^{\theta'}\right) \frac{\partial}{\partial\theta}\big|_{\theta=\theta'} T_i\left(z_i\left(T^{\theta'}\right),\theta\right) \mathrm{d}i < 0$. It follows from the global improvement principle (in Section II.B) that for all $\theta' \in (\theta_0, \theta_1)$, $T^{\theta_0} \prec^g T^{\theta'}$. This establishes the first claim in Proposition 1.

Next assume that $\int g_i\left(\theta_0\right) \frac{\partial}{\partial\theta}\big|_{\theta=\theta_0} T_i\left(z_i\left(\theta_0\right),\theta\right) \mathrm{d}i > 0$. Now define the parameterized family of tax policies, $\left(\tilde{T}^{\theta}\right)_{\theta\in\left[-\overline{\theta},-\underline{\theta}\right]}$ by $\tilde{T}^{\theta} = T^{-\theta}, \forall\theta \in \left[-\overline{\theta},-\underline{\theta}\right]$, and, using notation analogous to that introduced in Section II.A, let $\tilde{T}_i\left(z,\theta\right) = \tilde{T}_i^{\theta}\left(z\right)$. Then we have:

$$\begin{aligned}\int g_i\left(\tilde{T}^{-\theta_0}\right) \frac{\partial}{\partial\theta}\bigg|_{\theta=-\theta_0} \tilde{T}_i\left(z_i\left(\tilde{T}^{-\theta_0}\right),\theta\right) \mathrm{d}i &= \int g_i\left(T^{\theta_0}\right) \times \left(-\frac{\partial}{\partial\theta}\bigg|_{\theta=\theta_0} T_i\left(z_i\left(T^{\theta_0}\right),\theta\right)\right) \mathrm{d}i \\ &= -\int g_i\left(T^{\theta_0}\right) \frac{\partial}{\partial\theta}\bigg|_{\theta=\theta_0} T_i\left(z_i\left(T^{\theta_0}\right),\theta\right) \mathrm{d}i < 0,\end{aligned}$$

where the inequality follows from the assumption made at the beginning of the paragraph. It follows from the smoothness of welfare weights, utility functions and parameterized families of tax policies that if $-\theta_1 \in \left(-\overline{\theta}, -\theta_0\right)$ is sufficiently close to $-\theta_0$, then for all $\theta' \in [-\theta_1, -\theta_0]$, $\int g_i\left(\tilde{T}^{\theta'}\right) \frac{\partial}{\partial\theta}\big|_{\theta=\theta'} \tilde{T}_i\left(z_i\left(\tilde{T}^{\theta'}\right),\theta\right) \mathrm{d}i < 0$. So the global improvement principle implies that, for all $-\theta' \in (-\theta_1, -\theta_0)$, $\tilde{T}^{-\theta'} \prec^g \tilde{T}^{-\theta_0}$. So for all $\theta' \in \left(\theta_0, \theta_1\right), T^{\theta_0} \succ^g T^{\theta'}$. This establishes the second claim of Proposition 1. □

## A.3 Proof of Proposition 2

First assume that all agents are indifferent as $\theta$ varies in the interval $[\theta_0, \theta_1]$. Then, for all $\theta' \in [\theta_0, \theta_1]$ and agents $i$, $\frac{\mathrm{d}}{\mathrm{d}\theta}\big|_{\theta'=\theta} U_i\left(T^{\theta'}\right) = 0$. Hence, by (5), for all $\theta' \in [\theta_0, \theta_1]$ and agents $i$, $\frac{\partial}{\partial\theta}\big|_{\theta=\theta'} T_i\left(z_i\left(T^{\theta'}\right),\theta\right) = 0$. So, for all $\theta' \in [\theta_0, \theta_1]$, $\int g_i\left(T\right) \frac{\partial}{\partial\theta}\big|_{\theta=\theta'} T_i\left(z_i\left(T^{\theta'}\right),\theta\right) \mathrm{d}i = 0$. So by the global indifference principle (in Section II.B), $T^{\theta_0} \sim^g T^{\theta_1}$. This establishes Pareto indifference along paths. Weak Pareto along paths is similar, appealing again to (5), and using the global improvement principle (also in Section II.B) instead of the global indifference principle. □

## A.4 Definitions for and proof of Corollary 1

Consider a real-valued social welfare function $W\left(T\right)$, whose domain is the set of regular tax policies. Say the social welfare function is **sufficiently differentiable** if for all well-behaved families $\left(T^{\theta}\right)_{\theta\in\Theta}$ and $\theta_0 \in \Theta$, the derivative $\frac{\mathrm{d}}{\mathrm{d}\theta}\big|_{\theta=\theta_0} W\left(T^{\theta}\right)$ exists. Say that a social welfare function $W$ is **Paretian along paths** if for all well-behaved $\left(T^{\theta}\right)_{\theta\in\Theta}$ and all $\theta_0, \theta_1 \in \Theta$ with $\theta_0 < \theta_1$, $W$ satisfies the following properties:

1. **Pareto indifference along a path.** Suppose that all agents are indifferent among all tax policies $T^\theta$ for $\theta \in [\theta_0, \theta_1]$. Then $W\left(T^{\theta_0}\right) = W\left(T^{\theta_1}\right)$.

2. **Weak Pareto along paths.** Suppose that, for all $\hat{\theta} \in [\theta_0, \theta_1]$ and all agents $i$, $\frac{\mathrm{d}}{\mathrm{d}\theta} U_i\left(\hat{\theta}\right) > 0$. Then $W\left(T^{\theta_0}\right) < W\left(T^{\theta_1}\right)$.

Say that a system of welfare weights $g$ **implements social welfare function** $W$ if $W$ is sufficiently differentiable and for all well-behaved families $\left(T^\theta\right)_{\theta\in\Theta}$ and all $\theta' \in \Theta$,

$$\left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=\theta'} W\left(T^\theta\right) > 0 \Leftrightarrow \int g_i\left(T^{\theta'}\right) \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} T\left(z_i\left(T^{\theta'}\right), \theta\right) \mathrm{d}i < 0 \text{ and} \tag{A.1}$$

$$\left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=\theta'} W\left(T^\theta\right) = 0 \Leftrightarrow \int g_i\left(T^{\theta'}\right) \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} T\left(z_i\left(T^{\theta'}\right), \theta\right) \mathrm{d}i = 0. \tag{A.2}$$

The first condition says that increasing $\theta$ is good according to the social welfare function $W$ and this is detected by the $\theta$-derivative of $W\left(T^\theta\right)$ if and only if increasing $\theta$ is desirable according welfare weights $g$. The second condition says that the $\theta$-derivative of $W\left(T^\theta\right)$ does not detect any change in social welfare if and only if welfare weights do not detect any change in social welfare.

Having made the terms in the corollary precise, I now prove the corollary. Assume that the system of welfare weights $g$ implements social welfare function $W$. Let $\left(T^\theta\right)_{\theta\in\Theta}$ be well-behaved and let $\theta_0, \theta_1 \in \Theta$ with $\theta_0 < \theta_1$, and suppose that all agents are indifferent among all tax policies $T^{\theta'}$ for $\theta' \in [\theta_0, \theta_1]$. Then arguing as in the proof of Proposition 2, it follows that $\int g_i\left(T^{\theta'}\right) \frac{\partial}{\partial\theta}\big|_{\theta=\theta'} T\left(z_i\left(T^{\theta'}\right), \theta\right) \mathrm{d}i = 0$. So by (A.2), $\frac{\mathrm{d}}{\mathrm{d}\theta}\big|_{\theta=\theta'} W\left(T^\theta\right) = 0$, for all $\theta' \in [\theta_0, \theta_1]$. So $W\left(T^{\theta_0}\right) = W\left(T^{\theta_1}\right)$. So any social welfare function implemented by $g$ satisfies Pareto indifference along paths. The argument that any social welfare function $W$ implemented by welfare weights satisfies Weak Pareto along paths, proceeds similarly, using (A.1) in the place of (A.2) to derive $\frac{\mathrm{d}}{\mathrm{d}\theta}\big|_{\theta=\theta'} W\left(T^\theta\right) > 0$, for all $\theta' \in [\theta_0, \theta_1]$, and hence $W\left(T^{\theta_0}\right) < W\left(T^{\theta_1}\right)$. □

### A.5 Proof of Proposition 3

It is convenient to prove a stronger version of Proposition 3, which adds a third equivalent condition – condition 2 in Proposition A.1 below – to conditions 1 and 3. Recall that we have assumed that $g_i\left(c_i, z_i\right)$ is a smooth function of $\left(c_i, z_i\right)$.

**Proposition A.1** *Let $g$ and $\hat{g}$ be related as in (9). Then the following conditions are equivalent:*

1. *$g$ is structurally utilitarian.*

2. $\forall i \in I, \forall \hat{u}_i \in \mathbb{R},\ \forall z_i, z_i' \in Z,\ \ \hat{g}_i\left(\hat{u}_i, z_i\right) = \hat{g}_i\left(\hat{u}_i, z_i'\right).$

3. $\forall i \in I, \forall \hat{u}_i \in \mathbb{R}, \forall z_i \in Z, \frac{\partial}{\partial z_i}\hat{g}_i\left(\hat{u}_i, z_i\right) = 0.$

Proof. First I argue that condition 1 of the proposition implies condition 2. Assume that $g$ is structurally utilitarian. Now choose $i \in I, z_i, z_i' \in Z$, and $\hat{u}_i \in \mathbb{R}$. Define $c_i = \hat{u}_i + v_i\left(z_i\right)$ and

$c_i' = \hat{u}_i + v_i(z_i')$. Then observe that

$$c_i - v_i(z_i) = \hat{u}_i = c_i' - v_i\left(z_i'\right). \tag{A.3}$$

Then $\hat{g}_i(\hat{u}_i, z_i) = g_i(c_i, z_i) = g_i(c_i', z_i') = \hat{g}_i(\hat{u}_i, z_i')$, where the first and last equalities follow from (9), and the middle equality follows from (A.3) and the assumption that $g$ is structurally utilitarian. It follows that condition 2 of the proposition holds.

Next I argue that condition 2 implies condition 1. So assume condition 2. Choose $i \in I, c_i, c_i' \in \mathbb{R}, z_i, z_i' \in Z$ and $\hat{u}_i \in \mathbb{R}$ such that $\hat{u}_i = c_i - v_i(z_i) = c_i' - v_i(z_i')$. It follows that $g_i(c_i, z_i) = \hat{g}_i(\hat{u}_i, z_i) = \hat{g}_i(\hat{u}_i, z_i') = g_i(c_i', z_i')$, where the first and last equalities follow from (9), and the middle equality follows from condition 2 of the proposition. This establishes condition 1.

Finally, consider the equivalence of conditions 2 and 3. First observe that our smoothness assumptions imply that condition 2 implies: $\forall i \in I, \forall \hat{u}_i \in \mathbb{R}, \forall z_i \in Z, \frac{\partial}{\partial z_i}\hat{g}_i(\hat{u}_i, z_i) = 0$. Going in the other direction, the equivalence now follows from the fundamental theorem of calculus. □

### A.6 Proof of Theorem 1

First assume welfare weights arise from a generalized utilitarian social welfare function, meaning that they are of the form $g_i(c_i, z_i) = F_i'(U_i(c_i, z_i)) \frac{\partial}{\partial c_i} U_i(c_i, z_i)$. These weights are structurally utilitarian because, if, for all $c_i, c_i', z_i, z_i'$, if $c_i - v_i(z_i) = c_i' - v_i(z_i')$, then $U_i(c_i, z_i) = u(c_i - v_i(z_i)) = u(c_i' - v_i(z_i')) = U_i(c_i', z_i')$ and $\frac{\partial}{\partial c_i} U_i(c_i, z_i) = u'(c_i - v_i(z_i)) = u'(c_i' - v_i(z_i')) = \frac{\partial}{\partial c_i} U_i(c_i', z_i')$. So if $c_i - v_i(z_i) = c_i' - v_i(z_i')$, then $g_i(c_i, z_i) = g_i(c_i', z_i')$.

Going in the other direction, by Proposition 3, structural utilitarianism is equivalent to the requirement that, holding fixed agent characteristics $(x_i, y_i)$, welfare weights are a function of $\hat{u}_i = c_i - v_i(z_i)$, so that, assuming structural utilitarianism, we can write $g_i(c_i, z_i) = g(c_i, z_i, x_i, y_i) = \hat{g}(\hat{u}_i, x_i, y_i) = \hat{g}_i(\hat{u}_i)$. Define the function $w_i(\hat{u}_i) = w(\hat{u}_i, x_i, y_i)$ by $w_i\left(\hat{u}_i^0\right) = \int_0^{\hat{u}_i^0} \hat{g}_i(\hat{u}_i)\, \mathrm{d}\hat{u}_i$. Now define the Function $F : \mathbb{R} \times X \times Y \to \mathbb{R}$ by $F(v_i, x_i, y_i) = w\left(u^{-1}(v_i), x_i, y_i\right)$, where $u^{-1}(\cdot)$ is the inverse of $u(\cdot)$. If $x_i$ and $y_i$ are not discrete, the smoothness of $w$ and $u$ imply that $F$ is smooth. If $x_i$ and $y_i$ are discrete, $w$ is smooth in its first argument and hence $F$ is smooth in $v_i$. Let $F_i(v_i) = F(v_i, x_i, y_i)$ and define $W_i(c_i, z_i) = F_i(U_i(c_i, z_i))$. We have $W_i(c_i, z_i) = F_i(U_i(c_i, z_i)) = w_i\left(u^{-1}(u(c_i - v_i(z_i)))\right) = w_i(c_i - v_i(z_i))$. Note that, from the above, we have $g_i(c_i, z_i) = \hat{g}_i(c_i - v_i(z_i)) = w_i'(c_i - v_i(z_i)) = \frac{\partial}{\partial c_i} W_i(c_i, z_i) = F_i'(U_i(c_i, z_i)) \frac{\partial}{\partial c_i} U_i(c_i, z_i)$. So the weights arise from a generalized utilitarian social welfare function. □

### A.7 Proof of Corollary 2

Suppose that welfare weights $g$ are structurally utilitarian. It follows from Theorem 1 that welfare weights are of the form $g_i(c_i, z_i) = F_i'(U_i(c_i, z_i)) \frac{\partial}{\partial c_i} U_i(c_i, z_i)$ for $F_i(u_i) = F(u_i, x_i, y_i)$ for some $F$. So for the social welfare function $W(T) = -\int F_i(U_i(c_i(T), z_i(T)))\, \mathrm{d}i$, the envelope theorem implies that, for all well-behaved families $\left(T^\theta\right)_{\theta \in \Theta}$ and $\theta_0 \in \Theta$, $\frac{\mathrm{d}}{\mathrm{d}\theta}\big|_{\theta=\theta_0} W\left(T^\theta\right) = -\int g_i\left(T^{\theta_0}\right) \frac{\partial}{\partial \theta}\big|_{\theta=\theta_0} T_i\left(z_i\left(T^{\theta_0}\right), \theta\right) \mathrm{d}i$. □

### A.8 Proof of Theorem 2

#### A.8.1 Main argument

What follows is a more formal version of the argument in the main text. Assume that welfare weights are not structurally utilitarian. It follows from Proposition 3 that there exists $j \in I, \hat{u}^* \in \mathbb{R}, z^* \in Z$, such that $\frac{\partial}{\partial z_j}\hat{g}_j(\hat{u}^*, z^*) \neq 0$. Smoothness of the primitives implies that we can choose $z^*$ so that $z^* > 0$. Assume that $\frac{\partial}{\partial z_j}\hat{g}_j(\hat{u}^*, z^*) < 0$. (The argument would be similar if we assumed instead that $\frac{\partial}{\partial z_j}\hat{g}_j(\hat{u}^*, z^*) > 0$.) Our smoothness assumptions then imply that there exists a non-degenerate[1] closed interval of agents $S$, which is a proper subset of $I = [0,1]$, such that, for all agents $i \in S$, $\frac{\partial}{\partial z_i}\hat{g}_i(\hat{u}^*, z^*) < 0$.[2] Let $O$ and $Q$ be two other non-degenerate closed intervals contained in $[0,1]$, such that $S, O$, and $Q$ are pairwise disjoint. Now consider a doubly parameterized family of tax policies $\left(T^{\theta,\epsilon}\right)_{\theta\in\Theta,\epsilon\in E}$, where $\Theta = \left[\underline{\theta}, \overline{\theta}\right]$ for some $\underline{\theta} < \overline{\theta}$ and $E = [-\bar{\epsilon}, \bar{\epsilon}]$ for some $\bar{\epsilon} > 0$, and which takes the following form:

$$T_i^{\theta,\epsilon}(z_i) = \begin{cases} \tau_i(\theta) z_i + \kappa_i(\theta) + \epsilon t_S, & \text{if } i \in S, \\ -\epsilon t_O, & \text{if } i \in O, \\ \bar{\tau}(\theta,\epsilon) z_i + \bar{\kappa}_i(\theta,\epsilon), & \text{if } i \in Q, \\ 0, & \text{otherwise.} \end{cases} \tag{A.4}$$

Above $\tau_i(\theta)$ is a personalized marginal tax rate for agents in $i$ in $S$, and $\bar{\tau}(\theta,\epsilon)$ is a marginal tax rate which is not personalized on $Q$; both $\tau_i(\theta)$ and $\bar{\tau}(\theta,\epsilon)$ depend on parameter values. $\kappa_i(\theta)$ and $\bar{\kappa}_i(\theta,\epsilon)$ are personalized lumpsum taxes that depend on parameters. $t_S$ and $t_O$ are positive real numbers, so that $\epsilon t_S$ and $-\epsilon t_O$ are lumpsum taxes as well. I assume that the map $(i,\theta) \mapsto \tau_i(\theta)$ is smooth on the domain $S \times \Theta$ and that the map $(\theta,\epsilon) \mapsto \bar{\tau}(\theta,\epsilon)$ is smooth on the domain $\Theta \times E$. Moreover, I assume that there exists $\theta_0 \in \left(\underline{\theta}, \overline{\theta}\right)$ such that, for all $i \in S$, $\tau_i(\theta_0) = 1 - v_i'(z^*)$ and, for all $\theta \in \Theta$, $\tau_i'(\theta) > 0$.[3] In what follows, let $\hat{U}_i(\theta,\epsilon) = \hat{U}_i\left(T^{\theta,\epsilon}\right) = z_i(\theta,\epsilon) - T_i^{\theta,\epsilon}(z_i(\theta,\epsilon)) - v_i(z_i(\theta,\epsilon))$ be $i$'s utility in response to $T^{\theta,\epsilon}$, using the representation that omits the outer utility function $u(\cdot)$, and note that $g_i(\theta,\epsilon) = g_i\left(T^{\theta,\epsilon}\right) = \hat{g}_i\left(\hat{U}_i(\theta,\epsilon), z_i(\theta,\epsilon)\right)$. When an agent $i$ in $S$ faces tax policy $T^{\theta_0,0}$, they will solve the problem $\max_{z_i}(1-\tau_i(\theta_0)) z_i - \kappa_i(\theta_0) - v_i(z_i)$. It follows from the construction of $\tau_i(\theta_0)$ and the fact that $v_i(z_i)$ is strictly convex that $z_i = z^*$ uniquely satisfies the agent's first order condition when $(\theta,\epsilon) = (\theta_0, 0)$, namely, $(1 - \tau(\theta_0)) - v_i'(z_i) = 0$. Because agents' objective is strictly concave, it follows that $z_i = z^*$ is the unique optimum for all agents $i \in S$ when facing tax policy $T^{\theta_0,0}$, so that $z_i(\theta_0, 0) = z^*$ for all $i \in S$. For all $i \in S$, define the function $\kappa_i(\theta)$ in (A.4) to solve:

$$(1 - \tau_i(\theta)) z_i(\theta, 0) - v_i(z_i(\theta,0)) - \kappa_i(\theta) = \hat{u}^*, \quad \forall\theta \in \Theta. \tag{A.5}$$

[1] By a non-degenerate closed interval, I mean a closed interval which is not equal to a single point.

[2] Of course, it is possible that $\frac{\partial}{\partial z_i}\hat{g}_i(\hat{u}^*, z^*) < 0$ for all $i \in [0,1]$, but in this case there is also a closed interval $S$, which is a proper subset of $[0,1]$, on which this property holds.

[3] We allow for the possibility that $\tau_i(\theta_0) < 0$.

That is, the lumpsum tax $\kappa_i(\theta)$ is chosen so as the keep the agents' (in $S$) utility fixed at $\hat{u}^*$ when the agent faces tax policies of the form $T^{\theta,0}$ as $\theta$ changes – where we measure utility via the representation $\hat{U}_i\left(T^{\theta,0}\right)$ that excludes the outer utility function $u(\cdot)$. Note that we can freely define $\kappa_i(\theta)$ in this way because the optimal income $z_i(\theta,0)$ depends only on the marginal tax rate $\tau_i(\theta)$ and not on the lumpsum tax $\kappa_i(\theta)$. Note, moreover, that, for any $\epsilon \in E$, $\theta \in \Theta$, and $i \in S$, $i$'s utility, when facing $T^{\theta,\epsilon}$, is $\hat{U}_i(\theta,\epsilon) = \hat{u}^* - \epsilon t_S$, which does not depend on $\theta$. So, holding $\epsilon$ fixed, each agent $i \in S$ is indifferent as $\theta$ varies. Likewise, for all $i \in Q$, define $\bar{\kappa}_i(\theta,\epsilon)$ to satisfy the following equation:

$$\left(1-\bar{\tau}(\theta,\epsilon)\right) z_i(\theta,\epsilon) - v_i\left(z_i(\theta,\epsilon)\right) - \bar{\kappa}_i(\theta,\epsilon) = 0, \quad \forall \theta \in \Theta, \forall \epsilon \in E. \tag{A.6}$$

That is, the lumpsum tax $\bar{\kappa}_i(\theta,\epsilon)$ is selected to keep the utility $\hat{U}_i(\theta,\epsilon)$ of all agents $i \in Q$ equal to zero as $\theta$ and $\epsilon$ vary. Again, observe that $z_i(\theta,\epsilon)$ only depends on the marginal tax rate $\bar{\tau}(\theta,\epsilon)$ and not on the lumpsum tax $\bar{\kappa}_i(\theta,\epsilon)$. Given the above, it follows by construction that, holding $\epsilon$ fixed, *all* agents are indifferent, as $\theta$ varies in $T^{\theta,\epsilon}$. So, it follows from part 1 of Proposition 2 – Pareto indifference along paths – that

$$T^{\theta_0,\epsilon} \sim^g T^{\theta_1,\epsilon}, \quad \forall \epsilon \in E, \tag{A.7}$$

where $\theta_1$, satisfying $\theta_0 < \theta_1$, is a value of $\theta$ that we now select. In particular, it follows from the facts that $\frac{\partial}{\partial z_i}\hat{g}_i(z^*,\hat{u}^*) < 0$ and $z_i(\theta_0,0) = z^*$ for all $i \in S$ and the smoothness of the primitives of the model that if we choose $\theta_1$ sufficiently close to $\theta_0$,

$$\frac{\partial}{\partial z_i}\hat{g}_i\left(\hat{U}_i(\theta,0), z_i(\theta,0)\right) = \frac{\partial}{\partial z_i}\hat{g}_i\left(\hat{u}^*, z_i(\theta,0)\right) < 0, \quad \forall \theta \in [\theta_0,\theta_1], \forall i \in S. \tag{A.8}$$

So let us choose $\theta_1$ so that (A.8) is satisfied. Moreover, since $z_i(\theta_0,0) = z^* > 0, \forall i \in S$, we may assume that $\theta_1$ is chosen sufficiently close to $\theta_0$ that, for all $i \in S$ and $\theta \in [\theta_0,\theta_1], z_i(\theta,0) > 0$.

For any $\theta \in \Theta$, define $g_S(\theta,0) = \int_S g_i(\theta,0)\,\mathrm{d}i$ and $g_O(\theta,0) = \int_O g_i(\theta,0)\,\mathrm{d}i$. It follows from the fact that $\hat{U}_i(\theta,0) = \hat{u}^*, \forall \theta \in \Theta, \forall i \in S$, (A.8), and the assumption that $\tau_i'(\theta) > 0, \forall \theta \in \Theta, \forall i \in S$, which, given that $z_i(\theta,0) > 0, \forall \theta \in [\theta_0,\theta_1], \forall i \in S$, implies that $\frac{\partial}{\partial \theta} z_i(\theta,0) < 0, \forall \theta \in [\theta_0,\theta_1], \forall i \in S$, that

$$\frac{\partial}{\partial \theta} g_S(\theta,0) > 0, \quad \forall \theta \in [\theta_0,\theta_1]. \tag{A.9}$$

Choose $\theta' \in (\theta_0,\theta_1)$ and suppose that the positive numbers $t_S$ and $t_O$ in (A.4) were selected to satisfy

$$g_S(\theta',0)\, t_S = g_O(\theta',0)\, t_O. \tag{A.10}$$

Then, writing $T(z_i, \theta, \epsilon) = T^{\theta,\epsilon}(z_i)$, we have:

$$\begin{aligned}
&\int g_i(\theta_0, 0) \left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=0} T_i(z_i(\theta_0, 0), \theta_0, \epsilon)\, \mathrm{d}i \\
&= g_S(\theta_0, 0)\, t_S - g_O(\theta_0, 0)\, t_O + \int_Q \left( \left[ \left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=0} \bar{\tau}(\theta_0, \epsilon) \right] z_i(\theta_0, 0) + \left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=0} \bar{\kappa}_i(\theta_0, \epsilon) \right) \mathrm{d}i \\
&= g_S(\theta_0, 0)\, t_S - g_O(\theta_0, 0)\, t_O + \int_Q - \left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=0} \hat{U}_i(\theta_0, \epsilon)\, \mathrm{d}i \\
&= g_S(\theta_0, 0)\, t_S - g_O(\theta_0, 0)\, t_O < 0,
\end{aligned} \tag{A.11}$$

where the second equality follows from the envelope theorem, and the third equality follows from the fact that, by (A.6), the utility of all agents in $Q$ is held fixed as $\epsilon$ varies in $T^{\theta_0,\epsilon}$, so that, for all $i \in Q$, $\left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=0} \hat{U}_i(\theta_0, \epsilon) = 0$. The inequality follows from (A.10), and the facts that $g_O(\theta, 0)$ is constant in $\theta$, that, by (A.9), $g_S(\theta, 0)$ is increasing in $\theta$, and that $\theta_0 < \theta'$. Using similar arguments,

$$\int g_i(\theta_1, 0) \left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=0} T_i(z_i(\theta_1, 0), \theta_1, \epsilon)\, \mathrm{d}i = g_S(\theta_1, 0)\, t_S - g_O(\theta_1, 0)\, t_O > 0, \tag{A.12}$$

The reason that the the inequality in (A.12) points in the opposite direction of the inequality in (A.11) is that, whereas $\theta_0 < \theta'$, $\theta_1 > \theta'$. It follows from (A.11), (A.12), and the local improvement principle – Proposition 1 – that

$$\begin{aligned}
&T^{\theta_0,0} \prec^g T^{\theta_0,\epsilon}, \\
&T^{\theta_1,0} \succ^g T^{\theta_1,\epsilon},
\end{aligned} \quad \text{for sufficiently small } \epsilon > 0. \tag{A.13}$$

Putting (A.7) and (A.13), together, we have that for sufficiently small $\epsilon > 0$,

$$T^{\theta_0,0} \prec^g T^{\theta_0,\epsilon} \sim T^{\theta_1,\epsilon} \prec^g T^{\theta_1,0} \sim^g T^{\theta_0,0}. \tag{A.14}$$

So, on the assumption that welfare weights are not structurally utilitarian, we have constructed a social preference cycle.

The last step is to show that revenue can be held fixed across the tax policies in the cycle. This is achieved via the selection of $\bar{\tau}(\theta, \epsilon)$ in (A.4). For any marginal tax rate $\tau$, write $z_i(\tau)$ to be the income that $i$ would earn, if $i$ faces the tax policy $T(z) = \tau z$, or, in other words, if $i$ faces a constant marginal tax rate of $\tau$. It follows that, for all $i \in Q$, we can write $z_i(\bar{\tau}(\theta, \epsilon)) = z_i(\theta, \epsilon)$ because every agent $i \in Q$ faces the constant marginal tax rate $\bar{\tau}(\theta, \epsilon)$ under tax policy $T^{\theta,\epsilon}$. Let

$R_Q(\theta,\epsilon)$ be the revenue raised from agents in $Q$ by tax policy $T^{\theta,\epsilon}$. Then we have

$$\begin{aligned} R_Q(\theta,\epsilon) &= \int_Q T(z_i(\theta,\epsilon),\theta,\epsilon)\,\mathrm{d}i = \int_Q [\bar{\tau}(\theta,\epsilon) z_i(\theta,\epsilon) + \bar{\kappa}_i(\theta,\epsilon)]\,\mathrm{d}i \\ &= \int_Q [\bar{\tau}(\theta,\epsilon) z_i(\theta,\epsilon) + (1-\bar{\tau}(\theta,\epsilon)) z_i(\theta,\epsilon) - v_i(z_i(\theta,\epsilon))]\,\mathrm{d}i \\ &= \int_Q [z_i(\theta,\epsilon) - v_i(z_i(\theta,\epsilon))]\,\mathrm{d}i, \end{aligned} \tag{A.15}$$

where the third equality follows from (A.6). Next, for any marginal tax rate $\tau$, define $\tilde{R}_Q(\tau)$ by

$$\tilde{R}_Q(\tau) = \int_Q [z_i(\tau) - v_i(z_i(\tau))]\,\mathrm{d}i.$$

Then it follows from (A.15) and the fact that $z_i(\bar{\tau}(\theta,\epsilon)) = z_i(\theta,\epsilon)$ that $\tilde{R}_Q(\bar{\tau}(\theta,\epsilon)) = R_Q(\theta,\epsilon)$. Since we assume that, in the absence of taxes, all agents earn positive income (see Section I), there exists a positive marginal tax rate $\tau_0$, which is sufficiently small that, for all $i \in Q$, $z_i(\tau_0) > 0$.[4] From agent $i$'s first order condition, when facing marginal tax rate $\tau_0$, we have that, for all $i \in Q$, $0 = (1-\tau_0) - v_i'(z_i(\tau_0)) < 1 - v_i'(z_i(\tau_0))$. Assume that $\bar{\tau}(\theta_0,0) = \tau_0$. Define $R_{-Q}(\epsilon,\theta) = \int_{I\setminus Q} T_i(z_i(\theta,\epsilon),\theta,\epsilon)\,\mathrm{d}i$ to be the revenue raised by tax policy $T^{\theta,\epsilon}$ from all agents not in $Q$. Now consider the condition:

$$\tilde{R}_Q(\bar{\tau}(\theta,\epsilon)) + R_{-Q}(\theta,\epsilon) = \tilde{R}_Q(\tau_0) + R_{-Q}(\theta_0,0). \tag{A.16}$$

Observe that $\tilde{R}_Q'(\bar{\tau}(\theta_0,0)) = \int_Q z_i'(\tau_0)[1 - v_i'(z_i(\tau_0))]\,\mathrm{d}i < 0$.[5] It follows from the implicit function theorem that the function $\bar{\tau}(\theta,\epsilon)$ is uniquely determined in a neighborhood of $(\theta_0,0)$ by $\bar{\tau}(\theta_0,0) = \tau_0$ and (A.16). Redefining $\bar{\epsilon}$ to be sufficiently small and $\overline{\theta}$ and $\underline{\theta}$ to be sufficiently close to $\theta_0$ if necessary, and assuming that $\theta_1$ was chosen sufficiently close to $\theta_0$ so that $\theta_0 < \theta_1 < \bar{\theta}$ still holds, we may assume that we have thus defined $\bar{\tau}(\theta,\epsilon)$ on all of $\Theta \times E$, and moreover such that $z_i(\theta,\epsilon) > 0$ for all $i$ in $Q$, $\theta \in \Theta$, and $\epsilon \in E$ (since $z_i(\theta_0,0) = z_i(\tau_0) > 0, \forall i \in Q$ and $Q$ is compact). Note now that (A.16) implies that the revenue of $T^{\theta,\epsilon}$ is held constant as $\theta$ and $\epsilon$ vary. This completes the proof. □

### A.8.2 Well-behavedness of $(T^{\theta,\epsilon})$

Here I verify that the family $(T^{\theta,\epsilon})$ in (A.4) above is well-behaved (see Sections II.A and A.1), as this is required for Propositions 1 and 2. I begin by verifying the first condition for well-behavedness.

[4] The assumption that, in the absence of taxes, all agents earn positive income, is not necessary for the proof. In the absence of this assumption, we could instead select $\tau_0$ to be a sufficiently small negative marginal tax rate that, for all $i \in Q$, $z_i(\tau_0) > 0$. Then the proof would proceed in the same way as below except that $\tilde{R}_Q'(\bar{\tau}(\theta,\epsilon)) > 0$ rather than $\tilde{R}_Q'(\bar{\tau}(\theta,\epsilon)) < 0$. However what matters for the argument is only that $\tilde{R}_Q'(\bar{\tau}(\theta,\epsilon)) \neq 0$.

[5] This inequality follows from the facts that, by our assumptions above imply that, for all $i \in Q$ (i) $z_i(\tau_0) > 0$, so that $z_i'(\tau_0) < 0$, and that (ii) $1 - v_i'(z_i(\tau_0)) > 0$.

Existence and uniqueness of $z_i(\theta,\epsilon)$ are straightforward to establish.[6] That $z_i(\theta,\epsilon)>0$ for all $i$ in $S$ and $Q$ was established in the course of the proof (noting that $z_i(\theta,\epsilon)=z_i(\theta,0),\forall i\in S$), and for $i$ not in $S$ or $Q$, $z_i(\theta,\epsilon)>0$ follows from the assumption that, when facing a zero marginal tax rate, all agents select a positive income (see Section I.A). That each agent's second order condition holds with a strict inequality follows from the fact that $v_i''>0$ and $u'>0$ hold everywhere and that all agents face a tax policy that is linear in $z$ (possibly with a zero marginal tax rate) under $T^{\theta,\epsilon}$. This establishes that $\left(T^{\theta,\epsilon}\right)$ satisfies the first condition required for well-behavedness.

To establish the second condition, I appeal to the following observation.

**Observation A.2** *The maps $(i,\theta)\mapsto\tau_i(\theta)$ and $(i,\theta)\mapsto\kappa_i(\theta)$ are smooth on $S\times\Theta$; $(\theta,\epsilon)\mapsto\bar{\tau}(\theta,\epsilon)$ is smooth on $\Theta\times E$; and the map $(i,\theta,\epsilon)\mapsto\kappa_i(\theta,\epsilon)$ is smooth on $Q\times\Theta\times E$.*

The map $(i,\theta)\mapsto\tau_i(\theta)$ is smooth on $S\times\Theta$ by assumption.[7] The map $(i,\theta)\mapsto\kappa_i(\theta)$ is smooth on $S\times\Theta$ because it is defined by (A.5) and all of the other functions in (A.5) are smooth.[8] The map $(\theta,\epsilon)\mapsto\bar{\tau}(\theta,\epsilon)$ is smooth because it is defined by the implicit function theorem via equation (A.16) and the other functions in (A.16) are smooth. Finally, $(i,\theta,\epsilon)\mapsto\kappa_i(\theta,\epsilon)$ is smooth on $Q\times\Theta\times E$ because it is defined by (A.6) and the other functions in (A.6) are smooth.[9]

That, for all $i$, $(z,\theta,\epsilon)\mapsto T_i(z,\theta,\epsilon)$ is smooth follows from (A.4) and Observation A.2. Recall that $S,O$, and $Q$ are assumed in Section A.8.1 to be pairwise disjoint closed intervals. It then follows from (A.4) and Observation A.2 that the map $(i,z,\theta,\epsilon)\mapsto T_i(z,\theta,\epsilon)$ only fails to be smooth when $i$ is one of the six endpoints of these three intervals. This establishes the second condition required for the well-behavedness of $\left(T^{\theta,\epsilon}\right)$.

### A.9 Calculations from Section IV.B

That the revenue of $T^{\theta,\epsilon}$ is $\frac{1}{4}$, for all $\theta$ and $\epsilon$, is verified by the following calculation:

$$\begin{aligned}
R\left(T^{\theta,\epsilon}\right) &= \underbrace{\frac{1}{2}\left[z(\theta)\theta+\kappa(\theta)+\epsilon\right]}_{\text{revenue from type } A \text{ agents}}+\underbrace{\frac{1}{2}\left[z\left(\sqrt{1-\theta^2}\right)\sqrt{1-\theta^2}+\kappa(\theta)-\epsilon\right]}_{\text{revenue from type } B \text{ agents.}}\\
&=\frac{1}{2}\left[(1-\theta)\theta+\frac{1}{2}(1-\theta)^2\right]+\frac{1}{2}\left[\left(1-\sqrt{1-\theta^2}\right)\sqrt{1-\theta^2}+\frac{1}{2}\left(1-\sqrt{1-\theta^2}\right)^2\right]\\
&=\frac{1}{2}\left[\frac{1}{2}(1-\theta)(1+\theta)\right]+\frac{1}{2}\left[\frac{1}{2}\left(1-\sqrt{1-\theta^2}\right)\left(1+\sqrt{1-\theta^2}\right)\right]=\frac{1}{4}\left(1-\theta^2\right)+\frac{1}{4}\theta^2=\frac{1}{4}.
\end{aligned}$$

[6]Existence and uniqueness follow from the assumptions of Section I.A, the fact that when facing a linear tax policy, agents' objectives are strictly concave, the selection of the marginal tax rates $\tau_i(\theta_0)$ and $\tau_0$, and the construction of $\tau_i(\theta)$ and $\bar{\tau}(\theta,\epsilon)$ using the implicit function theorem.

[7]This is consistent with the other assumptions made on $\tau_i(\theta)$. In particular, I assumed that, for all $i\in S$, $\tau_i(\theta_0)=1-v_i'(z^*)$, and that, for all $\theta\in\Theta$, $\tau_i'(\theta)>0$. So for example, if I had specifically defined $\tau_i(\theta)=1-v_i'(z^*)+(\theta-\theta_0)$ on $S\times\Theta$, $(i,\theta)\mapsto\tau_i(\theta)$ would have satisfied these properties, and, moreover, would be smooth on $S\times\Theta$, since the assumptions of Section IV.A imply that $i\mapsto v_i'(z^*)$ is smooth.

[8] In particular, $(i,\theta)\mapsto z_i(\theta,0)$ is smooth because the latter is characterized by the implicit function theorem applied to $i$'s first order condition and the the functions that feature in the first order condition are smooth in $(i,\theta)$.

[9]Again, the map $(i,\theta,\epsilon)\mapsto z_i(\theta,\epsilon)$ is smooth for reasons similar to those explained in footnote 8 of the appendix.

The above calculation also implies that, at $T^{\theta,\epsilon}$, the total tax paid by a type $A$ agent is $\frac{1}{2}\left(1-\theta^2\right)+\epsilon$ and the total tax paid by a type $B$ agent is $\frac{1}{2}\theta^2-\epsilon$. So as $\theta$ rises from $\theta_0=\sqrt{\frac{1}{3}}$ to $\theta_1=\sqrt{\frac{2}{3}}$, the total tax paid by a type $A$ agent falls from $\frac{1}{3}+\epsilon$ to $\frac{1}{6}+\epsilon$ while the total tax paid by a type $B$ agent rises from $\frac{1}{6}-\epsilon$ to $\frac{1}{3}-\epsilon$.

A formal derivation that $T^{\theta_0,0}\succ^g T^{\theta_0,\epsilon}$ for sufficiently small $\epsilon>0$ is as follows.

$$\int_0^1 g_i\left(\theta_0,0\right)\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=0} T_i\left(z_i\left(T\left(\theta_0,0\right)\right),\theta_0,\epsilon\right)\mathrm{d}i=\underbrace{\int_0^{\frac{1}{2}}\left[\tilde{g}\left(\frac{1}{3}\right)\times 1\right]\mathrm{d}i}_{\text{type } A \text{ agents}}+\underbrace{\int_{\frac{1}{2}}^1\left[\tilde{g}\left(\frac{1}{6}\right)\times(-1)\right]\mathrm{d}i}_{\text{type } B \text{ agents}}$$
$$=\frac{1}{2}\tilde{g}\left(\frac{1}{3}\right)-\frac{1}{2}\tilde{g}\left(\frac{1}{6}\right)>0.$$

So by Proposition 1 – the local improvement principle – it follows that $T^{\theta_0,0}\succ^g T^{\theta_0,\epsilon}$ for sufficiently small $\epsilon>0$.

Similarly, $\int_0^1 g_i\left(\theta_1,0\right)\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=0} T_i\left(z_i\left(\theta_1,0\right),\theta_1,\epsilon\right)\mathrm{d}i=\frac{1}{2}\tilde{g}\left(\frac{1}{6}\right)-\frac{1}{2}\tilde{g}\left(\frac{1}{3}\right)<0$, and, again by Proposition 1, $T^{\theta_1,0}\prec^g T^{\theta_1,\epsilon}$, for sufficiently small $\epsilon>0$.

### A.10 Proof of Corollary 3

Assume that $g$ is not structurally utilitarian. It follows from Proposition 3 that there exists an agent $j\in(0,1)$, $z^*\in Z$ with $z^*>0$ and $\hat{u}^*\in\mathbb{R}$ and such that

$$\frac{\partial}{\partial z_j}\hat{g}_j\left(\hat{u}^*,z^*\right)\neq 0. \tag{A.17}$$

We can assume that $j$ is in the interior of $I=[0,1]$ and $z^*>0$ because of the smoothness of the primitives. Choose a smooth strictly convex tax policy $T$, with moreover $T''\left(z_i\right)>0,\forall z_i$, such that (i) $T'\left(z^*\right)=1-v_j'\left(z^*\right)$, (ii) $T'(0)$ is sufficiently small (or negative if $1-v_j'\left(z^*\right)<0$) such that all agents would earn a positive income in response to $T$ – recall that in the absence of taxes, all agents earn a positive income (see Section I.A) –, and (iii) $\lim_{z\to\infty}T'(z)>0$. These assumptions, together with the strict convexity of $v_i\left(z_i\right)$ and the assumption that $v_i'\left(z_i\right)>1$ for sufficiently large $z_i$ (see Section I.A), imply that $T$ is regular. (See Section A.1.2 for the requirements for regularity.) It follows from property (i) that $z_j(T)=z^*$. By the appropriate choice of a lumpsum transfer in $T$, we can ensure that $\hat{U}_j(T)=\hat{u}^*$. (A.17) together with the smoothness of the primitives and of $T$ now ensure that if we select a sufficiently small interval $\left(i_a,i_b\right)$ containing $j$, then either (18) or (19) holds. □

### A.11 Omitted details from the proof of Lemma 2

Here I present the details of the proof of Lemma 2 that were omitted in the main text: the expression for the overlapping term $C$ discussed in the text, and the proof of conditions (21)-(22). First, I present the expression for the term $C$, which I will prove is the overlapping term below:

$$C = \int \left( -\frac{\partial}{\partial \hat{u}_i} \hat{g}_i \left( \hat{U}_i(\theta_0, \epsilon_0), z_i(\theta_0, \epsilon_0) \right) \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=\epsilon_0} T(z_i(\theta_0, \epsilon_0), \theta_0, \epsilon) \frac{\partial}{\partial \theta}\bigg|_{\theta=\theta_0} T(z_i(\theta_0, \epsilon_0), \theta, \epsilon_0) \right.$$
$$+ g_i(\theta_0, \epsilon_0) \left[ -\frac{\frac{\partial^2}{\partial \theta \partial z_i}\Big|_{\theta=\theta_0, z_i=z_i(\theta_0,\epsilon_0)} T(z_i, \theta, \epsilon_0) \frac{\partial^2}{\partial \epsilon \partial z_i}\Big|_{\epsilon=\epsilon_0, z_i=z_i(\theta_0,\epsilon_0)} T(z_i, \theta_0, \epsilon)}{v_i''(z_i(\theta_0, \epsilon_0)) + \frac{\partial^2}{\partial z_i^2}\Big|_{z_i=z_i(\theta_0,\epsilon_0)} T(z_i, \theta_0, \epsilon_0)} \right.$$
$$\left. \left. + \frac{\partial^2}{\partial \epsilon \partial \theta}\bigg|_{\epsilon=\epsilon_0, \theta=\theta_0} T(z_i(\theta_0, \epsilon_0), \theta, \epsilon) \right] \right) \mathrm{d}i. \tag{A.18}$$

Next, I present some useful preliminary facts, which I use to establish (21)-(22). Observe that at $(\theta_0, \epsilon_0)$, agent $i$'s optimization problem is: $\max_{z_i} [z_i - v_i(z_i) - T(z_i, \theta_0, \epsilon_0)]$. The first-order condition is: $1 - v_i'(z_i) - \frac{\partial}{\partial z_i} T(z_i, \theta_0, \epsilon_0) = 0$. Applying the implicit function theorem to the first-order condition,[10] we have:

$$\frac{\partial}{\partial \theta}\bigg|_{\theta=\theta_0} z_i(\theta, \epsilon_0) = -\frac{\frac{\partial^2}{\partial \theta \partial z_i}\Big|_{\theta=\theta_0, z_i=z_i(\theta,\epsilon_0)} T(z_i, \theta, \epsilon_0)}{v_i''(z_i(\theta_0, \epsilon_0)) + \frac{\partial^2}{\partial z_i^2}\Big|_{z_i=z_i(\theta_0,\epsilon_0)} T(z_i, \theta_0, \epsilon_0)}, \tag{A.19}$$

$$\frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=\epsilon_0} z_i(\theta_0, \epsilon) = -\frac{\frac{\partial^2}{\partial \epsilon \partial z_i}\Big|_{\epsilon=\epsilon_0, z_i=z_i(\theta_0,\epsilon_0)} T(z_i, \theta_0, \epsilon)}{v_i''(z_i(\theta_0, \epsilon_0)) + \frac{\partial^2}{\partial z_i^2}\Big|_{z_i=z_i(\theta_0,\epsilon_0)} T(z_i, \theta_0, \epsilon_0)}. \tag{A.20}$$

I am now ready to establish (21)-(22). First, I establish (21):

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\bigg|_{\epsilon=\epsilon_0} \int g_i(\theta_0, \epsilon) \frac{\partial}{\partial \theta}\bigg|_{\theta=\theta_0} T(z_i(\theta_0, \epsilon), \theta, \epsilon) \,\mathrm{d}i$$
$$= \int \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=\epsilon_0} \left[ g_i(\theta_0, \epsilon) \frac{\partial}{\partial \theta}\bigg|_{\theta=\theta_0} T(z_i(\theta_0, \epsilon), \theta, \epsilon) \right] \mathrm{d}i \tag{A.21}$$
$$= \int \left( \left[ \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=\epsilon_0} g_i(\theta_0, \epsilon) \right] \frac{\partial}{\partial \theta}\bigg|_{\theta=\theta_0} T(z_i(\theta_0, \epsilon_0), \theta, \epsilon_0) \right.$$
$$\left. + g_i(\theta_0, \epsilon_0) \frac{\mathrm{d}}{\mathrm{d}\epsilon}\bigg|_{\epsilon=\epsilon_0} \frac{\partial}{\partial \theta}\bigg|_{\theta=\theta_0} T(z_i(\theta_0, \epsilon), \theta, \epsilon) \right) \mathrm{d}i \tag{A.22}$$

[10] As $\left(T^{\theta,\epsilon}\right)_{\theta\in\Theta,\epsilon\in E}$ is well-behaved, it follows that the first-order condition uniquely characterizes agent $i$'s optimal income $z_i(\theta, \epsilon)$.

$$= \int \Bigg( \Bigg[ -\frac{\partial}{\partial \hat{u}_i} \hat{g}_i \left( \hat{U}_i(\theta_0, \epsilon_0), z_i(\theta_0, \epsilon_0) \right) \frac{\partial}{\partial \epsilon}\Big|_{\epsilon=\epsilon_0} T(z_i(\theta_0, \epsilon_0), \theta_0, \epsilon)$$

$$+ \frac{\partial}{\partial z_i} \hat{g}_i \left( \hat{U}_i(\theta_0, \epsilon_0), z_i(\theta_0, \epsilon_0) \right) \frac{\partial}{\partial \epsilon}\Big|_{\epsilon=\epsilon_0} z_i(\theta_0, \epsilon) \Bigg] \frac{\partial}{\partial \theta}\Big|_{\theta=\theta_0} T(z_i(\theta_0, \epsilon_0), \theta, \epsilon_0) \tag{A.23}$$

$$+ g_i(\theta_0, \epsilon_0) \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=\epsilon_0} \frac{\partial}{\partial \theta}\Big|_{\theta=\theta_0} T(z_i(\theta_0, \epsilon), \theta, \epsilon) \Bigg) \mathrm{d}i$$

$$= \int \Bigg( \Bigg[ -\frac{\partial}{\partial \hat{u}_i} \hat{g}_i \left( \hat{U}_i(\theta_0, \epsilon_0), z_i(\theta_0, \epsilon_0) \right) \frac{\partial}{\partial \epsilon}\Big|_{\epsilon=\epsilon_0} T(z_i(\theta_0, \epsilon_0), \theta_0, \epsilon)$$

$$+ \frac{\partial}{\partial z_i} \hat{g}_i \left( \hat{U}_i(\theta_0, \epsilon_0), z_i(\theta_0, \epsilon_0) \right) \frac{\partial}{\partial \epsilon}\Big|_{\epsilon=\epsilon_0} z_i(\theta_0, \epsilon) \Bigg] \frac{\partial}{\partial \theta}\Big|_{\theta=\theta_0} T(z_i(\theta_0, \epsilon_0), \theta, \epsilon_0)$$

$$+ g_i(\theta_0, \epsilon_0) \Bigg[ \frac{\partial^2}{\partial z_i \partial \theta}\Big|_{z_i = z_i(\theta_0, \epsilon_0), \theta=\theta_0} T(z_i, \theta, \epsilon_0) \frac{\partial}{\partial \epsilon}\Big|_{\epsilon=\epsilon_0} z_i(\theta_0, \epsilon) \tag{A.24}$$

$$+ \frac{\partial^2}{\partial \epsilon \partial \theta}\Big|_{\epsilon=\epsilon_0, \theta=\theta_0} T(z_i(\theta_0, \epsilon_0), \theta, \epsilon) \Bigg] \Bigg) \mathrm{d}i$$

$$= \int \Bigg( \Bigg[ -\frac{\partial}{\partial \hat{u}_i} \hat{g}_i \left( \hat{U}_i(\theta_0, \epsilon_0), z_i(\theta_0, \epsilon_0) \right) \frac{\partial}{\partial \epsilon}\Big|_{\epsilon=\epsilon_0} T(z_i(\theta_0, \epsilon_0), \theta_0, \epsilon)$$

$$+ \frac{\partial}{\partial z_i} \hat{g}_i (U_i(\theta_0, \epsilon_0), z_i(\theta_0, \epsilon_0)) \frac{\partial}{\partial \epsilon}\Big|_{\epsilon=\epsilon_0} z_i(\theta_0, \epsilon) \Bigg] \frac{\partial}{\partial \theta}\Big|_{\theta=\theta_0} T(z_i(\theta_0, \epsilon_0), \theta, \epsilon_0)$$

$$+ g_i(\theta_0, \epsilon_0) \Bigg[ -\frac{\partial^2}{\partial z_i \partial \theta}\Big|_{z_i = z_i(\theta_0, \epsilon_0), \theta=\theta_0} T(z_i, \theta, \epsilon_0) \frac{\frac{\partial^2}{\partial \epsilon \partial z_i}\Big|_{\epsilon=\epsilon_0, z_i = z_i(\theta_0, \epsilon_0)} T(z_i, \theta_0, \epsilon)}{v_i''(z_i(\theta_0, \epsilon_0)) + \frac{\partial^2}{\partial z^2}\Big|_{z_i = z_i(\theta_0, \epsilon_0)} T(z_i, \theta_0, \epsilon_0)}$$

$$+ \frac{\partial^2}{\partial \epsilon \partial \theta}\Big|_{\epsilon=\epsilon_0, \theta=\theta_0} T(z_i(\theta_0, \epsilon_0), \theta, \epsilon) \Bigg] \Bigg) \mathrm{d}i, \tag{A.25}$$

$$= A + C \tag{A.26}$$

where (A.23) analyzes the term $\frac{\partial}{\partial \epsilon}\big|_{\epsilon=\epsilon_0} g_i(\theta_0, \epsilon)$ and appeals to the fact that, by the envelope theorem, $\frac{\partial}{\partial \epsilon}\big|_{\epsilon=\epsilon_0} \hat{U}_i(\theta_0, \epsilon) = -\frac{\partial}{\partial \epsilon}\big|_{\epsilon=\epsilon_0} T(z_i(\theta_0, \epsilon_0), \theta_0, \epsilon)$, and (A.25) follows from (A.20), $A$ is defined as in the proof outline of Lemma 2 in the main text and $C$ is defined by (A.18). This establishes (21).

As the derivation of (22) is similar, I present it in an abbreviated form:

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\Big|_{\theta=\theta_0} \int g_i(\theta, \epsilon_0) \frac{\partial}{\partial \epsilon}\Big|_{\epsilon=\epsilon_0} T(z_i(\theta, \epsilon_0), \theta, \epsilon) \, \mathrm{d}i$$

$$= \int \Bigg( \Bigg[ -\frac{\partial}{\partial \hat{u}_i} \hat{g}_i \left( \hat{U}_i(\theta_0, \epsilon_0), z_i(\theta_0, \epsilon_0) \right) \frac{\partial}{\partial \theta}\Big|_{\theta=\theta_0} T(z_i(\theta_0, \epsilon_0), \theta, \epsilon_0)$$

$$
\begin{aligned}
&+\frac{\partial}{\partial z_i}\hat{g}_i\left(\hat{U}_i\left(\theta_0,\epsilon_0\right),z_i\left(\theta_0,\epsilon_0\right)\right)\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} z_i\left(\theta,\epsilon_0\right)\Bigg]\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(\theta_0,\epsilon_0\right),\theta_0,\epsilon\right)\\
&+g_i\left(\theta_0,\epsilon_0\right)\left[\left.\frac{\partial^2}{\partial z_i\partial\epsilon}\right|_{z_i=z_i(\theta_0,\epsilon_0),\epsilon=\epsilon_0} T\left(z_i,\theta_0,\epsilon\right)\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} z_i\left(\theta,\epsilon_0\right)\right.\\
&\left.\left.+\left.\frac{\partial^2}{\partial\theta\partial\epsilon}\right|_{\theta=\theta_0,\epsilon=\epsilon_0} T\left(z_i\left(\theta_0,\epsilon_0\right),\theta,\epsilon\right)\right]\right)\mathrm{d}i\\
=&\int\left(\left[-\frac{\partial}{\partial\hat{u}_i}g_i\left(\hat{U}_i\left(\theta_0,\epsilon_0\right),z_i\left(\theta_0,\epsilon_0\right)\right)\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} T\left(z_i\left(\theta_0,\epsilon_0\right),\theta,\epsilon_0\right)\right.\right.\\
&+\frac{\partial}{\partial z}\hat{g}_i\left(\hat{U}_i\left(\theta_0,\epsilon_0\right),z_i\left(\theta_0,\epsilon_0\right)\right)\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} z_i\left(\theta,\epsilon_0\right)\Bigg]\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(\theta_0,\epsilon_0\right),\theta_0,\epsilon\right)\\
&+\hat{g}_i\left(\theta_0,\epsilon_0\right)\left[-\left.\frac{\partial^2}{\partial z_i\partial\epsilon}\right|_{z_i=z_i(\theta_0,\epsilon_0),\epsilon=\epsilon_0} T\left(z_i,\theta_0,\epsilon\right)\frac{\left.\frac{\partial^2}{\partial\theta\partial z_i}\right|_{\theta=\theta_0,z=z_i(\theta_0,\epsilon_0)} T\left(z_i,\theta_0,\epsilon\right)}{v_i''\left(z_i\left(\theta_0,\epsilon_0\right)\right)+\left.\frac{\partial^2}{\partial z_i^2}\right|_{z_i=z_i(\theta,\epsilon)} T\left(z_i,\theta_0,\epsilon_0\right)}\right.\\
&\left.\left.+\left.\frac{\partial^2}{\partial\theta\partial\epsilon}\right|_{\theta=\theta_0,\epsilon=\epsilon_0} T\left(z_i\left(\theta_0,\epsilon_0\right),\theta,\epsilon\right)\right]\right)\mathrm{d}i\\
=&B+C.
\end{aligned}
$$

The justification is similar to the justification for (A.21)-(A.26), using (A.19) instead of (A.20). This establishes (22). □

## A.12 Proof of Lemma 3

The main argument proving Lemma 3 is presented in Section A.12.1. The proofs of a supporting lemma and some related material are presented in the subsequent subsections.

### A.12.1 Main argument

Choose a regular tax policy $T$. (See Section A.1.2 for the requirements for a regular tax policy when taxes are not individualized.) To establish the lemma, I construct a well-behaved doubly parameterized family of tax policies $\left(T^{\theta,\epsilon}\right)_{\theta\in\Theta,\epsilon\in E}$ satisfying (15), (16) and (23), and such that, for the $\theta_0\in\left(\underline{\theta},\overline{\theta}\right),\epsilon_0\in\left(\underline{\epsilon},\overline{\epsilon}\right)$ that feature in the preceding conditions, $T^{\theta_0,\epsilon_0}=T$.

Recall that the *support* of a function $h$ with argument $x$ is the closure of $\{x:h\left(x\right)\neq 0\}$.

To construct $\left(T^{\theta,\epsilon}\right)_{\theta\in\Theta,\epsilon\in E}$, I consider four smooth tax reforms $\mu_1,\mu_2,\eta_1,\eta_2$. Let $i_k,k=1,2,3,5$ be elements of $(0,1)$ be such that $i_1<i_2<i_3=i_a<i_5=i_b$. The reader will notice that we have skipped $i_4$; this term will be introduced below (see Lemma A.1). If we let $\hat{z}_k=z_{i_k}\left(T\right)$ for $k=1,2,3,5$, it follows from assumptions on $v$ and $i\mapsto y_i$ in Section V.A that $\hat{z}_1<\hat{z}_2<\hat{z}_3<\hat{z}_5$. I assume that $\mu_1\left(z\right)=0$ when $z\leq\hat{z}_3$, $\mu_1\left(z\right)$ is increasing in $z$ on the interval $(\hat{z}_3,\hat{z}_5)$, and $\mu_1\left(z\right)$ remains constant at some positive number thereafter. I assume that $\mu_2\left(z\right)=0$ when $z\leq\hat{z}_2$, $\mu_2\left(z\right)$ is increasing in $z$ on the interval $(\hat{z}_2,\hat{z}_3)$ and $\mu_2\left(z\right)=1$ when $z\geq\hat{z}_3$. Assume, moreover, that $\mu_1$

and $\mu_2$ are chosen such that:

$$\int_0^1 g_i(T)\,\mu_1(z_i(T))\,\mathrm{d}i = \int_0^1 g_i(T)\,\mu_2(z_i(T))\,\mathrm{d}i. \tag{A.27}$$

That is, both tax reforms $\mu_1$ and $\mu_2$ have the same marginal effect on social welfare, when benefits are weighted by welfare weights. The above assumptions imply the following lemma, which is proved in Section A.12.2 of the Appendix.

**Lemma A.1** *There exists $i_4 \in (i_3, i_5)$ such that $\mu_1(z_{i_4}(T)) = 1$.*

If we define $\hat{z}_4 = z_{i_4}(T)$, it follows from the fact that $i_3 < i_4 < i_5$ that $\hat{z}_3 < \hat{z}_4 < \hat{z}_5$. So Lemma A.1 says that there is some income level $\hat{z}_4$, between $\hat{z}_3$ and $\hat{z}_5$, such that $\mu_1(\hat{z}_4) = 1$, and moreover income level $\hat{z}_4$ is chosen by some agent $i_4$ when facing tax policy $T$.

Assume that $\eta_1$ has support $[\hat{z}_3, \hat{z}_5]$, and that $\eta_1$ is increasing on $(\hat{z}_3, \hat{z}_4)$ and decreasing on $(\hat{z}_4, \hat{z}_5)$, which implies that $\eta_1(z) > 0, \forall z \in (\hat{z}_3, \hat{z}_5)$. Assume that the support of $\eta_2$ is $[\hat{z}_1, \hat{z}_2]$, that $\eta_2(z) < 0, \forall z \in (\hat{z}_1, \hat{z}_2)$, and that

$$\int_0^1 g_i(T)\,\eta_1(z_i(T))\,\mathrm{d}i = -\int_0^1 g_i(T)\,\eta_2(z_i(T))\,\mathrm{d}i. \tag{A.28}$$

In other words the marginal welfare effect of reform $\eta_1$ is the negative of the marginal welfare effect of reform $\eta_2$, so that the two cancel out.

For any real numbers, $\theta$ and $\epsilon$, define $T_*^{\theta,\epsilon}$ by:

$$T_*^{\theta,\epsilon} = T + \theta\mu_1 + \epsilon(\eta_1 + \eta_2). \tag{A.29}$$

It follows from the Picard-Lindelöf theorem (see Section A.14.1 for a more explicit formulation) that there exist real numbers $\underline{\theta}, \overline{\theta}, \underline{\epsilon}, \overline{\epsilon}$ such that $\underline{\theta} < 0 < \overline{\theta}, \underline{\epsilon} < 0 < \overline{\epsilon}$, and such that we can define the real-valued function $\zeta(\theta, \epsilon)$ on $\Theta \times E$, where $\Theta = \left[\underline{\theta}, \overline{\theta}\right]$ and $E = [\underline{\epsilon}, \overline{\epsilon}]$, by

$$\zeta(0,\epsilon) = 0, \quad \forall \epsilon \in E, \tag{A.30}$$

$$\begin{aligned} &\int g_i\left(T_*^{\theta,\epsilon} - \zeta(\theta,\epsilon)\,\mu_2\right) \\ &\times \left[\mu_1\left(z_i\left(T_*^{\theta,\epsilon} - \zeta(\theta,\epsilon)\,\mu_2\right)\right) - \frac{\partial}{\partial\theta}\zeta(\theta,\epsilon)\,\mu_2\left(z_i\left(T_*^{\theta,\epsilon} - \zeta(\theta,\epsilon)\,\mu_2\right)\right)\right]\mathrm{d}i = 0, \\ &\forall \theta \in \Theta, \forall \epsilon \in E. \end{aligned} \tag{A.31}$$

Next, for all $\theta \in \Theta$ and $\epsilon \in E$, define

$$T^{\theta,\epsilon} = T + [\theta \times \mu_1] - [\zeta(\theta,\epsilon) \times \mu_2] + [\epsilon \times (\eta_1 + \eta_2)] \tag{A.32}$$

$$= T_*^{\theta,\epsilon} - \zeta(\theta,\epsilon)\,\mu_2. \tag{A.33}$$

In Section A.14, I establish that if $\underline{\theta}, \overline{\theta}, \underline{\epsilon}$, and $\overline{\epsilon}$ are all chosen sufficiently close to 0, then $\left(T^{\theta,\epsilon}\right)_{\theta\in\Theta,\epsilon\in E}$

is well-behaved.

So now consider the parameterized family of tax policies $\left(T^{\theta,\epsilon}\right)_{\theta\in\Theta,\epsilon\in E}$, for which we will verify the properties required in the lemma. Let $\theta_0 = 0$ and $\epsilon_0 = 0$. Then note that $T^{\theta_0,\epsilon_0} = T$, as required for the result. Let $S = \{i \in I : z_i(T) \in (\hat{z}_3, \hat{z}_5)\}$ and $O = \{i \in I : z_i(T) \in (\hat{z}_1, \hat{z}_2)\}$, so that, as described in Section V.C, starting at $\theta = \theta_0 = 0$ and $\epsilon = \epsilon_0 = 0$, as $\epsilon$ increases, taxes on the incomes earned by agents in $S$ rise and taxes on incomes earned by agents in $O$ fall.

Recalling that $T(z,\theta,\epsilon) = T^{\theta,\epsilon}(z)$, it follows from (A.32) that, for all $i \in I, \epsilon \in E$, and $\theta' \in \left(\underline{\theta}, \overline{\theta}\right)$,

$$\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} T\left(z_i\left(\theta',\epsilon\right),\theta,\epsilon\right) = \mu_1\left(z_i\left(\theta',\epsilon\right)\right) - \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} \zeta\left(\theta',\epsilon\right)\mu_2\left(z_i\left(\theta',\epsilon\right)\right), \tag{A.34}$$

and it follows from (A.32), (A.30), and the fact that $\theta_0 = 0$, that, for all $i \in I$,

$$\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(\theta_0,\epsilon_0\right),\theta_0,\epsilon\right) = \eta_1\left(z_i(T)\right) + \eta_2\left(z_i(T)\right). \tag{A.35}$$

It follows from (A.31), (A.33), and (A.34) that $\left(T^{\theta,\epsilon}\right)_{\theta\in\Theta,\epsilon\in E}$ satisfies (15), and from (A.28) and (A.35) that $\left(T^{\theta,\epsilon}\right)_{\theta\in\Theta,\epsilon\in E}$ satisfies (16).

Next I seek to establish that $\left(T^{\theta,\epsilon}\right)_{\theta\in\Theta,\epsilon\in E}$ satisfies (23). In the special case in which $\theta = \theta_0$ and $\epsilon = \epsilon_0$ (recall that $\theta_0 = \epsilon_0 = 0$), the general statement in (A.31) reduces to

$$\int g_i(T)\left[\mu_1\left(z_i(T)\right) - \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0}\zeta\left(\theta,\epsilon_0\right)\mu_2\left(z_i(T)\right)\right]\mathrm{d}i = 0.$$

Solving for $\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0}\zeta\left(\theta,\epsilon_0\right)$ from the above equation, it follows that

$$\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0}\zeta\left(\theta,\epsilon_0\right) = \frac{\int g_i(T)\,\mu_1\left(z_i(T)\right)\mathrm{d}i}{\int g_i(T)\,\mu_2\left(z_i(T)\right)\mathrm{d}i} = 1, \tag{A.36}$$

where the second equality follows from (A.27).

Consider the type $i$ agent's optimization problem when facing tax policy $T^{\theta,\epsilon}$–that is, of choosing $z$ so as to maximize $z - v_i(z) - T^{\theta,\epsilon}(z)$. It follows from the implicit function theorem applied to the first order condition for this optimization problem at $(\theta,\epsilon) = (\theta_0,\epsilon_0)$ that

$$\begin{aligned}
\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} z_i\left(\theta,\epsilon_0\right) &= -\frac{\mu_1'\left(z_i(T)\right) - \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0}\zeta\left(\theta,\epsilon_0\right)\mu_2'\left(z_i(T)\right)}{T''\left(z_i(T)\right) + v_i''\left(z_i(T)\right)}, \\
&= -\frac{\mu_1'\left(z_i(T)\right) - \mu_2'\left(z_i(T)\right)}{T''\left(z_i(T)\right) + v_i''\left(z_i(T)\right)} \qquad \forall i \in [0,1], \\
\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} z_i\left(\theta_0,\epsilon\right) &= -\frac{\eta_1'\left(z_i(T)\right) + \eta_2'\left(z_i(T)\right)}{T''\left(z_i(T)\right) + v_i''\left(z_i(T)\right)},
\end{aligned} \tag{A.37}$$

where the second equality for the term $\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} z_i\left(\theta,\epsilon_0\right)$ uses (A.36), and the equality for the term

$\frac{\partial}{\partial\epsilon}\big|_{\epsilon=\epsilon_0} z_i(\theta_0,\epsilon)$ uses the fact that $\frac{\partial}{\partial\epsilon}\big|_{\epsilon=\epsilon_0} \zeta(\theta_0,\epsilon)=0$, which follows from (A.30) and the assumption that $\theta_0=0$. These equations simplify when $i\in[i_3,i_5]$. In particular,

$$\begin{aligned}\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} z_i(\theta,\epsilon_0) &= -\frac{\mu_1'(z_i(T))}{T''(z_i(T))+v_i''(z_i(T))},\\ \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} z_i(\theta_0,\epsilon) &= -\frac{\eta_1'(z_i(T))}{T''(z_i(T))+v_i''(z_i(T))},\end{aligned} \qquad \forall i\in[i_3,i_5]. \tag{A.38}$$

This simplification is explained by the observations that (i) since $\mu_2(z_i(T))=1$ when $i\in[i_3,i_5]$, $\mu_2'(z_i(T))=0$ when $i\in[i_3,i_5]$, and (ii) the support of $\eta_2$ is $[\hat{z}_1,\hat{z}_2]$, so that $\eta_2'(z_i(T))=0$ when $i\in[i_3,i_5]$.

When $(\theta,\epsilon)=(\theta_0,\epsilon_0)$ and $i\in[i_3,i_5]$, (A.34)-(A.35) also simplify:

$$\begin{aligned}\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} T(z_i(\theta_0,\epsilon_0),\theta,\epsilon_0) &= \mu_1(z_i(T))-1,\\ \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} T(z_i(\theta_0,\epsilon_0),\theta_0,\epsilon) &= \eta_1(z_i(T)),\end{aligned} \qquad \forall i\in[i_3,i_5], \tag{A.39}$$

where the first equality uses (A.36) and the fact that $\mu_2(z)=1$ when $z\in[\hat{z}_3,\hat{z}_5]$, and the second equality uses the fact that $\eta_2(z)=0$ when $z\in[\hat{z}_3,\hat{z}_5]$.

Recalling that $i_a=i_3$ and $i_b=i_5$, it follows from (A.38) and (A.39) that

$$\begin{aligned}&\forall i\in(i_a,i_b),\\ &\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} z_i(\theta,\epsilon_0)\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} T(z_i(\theta_0,\epsilon_0),\theta_0,\epsilon)-\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} z_i(\theta_0,\epsilon)\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} T(z_i(\theta_0,\epsilon_0),\theta,\epsilon_0)\\ &=\frac{-\overbrace{\mu_1'(z_i(T))}^{+}\overbrace{\eta_1(z_i(T))}^{+}+\left(\overbrace{\eta_1'(z_i(T))}^{+\text{ on }(i_3,i_4),-\text{ on }(i_4,i_5)}\times\overbrace{[\mu_1(z_i(T))-1]}^{-\text{ on }(i_3,i_4),+\text{ on }(i_4,i_5)}\right)}{\underbrace{T''(z_i(T))+v_i''(z_i(T))}_{+}}<0.\end{aligned} \tag{A.40}$$

where the signs are derived from the assumptions we made above about $\eta_1$ and $\mu_1$ – in particular note that $\mu_1(z)>0$ is increasing on $(\hat{z}_3,\hat{z}_5)$ and, by Lemma A.1, $\mu_1(\hat{z}_4)=1$ – as well as the fact that because $T$ is regular, $0>\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}\big|_{z_i=z_i(T)} u(z_i-T(z_i)-v_i(z_i))=-u'(z_i(T)-T_i(z_i(T))-v_i(z_i(T)))\times[v_i''(z_i(T))+T''(z_i(T))],\forall i\in I$[11] (see Section A.1.2), so that $v_i''(z_i(T))+T''(z_i(T))>0,\forall i\in I$. Next observe that:

- The support of $i\mapsto\frac{\partial}{\partial\theta}\big|_{\theta=\theta_0} z_i(\theta,\epsilon_0)$ is $[i_2,i_5]$.

[11]Observe that $\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}\big|_{z_i=z_i(T)} u(z_i-T(z_i)-v_i(z_i)) = u''(z_i-T(z_i)-v_i(z_i))\times\underbrace{[1-v_i'(z_i(T))-T'(z_i(T))]^2}_{=0}-u'(z_i(T)-T_i(z_i(T))-v_i(z_i(T)))\times[v_i''(z_i(T))+T''(z_i(T))] = -u'(z_i(T)-T_i(z_i(T))-v_i(z_i(T)))\times[v_i''(z_i(T))+T''(z_i(T))]$.

- The support of $i \mapsto \frac{\partial}{\partial\epsilon}\big|_{\epsilon=\epsilon_0} T\left(z_i\left(\theta_0,\epsilon_0\right),\theta_0,\epsilon\right)$ is $[i_1,i_2]\cup[i_3,i_5]$.

Recalling that $i_a = i_3$ and $i_b = i_5$, it follows that

$$\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} z_i\left(\theta,\epsilon_0\right)\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(\theta_0,\epsilon_0\right),\theta_0,\epsilon\right)=0,\quad \forall i\notin\left(i_a,i_b\right). \tag{A.41}$$

To understand why the above expression is equal to zero when $i\in\{i_2,i_3,i_5\}$, note that the expressions $\frac{\partial}{\partial\theta}\big|_{\theta=\theta_0} z_i\left(\theta,\epsilon_0\right)$ and $\frac{\partial}{\partial\epsilon}\big|_{\epsilon=\epsilon_0} T\left(z_i\left(\theta_0,\epsilon_0\right),\theta_0,\epsilon\right)$ are equal to zero on the boundaries of their supports.

- The support of $i\mapsto\frac{\partial}{\partial\epsilon}\big|_{\epsilon=\epsilon_0} z_i\left(\theta_0,\epsilon\right)$ is contained in $[i_1,i_2]\cup[i_3,i_5]$.
- The support of $i\mapsto\frac{\partial}{\partial\theta}\big|_{\theta=\theta_0} T\left(z_i\left(\theta_0,\epsilon_0\right),\theta,\epsilon_0\right)$ is $[i_2,1]$.

It follows that

$$\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} z_i\left(\theta_0,\epsilon\right)\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} T\left(z_i\left(\theta_0,\epsilon_0\right),\theta,\epsilon_0\right)=0,\quad \forall i\notin\left(i_a,i_b\right). \tag{A.42}$$

Again, the above condition uses the fact that $\frac{\partial}{\partial\epsilon}\big|_{\epsilon=\epsilon_0} z_i\left(\theta_0,\epsilon\right)$ and $\frac{\partial}{\partial\theta}\big|_{\theta=\theta_0} T\left(z_i\left(\theta_0,\epsilon_0\right),\theta,\epsilon_0\right)$ are equal to zero on the boundaries of their supports. Together (A.41), (A.42), and the inequality established in (A.40) are equivalent to (23). We have now established that the family $\left(T^{\theta,\epsilon}\right)$ satisfies all of the conditions required by the lemma. □

### A.12.2 Proof of Lemma A.1

Assume, for contradiction, that, for all $i\in\left(i_3,i_5\right), \mu_1\left(z_i\left(T\right)\right)\neq 1$. Then, since the function $i\mapsto\mu_1\left(z_i\left(T\right)\right)$ is smooth (this follows from the assumed smoothness of relevant functions and the implicit function theorem), $\mu_1\left(z_{i_3}\left(T\right)\right)=0$, and $i\mapsto\mu_1\left(z_i\left(T\right)\right)$ is a constant function on $[i_5,1]$, it follows from the intermediate value theorem that $\mu_1\left(z_i\left(T\right)\right)<1,\forall i\in[i_3,1]$. So

$$\begin{aligned}\int_0^1 g_i\left(T\right)\mu_1\left(z_i\left(T\right)\right)\mathrm{d}i &= \int_{i_3}^1 g_i\left(T\right)\mu_1\left(z_i\left(T\right)\right)\mathrm{d}i < \int_{i_3}^1 g_i\left(T\right)\mathrm{d}i \\ &= \int_{i_3}^1 g_i\left(T\right)\mu_2\left(z_i\left(T\right)\right)\mathrm{d}i < \int_0^1 g_i\left(T\right)\mu_2\left(z_i\left(T\right)\right)\mathrm{d}i,\end{aligned} \tag{A.43}$$

where the first equality follows from the fact that the support of $\mu_1$ is $[\hat{z}_3,\bar{z}]$; the first inequality from the our conclusion that $\mu_1\left(z_i\left(T\right)\right)<1,\forall i\in[i_3,1]$ and the fact that $g_i\left(T\right)>0,\forall i\in[0,1]$; the second equality form the fact that $\mu_2\left(z\right)=1$ for all $z\in[\hat{z}_3,\bar{z}]$, and the last inequality from the fact that the $\mu_2$ is nonnegative everywhere and $\mu_2\left(z\right)>0$ for $z\in\left(\hat{z}_2,\hat{z}_3\right)$. However, (A.43) contradicts (A.27). So the assumption that $\mu_1\left(z_i\left(T\right)\right)$ is never equal to 1 on $\left(i_3,i_5\right)$ leads to a contradiction, completing the proof. □

### A.12.3 A variant of Lemma 3

This section discusses the proof of a variant of Lemma 3; I appeal to this variant in the proof of Lemma 4.

**Lemma A.2** *Let $T$ be a regular tax policy and let $i_a, i_b \in (0,1)$ be such that $i_a < i_b$. Then there exists a well-behaved family $\left(T^{\theta,\epsilon}\right)$ with $T^{\theta_0,\epsilon_0} = T$ for some interior parameter values $\theta_0, \epsilon_0$ and that satisfies (15), (16), and*

$$\begin{array}{l} \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} z_i(\theta,\epsilon_0) \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i(\theta_0,\epsilon_0),\theta_0,\epsilon\right) \\ -\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} z_i(\theta_0,\epsilon) \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} T\left(z_i(\theta_0,\epsilon_0),\theta,\epsilon_0\right) \end{array} \begin{cases} > 0, & \text{if } i \in (i_a, i_b), \\ = 0, & \text{if } i \notin (i_a, i_b). \end{cases} \tag{A.44}$$

This lemma differs from Lemma 3 only in that the inequality in (A.44) points in the opposite direction to (23). If one modifies the construction in the proof of Lemma 3 only by assuming that $\eta_1$ is decreasing (rather than increasing) on $(\hat{z}_3, \hat{z}_4)$ and increasing (rather than decreasing) on $(\hat{z}_4, \hat{z}_5)$, so that $\eta_1(z) < 0$ (rather than $\eta_1(x) > 0$) on $(\hat{z}_3, \hat{z}_5)$, and correspondingly if one assumes that $\eta_2(z) > 0$ on $(\hat{z}_1, \hat{z}_2)$ (rather than $\eta_2(z) < 0$), then one flips the inequality in (23), and so attains (A.44). □

## A.13 Proof of Lemma 5

Assume that welfare weights $g$ are not structurally utilitarian. By Lemma 4, in this case, we may choose a well behaved family $\left(T^{\theta,\epsilon}\right)_{\theta\in\Theta,\epsilon\in E}$ satisfying (15)-(17). The construction of this family is presented in the proofs of Lemmas 3 and 4. Let us consider again the construction of $\left(T^{\theta,\epsilon}\right)$. First, by Corollary 3, since $g$ is not structurally utilitarian we can select a regular tax policy $T$ such that for some such that for some $i_a, i_b \in (0,1)$ with $i_a < i_b$, either condition (18) or (19) is satisfied. An examination of the construction of the proof of Corollary 3 shows that it is possible to select $T$ such that

$$T'\left(z_0(T)\right) \neq 0. \tag{A.45}$$

We did not previously assume property (A.45) but let us assume henceforth that (A.45) is satisfied. Next we a use tax policy $T$ and $i_a$ and $i_b$ with the above properties to construct a family of tax polices, $\left(T^{\theta,\epsilon}\right)$, as in the proof of Lemma 3, of the form $T^{\theta,\epsilon} = T + [\theta \times \mu_1] - [\zeta(\theta,\epsilon) \times \mu_2] + [\epsilon \times (\eta_1 + \eta_2)]$ (see (A.32)). The proof of Lemma 4 shows that such a family satisfies (15)-(17). It follows from their construction in the proof of Lemma 3 that the supports of the functions $\mu_1, \mu_2, \eta_1$, and $\eta_2$ are all contained in the set $[\hat{z}_1, +\infty)$, where $\hat{z}_1 = z_{i_1}(T)$ was defined in the beginning of the proof of Lemma 3. As $0 < i_1$, it follows from assumptions in Section V.A, that

$$z_0(T) < z_{i_1}(T) = \hat{z}_1. \tag{A.46}$$

It follows that

$$T^{\theta,\epsilon}(z) = T(z), \quad \forall z \in [0, \hat{z}_1], \forall \theta \in \Theta, \forall \epsilon \in E. \tag{A.47}$$

So $T^{\theta,\epsilon}(z) = T(z, \theta, \epsilon)$ does not depend on $\theta$ or $\epsilon$ for $z$ below $\hat{z}_1$. Recall that in the construction of $\left(T^{\theta,\epsilon}\right)$, we assumed that $\theta_0 = 0$ and $\epsilon_0 = 0$, so that, by (A.32), $T^{\theta_0,\epsilon_0} = T$.

**Lemma A.3** *There exists a family of tax reforms $\left(\Delta T^{\xi}\right)_{\xi \in \Xi}$, where $\Xi = \left[\underline{\xi}, \overline{\xi}\right]$ for real numbers $\underline{\xi}, \overline{\xi}$ satisfying $\underline{\xi} < 0 < \overline{\xi}$, and such that $\Delta T^0 \equiv 0$, the support of $\Delta T^{\xi}$ is contained in $[0, \hat{z}_1]$ for all $\xi \in \Xi$, the map $(z, \xi) \mapsto \Delta T^{\xi}(z)$ is smooth, and for some sets $\Theta' = \left[\underline{\theta}', \overline{\theta}'\right] \subseteq \Theta$, $E' = [\underline{\epsilon}', \overline{\epsilon}'] \subseteq E$, with $\underline{\theta}' < 0 < \overline{\theta}'$ and $\underline{\epsilon}' < 0 < \overline{\epsilon}'$,*

$$\int g_i\left(T^{\theta,\epsilon} + \Delta T^{\xi}\right) \left.\frac{\partial}{\partial \xi}\right|_{\xi=\xi'} \Delta T\left(z_i\left(T^{\theta,\epsilon} + \Delta T^{\xi'}\right), \xi\right) \mathrm{d}i = 0, \quad \forall \theta \in \Theta', \forall \epsilon \in E', \forall \xi' \in \Xi, \tag{A.48}$$

$$\left.\frac{\mathrm{d}}{\mathrm{d}\xi}\right|_{\xi=\xi'} R\left(T^{\theta,\epsilon} + \Delta T^{\xi}\right) \neq 0, \quad \forall \theta \in \Theta', \forall \epsilon \in E', \forall \xi' \in \Xi. \tag{A.49}$$

*where in (A.48) we use the notation $\Delta T(z_i, \xi) = \Delta T^{\xi}(z_i)$. Moreover, $\left(\Delta T^{\xi}\right)_{\xi \in \Xi}$ can be constructed so that $T^{\theta,\epsilon} + \Delta T^{\xi}$ is regular, for all $\xi \in \Xi, \theta \in \Theta'$, and $\epsilon \in E'$.*

To understand this lemma, first recall that $T^{\theta_0,\epsilon_0} = T$, and note that, by construction, all tax policies $T^{\theta,\epsilon}$ are equal to $T$ on the interval $[0, \hat{z}_1]$, which contains the support of all tax reforms $\Delta T^{\xi}$. Lemma A.3 says that the family of reforms $\left(\Delta T^{\xi}\right)$ is such that varying $\xi$ in $T^{\theta,\epsilon} + \Delta T^{\xi}$ has no effect on welfare according to welfare weights (see (A.48)), but does have an effect on revenue (see (A.49)). Obviously, if $T^{\theta,\epsilon}$ were an optimal tax policy, it would not be possible to do this. However note that $T$, which coincides with all policies $T^{\theta,\epsilon}$ at the bottom of the income distribution, is such that marginal tax rate at at the income $z_0(T)$ at the bottom of the income distribution is non-zero, and, moreover, since $T$ is regular, $z_0(T) > 0$ (see Section A.1.2), and hence, none of the tax policies $T^{\theta,\epsilon}$ are optimal. As shown by Saez and Stantcheva (2016), (see Section A.2 of their Online Appendix), at an optimal tax policy in the generalized social welfare weights framework, the marginal tax rate for the bottom earner is zero if the bottom earner has a positive income. Lemma A.3 is proven in Section B.1 of the Appendix.

So let us assume that a family $\left(\Delta T^{\xi}\right)$ with the properties in Lemma A.3 is chosen. Noting that $\Delta T^0 \equiv 0$, it follows from (A.49) and the implicit function theorem that there exists $\underline{\theta}'', \overline{\theta}'' \in \Theta'$ with $\underline{\theta}'' < 0 < \overline{\theta}''$ and $\underline{\epsilon}'', \overline{\epsilon}'' \in E'$ with $\underline{\epsilon}'' < 0 < \overline{\epsilon}''$ and a function $\hat{\xi} : \Theta'' \times E'' \to \Xi$, where $\Theta'' = \left[\underline{\theta}'', \overline{\theta}''\right]$ and $E'' = [\underline{\epsilon}'', \overline{\epsilon}'']$, satisfying:

$$\hat{\xi}(\theta_0, \epsilon_0) = 0, \tag{A.50}$$

$$R\left(T^{\theta,\epsilon} + \Delta T^{\hat{\xi}(\theta,\epsilon)}\right) = R\left(T^{\theta_0,\epsilon_0}\right), \quad \forall \theta \in \Theta', \forall \epsilon \in E'', \tag{A.51}$$

where, in (A.50), $\Delta T^{\hat{\xi}(\theta,\epsilon)}$ is $\Delta T^{\xi}$ evaluated at $\xi = \hat{\xi}(\theta, \epsilon)$. Because the other functions occurring in (A.51) are smooth, it follows that $\hat{\xi}(\theta,\epsilon)$ is smooth.

Define the doubly parameterized family of tax policies $\left(\hat{T}^{\theta,\epsilon}\right)_{\theta\in\Theta'',\epsilon\in E''}$ by

$$\hat{T}^{\theta,\epsilon} = T^{\theta,\epsilon} + \Delta T^{\hat{\xi}(\theta,\epsilon)}, \quad \forall\theta\in\Theta'', \forall\epsilon\in E''. \tag{A.52}$$

It follows from Lemma A.3, the fact that $\left(T^{\theta,\epsilon}\right)$ is well-behaved, the fact that $\hat{\xi}(\theta,\epsilon)$ is smooth, and Lemma C.2 that, if above $\underline{\theta}''$, $\overline{\theta}''$, $\underline{\epsilon}''$, and $\overline{\epsilon}''$ are selected sufficiently close to zero, then $\left(\hat{T}^{\theta,\epsilon}\right)_{\theta\in\Theta'',\epsilon\in E''}$ is well-behaved. The well-behavedness of $\left(\hat{T}^{\theta,\epsilon}\right)$ is elaborated in greater detail in Section A.14, and specifically Section A.14.2.

**Lemma A.4** $\left(\hat{T}^{\theta,\epsilon}\right)_{\theta\in\Theta'',\epsilon\in E''}$ *satisfies (15)-(17).*

It is straightforward to verify that $\left(\hat{T}^{\theta,\epsilon}\right)_{\theta\in\Theta'',\epsilon\in E''}$ inherits properties (15)-(17) from $\left(T^{\theta,\epsilon}\right)_{\theta\in\Theta,\epsilon\in E}$. The calculations verifying Lemma A.4 are in Section B.2. Moreover it follows from (A.52) and (A.51) that $\left(\hat{T}^{\theta,\epsilon}\right)_{\theta\in\Theta'',\epsilon\in E''}$ has constant revenue. Thus, we have constructed a family of tax policies with the desired properties, which completes the proof. □

### A.14 Well-behavedness of families $\left(T^{\theta,\epsilon}\right)$ and $\left(\hat{T}^{\theta,\epsilon}\right)$ in the proof of Theorem 3

This section explains why the families of tax policies $\left(T^{\theta,\epsilon}\right)$ and $\left(\hat{T}^{\theta,\epsilon}\right)$ constructed in the proof of Theorem 3 are well-behaved. Well-behavedness consists of conditions on agents' optimization problems when facing the tax policies as well as smoothness conditions. (See Section A.1.2.) At a high level, the reason that $\left(T^{\theta,\epsilon}\right)$ and $\left(\hat{T}^{\theta,\epsilon}\right)$ satisfy the smoothness conditions is that these tax policies are constructed by combining functions that are assumed to be smooth in ways that preserve smoothness. More specifically, smoothness follows because relevant functions are derived from the implicit function theorem applied to smooth functions, which preserves smoothness (see Theorem 1.37 on p. 30 of Warner (2013)) or from the fact that solution to a parameterized initial value problem (whose existence and uniqueness are guaranteed by the Picard-Lindelöff theorem) is smooth when the parameterized initial value problem is appropriately constructed out of smooth functions (see Corollary 4.1 on p. 101 of Hartman (1982)). I give a more detailed argument below.

#### A.14.1 The family $\left(T^{\theta,\epsilon}\right)$

In the proof of Lemma 3 in Section A.12.1, I wrote that if $\underline{\theta}, \overline{\theta}, \underline{\epsilon}$, and $\overline{\epsilon}$, with $\underline{\theta} < 0 < \overline{\theta}, \underline{\epsilon} < 0 < \overline{\epsilon}$ are all chosen sufficiently close to 0, then the family $\left(T^{\theta,\epsilon}\right)_{\theta\in[\underline{\theta},\overline{\theta}],\epsilon\in[\underline{\epsilon},\overline{\epsilon}]}$ is well-behaved. I now substantiate that claim. First, for easy reference, recall definitions (A.29) and (A.33):

$$T_*^{\theta,\epsilon} = T + \theta\mu_1 + \epsilon\left(\eta_1+\eta_2\right), \tag{A.53}$$

$$T^{\theta,\epsilon} = T_*^{\theta,\epsilon} - \zeta(\theta,\epsilon)\,\mu_2. \tag{A.54}$$

As stated in Section A.1.2, a (non-individualized) tax policy is regular if the tax policy is smooth in income, and, for each agent $i$, when facing the tax policy, there is a unique optimal income, and at this optimum, $i$'s income is non-negative and $i$'s second order condition holds with a strict inequality. Recall that the tax policy $T$ in the definition of $T_*^{\theta,\epsilon}$ was assumed to be regular. Now consider a tax policy of the form $T_*^{\theta,\epsilon} - \zeta\mu_2$, where $\zeta$ is a real number. If $\theta = \epsilon = \zeta = 0$, then $T_*^{\theta,\epsilon} - \zeta\mu_2 = T$. Since $\mu_1, \mu_2, \eta_1$, and $\eta_2$ are all assumed to be smooth in $z$, and $T_*^{\theta,\epsilon} - \zeta\mu_2$ varies smoothly in $(\theta, \epsilon, \zeta)$, it follows that if $\theta, \epsilon$, and $\zeta$, are sufficiently close to zero, then $T_*^{\theta,\epsilon} - \zeta\mu_2$ continues to be regular: for each agent $i$, the optimum continues to be unique and positive, the second order condition continues to hold with a strict inequality, and the tax policy continues to be smooth in income. To state this formally, Lemma C.2 implies that it is possible to choose $\theta^* > 0, \epsilon^* > 0, \zeta^* > 0$ sufficiently small that,

$$\text{for all } \theta, \epsilon, \zeta, \text{ if } |\theta| \leq \theta^*, |\epsilon| \leq \epsilon^*, |\zeta| \leq \zeta^*, \text{ then } T_*^{\theta,\epsilon} - \zeta\mu_2 \text{ is regular.} \tag{A.55}$$

Next, I establish that the function $\zeta(\theta, \epsilon)$ is smooth in its arguments. Define the functions $f_1, f_2 : I \times [-\theta^*, \theta^*] \times [-\zeta^*, \zeta^*] \times [-\epsilon^*, \epsilon^*] \to \mathbb{R}$ by

$$\begin{aligned} f_1(i, \theta, \zeta, \epsilon) &= g_i\left(T_*^{\theta,\epsilon} - \zeta\mu_2\right)\mu_1\left(z_i\left(T_*^{\theta,\epsilon} - \zeta\mu_2\right)\right), \\ f_2(i, \theta, \zeta, \epsilon) &= g_i\left(T_*^{\theta,\epsilon} - \zeta\mu_2\right)\mu_2\left(z_i\left(T_*^{\theta,\epsilon} - \zeta\mu_2\right)\right). \end{aligned}$$

For $\epsilon \in [-\epsilon^*, \epsilon^*]$, define $F_\epsilon(\theta, \zeta) : [-\theta^*, \theta^*] \times [-\zeta^*, \zeta^*] \to \mathbb{R}$ by

$$F_\epsilon(\theta, \zeta) = \frac{\int_0^1 f_1(i, \theta, \zeta, \epsilon)\, \mathrm{d}i}{\int_0^1 f_2(i, \theta, \zeta, \epsilon)\, \mathrm{d}i}. \tag{A.56}$$

Choose a real number $M$ satisfying $M \geq |F_\epsilon(\theta, \zeta)|$, for all $\epsilon \in [-\epsilon^*, \epsilon^*], \forall\theta \in [-\theta^*, \theta^*], \forall\zeta \in [-\zeta^*, \zeta^*]$. Let $\overline{\theta} = \min\left\{\theta^*, \frac{\zeta^*}{M}\right\}$. Since both the functions $f_1$ and $f_2$ are smooth in their arguments,[12] the integrals in both the numerator and the denominator of the right hand side of (A.56) are smooth in $(\theta, \epsilon, \zeta)$, and since the denominator is never equal to zero, it follows that $(\theta, \zeta, \epsilon) \mapsto F_\epsilon(\theta, \zeta)$ is smooth. The smoothness of $F_\epsilon(\theta, \zeta)$ implies that, in particular, $F_\epsilon(\theta, \zeta)$ is continuous in $\theta$ and uniformly Lipschitz continuous in $\zeta$. It now follows from the Picard-Lindelöf theorem (see Theorem 1.1 on p. 8 of Hartman (1982)) that, for all $\epsilon \in [-\epsilon^*, \epsilon^*]$, there exists a unique function $\zeta_\epsilon(\theta) : \left[-\overline{\theta}, \overline{\theta}\right] \to [-\zeta^*, \zeta^*]$ satisfying

$$\begin{aligned} \zeta_\epsilon(0) &= 0, \\ \frac{\mathrm{d}}{\mathrm{d}\theta}\zeta_\epsilon(\theta) &= F_\epsilon(\theta, \zeta_\epsilon(\theta)), \quad \forall\theta \in \left[-\overline{\theta}, \overline{\theta}\right]. \end{aligned} \tag{A.57}$$

[12] Note in particular that $z_i\left(T_*^{\theta,\epsilon} - \zeta\mu_2\right)$ is smooth in $(i, \theta, \zeta, \epsilon)$ because, as it has been established above that $T_*^{\theta,\epsilon} - \zeta\mu_2$ is regular, $z_i\left(T_*^{\theta,\epsilon} - \zeta\mu_2\right)$ is characterized via the implicit function theorem from the agent's first order condition, and the other functions featuring in this condition are smooth.

We can write $\zeta(\theta,\epsilon) = \zeta_\epsilon(\theta)$. Note that (A.57) is equivalent to (A.30)-(A.31). Corollary 4.1 on p. 101 of Hartman (1982) implies that, if, in a parameterized initial value problem, such as (A.57), the map $(\theta,\zeta,\epsilon) \mapsto F_\epsilon(\theta,\zeta)$ is smooth, then the parameterized solution $(\theta,\epsilon) \mapsto \zeta(\theta,\epsilon)$ is smooth as well, establishing the desired smoothness of $\zeta(\theta,\epsilon)$.

It now follows from the fact that the range of $\zeta(\theta,\epsilon)$, on $\left[-\underline{\theta},\overline{\theta}\right] \times [-\epsilon^*,\epsilon^*]$, is contained in $[-\zeta^*,\zeta^*]$ (see the preceding paragraph), combined with (A.54) and (A.55), that, for all $(\theta,\epsilon) \in \left[-\underline{\theta},\overline{\theta}\right] \times [-\epsilon^*,\epsilon^*]$, $T^{\theta,\epsilon}$ is regular. Next, it follows from the smoothness of $\zeta(\theta,\epsilon)$, established above, together with (A.53) and (A.54) (and the smoothness of the functions on the right hand side of (A.53)) that $(z,\theta,\epsilon) \mapsto T(z,\theta,\epsilon)$ is smooth. It now follows from Observation A.1, which says that, for non-individualized tax policies, regularity of each $T^{\theta,\epsilon}$ and smoothness of $(z,\theta,\epsilon) \mapsto T(z,\theta,\epsilon)$ is equivalent to well-behavedness, that, if we set $\underline{\theta} = -\overline{\theta}, \overline{\epsilon} = \epsilon^*, \underline{\epsilon} = -\epsilon^*$, then $\left(T^{\theta,\epsilon}\right)_{\theta\in\left[\underline{\theta},\overline{\theta}\right],\epsilon\in[\underline{\epsilon},\overline{\epsilon}]}$ is well-behaved.

### A.14.2 The family $\left(\hat{T}^{\theta,\epsilon}\right)$

This section shows that if $\underline{\theta}'', \overline{\theta}'', \underline{\epsilon}''$, and $\overline{\epsilon}''$, with $\underline{\theta}'' < 0 < \overline{\theta}'', \underline{\epsilon}'' < 0 < \overline{\epsilon}''$ are all chosen sufficiently close to 0, then the family $\left(T^{\theta,\epsilon}\right)_{\theta\in\left[\underline{\theta}'',\overline{\theta}''\right],\epsilon\in[\underline{\epsilon}'',\underline{\epsilon}'']}$ is well-behaved. Recall from (A.52) that

$$\hat{T}^{\theta,\epsilon} = T^{\theta,\epsilon} + \Delta T^{\hat{\xi}(\theta,\epsilon)}. \tag{A.58}$$

First I explain why $(z,\theta,\epsilon) \mapsto \Delta T^{\hat{\xi}(\theta,\epsilon)}(z)$ is smooth. Note that $(z,\theta,\epsilon) \mapsto \Delta T^{\hat{\xi}(\theta,\epsilon)}(z)$ is the composition of the maps $(z,\xi) \mapsto \Delta T^{\xi}(z)$ and $(\theta,\epsilon) \mapsto \hat{\xi}(\theta,\epsilon)$. The smoothness of $(z,\xi) \mapsto \Delta T^{\xi}(z)$ is established by Lemma A.3. (See in particular the discussion following (B.17) in Section B.1.3.) The function $(\theta,\epsilon) \mapsto \hat{\xi}(\theta,\epsilon)$ is defined by (A.50)-(A.51) via the implicit function theorem and the fact that it is smooth follows from the fact that the other functions in (A.51) are smooth.[13] This establishes the smoothness of $(z,\theta,\epsilon) \mapsto \Delta T^{\hat{\xi}(\theta,\epsilon)}(z)$. The regularity of $\hat{T}^{\theta,\epsilon}$, for each $\theta$ and $\epsilon$, given that $\underline{\theta}'', \overline{\theta}'', \underline{\epsilon}'', \overline{\epsilon}''$ are selected sufficiently close to zero, now follows from a similar argument as that for the regularity of $T_*^{\theta,\epsilon} - \zeta\mu_2$ in the previous section, again appealing to and Lemma C.2 and the fact that $\hat{T}^{\theta_0,\epsilon_0} = T$.[14]

The smoothness, established above, of $(z,\theta,\epsilon) \mapsto T(z,\theta,\epsilon)$ and $(z,\theta,\epsilon) \mapsto \Delta T^{\hat{\xi}(\theta,\epsilon)}(z)$, together with (A.58) now implies the smoothness of $(z,\theta,\epsilon) \mapsto \hat{T}^{\theta,\epsilon}(z)$, which, appealing again to Observation A.1, completes the argument that $\left(\hat{T}^{\theta,\epsilon}\right)$ is well-behaved.

[13] We have already established the smoothness of $T^{\theta,\epsilon}$ and $\Delta T^{\xi}$ above, and, noting that each agent's optimal income varies smoothly in response to smooth changes in tax policy, tax revenue also varies smoothly in response to such smooth changes.

[14] That $\hat{T}^{\theta_0,\epsilon_0} = T$ follows from (A.58) and the facts that $T^{\theta_0,\epsilon_0} = T$ and $\Delta T^{\hat{\xi}(\theta_0,\epsilon_0)} \equiv 0$; see Section A.13 for this last point.

### A.15 Proof of Proposition 5

In the poverty alleviation model of Section V.D, condition (24) is equivalent to condition (15). Given that $\kappa(\theta_0, \epsilon) = 0, \forall \epsilon$, it follows that $\frac{\partial}{\partial \epsilon}\big|_{\epsilon=\epsilon_0} \kappa(\theta_0, \epsilon) = 0$, which implies that the assumption that

$$\int g(\theta_0, \epsilon_0)\left[z_i(\theta_0, \epsilon_0) - \alpha\right] \mathrm{d}i = 0 \tag{A.59}$$

is equivalent to (16).

I now establish some facts that will be useful for establishing (17). First, using (24), we have

$$\begin{aligned}
\frac{\partial}{\partial \theta}\bigg|_{\theta=\theta_0} \kappa(\theta, \epsilon_0) &= \int \frac{g_i(\theta_0, \epsilon_0)}{\int g_j(\theta_0, \epsilon_0)\,\mathrm{d}j}\left[z_i(\theta_0, \epsilon_0) + f(z_i(\theta_0, \epsilon_0))\right] \mathrm{d}i \\
&= \int \frac{g_i(\theta_0, \epsilon_0)}{\int g_j(\theta_0, \epsilon_0)\,\mathrm{d}j} z_i(\theta_0, \epsilon_0)\,\mathrm{d}i + \int \frac{g_i(\theta_0, \epsilon_0)}{\int g_j(\theta_0, \epsilon_0)\,\mathrm{d}j} f(z_i(\theta_0, \epsilon_0))\,\mathrm{d}i \\
&= \alpha + \underbrace{\int \frac{g_i(\theta_0, \epsilon_0)}{\int g_j(\theta_0, \epsilon_0)\,\mathrm{d}j} f(z_i(\theta_0, \epsilon_0))\,\mathrm{d}i}_{\beta},
\end{aligned} \tag{A.60}$$

where the third equality follows from (A.59), and $\beta$ is a label for the last integral. It follows from the assumptions of Section V.D that $\beta > 0$. Let $\bar{i}$ be the unique agent satisfying $z_{\bar{i}}(\theta_0, \epsilon_0) = \bar{z}$. That such a $\bar{i}$ exists and is unique follows from Lemma C.1. It also follows from Lemma C.1 and the assumptions in Section V.D that all agents in the interval $[0, \bar{i})$ earn an income less than $\bar{z}$ when facing tax policy $T^{\theta_0, \epsilon_0}$. It follows from assumptions on $f$ in Section V.D that, for all incomes $z$ earned by agents in the interval $[0, \bar{i}]$, when facing tax policy $T^{\theta_0, \epsilon_0}$, $f(z) = 0$, so that for all $i \in [0, \bar{i}]$, $f'(z_i(\theta_0, \epsilon_0)) = 0$ and $f''(z_i(\theta_0, \epsilon_0)) = 0$. Taking this into account, and applying the implicit function theorem to the first order condition for agents' optimization problem when facing tax policy $T^{\theta_0, \epsilon_0}$, we have

$$\left.\begin{aligned}
\frac{\partial}{\partial \theta}\bigg|_{\theta=\theta_0} z_i(\theta, \epsilon_0) &= -\frac{1}{v_i''(z_i(\theta_0, \epsilon_0))}, \\
\frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=\epsilon_0} z_i(\theta_0, \epsilon) &= \frac{1}{v_i''(z_i(\theta_0, \epsilon_0))},
\end{aligned}\right. \quad \forall i \in [0, \bar{i}]. \tag{A.61}$$

Again, using the fact that $f'(z_i(\theta_0, \epsilon_0)) = 0, \forall i \in [0, \bar{i}]$ and (A.60), and the fact that $\frac{\partial}{\partial \epsilon}\big|_{\epsilon=\epsilon_0} \kappa(\theta_0, \epsilon) = 0$, we have

$$\left.\begin{aligned}
\frac{\partial}{\partial \theta}\bigg|_{\theta=\theta_0} T(z_i(\theta_0, \epsilon_0), \theta, \epsilon_0) &= z_i(\theta_0, \epsilon_0) - (\alpha + \beta), \\
\frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=\epsilon_0} T(z_i(\theta_0, \epsilon_0), \theta_0, \epsilon) &= -z_i(\theta_0, \epsilon_0) + \alpha,
\end{aligned}\right. \quad \forall i \in [0, \bar{i}]. \tag{A.62}$$

Using (A.61) and (A.62), we have that, for all $i \in [0, \bar{i}]$,

$$\begin{aligned}
&\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} z_i(\theta, \epsilon_0) \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} T(z_i(\theta_0, \epsilon_0), \theta_0, \epsilon) - \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} z_i(\theta_0, \epsilon) \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} T(z_i(\theta_0, \epsilon_0), \theta, \epsilon_0) \\
&= \left(-\frac{1}{v_i''(z_i(\theta_0, \epsilon_0))}\right)(-z_i(\theta_0, \epsilon_0) + \alpha) - \left(\frac{1}{v_i''(z_i(\theta_0, \epsilon_0))}\right)(z_i(\theta_0, \epsilon_0) - (\alpha + \beta)) \\
&= \frac{1}{v_i''(z_i(\theta_0, \epsilon_0))}\beta > 0.
\end{aligned}$$

Next recall the relationship between the variables $c_i, \hat{u}_i$ and $z_i$ from Section III: $c_i = \hat{u}_i + v_i(z_i)$. It follows that $\hat{g}_i(\hat{u}_i, z_i) = \tilde{g}(\hat{u}_i + v_i(z_i))$, and hence $\frac{\partial}{\partial z_i}\hat{g}_i(\hat{u}_i, z_i) = \tilde{g}'(\hat{u}_i + v_i(z_i))\, v_i'(z_i) = \tilde{g}'(c_i)\, v_i'(z_i)$. It follows from the above that:

$$\begin{aligned}
&\int \frac{\partial}{\partial z_i}\hat{g}_i\left(\hat{U}_i(\theta_0, \epsilon_0), z_i(\theta_0, \epsilon_0)\right)\left[\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} z_i(\theta, \epsilon_0) \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} T(z_i(\theta_0, \epsilon_0), \theta_0, \epsilon)\right. \\
&\qquad \left. - \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} z_i(\theta_0, \epsilon) \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} T(z_i(\theta_0, \epsilon_0), \theta, \epsilon_0)\right] \mathrm{d}i \qquad \text{(A.63)} \\
&= \beta \int_0^{\bar{i}} \tilde{g}'(c_i(\theta_0, \epsilon_0)) \frac{v_i'(z_i(\theta_0, \epsilon_0))}{v_i''(z_i(\theta_0, \epsilon_0))} \mathrm{d}i < 0,
\end{aligned}$$

where the upper bound of integration in the second integral follows from that fact that all $i \in (\bar{i}, 1]$ are above the poverty line when facing tax policy $T^{\theta_0, \epsilon_0}$ and hence $\tilde{g}'(c_i(\theta_0, \epsilon_0)) = 0$ for all such agents. The inequality follows from the fact that $v_i'(z_i) > 0$ and $v_i''(z_i) > 0$, for all $z_i$, $\tilde{g}'(c_i) \leq 0$, for all $c_i$, and, since a positive measure of agents in the interval $[0, \bar{i}]$ is beneath the poverty line at tax policy $T^{\theta_0, \epsilon_0}$, $\tilde{g}'(c_i(\theta_0, \epsilon_0)) < 0$ for a positive measure of agents in $[0, \bar{i}]$. It now follows from Lemma 2 that the family $(T^{\theta, \epsilon})$ in the poverty alleviation model of Section V.D satisfies (17). (Note that the proof of Lemma 2 also establishes that the first integral in (A.63) is equal to $\frac{\mathrm{d}}{\mathrm{d}\theta}\big|_{\theta=\theta_0} \int g_i(\theta, \epsilon_0) \frac{\partial}{\partial\epsilon}\big|_{\epsilon=\epsilon_0} T(z_i(\theta, \epsilon_0), \theta, \epsilon)\, \mathrm{d}i$, which shows that (A.63) justifies the inequality that was said to be the key calculation corresponding to (17) in Section V.D of the main text.) $\square$

## B Lemmas supporting Lemma 5

This section proves Lemmas A.3 and A.4, to which I appealed in the proof of Lemma 5.

### B.1 Proof of Lemma A.3

I begin the proof by establishing a pair of lemmas and then proceed to complete the proof.

#### B.1.1 Lemma B.1

The following lemma establishes the linearity of the function $f(\Delta T) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\big|_{\varepsilon=0} R(T + \varepsilon\Delta T)$.

**Lemma B.1** *Let $T$ be a regular tax policy. Let $\Delta T_1$ and $\Delta T_2$ be smooth tax reforms and let $r_1$ and $r_2$ be real numbers. Then* $\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\big|_{\varepsilon=0} R\left(T+\varepsilon\left(r_1 \Delta T_1+r_2 \Delta T_2\right)\right)=r_1 \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\big|_{\varepsilon=0} R\left(T+\varepsilon \Delta T_1\right)+$ $r_2 \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\big|_{\varepsilon=0} R\left(T+\varepsilon \Delta T_1\right)$.

Proof. Let $\Delta T_1$ and $\Delta T_2$ be smooth tax reforms and let $r_1$ and $r_2$ be real numbers. Then

$$
\begin{aligned}
& \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} z_i\left(T+\varepsilon\left(r_1 \Delta T^\gamma+r_2 \Delta T_2\right)\right) \\
& =-\frac{r_1 \Delta T_1^{\prime}\left(z_i(T)\right)+r_2 \Delta T_2^{\prime}\left(z_i(T)\right)}{T^{\prime \prime}\left(z_i(T)\right)+v_i^{\prime \prime}\left(z_i(T)\right)} \\
& =r_1\left(-\frac{\Delta T_1^{\prime}\left(z_i(T)\right)}{T^{\prime \prime}\left(z_i(T)\right)+v_i^{\prime \prime}\left(z_i(T)\right)}\right)+r_2\left(-\frac{\Delta T_2^{\prime}\left(z_i(T)\right)}{T^{\prime \prime}\left(z_i(T)\right)+v_i^{\prime \prime}\left(z_i(T)\right)}\right) \\
& =r_1 \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} z_i\left(T+\varepsilon \Delta T_1\right)+r_2 \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} z_i\left(T+\varepsilon \Delta T_2\right),
\end{aligned} \tag{B.1}
$$

where the first and third equalities follow from applying the implicit function theorem to the first order conditions of agent $i$'s optimization problem when facing tax policies $T+\varepsilon\left(r_1 \Delta T_1+r_2 \Delta T_2\right)$, $T+\varepsilon \Delta T_1$, and $T+\varepsilon \Delta T_2$, and $\Delta T_1^{\prime}(z)$ and $\Delta T_2^{\prime}(z)$, are, respectively, the derivatives of $\Delta T_1(z)$ and $\Delta T_2(z)$. Next, observe that

$$
\begin{aligned}
& \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R\left(T+\varepsilon\left(r_1 \Delta T_1+r_2 \Delta T_2\right)\right) \\
= & \int\left[r_1 \Delta T_1\left(z_i(T)\right)+r_2 \Delta T_2\left(z_i(T)\right)\right] \mathrm{d} i+\int T^{\prime}\left(z_i(T)\right) \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} z_i\left(T+\varepsilon\left(r_1 \Delta T_1+r_2 \Delta T_2\right)\right) \mathrm{d} i \\
= & \int\left[r_1 \Delta T_1\left(z_i(T)\right)+r_2 \Delta T_2\left(z_i(T)\right)\right] \mathrm{d} i \\
& +\int T^{\prime}\left(z_i(T)\right)\left[r_1 \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} z_i\left(T+\varepsilon \Delta T_1\right)+r_2 \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} z_i\left(T+\varepsilon \Delta T_2\right)\right] \mathrm{d} i \\
= & r_1\left[\int \Delta T_1\left(z_i(T)\right) \mathrm{d} i+\int T^{\prime}\left(z_i(T)\right) \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} z_i\left(T+\varepsilon \Delta T_1\right) \mathrm{d} i\right] \\
& +r_2\left[\int \Delta T_2\left(z_i(T)\right) \mathrm{d} i+\int T^{\prime}\left(z_i(T)\right) \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} z_i\left(T+\varepsilon \Delta T_2\right) \mathrm{d} i\right] \\
= & r_1 \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R\left(T+\varepsilon \Delta T_1\right)+r_2 \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R\left(T+\varepsilon \Delta T_2\right),
\end{aligned}
$$

where the second equality follows from (B.1). □

### B.1.2 Lemma B.2

Under the assumption that the lowest earned income is positive and the marginal tax rate at the bottom of the income distribution is nonzero, the following lemma establishes the existence of a desirable revenue-neutral tax reform in the generalized welfare weights framework. This mirrors a standard result in traditional (utilitarian) optimal tax theory. Saez and Stantcheva (2016) present a very closely related result, namely an optimal tax formula that, as they observe in their Online

Appendix, implies that when the lowest earned income is positive, the marginal tax rate at the bottom of the income distribution is zero in the generalized welfare weights framework, as in the traditional framework. Here, I prove a slightly stronger result than that there is a desirable reform when the bottom rate is nonzero: I also establish that the desirable reform can be assumed to have certain additional properties that are useful for our purposes.

**Lemma B.2** *Let $T$ be a regular tax policy (so that in particular $z_0(T) > 0$), and suppose that $T'(z_0(T)) \neq 0$. Let $z_*$ be such that $z_0(T) < z_* \leq z_1(T)$. Then there exists a desirable revenue neutral tax reform $\Delta T$ with support contained in $[0, z_*]$; formally, there exists a smooth tax reform $\Delta T$ with support contained in $[0, z_*]$ such that $\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\big|_{\varepsilon=0} R(T + \varepsilon\Delta T) = 0$ and $\int g_i(T)\,\Delta T(z_i(T))\,\mathrm{d}i < 0$. Moreover, there exist smooth tax reforms $\Delta T_1, \Delta T_2$, with supports contained in $[0, z_*]$ such that $\Delta T = \Delta T_1 - \Delta T_2, \Delta T_2(z) \geq 0, \forall z,$ and $\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\big|_{\varepsilon=0} R(T + \varepsilon\Delta T_2) \neq 0$.*

Recall that the set of agents is $I = [0, 1]$, $z_0(T)$ and $z_1(T)$ are the optimal responses to $T$ for agents 0 and 1 respectively. By the assumptions of Section V.A, $z_0(T)$ and $z_1(T)$ are respectively the bottom and top of the income distributions earned in response to $T$ (see also Lemma C.1).

I begin by stating some useful background facts and then proceed to prove the lemma.

#### B.1.2.1 Background facts

Choose a regular tax policy $T$, let $z_0 = z_0(T)$, and $z_1 = z_1(T)$. Define the function $\zeta : I \to Z$ by $\zeta(i) = z_i(T)$, and let $\iota = \zeta^{-1}$ be the inverse of $\zeta$ so that $\iota(z) = i$ if and only if $z_i(T) = z$. It follows from our assumptions in Section V.A that $\zeta(0) = z_0(T) > 0$ and $\zeta(i)$ is strictly increasing in $i$. Let $H$ be the cumulative distribution over incomes induced by tax policy $T$. Then, recalling that agents are uniformly distributed on the interval $I = [0, 1]$, it follows that $H(z) = 0$ for all $z \in Z$ such that $z < z_0$; $H(z) = \iota(z)$ for all $z \in Z$ with $z_0 \leq z \leq z_1$; and $H(z) = 1$ for all $z \in Z$ with $z_0 < z$. So if $h$ is the density corresponding to the cumulative distribution $H$, we have $h(z) = H'(z) = \iota'(z) = \frac{1}{\zeta'(\iota(z))}$ for all $z \in [z_0, z_1]$;[15] and $h(z) = 0$ for all $z \in Z$ with $z \notin [z_0, z_1]$. Observe, using a change of variables, that given a smooth tax reform $\Delta T$:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} R(T + \varepsilon\Delta T) &= \int_0^1 \Delta T(z_i(T))\,\mathrm{d}i + \int_0^1 T'(z_i(T))\,\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} z_i(T + \varepsilon\Delta T)\,\mathrm{d}i \\
&= \int_0^1 \Delta T(z_i(T))\,\mathrm{d}i - \int_0^1 T'(z_i(T))\,\frac{\Delta T'(z_i(T))}{T''(z_i(T)) + v_i''(z_i(T))}\mathrm{d}i \\
&= \int_{z_0}^{z_1} \Delta T(z)\,h(z)\,\mathrm{d}z - \int_{z_0}^{z_1} T'(z)\,\frac{\Delta T'(z)}{T''(z) + v_{\iota(z)}''(z)}h(z)\,\mathrm{d}z \\
&= \int_{z_0}^{z_1} \Delta T(z)\,h(z)\,\mathrm{d}z - \int_{z_0}^{z_1} \Delta T'(z)\,k_T(z)\,\mathrm{d}z
\end{aligned} \tag{B.2}$$

[15] Strictly speaking, $h(z)$ is, respectively, the right- and left-derivative of $H(z)$ at $z = z_0$ and $z = z_1$, and we have $h(z_0) = \frac{1}{\zeta'(\iota(z_0))} = \frac{1}{\zeta'(0)}$ and $h(z_1) = \frac{1}{\zeta'(1)}$.

where $\Delta T'$ is the derivative of $\Delta T$, $v''_{\iota(z)}(z)$ is $v''_i(z)$ evaluated at $i = \iota(z)$, and the second equality follows from applying the implicit function theorem to the first order condition for an agent's optimization problem when facing tax policy $T + \varepsilon \Delta T(z)$ around $\varepsilon = 0$ and

$$k_T(z) = \frac{T'(z) h(z)}{T''(z) + v''_{\iota(z)}(z)}, \quad \forall z \in [z_0, z_1]. \tag{B.3}$$

Note that we include the subscript $T$ in $k_T$ to express the dependence of $k_T$ on the tax policy $T$ through the terms $T'(z)$ and $T''(z)$. It follows from the fact that, for all regular $T$ and all $i \in I$, $\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}\Big|_{z_i = z_i(T)} u(z_i - T(z_i) - v_i(z_i)) < 0$ (see Section A.1.2), that the denominator on the right hand side of (B.3) is positive for all $z \in [z_0, z_1]$. Moreover the assumptions on $v_i$ and $y_i$ in Section V.A imply that $\zeta'(i) > 0, \forall i \in I$, and hence that $h(z) > 0, \forall z \in [z_0, z_1]$. It then follows that:

$$T'(z_0) \neq 0 \Rightarrow k_T(z_0) \neq 0. \tag{B.4}$$

#### B.1.2.2 Main argument

Again, let $z_0 = z_0(T)$ and $z_1 = z_1(T)$. Choose $z_*$ such that $z_0 < z_* \leq z_1$. Let $z_0 = z_0(T)$. As $T$ is regular, it follows that $z_0 > 0$ (see Section A.1.2). Choose $z_-$ so that $0 < z_- < z_0$. Consider a smooth tax reform $\Delta\hat{T}_1$ with $\Delta\hat{T}_1(z) = 2, \forall z \in [0, z_-], \Delta\hat{T}'_1(z) < 0, \forall z \in (z_-, z_*), \Delta\hat{T}_1(z_0) = 1$, and $\Delta\hat{T}_1(z) = 0, \forall z \in [z_*, +\infty)$. So the smooth tax reform $\Delta\hat{T}_1$ equal to 2 until $z = z_-$, at which point it falls, passing through $\Delta\hat{T}_1 = 1$ when $z = z_0$, and reaching $\Delta\hat{T}_1 = 0$ at $z = z_*$ and remains at zero thereafter.

For each $\gamma \in [1, +\infty)$, define $z_-^\gamma, z_*^\gamma$ by $\gamma(z_-^\gamma - z_0) + z_0 = z_-, \gamma(z_*^\gamma - z_0) + z_0 = z_*$. For $\gamma \geq 1$, we have $z_-^\gamma < z_0 < z_*^\gamma$; and $z_-^\gamma \uparrow z_0$ and $z_*^\gamma \downarrow z_0$ as $\gamma \uparrow +\infty$. Define $i^\gamma$ by the condition $z_{i^\gamma}(T) = z_*^\gamma$; that such an $i_\gamma$ exists follows from Lemma C.1. Using assumptions in Section V.A, we have $i^\gamma \downarrow 0$ as $\gamma \uparrow +\infty$.

Define the tax reform $\Delta T_1^\gamma$ by

$$\Delta T_1^\gamma(z) = \begin{cases} 2, & \text{if } z \in [0, z_-^\gamma], \\ \Delta\hat{T}_1(\gamma(z - z_0) + z_0), & \text{if } z \in (z_-^\gamma, z_*^\gamma), \\ 0, & \text{if } z \in [z_*^\gamma, +\infty). \end{cases} \tag{B.5}$$

Using the properties of $\Delta\hat{T}_1$, it is straightforward to verify that, for all $\gamma \geq 1$, $\Delta T_1^\gamma$ is a smooth function of $z$. So $\Delta T_1^\gamma$ is similar to $\Delta\hat{T}_1$, except that in the former $z_-^\gamma$ and $z_*^\gamma$ play the roles of $z_-$ and $z_*$ in the latter. For $\gamma > 1, \Delta T_1^\gamma$ falls more steeply than $\Delta\hat{T}_1$ near $z = z_0$. Observe that, for all $\gamma \geq 1$, $[0, z_*^\gamma]$ is the support of both $\Delta T_1^\gamma$, so that the support of $\Delta T_1^\gamma$ is contained in $[0, z_*]$.

**Lemma B.3** *Assume, as above, that $T'(z_0) \neq 0$. Then $\lim_{\gamma \to \infty} \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\big|_{\varepsilon=0} R(T + \varepsilon \Delta T_1^\gamma) \neq 0$.*

Lemma B.3 is proven in Section B.1.2.3.

Since we are assuming that $T'(z_0) \neq 0$, it follows from Lemma B.3 that there exists a tax reform $\Delta\hat{T}_2$ with support contained in $[0, z^*]$ such that $\Delta\hat{T}_2(z) \geq 0, \forall z \in Z$, and

$$\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R\left(T + \varepsilon\Delta\hat{T}_2\right) \neq 0. \tag{B.6}$$

In particular, we can choose $\Delta\hat{T}_2 = \Delta\hat{T}_1^{\gamma_0}$ for some sufficiently large $\gamma_0$. However, for our purposes, it is not important whether $\Delta\hat{T}_2 = \Delta\hat{T}_1^{\gamma_0}$ for some sufficiently large (fixed) $\gamma_0$; it matters only that is has the properties we have just ascribed to it.

It follows from Lemma B.1 and (B.6) that, for each $\gamma > 1$, there exists $r_\gamma$ such that

$$\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R\left(T + \varepsilon\left(\Delta T_1^\gamma - r_\gamma\Delta\hat{T}_2\right)\right) = 0. \tag{B.7}$$

It follows from Lemma B.1, (B.6), and (B.7) that

$$r_\gamma = \frac{\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R\left(T + \varepsilon\Delta T_1^\gamma\right)}{\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R\left(T + \varepsilon\Delta\hat{T}_2\right)}. \tag{B.8}$$

The negative of the marginal welfare effect of a small tax reform in direction $\Delta T_1^\gamma - r_\gamma\Delta\hat{T}_2$ is

$$W_\gamma = \int g_i(T)\,\Delta T^\gamma(z_i(T))\,\mathrm{d}i - r_\gamma \int g_i(T)\,\Delta\hat{T}_2(z_i(T))\,\mathrm{d}i.$$

Observe that because (i) the function $i \mapsto \Delta T^\gamma(z_i(T))$, whose domain is $[0,1]$, has support $[0, i^\gamma]$ and $i^\gamma \downarrow 0$ as $\gamma$ approaches infinity and (ii) $\Delta T^\gamma(z)$ is bounded between 0 and 2, for all $z$, it follows that $\int g_i(T)\,\Delta T^\gamma(z_i(T))\,\mathrm{d}i \to 0$ as $\gamma \to \infty$. Note that because $\Delta\hat{T}_2(z) \geq 0, \forall z \in Z$, and $\Delta\hat{T}_2$ satisfies (B.6), there must be a positive measure set of agents $i$ such that $\Delta\hat{T}_2(z_i(T)) > 0$. It follows that $\int g_i(T)\,\Delta\hat{T}_2(z_i(T))\,\mathrm{d}i > 0$. Lemma B.3 and (B.8) imply that $\lim_{\gamma\to\infty} r_\gamma \neq 0$.[16] It now follows from the results of the previous paragraph that if $\gamma$ is sufficiently large then $W_\gamma \neq 0$. Then choose such a sufficiently large $\gamma$. If $W_\gamma < 0$, then define $\Delta T_1 = \Delta T_1^\gamma$ and $\Delta T_2 = r_\gamma\Delta\hat{T}_2$; and if $W_\gamma > 0$, define $\Delta T_1 = -\Delta T_1^\gamma$ and $\Delta T_2 = -r_\gamma\Delta\hat{T}_2$. In either case define $\Delta T = \Delta T_1 - \Delta T_2$. In both cases, we have $\int g_i(T)\,\Delta T(z_i(T))\,\mathrm{d}i < 0$ and, appealing to Lemma B.1, (B.6), and (B.7), $\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R(T + \varepsilon\Delta T) = 0$ and $\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R(T + \varepsilon\Delta T_2) \neq 0$. Finally note the support of $\Delta T$, $\Delta T_1$, and $\Delta T_2$ are all contained in $[0, z_*]$. We have now established all of the properties required by Lemma B.2. □

[16] Observe that, from (B.2), $\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R\left(T + \varepsilon\Delta\hat{T}_2\right) = \int_{z_0}^{z_1} \Delta\hat{T}_2(z)\,h(z)\,\mathrm{d}z - \int_{z_0}^{z_1} \Delta T'(z)\,k_T(z)\,\mathrm{d}z$, which is finite, so the denominator in (B.8) is finite as well.

#### B.1.2.3 Proof of Lemma B.3.

It follows from (B.2) and the fact that the support of $\Delta T_1^\gamma$ is $[0, z_*^\gamma]$ that

$$\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R\left(T+\varepsilon \Delta T_1^\gamma\right)=\int_{z_0}^{z_*^\gamma} \Delta T_1^\gamma(z)\, h(z)\, \mathrm{d} z-\int_{z_0}^{z_*^\gamma} \frac{\mathrm{d}}{\mathrm{d} z} \Delta T_1^\gamma(z)\, k_T(z)\, \mathrm{d} z, \tag{B.9}$$

where $\frac{\mathrm{d}}{\mathrm{d}z}\Delta T_1^\gamma(z)$ is the derivative of $\Delta T_1^\gamma(z)$. Because $z_*^\gamma \downarrow z_0$ as $\gamma \to \infty$ and $\Delta T^\gamma(z)$ is bounded between 0 and 2 for all $z$,

$$\lim_{\gamma \rightarrow \infty} \int_{z_0}^{z_*^\gamma} \Delta T_1^\gamma(z)\, h(z)\, \mathrm{d} z=0. \tag{B.10}$$

Since $\frac{\mathrm{d}}{\mathrm{d}z}\Delta T_1^\gamma(z) \leq 0, \forall z \in Z$, it follows from the preceding that, if $\gamma$ is sufficiently large, we have:

$$\begin{aligned}\left(\max_{z \in\left[z_0, z_*^\gamma\right]} k_T(z)\right) \times \int_{z_0}^{z_*^\gamma} \frac{\mathrm{d}}{\mathrm{d} z} \Delta T_1^\gamma(z)\, \mathrm{d} z &\leq \int_{z_0}^{z_*^\gamma} \frac{\mathrm{d}}{\mathrm{d} z} \Delta T_1^\gamma(z)\, k_T(z)\, \mathrm{d} z \\ &\leq\left(\min_{z \in\left[z_0, z_*^\gamma\right]} k_T(z)\right) \times \int_{z_0}^{z_*^\gamma} \frac{\mathrm{d}}{\mathrm{d} z} \Delta T_1^\gamma(z)\, \mathrm{d} z.\end{aligned} \tag{B.11}$$

Next observe that

$$\int_{z_0}^{z_*^\gamma} \frac{\mathrm{d}}{\mathrm{d} z} \Delta T_1^\gamma(z)\, \mathrm{d} z=\int_{z_0}^{z_*^\gamma} \gamma \Delta \hat{T}_1^{\prime}\left(\gamma\left[z-z_0\right]+z_0\right) \mathrm{d} z=\int_{z_0}^{z_*} \Delta \hat{T}_1^{\prime}(\tilde{z})\, \mathrm{d} \tilde{z}=\Delta \hat{T}_1\left(z_*\right)-\hat{T}_1\left(z_0\right)=-1, \tag{B.12}$$

where $\Delta\hat{T}_1'(\tilde{z})$ is the derivative of $\Delta\hat{T}_1(\tilde{z})$ and the second equality uses the change of variables $z \mapsto \tilde{z} = \gamma[z - z_0] + z_0$. Next observe that as $k$ is smooth and $z^\gamma \downarrow z_0$ and $\gamma \to \infty$,

$$\lim_{\gamma \rightarrow \infty} \max_{z \in\left[z_0, z^\gamma\right]} k_T(z)=k_T\left(z_0\right) \text { and } \lim_{\gamma \rightarrow \infty} \min_{z \in\left[z_0, z^\gamma\right]} k_T(z)=k_T\left(z_0\right). \tag{B.13}$$

It follows from (B.9), (B.10), (B.11), (B.12), and (B.13) that

$$\lim_{\gamma \rightarrow \infty} \left.\frac{\mathrm{d}}{\mathrm{d} \varepsilon}\right|_{\varepsilon=0} R\left(T+\varepsilon \Delta T_1^\gamma\right)=k_T\left(z_0\right).$$

It follows from (B.4) and the assumption that $T'(z_0) \neq 0$ that $k_T(z_0) \neq 0$, which completes the proof. □

### B.1.3 Completion of the proof of Lemma A.3

Lemma B.2 established that, under certain conditions, there exists a desirable revenue neutral tax reform. It is intuitive that, starting from such a reform, and adding an appropriate lumpsum tax, one can attain a welfare-neutral reform that raises revenue. Lemma A.3 establishes the existence of something similar: a parameterized family of tax reforms (at a subset of tax policies in the family $\left(T^{\theta,\epsilon}\right)$) such that varying the parameter affects revenue but is socially indifferent according to welfare weights. This family is not constructed by modifying a desirable revenue neutral reform via a lumpsum tax, which would affect all taxpayers, but rather by local change in taxes that affects only taxpayers at the bottom of the income distribution.

I now use Lemmas B.1 and B.2 to prove Lemma A.3. Recall from Section A.13 that $T^{\theta_0,\epsilon_0}=T$ for a regular tax policy $T$ satisfying $T'\left(z_0\left(T\right)\right)\neq 0$. Recall also from Section A.13 that $\hat{z}_1=z_{i_1}\left(T\right)$, and moreover, $z_0\left(T\right)<\hat{z}_1<z_1\left(T\right)$. So letting $\hat{z}_1$ play the role of $z_*$ in Lemma B.2, there exist tax reforms $\Delta T_1,\Delta T_2$, and $\Delta T$, all with supports contained in $[0,\hat{z}_1]$ and satisfying the properties in Lemma B.2 in relation to the tax policy $T=T^{\theta_0,\epsilon_0}$. Define the function[17]

$$F\left(\xi,r\right)=\frac{\int g_i\left(T+\xi\Delta T_1-r\Delta T_2\right)\Delta T_1\left(z_i\left(T+\xi\Delta T_1-r\Delta T_2\right)\right)\mathrm{d}i}{\int g_i\left(T+\xi\Delta T_1-r\Delta T_2\right)\Delta T_2\left(z_i\left(T+\xi\Delta T_1-r\Delta T_2\right)\right)\mathrm{d}i}. \tag{B.14}$$

It follows from the properties stated in Lemma B.2 (which apply to $T,\Delta T_1$ and $\Delta T_2$) and the smoothness of the relevant functions that if $\xi$ and $r$ are sufficiently close to zero, then the denominator in the above expression is nonzero.[18] Note that $F$ is smooth in its arguments. It follows from the Picard-Lindelöf theorem that there exist real numbers $\underline{\xi},\overline{\xi}$ with $\underline{\xi}<0<\overline{\xi}$ and a function $s:\left[\underline{\xi},\overline{\xi}\right]\to\mathbb{R}$ satisfying

$$s\left(0\right)=0, \tag{B.15}$$

$$s'\left(\xi\right)=F\left(\xi,s\left(\xi\right)\right),\quad\forall\xi\in\Xi, \tag{B.16}$$

where $\Xi=\left[\underline{\xi},\overline{\xi}\right]$. Define the family of tax reforms $\left(\Delta T^{\xi}\right)_{\xi\in\Xi}$ by the condition

$$\Delta T^{\xi}=\xi\Delta T_1-s\left(\xi\right)\Delta T_2,\quad\forall\xi\in\Xi. \tag{B.17}$$

Observe that $\Delta T^0\equiv 0$ and, for all $\xi\in\Xi$, the support of $\Delta T^{\xi}$ is contained in $[0,\hat{z}_1]=[0,z_*]$ because the supports of $\Delta T_1$ and $\Delta T_2$ are contained in $[0,\hat{z}_1]$. It follows from the smoothness of the function $F\left(\xi,r\right)$, and Corollary 4.1 on p. 101 of Hartman (1982) that the function $s\left(\xi\right)$ is smooth, and hence also, given the smoothness of $\Delta T_1$ and $\Delta T_2$, that the map $\left(z,\xi\right)\mapsto\Delta T^{\xi}\left(z\right)$ is smooth. Lemma C.2 implies that it is possible to choose $\underline{\xi}$ and $\overline{\xi}$, and also $\underline{\theta}',\overline{\theta}'\in\Theta,\underline{\epsilon}',\overline{\epsilon}'\in E$ with $\underline{\theta}'<0<\overline{\theta}',\underline{\epsilon}'<0<\overline{\epsilon}'$ so that $T^{\theta,\epsilon}+\Delta T^{\xi}=T+\left[\theta\times\mu_1\right]-\left[\zeta\left(\theta,\epsilon\right)\times\mu_2\right]+\left[\epsilon\times\left(\eta_1+\eta_2\right)\right]+\Delta T^{\xi}$ is regular, for all $\theta\in\Theta'=\left[\underline{\theta}',\overline{\theta}'\right],\epsilon\in E'=\left[\underline{\epsilon}',\overline{\epsilon}'\right]$, and $\xi\in\Xi=\left[\underline{\xi},\overline{\xi}\right]$. So let us assume that $\underline{\xi},\overline{\xi},\underline{\theta}',\overline{\theta}',\underline{\epsilon}'$, and $\overline{\epsilon}'$ are so chosen.

[17] It follows from Lemma C.2 that if if $\xi$ and $r$ are sufficiently close to 0, then $T+\xi\Delta T_1-r\Delta T_2$ is regular, and hence $z_i\left(T+\xi\Delta T_1-r\Delta T_2\right)$ is uniquely defined, and so $g_i\left(T+\xi\Delta T_1-r\Delta T_2\right)$ is also uniquely defined.

[18] In particular, the facts that $\Delta T_2\left(z\right)\geq 0,\forall z$, and $\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\big|_{\varepsilon=0}R\left(T+\varepsilon\Delta T_2\right)\neq 0$ imply that there exists a positive measure of agents $i$ such that $\Delta T_2\left(z_i\left(T\right)\right)>0$, hence, invoking again $\Delta T_2\left(z\right)\geq 0,\forall z$, it follows that from the fact that welfare weights are positive $\int g_i\left(T\right)\Delta T_2\left(z_i\left(T\right)\right)\mathrm{d}i>0$. That the denominator of (B.14) is positive now follows from our smoothness assumptions.

Next observe that

$$\frac{\partial}{\partial \xi} \Delta T(z, \xi) = \Delta T_1(z) - s'(\xi) \Delta T_2(z), \quad \forall \xi \in \Xi, \forall z \in Z, \tag{B.18}$$

where $\Delta T(z, \xi) = \Delta T^{\xi}(z)$. Recalling that $T^{\theta_0, \epsilon_0} = T$, and using (B.14), (B.17), and (B.18), it follows that (B.16) is equivalent to

$$\int g_i\left(T + \Delta T^{\xi}\right) \left.\frac{\partial}{\partial \xi}\right|_{\xi=\xi'} \Delta T\left(z_i\left(T + \Delta T^{\xi'}\right), \xi\right) \mathrm{d}i = 0, \quad \forall \xi' \in \Xi. \tag{B.19}$$

It follows that the family $\left(\Delta T^{\xi}\right)$ satisfies (B.19).

Taking the derivative of the relation (B.18) with respect to $z$ yeilds:

$$\frac{\partial^2}{\partial \xi \partial z} \Delta T(z, \xi) = \Delta T_1'(z) - s'(\xi) \Delta T_2'(z), \quad \forall \xi \in \Xi, \forall z \in Z, \tag{B.20}$$

Now consider the tax reform $\Delta T_1 - s'(0) \Delta T_2$. This is just the tax reform $\Delta T_1 - r \Delta T_2$ in the special case in which $r = s'(0)$. When facing the tax policies $T + \varepsilon(\Delta T_1 - s'(0) \Delta T_2)$ and $T + \Delta T^{\xi}$, agent $i$ faces, respectively, optimization problems $\max_{z_i} [z_i - T(z_i) - \varepsilon(\Delta T_1(z_i) - s'(0) \Delta T_2(z_i)) - v_i(z_i)]$ and $\max_{z_i} \left[z_i - T(z_i) - \Delta T^{\xi}(z_i) - v_i(z_i)\right]$. Note that, because $T$ is regular, when $\varepsilon$ and $\xi$ are sufficiently small, all agents select an interior income (see also Lemma C.2). Applying the implicit function theorem to the agent's first order conditions for these two problems, we have:

$$\begin{aligned} &\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} z_i\left(T + \varepsilon\left(\Delta T_1 - s'(0) \Delta T_2\right)\right) = -\frac{\Delta T_1'(z_i(T)) - s'(0) \Delta T_2'(z_i(T))}{T''(z_i(T)) + v_i''(z_i(T))} \\ &= -\frac{\left.\frac{\partial^2}{\partial \xi \partial z}\right|_{\xi=0, z=z_i(T)} \Delta T(z, \xi)}{T''(z_i(T)) + v_i''(z_i(T))} = \left.\frac{\mathrm{d}}{\mathrm{d}\xi}\right|_{\xi=0} z_i\left(T + \Delta T^{\xi}\right), \end{aligned} \tag{B.21}$$

where the second equality follows from (B.20). This, in turn, implies that

$$\begin{aligned} &\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R\left(T + \varepsilon\left(\Delta T_1 - s'(0) \Delta T_2\right)\right) \\ &= \int \left[\Delta T_1(z_i(T)) - s'(0) \Delta T_2(z_i(T))\right] \mathrm{d}i + \int T'(z_i(T)) \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} z_i\left(T + \varepsilon\left(\Delta T_1 - s'(0) \Delta T_2\right)\right) \mathrm{d}i \\ &= \int \left.\frac{\partial}{\partial \xi}\right|_{\xi=0} \Delta T(z_i(T), \xi) \mathrm{d}i + \int T'(z_i(T)) \left.\frac{\mathrm{d}}{\mathrm{d}\xi}\right|_{\xi=0} z_i\left(T + \Delta T^{\xi}\right) \mathrm{d}i \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}\xi}\right|_{\xi=0} R\left(T + \Delta T^{\xi}\right), \end{aligned} \tag{B.22}$$

where the second equality uses (B.18) and (B.21).

Next observe that, by the properties implied by Lemma B.2, $0 > \int g_i(T) \Delta T(z_i(T)) \mathrm{d}i = \int g_i(T) \Delta T_1(z_i(T)) \mathrm{d}i - \int g_i(T) \Delta T_2(z_i(T)) \mathrm{d}i$. So, since, again by the properties in Lemma

B.2, $\int g_i(T)\,\Delta T_2(z_i(T))\,\mathrm{d}i > 0$ (see footnote 18 of the Appendix), it follows that $F(0,0) = \frac{\int g_i(T)\Delta T_1(z_i(T))\mathrm{d}i}{\int g_i(T)\Delta T_2(z_i(T))\mathrm{d}i} < 1$. So, by (B.14) and (B.16), $s'(0) < 1$. It follows from Lemma B.1 and the properties of Lemma B.2 that

$$\begin{aligned}
0 = & \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R(T+\varepsilon\Delta T) = \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R(T+\varepsilon(\Delta T_1 - \Delta T_2)) \\
= & \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R(T+\varepsilon\Delta T_1) - \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R(T+\varepsilon\Delta T_2) \\
\neq & \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R(T+\varepsilon\Delta T_1) - s'(0)\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R(T+\varepsilon\Delta T_2) \\
= & \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R\left(T+\varepsilon\left(\Delta T_1 - s'(0)\,\Delta T_2\right)\right) \\
= & \left.\frac{\mathrm{d}}{\mathrm{d}\xi}\right|_{\xi=0} R\left(T+\Delta T^{\xi}\right),
\end{aligned}$$

where the non-equality $\neq$ in the above derivation follows from the facts that $s'(0) \neq 1$ and $\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R(T+\varepsilon\Delta T_2) \neq 0$ (see Lemma B.2 for the latter). So, to summarize, $\left.\frac{\mathrm{d}}{\mathrm{d}\xi}\right|_{\xi=0} R\left(T+\Delta T^{\xi}\right) \neq 0$. By our smoothness assumptions, if $\underline{\xi}$ and $\overline{\xi}$ in $\Xi = \left[\underline{\xi}, \overline{\xi}\right]$ are selected so as to be sufficiently close to 0,

$$\left.\frac{\mathrm{d}}{\mathrm{d}\xi}\right|_{\xi=\xi'} R\left(T+\Delta T^{\xi}\right) \neq 0, \quad \forall \xi' \in \Xi. \tag{B.23}$$

Let us assume that $\underline{\xi}$ and $\overline{\xi}$ are so chosen.

Using the facts that, by the construction of $\left(T^{\theta,\epsilon}\right)$, $T^{\theta,\epsilon}(z) + \Delta T^{\xi}(z) = T(z) + \Delta T^{\xi}(z), \forall z \in [0,\hat{z}_1], \forall \theta \in \Theta', \forall \epsilon \in E', \forall \xi \in \Xi$, and that, for all $\xi \in \Xi$, the support of $\Delta T^{\xi}$ is contained in $[0,\hat{z}_1]$ (see in particular Lemma B.4 and (B.26)-(B.27) of Lemma B.5[19] and note that $T^{\theta_0,\epsilon_0} = T$), it follows that, for all $\theta \in \Theta'$, for all $\epsilon \in E'$, and for all $\xi' \in \Xi$,

$$\begin{aligned}
& \int g_i\left(T+\Delta T^{\xi}\right) \left.\frac{\partial}{\partial\xi}\right|_{\xi=\xi'} \Delta T\left(z_i\left(T+\Delta T^{\xi'}\right),\xi\right)\mathrm{d}i \\
= & \int g_i\left(T^{\theta,\epsilon}+\Delta T^{\xi}\right) \left.\frac{\partial}{\partial\xi}\right|_{\xi=\xi'} \Delta T\left(z_i\left(T^{\theta,\epsilon}+\Delta T^{\xi'}\right),\xi\right)\mathrm{d}i,
\end{aligned} \tag{B.24}$$

$$\left.\frac{\mathrm{d}}{\mathrm{d}\xi}\right|_{\xi=\xi'} R\left(T+\Delta T^{\xi}\right) = \left.\frac{\mathrm{d}}{\mathrm{d}\xi}\right|_{\xi=\xi'} R\left(T^{\theta,\epsilon}+\Delta T^{\xi}\right) \tag{B.25}$$

Conditions (A.48) and (A.49) follow from (B.19), (B.23), (B.24), and (B.25). We have now proven all the properties required by Lemma A.3. □

[19] The proof of Lemmas B.4 and B.5 depend on the fact that, for all $\xi \in \Xi$, the support of $\Delta T^{\xi}$ is contained in $[0,\hat{z}_1]$, but not the more detailed properties established in the current lemma, Lemma A.3.

### B.2 Proof of Lemma A.4

I now prove a pair of supporting lemmas, and then proceed to verify the properties required by Lemma A.4. The following sections appeal to the notation and terminology used in Section A.13.

#### B.2.1 Supporting lemmas

Here I establish a pair of lemmas that collect some properties that follow fairly immediately from above definitions.

**Lemma B.4** *For all $\theta \in \Theta', \epsilon \in E'$, and $\xi \in \Xi$,*

$$z_i \left(T^{\theta,\epsilon} + \Delta T^{\xi}\right) \begin{cases} \in [0, \hat{z}_1], & \text{if } i \in [0, i_1], \\ \in [\hat{z}_1, +\infty), & \text{if } i \in [i_1, 1]. \end{cases}$$

**Lemma B.5** *For all $i \in [0, i_1], \theta \in \Theta', \epsilon \in E'$, and $\xi \in \Xi$,*

$$z_i \left(T^{\theta_0,\epsilon_0} + \Delta T^{\xi}\right) = z_i \left(T^{\theta,\epsilon} + \Delta T^{\xi}\right), \tag{B.26}$$

$$g_i \left(T^{\theta_0,\epsilon_0} + \Delta T^{\xi}\right) = g_i \left(T^{\theta,\epsilon} + \Delta T^{\xi}\right), \tag{B.27}$$

*For all $i \in [i_1, 1], \theta \in \Theta''$, and $\epsilon \in E''$,*

$$z_i \left(\hat{T}^{\theta,\epsilon}\right) = z_i \left(T^{\theta,\epsilon}\right), \tag{B.28}$$

$$g_i \left(\hat{T}^{\theta,\epsilon}\right) = g_i \left(T^{\theta,\epsilon}\right). \tag{B.29}$$

I now proceed to prove both lemmas. It is useful to state a pair of facts, which follow from, respectively, the construction of $T^{\theta,\epsilon}$ (Fact B.1) and the characterization of regular tax policies in Section A.1.2 (Fact B.2). For Fact B.2 and the remainder of this section, it is also convenient to introduce the following notation: For any tax policy $T$ and agent $i$ and income $z_i$, let $U_i^T(z_i) = u(z_i - T(z_i) - v_i(z_i))$, be $i$'s utility when facing tax policy $T$ and choosing income level $z_i$.

**Fact B.1** *$T^{\theta,\epsilon}(z)$ does not depend on $\theta$ and $\epsilon$ when $z \in [0, \hat{z}_1]$; that is $T^{\theta,\epsilon}(z) = T^{\theta_0,\epsilon_0}(z), \forall z \in [0, \hat{z}_1], \forall \theta \in \Theta, \forall \epsilon \in E$.*

**Fact B.2** *For all regular tax policies $T$, and for all agents $i$, there exists a unique optimal income $z_i(T)$ for $i$ when facing $T$ and $z_i(T)$ is characterized by $i$'s first order condition in the sense that if $\frac{\mathrm{d}}{\mathrm{d}z} U_i^T(z) = 0$, then $z = z_i(T)$.*

To simplify notation, I write $\bar{T}^{\theta,\epsilon,\xi} = T^{\theta,\epsilon} + \Delta T^{\xi}$. Fix some $\theta' \in \Theta', \epsilon' \in E'$, and $\xi' \in \Xi$. Recall that $\hat{z}_1 \in \left(z_0\left(T^{\theta_0,\epsilon_0}\right), z_1\left(T^{\theta_0,\epsilon_0}\right)\right)$, and that $i_1$ is the unique agent in $I$ such that $z_{i_1}\left(T^{\theta_0,\epsilon_0}\right) = \hat{z}_1$. Let $I_0 := [0, i_1)$ and $I_1 := (i_1, 1]$.

By Fact B.1 and because $T^{\theta_0,\epsilon_0}$ is smooth, $\frac{\mathrm{d}}{\mathrm{d}z}T^{\theta_0,\epsilon_0}(\hat{z}_1) = \frac{\mathrm{d}}{\mathrm{d}z}T^{\theta',\epsilon'}(\hat{z}_1)$. So, because $\hat{z}_1 = z_{i_1}\left(T^{\theta_0,\epsilon_0}\right)$,

$$\frac{\mathrm{d}}{\mathrm{d}z}U_{i_1}^{T^{\theta',\epsilon'}}(\hat{z}_1) = \frac{\mathrm{d}}{\mathrm{d}z}U_{i_1}^{T^{\theta_0,\epsilon_0}}(\hat{z}_1) = 0. \tag{B.30}$$

Again, by Fact B.1, and the fact that the support of $\Delta T^{\xi}$ is contained in $[0, \hat{z}_1]$, it follows that

$$\forall z \leq \hat{z}_1, \quad \bar{T}^{\theta',\epsilon',\xi'}(z) = \bar{T}^{\theta_0,\epsilon_0,\xi'}(z), \tag{B.31}$$

$$\forall z \geq \hat{z}_1, \quad \bar{T}^{\theta',\epsilon',\xi'}(z) = T^{\theta',\epsilon'}(z). \tag{B.32}$$

Using (B.32) and the smoothness of the tax policies $\bar{T}^{\theta',\epsilon',\xi'}$ and $T^{\theta',\epsilon'}$, it follows that

$$\frac{\mathrm{d}}{\mathrm{d}z}U_{i_1}^{T^{\theta',\epsilon'}}(\hat{z}_1) = \frac{\mathrm{d}}{\mathrm{d}z}U_{i_1}^{\bar{T}^{\theta',\epsilon',\xi'}}(\hat{z}_1). \tag{B.33}$$

It follows from Fact B.2, (B.33), and (B.30), and the fact that $T^{\theta_0,\epsilon_0}$ and $\bar{T}^{\theta',\epsilon',\xi'}$ are regular (the latter was established in Section B.1.3) that

$$\hat{z}_1 = z_{i_1}\left(T^{\theta_0,\epsilon_0}\right) = z_{i_1}\left(\bar{T}^{\theta',\epsilon',\xi'}\right). \tag{B.34}$$

Because $\bar{T}^{\theta',\epsilon',\xi'}$ is regular, (B.34) and the fact that, for all regular tax policies $T$, the map $i \mapsto z_i(T)$ is strictly increasing in $i$ (see Lemma C.1) together establish Lemma B.4.

It follows from (B.31) and Lemma B.4 that, for all $i \in [0, i_1]$, $\frac{\mathrm{d}}{\mathrm{d}z}U_i^{\bar{T}^{\theta',\epsilon',\xi'}}\left(z_i\left(\bar{T}^{\theta_0,\epsilon_0,\xi'}\right)\right) = \frac{\mathrm{d}}{\mathrm{d}z}U_i^{\bar{T}^{\theta_0,\epsilon_0,\xi'}}\left(z_i\left(\bar{T}^{\theta_0,\epsilon_0,\xi'}\right)\right) = 0$. So, using the fact that $\bar{T}^{\theta',\epsilon',\xi'}$ and $\bar{T}^{\theta_0,\epsilon_0,\xi'}$ are regular, it follows from Fact B.2 that (B.26) holds. Similarly, it follows from (B.32) and Lemma B.4[20] that, for all $i \in [i_1, 1]$, $\frac{\mathrm{d}}{\mathrm{d}z}U_i^{\bar{T}^{\theta',\epsilon',\xi'}}\left(z_i\left(T^{\theta',\epsilon'}\right)\right) = \frac{\mathrm{d}}{\mathrm{d}z}U_i^{T^{\theta',\epsilon'}}\left(z_i\left(T^{\theta',\epsilon'}\right)\right) = 0$. So, using the fact that $\bar{T}^{\theta',\epsilon',\xi'}$ and $\bar{T}^{\theta',\epsilon'}$ are regular, it follows from Fact B.2 that (B.28) holds. If follows immediately from the definition of $g_i(T)$ (see Sections I.A-I.B), (B.26), and (B.28) that (B.27) and (B.29) hold. We have now established Lemma B.5. □

[20] Observe that $\Delta T^0 \equiv 0$, so that, setting $\xi = 0$, Lemma B.4 implies that, when $i \in [i_1, 1]$, $z_i\left(T^{\theta,\epsilon}\right) \geq \hat{z}_1$.

### B.2.2 Verification of properties required by Lemma A.4

I now proceed with the proof of Lemma A.4. Choose $\epsilon \in (\underline{\epsilon}'', \overline{\epsilon}'')$ and $\theta' \in \left(\underline{\theta}'', \overline{\theta}''\right)$. We have:

$$\begin{aligned}
&\int_0^1 g_i\left(\hat{T}^{\theta',\epsilon}\right) \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} \hat{T}\left(z_i\left(\hat{T}^{\theta',\epsilon}\right), \theta, \epsilon\right) \mathrm{d}i \\
&= \underbrace{\int_0^{i_1} g_i\left(\hat{T}^{\theta',\epsilon}\right) \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} \hat{T}\left(z_i\left(\hat{T}^{\theta',\epsilon}\right), \theta, \epsilon\right) \mathrm{d}i}_{A} \\
&\quad + \underbrace{\int_{i_1}^1 g_i\left(\hat{T}^{\theta',\epsilon}\right) \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} \hat{T}\left(z_i\left(\hat{T}^{\theta',\epsilon}\right), \theta, \epsilon\right) \mathrm{d}i}_{B},
\end{aligned} \tag{B.35}$$

where $A$ and $B$ are simply labels for the two integrals on the right-hand side, and we use the notation $\hat{T}(z_i, \theta, \epsilon) = \hat{T}^{\theta,\epsilon}(z_i)$, as in Section II.A. Next observe that

$$\begin{aligned}
A &= \int_0^{i_1} g_i\left(\hat{T}^{\theta',\epsilon}\right) \left[\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} T\left(z_i\left(\hat{T}^{\theta',\epsilon}\right), \theta, \epsilon\right)\right. \\
&\qquad \left. + \left.\frac{\partial}{\partial\xi}\right|_{\xi=\hat{\xi}(\theta',\epsilon)} \Delta T\left(z_i\left(\hat{T}^{\theta',\epsilon}\right), \xi\right) \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} \hat{\xi}(\theta,\epsilon)\right] \mathrm{d}i \\
&= \left[\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} \hat{\xi}(\theta,\epsilon)\right] \int_0^{i_1} g_i\left(\hat{T}^{\theta',\epsilon}\right) \left.\frac{\partial}{\partial\xi}\right|_{\xi=\hat{\xi}(\theta',\epsilon)} \Delta T\left(z_i\left(\hat{T}^{\theta',\epsilon}\right), \xi\right) \mathrm{d}i \\
&= \left[\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} \hat{\xi}(\theta,\epsilon)\right] \int_0^{i_1} g_i\left(T^{\theta',\epsilon} + \Delta T^{\hat{\xi}(\theta',\epsilon)}\right) \left.\frac{\partial}{\partial\xi}\right|_{\xi=\hat{\xi}(\theta',\epsilon)} \Delta T^{\xi}\left(z_i\left(T^{\theta',\epsilon} + \Delta T^{\hat{\xi}(\theta',\epsilon)}\right), \xi\right) \mathrm{d}i \\
&= \left[\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} \hat{\xi}(\theta,\epsilon)\right] \int_0^{i_1} g_i\left(T^{\theta_0,\epsilon_0} + \Delta T^{\hat{\xi}(\theta',\epsilon)}\right) \left.\frac{\partial}{\partial\xi}\right|_{\xi=\hat{\xi}(\theta',\epsilon)} \Delta T^{\xi}\left(z_i\left(T^{\theta_0,\epsilon_0} + \Delta T^{\hat{\xi}(\theta',\epsilon)}\right), \xi\right) \mathrm{d}i \\
&= \left[\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} \hat{\xi}(\theta,\epsilon)\right] \int_0^{1} g_i\left(T^{\theta_0,\epsilon_0} + \Delta T^{\hat{\xi}(\theta',\epsilon)}\right) \left.\frac{\partial}{\partial\xi}\right|_{\xi=\hat{\xi}(\theta',\epsilon)} \Delta T^{\xi}\left(z_i\left(T^{\theta_0,\epsilon_0} + \Delta T^{\hat{\xi}(\theta',\epsilon)}\right), \xi\right) \mathrm{d}i \\
&= 0,
\end{aligned} \tag{B.36}$$

where the first equality follows from the definition (A.52) of $\hat{T}^{\theta,\epsilon}$; the second equality follows from Fact B.1 and Lemma B.4, which imply that, when $i \in [0, i_1]$, $\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} T\left(z_i\left(\hat{T}^{\theta',\epsilon}\right), \theta, \epsilon\right) = 0$; the third equality follows again follows from (A.52); the fourth equality follows from (B.26)-(B.27); the fifth equality follows from the fact that, for all $\xi \in \Xi$, the support of $\Delta T^{\xi}$ is contained in $[0, \hat{z}_1]$ and Lemma B.4, so that the integrand in the expression following the third equality is equal to zero when $i \in [i_1, 1]$; and the last equality follows from (A.48).

Next, observe that

$$
\begin{aligned}
B &= \int_{i_1}^{1} g_i\left(\hat{T}^{\theta',\epsilon}\right) \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} T\left(z_i\left(\hat{T}^{\theta',\epsilon}\right),\theta,\epsilon\right) \mathrm{d}i \\
&= \int_{i_1}^{1} g_i\left(T^{\theta',\epsilon}\right) \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} T\left(z_i\left(T^{\theta',\epsilon}\right),\theta,\epsilon\right) \mathrm{d}i \\
&= \int_{0}^{1} g_i\left(T^{\theta',\epsilon}\right) \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta'} T\left(z_i\left(\hat{T}^{\theta',\epsilon}\right),\theta,\epsilon\right) \mathrm{d}i \\
&= 0,
\end{aligned}
\tag{B.37}
$$

where the first equality follows from Lemma B.4, (A.52), and the fact that, for all $\xi \in \Xi$, the support of $\Delta T^{\xi}$ is contained in $[0, \hat{z}_1]$, so that $T\left(z_i\left(\hat{T}^{\theta',\epsilon}\right),\theta,\epsilon\right) = \hat{T}\left(z_i\left(\hat{T}^{\theta',\epsilon}\right),\theta,\epsilon\right)$ when $i \in [i_1, 1]$; the second equality follows from (B.28)-(B.29); the third equality follows from the fact that, by Fact B.1, $T\left(z_i\left(T^{\theta',\epsilon}\right),\theta,\epsilon\right)$ does not depend on $\theta$ when $i \in [0, i_1]$, so the integrand in the expression following the second equality is zero when $i \in [0, i_1]$; and the last equality follows from the fact that $\left(T^{\theta,\epsilon}\right)$ satisfies (15).

Putting together (B.35), (B.36), and (B.37), it follows that $\left(\hat{T}^{\theta,\epsilon}\right)_{\theta\in\Theta'',\epsilon\in E''}$ satisfies (15).

Next observe that:

$$
\begin{aligned}
&\int_{0}^{1} g_i\left(\hat{T}^{\theta_0,\epsilon_0}\right) \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} \hat{T}\left(z_i\left(\hat{T}^{\theta_0,\epsilon_0}\right),\theta_0,\epsilon\right) \mathrm{d}i \\
=& \underbrace{\int_{0}^{i_1} g_i\left(\hat{T}^{\theta_0,\epsilon_0}\right) \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} \hat{T}\left(z_i\left(\hat{T}^{\theta_0,\epsilon_0}\right),\theta_0,\epsilon\right) \mathrm{d}i}_{C} \\
&+ \underbrace{\int_{i_1}^{1} g_i\left(\hat{T}^{\theta_0,\epsilon_0}\right) \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} \hat{T}\left(z_i\left(\hat{T}^{\theta_0,\epsilon_0}\right),\theta_0,\epsilon\right) \mathrm{d}i}_{D}.
\end{aligned}
\tag{B.38}
$$

Analyzing the first term:

$$
\begin{aligned}
C &= \int_0^{i_1} g_i\left(\hat{T}^{\theta_0,\epsilon_0}\right)\left[\left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(\hat{T}^{\theta_0,\epsilon_0}\right),\theta_0,\epsilon\right)\right. \\
&\quad \left. + \left.\frac{\partial}{\partial \xi}\right|_{\xi=\hat{\xi}(\theta_0,\epsilon_0)} \Delta T^{\xi}\left(z_i\left(\hat{T}^{\theta_0,\epsilon_0}\right)\right)\left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=\epsilon_0}\hat{\xi}\left(\theta_0,\epsilon\right)\right]\mathrm{d}i \\
&= \left[\left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=\epsilon_0}\hat{\xi}\left(\theta_0,\epsilon\right)\right]\int_0^{i_1} g_i\left(\hat{T}^{\theta_0,\epsilon_0}\right)\left.\frac{\partial}{\partial \xi}\right|_{\xi=\hat{\xi}(\theta_0,\epsilon_0)}\Delta T^{\xi}\left(z_i\left(\hat{T}^{\theta_0,\epsilon_0}\right),\xi\right)\mathrm{d}i \\
&= \left[\left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=\epsilon_0}\hat{\xi}\left(\theta_0,\epsilon\right)\right]\int_0^{i_1} g_i\left(T^{\theta_0,\epsilon_0}+\Delta T^{\hat{\xi}(\theta_0,\epsilon_0)}\right)\left.\frac{\partial}{\partial \xi}\right|_{\xi=\hat{\xi}(\theta_0,\epsilon_0)}\Delta T^{\xi}\left(z_i\left(T^{\theta_0,\epsilon_0}+\Delta^{\hat{\xi}(\theta_0,\epsilon_0)}\right),\xi\right)\mathrm{d}i \\
&= \left[\left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=\epsilon_0}\hat{\xi}\left(\theta_0,\epsilon\right)\right]\int_0^{1} g_i\left(T^{\theta_0,\epsilon_0}+\Delta T^{\hat{\xi}(\theta_0,\epsilon_0)}\right)\left.\frac{\partial}{\partial \xi}\right|_{\xi=\hat{\xi}(\theta_0,\epsilon_0)}\Delta T^{\xi}\left(z_i\left(T^{\theta_0,\epsilon_0}+\Delta T^{\hat{\xi}(\theta_0,\epsilon_0)}\right),\xi\right)\mathrm{d}i \\
&= 0,
\end{aligned}
\tag{B.39}
$$

where the first equality follows from (A.52); the second equality from Fact B.1 and Lemma B.4, which imply that, when $i \in [0, i_1]$, $\left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(\hat{T}^{\theta_0,\epsilon_0}\right),\theta_0,\epsilon\right) = 0$; the third equality follows from (A.52); the fourth equality follows from Lemma B.4 and the fact that, for all $\xi \in \Xi$, the support of $\Delta T^{\xi}$ is contained in $[0, \hat{z}_1]$, so that the integrand in the expression following the fourth equality is zero when $i \in [i_1, 1]$; and the last equality follows from (A.48).

Analyzing the second term:

$$
\begin{aligned}
D &= \int_{i_1}^{1} g_i\left(\hat{T}^{\theta_0,\epsilon_0}\right)\left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(\hat{T}^{\theta_0,\epsilon_0}\right),\theta_0,\epsilon\right)\mathrm{d}i \\
&= \int_{i_1}^{1} g_i\left(T^{\theta_0,\epsilon_0}\right)\left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(T^{\theta_0,\epsilon_0}\right),\theta_0,\epsilon\right)\mathrm{d}i \\
&= \int_{0}^{1} g_i\left(T^{\theta_0,\epsilon_0}\right)\left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(T^{\theta_0,\epsilon_0}\right),\theta_0,\epsilon\right)\mathrm{d}i \\
&= 0,
\end{aligned}
\tag{B.40}
$$

where the first equality follows from Lemma B.4, (A.52), and the fact that, for all $\xi \in \Xi$, the support of $\Delta T^{\xi}$ is contained in $[0, \hat{z}_1]$, so that $T\left(z_i\left(\hat{T}^{\theta_0,\epsilon_0}\right),\theta_0,\epsilon\right) = \hat{T}\left(z_i\left(\hat{T}^{\theta_0,\epsilon_0}\right),\theta_0,\epsilon\right)$ when $i \in [i_1, 1]$; the second follows from the fact that, by (A.50) and $\Delta T^0 \equiv 0$, $T^{\theta_0,\epsilon_0} = \hat{T}^{\theta_0,\epsilon_0}$; the third equality follows from the fact that, by Fact B.1 and Lemma B.4, $\left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(T^{\theta_0,\epsilon_0}\right),\theta_0,\epsilon\right) = 0$ when $i \in [0, i_1]$; and the last equality follows from the fact that $\left(T^{\theta,\epsilon}\right)$ satisfies (16).

Putting together (B.38), (B.39), and (B.40), we see that $\left(\hat{T}^{\theta,\epsilon}\right)_{\theta\in\Theta'',\epsilon\in E''}$ satisfies (16).

Next observe that:

$$
\begin{aligned}
&\left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=\theta_0} \int_0^1 g_i\left(\hat{T}^{\theta,\epsilon_0}\right) \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} \hat{T}\left(z_i\left(\hat{T}^{\theta,\epsilon_0}\right),\theta,\epsilon\right)\mathrm{d}i \\
&= \underbrace{\left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=\theta_0} \int_0^{i_1} g_i\left(\hat{T}^{\theta,\epsilon_0}\right) \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} \hat{T}\left(z_i\left(\hat{T}^{\theta,\epsilon_0}\right),\theta,\epsilon\right)\mathrm{d}i}_{E} \\
&\quad + \underbrace{\left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=\theta_0} \int_{i_1}^1 g_i\left(\hat{T}^{\theta,\epsilon_0}\right) \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} \hat{T}\left(z_i\left(\hat{T}^{\theta,\epsilon_0}\right),\theta,\epsilon\right)\mathrm{d}i}_{F}.
\end{aligned}
\tag{B.41}
$$

Analyzing the first term:

$$
\begin{aligned}
E &= \left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=\theta_0} \int_0^{i_1} g_i\left(\hat{T}^{\theta,\epsilon_0}\right) \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(\hat{T}^{\theta,\epsilon_0}\right),\theta,\epsilon\right)\mathrm{d}i \\
&\quad + \left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=\theta_0} \left[\left(\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} \hat{\xi}(\theta,\epsilon)\right) \int_0^{i_1} g_i\left(\hat{T}^{\theta,\epsilon_0}\right) \left.\frac{\partial}{\partial\xi}\right|_{\xi=\hat{\xi}(\theta,\epsilon_0)} \Delta T^{\xi}\left(z_i\left(\hat{T}^{\theta,\epsilon_0}\right),\xi\right)\mathrm{d}i\right] \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=\theta_0} \left[\left(\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} \hat{\xi}(\theta,\epsilon)\right) \int_0^{i_1} g_i\left(\hat{T}^{\theta,\epsilon_0}\right) \left.\frac{\partial}{\partial\xi}\right|_{\xi=\hat{\xi}(\theta,\epsilon_0)} \Delta T^{\xi}\left(z_i\left(\hat{T}^{\theta,\epsilon_0}\right),\xi\right)\mathrm{d}i\right] \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=\theta_0} \left[\left(\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} \hat{\xi}(\theta,\epsilon)\right) \int_0^{i_1} g_i\left(T^{\theta,\epsilon_0}+\Delta T^{\hat{\xi}(\theta,\epsilon_0)}\right)\right. \\
&\quad \left.\times \left.\frac{\partial}{\partial\xi}\right|_{\xi=\hat{\xi}(\theta,\epsilon_0)} \Delta T^{\xi}\left(z_i\left(T^{\theta,\epsilon_0}+\Delta T^{\hat{\xi}(\theta,\epsilon_0)}\right),\xi\right)\mathrm{d}i\right] \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=\theta_0} \left[\left(\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} \hat{\xi}(\theta,\epsilon)\right) \int_0^{1} g_i\left(T^{\theta,\epsilon_0}+\Delta T^{\hat{\xi}(\theta,\epsilon_0)}\right)\right. \\
&\quad \left.\times \left.\frac{\partial}{\partial\xi}\right|_{\xi=\hat{\xi}(\theta,\epsilon_0)} \Delta T^{\xi}\left(z_i\left(T^{\theta,\epsilon_0}+\Delta T^{\hat{\xi}(\theta,\epsilon_0)}\right),\xi\right)\mathrm{d}i\right] \\
&= 0,
\end{aligned}
\tag{B.42}
$$

where the first equality follows from (A.52); the second equality from Fact B.1 and Lemma B.4, which imply that $\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(\hat{T}^{\theta,\epsilon_0}\right),\theta,\epsilon\right) = 0, \forall\theta\in\Theta''$, when $i\in[0,i_1]$; the third equality follows from (A.52); the fourth equality follows from Lemma B.4 and the fact that, for all $\xi\in\Xi$, the support of $\Delta T^{\xi}$ is contained in $[0,\hat{z}_1]$, so that the integrand in the expression following the fourth equality is zero, for all values of $\theta$ in $\Theta''$, when $i\in[i_1,1]$; and the last equality follows from (A.48).

Analyzing the second term:

$$
\begin{aligned}
F &= \left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=\theta_0} \int_{i_1}^{1} g_i\left(\hat{T}^{\theta,\epsilon_0}\right) \left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(\hat{T}^{\theta,\epsilon_0}\right),\theta,\epsilon\right)\mathrm{d}i \\
&\quad + \left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=\theta_0} \left[\left(\left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=\epsilon_0} \hat{\xi}\left(\theta,\epsilon\right)\right) \int_{i_1}^{1} g_i\left(\hat{T}^{\theta,\epsilon_0}\right) \left.\frac{\partial}{\partial \xi}\right|_{\xi=\hat{\xi}(\theta_0,\epsilon_0)} \Delta T^{\xi}\left(z_i\left(\hat{T}^{\theta,\epsilon_0}\right),\xi\right)\mathrm{d}i\right] \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=\theta_0} \int_{i_1}^{1} g_i\left(\hat{T}^{\theta,\epsilon_0}\right) \left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(\hat{T}^{\theta,\epsilon_0}\right),\theta,\epsilon\right)\mathrm{d}i \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=\theta_0} \int_{i_1}^{1} g_i\left(T^{\theta,\epsilon_0}\right) \left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(T^{\theta,\epsilon_0}\right),\theta,\epsilon\right)\mathrm{d}i \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=\theta_0} \int_{0}^{1} g_i\left(T^{\theta,\epsilon_0}\right) \left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(T^{\theta,\epsilon_0}\right),\theta,\epsilon\right)\mathrm{d}i \\
&< 0,
\end{aligned}
\tag{B.43}
$$

where the first equality follows from (A.52); the second equality follows from Lemma B.4, (A.52), and the fact that, for all $\xi \in \Xi$, the support of $\Delta T^{\xi}$ is contained in $[0, \hat{z}_1]$, so that $\left.\frac{\partial}{\partial \xi}\right|_{\xi=\hat{\xi}(\theta,\epsilon_0)} \Delta T^{\xi}\left(z_i\left(\hat{T}^{\theta,\epsilon_0}\right),\xi\right) = 0, \forall \theta \in \Theta''$, when $i \in [i_1, 1]$; the third equality follows from Lemma B.4, (A.52) and (B.28)-(B.29); the fourth equality follows from the fact that, by Fact B.1, $\left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(T^{\theta,\epsilon_0}\right),\theta,\epsilon\right) = 0, \forall \theta \in \Theta''$, when $i \in [0, i_1]$; and the inequality follows from the fact that $\left(T^{\theta,\epsilon}\right)$ satisfies (17).

Putting together (B.41), (B.42), and (B.43), it follows that $\left(\hat{T}^{\theta,\epsilon}\right)_{\theta\in\Theta'',\epsilon\in E''}$ satisfies (17).

We have now established that $\left(\hat{T}^{\theta,\epsilon}\right)_{\theta\in\Theta'',\epsilon\in E''}$ satisfies (15)-(17), completing the proof of Lemma A.4. □

## C Additional lemmas

The lemmas in this section apply to tax policies that are not individualized.

**Lemma C.1** *For all regular tax policies $T$, $\{z_i(T) : i \in I\} = [z_0(T), z_1(T)]$, and the map $i \mapsto z_i(T)$ is strictly increasing.*

Proof. Let $T$ be a regular tax policy. It follows from our assumptions (see Section A.1.2) that $z_i(T)$ is characterized by the first order condition $1 - T'(z_i(T)) - v_i'(z_i(T)) = 0$. The smoothness of $T$ and $(z,i) \mapsto v_i(z)$ imply that the function $i \mapsto z_i(T)$ is smooth. It follows from the facts that (i) $v_i(z) = v(z, y_i)\,\forall i, \forall z$, (ii) $\frac{\partial^2}{\partial z \partial y} v(z,y) < 0, \forall z, \forall y$, and (iii) $\frac{\mathrm{d}}{\mathrm{d}i} y_i > 0, \forall i$, that the map $i \mapsto z_i(T)$ is strictly increasing (see Section V.A for the preceding assumptions). Since $i \mapsto z_i(T)$ is continuous and strictly increasing on $I = [0,1]$, $\{z_i(T) : i \in I\} = [z_0(T), z_1(T)]$. □

**Lemma C.2** *Let $T$ be a regular tax policy. For $j = 1, \ldots, n$, let $\Theta_j = \left[-\bar{\theta}_j, \bar{\theta}_j\right] \subseteq \mathbb{R}$, where $\theta_j > 0$, and let $\bar{\Theta} = \times_{j=1}^{n} \Theta_j$. Write $\bar{\theta} = (\theta_1, \ldots, \theta_j, \ldots, \theta_n)$. Let $\left(\Delta T^{\bar{\theta}}\right)_{\bar{\theta}\in\bar{\Theta}}$ be a family of tax reforms such* *that the map* $\left(z,\bar{\theta}\right)\mapsto\Delta T^{\bar{\theta}}\left(z\right)$ *is smooth and* $\Delta T^{(0,0,\ldots,0)}\equiv 0$*. Then there exist* $\theta_j^*\in\Theta_j$ *with* $\theta_j^*>0$ *for* $j=1,\ldots,n$ *such that, for all* $\bar{\theta}=\left(\theta_1,\ldots,\theta_j,\ldots,\theta_n\right)\in\bar{\Theta}$*, if* $|\theta_j|\leq\theta_j^*$ *for* $j=1,\ldots,n$*, then* $T+\Delta T^{\bar{\theta}}$ *is regular.*

Proof. I use the following notation: for any tax policy $T$ and income $z_i$, define $U_i^T\left(z_i\right)=u\left(z_i-T\left(z_i\right)-v_i\left(z_i\right)\right)$. Now let $T$ be a regular tax policy. It follows that, for all agents $i$, $\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}U_i^T\left(z_i\left(T\right)\right)<0$ (see Section A.1.2). Also, since $T$ is regular, $z_i\left(T\right)>0$, for all agents $i$. By the smoothness of the primitives and $T$, it follows that there is a neighborhood $N_i$ of the income $z_i\left(T\right)$ such that, for all $z_i\in N_i$, $\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}U_i^T\left(z_i\right)<0$ and $z_i>0$. For each $i$, let

$$\delta_i=\sup\left\{\delta>0:z_i\left(T\right)-\delta>0,\forall z_i\in\left(z_i\left(T\right)-\delta,z_i\left(T\right)+\delta\right),\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}U_i^T\left(z_i\right)<0\right\}.$$

We have $\delta_i>0,\forall i$, and, moreover, the smoothness of the primitives and of $T$ implies that $i\mapsto\delta_i$ is smooth. Since a continuous function attains its minimum on a compact set, it follows that $\delta^*=\min\left\{\delta_i:i\in[0,1]\right\}$ exists and $\delta^*>0$. For each $i$, define the neighborhood $N_i'=\left(z_i\left(T\right)-\frac{1}{2}\delta^*,z_i\left(T\right)+\frac{1}{2}\delta^*\right)$ and let $\bar{N}_i'$ be the closure of $N_i'$. For each $\bar{\theta}\in\bar{\Theta}$, define $T^{\bar{\theta}}=T+\Delta T^{\bar{\theta}}$. Define $\gamma_i^{\bar{\theta}}=U^{T^{\bar{\theta}}}\left(z_i\left(T\right)\right)-\max_{z_i\in Z\setminus N_i'}U_i^{T^{\bar{\theta}}}\left(z_i\right)$ and $\gamma^{\bar{\theta}}=\min_{i\in[0,1]}\gamma_i^{\bar{\theta}}$. As $T^{(0,0,\ldots,0)}=T+\Delta T^{(0,0,\ldots,0)}=T$, and, as $T$ is regular, $U_i^{T^{(0,0,\ldots,0)}}\left(z_i\right)$ has a unique maximizer $z_i\left(T\right)$, it follows that, for all $i$, $\gamma_i^{(0,0,\ldots,0)}>0$, and hence, again because a continuous function attains its minimum on a compact set, $\gamma^{(0,0,\ldots,0)}>0$. It follows from our smoothness assumptions that there exist $\theta_j'\in\Theta_j$ with $\theta_j'>0$ for $j=1,\ldots,n$ such for all $\bar{\theta}=\left(\theta_1,\ldots,\theta_j,\ldots,\theta_n\right)\in\bar{\Theta}$, if $|\theta_j|\leq\theta_j'$ for $j=1,\ldots,n$, $\gamma^{\bar{\theta}}>0$, so that, for all such $\bar{\theta}$, $U_i^{T^{\bar{\theta}}}\left(z_i\right)$ does not have any maximizers $z_i$ outside of $N_i'$. Note that we have: $\forall i\in I,\forall z_i\in\bar{N}_i'$, $\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}U_i^{T^{(0,0,\ldots,0)}}\left(z_i\right)<0$. So $\max_{i\in[0,1],z_i\in\bar{N}_i'}\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}U_i^{T^{(0,0,\ldots,0)}}<0$. As the map $\bar{\theta}\mapsto\max_{i\in[0,1],z_i\in\bar{N}_i'}\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}U_i^{T^{\bar{\theta}}}\left(z_i\right)$ is continuous, it follows that there exist $\theta_j''\in\Theta_j$ with $\theta_j''>0$ for $j=1,\ldots,n$ such for all $\bar{\theta}=\left(\theta_1,\ldots,\theta_j,\ldots,\theta_n\right)\in\bar{\Theta}$, if $|\theta_j|\leq\theta_j''$ for $j=1,\ldots,n$, then, for all agents $i$ and all $z_i\in\bar{N}_i'$, $\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}U_i^{T^{\bar{\theta}}}\left(z_i\right)<0$, so that $U_i^{T^{\bar{\theta}}}\left(z_i\right)$ is strictly convex on $\bar{N}_i'$, implying that $U_i^{T^{\bar{\theta}}}\left(z_i\right)$ has a unique maximizer on $\bar{N}_i'$. It follows that if $\theta_j^*=\min\left\{\theta_j',\theta_j''\right\}$ for $j=1,\ldots,n$, then, then for all $\bar{\theta}=\left(\theta_1,\ldots,\theta_j,\ldots,\theta_n\right)\in\bar{\Theta}$, if $|\theta_j|\leq\theta_j^*$ for $j=1,\ldots,n$, then for, all agents $i$, $U_i^{T^{\bar{\theta}}}\left(z_i\right)$ has a unique maximizer $z_i\left(T^{\bar{\theta}}\right)>0$, and, moreover, $\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}U_i^{T^{\bar{\theta}}}\left(z_i\left(T^{\bar{\theta}}\right)\right)<0$, so that $T^{\bar{\theta}}$ is regular. □

## D Theorems 1 and 3 without quasilinear utility

### D.1 Preliminaries

This Appendix explains how Theorems 1 and 3 are still valid without the assumption of quasilinearity. In particular, I describe how the proofs of the theorems must be modified if the assumption of quasilinearity is removed. I assume that utility takes the form $U_i\left(c_i,z_i\right)=U\left(c_i,z_i;x_i,y_i\right)$, where $U\left(c_i,z_i;x_i,y_i\right)$ is smooth in $\left(c_i,z_i;x_i,y_i\right)$ unless $\left(x_i,y_i\right)$ are discrete, in which case $U\left(c_i,z_i;x_i,y_i\right)$ is smooth in $\left(c_i,z_i\right)$. I assume that $U_i\left(c_i,z_i\right)$ is strictly increasing in $c_i$ (with a strictly positive partial derivative everywhere), strictly decreasing in $z_i$, and strictly concave in $(c_i, z_i)$. I assume for simplicity that $c_i$ can take on any real value and that for any income level $z_i$, the range of $c_i \mapsto U_i(c_i, z_i)$ is the entire real line. I continue to assume that, in the absence of taxes, all agents earn a positive income.

In what follows it will be useful to define the function $\tilde{c}_i(u_i, z_i)$ by the following condition:

$$U_i(\tilde{c}_i(u_i, z_i), z_i) = u_i, \quad \forall z_i \in Z, \forall u_i \in \mathbb{R}. \tag{D.1}$$

So, $\tilde{c}_i(u_i, z_i)$ is the level of consumption that is necessary to give $i$ utility $u_i$ given income level $z_i$; $\tilde{c}_i(u_i, z_i)$ is well-defined because $U_i$ is strictly increasing in $c_i$. It follows from the implicit function theorem that:

$$\frac{\partial}{\partial z_i}\tilde{c}_i(u_i, z_i) = -\frac{\frac{\partial}{\partial z_i}U_i(\tilde{c}_i(u_i, z_i), z_i)}{\frac{\partial}{\partial c_i}U_i(\tilde{c}_i(u_i, z_i), z_i)}, \quad \forall u_i, \forall z_i. \tag{D.2}$$

I assume that along any $(c_i, z_i)$-indifference curve, the marginal rate of substitution of consumption for avoiding the effort of earning income exceeds 1 as $z$ becomes large:

$$\forall u_i, \lim_{z_i \to +\infty} \frac{\partial}{\partial z_i}\tilde{c}_i(u_i, z_i) > 1.$$

In other words, as one increases both income and consumption along an indifference curve, it is eventually necessary to compensate an agent by more than a dollar in order to bear the cost of earning another dollar of income. This has the consequence that, whenever facing a tax policy under which marginal tax rates become nonnegative once income is sufficiently large, the agent optimally selects some finite income and does not want to increase their income without bounds.

Note that when $U_i(c_i, z_i) = u(c_i - v_i(z_i))$ where $u' > 0$ and $u'' < 0$ everywhere, and all the other assumptions of Section I.A are satisfied, then all of the above assumptions are satisfied, so the assumption here in essence generalize the assumptions made for the quasilinear case. I also carry over other assumptions (or analogous assumptions) and notation from the quasilinear case.

The key preliminary definitions and results supporting the main results continue to hold in this more general framework. Observe first that, even without quasilinear preferences, the envelope theorem still implies that for any well behaved parameterized family of tax policies $(T^\theta)$, $\frac{\mathrm{d}}{\mathrm{d}\theta}U_i(T^\theta) = -\frac{\partial}{\partial c_i}U_i(c_i(\theta), z_i(\theta))\frac{\partial}{\partial \theta}T_i(z_i(T), \theta)$. It follows that the local and global improvement principles are still valid when welfare weights are utilitarian. So the justification for the global and local improvement principles that was given in Section II.B, on analogy with the utilitarian case, still applies without quasilinearity. Likewise, the supporting Proposition 2 on Pareto indifference and weak Pareto along paths is unchanged, and so the result continues to hold.

We can no longer define $\hat{g}_i(\hat{u}_i, z_i)$ as we did in (9) in Section III because that definition depended on the assumption of quasilinear utility. Instead we define $\tilde{g}_i(u_i, z_i)$ which is a function of the variable $u_i = U_i(c_i, z_i)$ and $z_i$, as follows:

$$\tilde{g}_i(u_i, z_i) = g_i(\tilde{c}_i(u_i, z_i), z_i), \quad \forall z_i \in Z, \forall u_i \in \mathbb{R}. \tag{D.3}$$

Next define:

$$k_i(u_i, z_i) = \frac{\partial}{\partial c_i} U_i(\tilde{c}_i(u_i, z_i), z_i), \tag{D.4}$$

$$h_i(u_i, z_i) = \frac{\tilde{g}_i(u_i, z_i)}{\frac{\partial}{\partial c_i} U_i(\tilde{c}_i(u_i, z_i), z_i)}. \tag{D.5}$$

Then observe that

$$\tilde{g}_i(u_i, z_i) = k_i(u_i, z_i) h_i(u_i, z_i). \tag{D.6}$$

Now choose $z_i, z_i'$ and $u_i$ and observe that it follows from (D.1) that

$$U_i(\tilde{c}_i(u_i, z_i), z_i) = U_i\left(\tilde{c}_i\left(u_i, z_i'\right), z_i'\right). \tag{D.7}$$

Then if $g_i$ is structurally utilitarian, $h_i(u_i, z_i) = \frac{\tilde{g}_i(u_i,z_i)}{\frac{\partial}{\partial c_i} U_i(\tilde{c}_i(u_i,z_i),z_i)} = \frac{\tilde{g}_i\left(u_i,z_i'\right)}{\frac{\partial}{\partial c_i} U_i\left(\tilde{c}_i\left(u_i,z_i'\right),z_i'\right)} = h_i(u_i, z_i')$, where the second equality follows from (D.7) and the definition of structural utilitarianism without quasilinearity (Definition 2). So for structurally utilitarian weights, $h_i(u_i, z_i)$ does not depend on $z_i$. It is also easy to see that, if $h_i(u_i, z_i)$ does not depend on $z_i$, then the corresponding welfare weights are structurally utilitarian. This can be summarized in a form a proposition which is the non-quasilinear analog of Proposition 3.

**Proposition D.1** *Let $g$ and $\tilde{g}$ be related as in (D.3) and let $h$ be defined in terms of $\tilde{g}$ as in (D.5). Then welfare weights $g$ are structurally utilitarian if and only if* $\forall i \in I, \forall u_i \in \mathbb{R}, \forall z_i \in Z$, $\frac{\partial}{\partial z_i} h_i(u_i, z_i) = 0$.

### D.2 Proof of Theorem 1 without quasilinearity

I now present the proof of Theorem 1 without assuming quasilinearity using the more general definition of structural utilitarianism (Definition 2). First assume welfare weights are generalized utilitarian. This means that welfare weights are of the form $g_i(c_i, z_i) = F_i'(U_i(c_i, z_i)) \frac{\partial}{\partial c_i} U(c_i, z_i)$. Assume that $U_i(c_i, z_i) = U_i(c_i', z_i')$. Then $F_i'(U_i(c_i, z_i)) = F_i'(U_i(c_i', z_i'))$. So

$$\frac{\frac{\partial}{\partial c_i} U_i(c_i, z_i)}{\frac{\partial}{\partial c_i} U_i(c_i', z_i')} = \frac{F_i'(U_i(c_i, z_i)) \frac{\partial}{\partial c_i} U_i(c_i, z_i)}{F_i'(U_i(c_i', z_i')) \frac{\partial}{\partial c_i} U_i(c_i', z_i')} = \frac{g_i(c_i, z_i)}{g_i(c_i', z_i')}.$$

So, welfare weights are structurally utilitarian.

Going in the other direction, assume that welfare weights $g$ are structurally utilitarian. Let $h_i$ be defined from $g_i$ via (D.3) and (D.5). It follows from Proposition D.1 that $h_i(u_i, z_i)$ does not depend on $z_i$, and hence we can write this function as $h_i(u_i)$, without the argument $z_i$. Now define $F_i(u_i) = \int_0^{u_i} h_i(u_i)\,\mathrm{d}u_i$. Appealing to (D.1), (D.3) and (D.5), note that because $g_i(c_i, z_i) = g(c_i, z_i; x_i, y_i)$ and $U_i(c_i, z_i) = U(c_i, z_i; x_i, y_i)$, we can write $F_i(u_i) = F(u_i; x_i, y_i)$ and $F$ inherits the appropriate smoothness properties from $g$ and $U$. If follows from the above construction that $F_i'(U_i(c_i, z_i)) \frac{\partial}{\partial c_i} U_i(c_i, z_i) = h_i(U_i(c_i, z_i)) \frac{\partial}{\partial c_i} U_i(c_i, z_i) = \frac{g_i(c_i, z_i)}{\frac{\partial}{\partial c_i} U_i(c_i, z_i)} \frac{\partial}{\partial c_i} U_i(c_i, z_i) = g_i(c_i, z_i)$. So welfare weights are generalized utilitarian. □

### D.3 An example

In this section, I present as informal example with individualized taxes, similar to the examples in Section IV, that illustrates that, in the non-quasilinear case, if structural utilitarianism in the sense of Definition 2 is violated, then it is possible to construct a social preference cycle. In this example, I will not be concerned with holding revenue constant because that can be achieved with a modification of the example by means similar to that presented in Section IV.A. The purpose of this section is to provide the reader with intuition and an understanding of the essence of the argument that a failure of structural utilitarianism leads to a social preference cycle in the non-quasilinear case.

Suppose that there is just a single observable binary characteristic $x_i$ such that $x_i = A$ if $i \in \left[0, \frac{1}{2}\right]$ and $x_i = B$ if $i \in \left(\frac{1}{2}, 1\right]$, and taxes are conditioned on this characteristic. There are no unobservable characteristics. All agents of type $A$ are identical with one another and all agents of type $B$ are identical with one another as well. I write $U_A(c, z) = U(c, z, A)$ and $U_B(c, z) = U(c, z, B)$ for the utility functions of agents with characteristics $A$ and $B$ respectively. Likewise, I write $g_A(c, z) = g(c, z, A)$ and $g_B(c, z) = g(c, z, B)$ for the welfare weights of types $A$ and $B$ respectively.

Suppose that welfare weights for type $A$ agents are not structurally utilitarian. It follows from Definition 2 that there exist allocations $(c_0, z_0)$, $(c_1, z_1)$ such that

$$U_A(c_0, z_0) = U_A(c_1, z_1) = u^* \tag{D.8}$$

but $\frac{g_A(c_0,z_0)}{\frac{\partial}{\partial c} U_A(c_0,z_0)} \neq \frac{g_A(c_1,z_1)}{\frac{\partial}{\partial c} U_A(c_1,z_1)}$. Assume without loss of generality that $\frac{g_A(c_0,z_0)}{\frac{\partial}{\partial c} U_A(c_0,z_0)} < \frac{g_A(c_1,z_1)}{\frac{\partial}{\partial c} U_A(c_1,z_1)}$. Then there exists a number $b$ such that

$$\frac{g_A(c_0, z_0)}{\frac{\partial}{\partial c} U_A(c_0, z_0)} < b < \frac{g_A(c_1, z_1)}{\frac{\partial}{\partial c} U_A(c_1, z_1)}. \tag{D.9}$$

Because utility functions are strictly concave, and hence upper contour sets are strictly convex, for any consumption-income bundle $(c^*, z^*)$, it is possible to construct a linear tax policy (linear in $z$) $\bar{T}^{c^*,z^*}(z) = \tau(c^*, z^*) z + \kappa(c^*, z^*)$ such that type $A$'s optimal consumption and income in response to $\bar{T}^{c^*,z^*}$ is $(c^*, z^*)$. As above, let $\tilde{c}_A(u, z)$ be the level of consumption that gives agents of type $A$ a utility of $u$ when their income is $z$. Let $(\hat{c}, \hat{z})$ be type $B$'s optimal consumption and income in the absence of taxes. (Of course $\hat{c} = \hat{z}$).

Now consider a family of tax policies $\left(T^{\zeta,u}\right)$ parameterized by real numbers $\zeta$ and $u$, where $\zeta \geq 0$, defined by

$$T^{\zeta,u}(z,x) = \begin{cases} \bar{T}^{\tilde{c},\zeta}(z) \text{ with } \tilde{c} = \tilde{c}_A(u,\zeta), & \text{if } x = A, \\ \frac{b}{g_B(\hat{c},\hat{z})}(u-u^*), & \text{if } x = B. \end{cases}$$

I now explain this tax policy. First consider type $A$ agents. At $T^{\zeta,u}$, type $A$ agents face linear tax policy of the form $\bar{T}^{c^*,z^*}$ where $c^* = \tilde{c}_A(u,\zeta)$ and $z^* = \zeta$. As explained above, this leads type $A$ agents to select consumption $\tilde{c}_A(u,\zeta)$ and income $\zeta$, and hence to attain utility $u$. Type $B$ agents face only a lumpsum tax $\frac{b}{g_B(\hat{c},\hat{z})}(u-u^*)$, where $u^*$ is defined by (D.8) and $b$ satisfies (D.9).

By construction, holding fixed $u$ and varying $\zeta$ in $T^{\zeta,u}$, type $A$ agents' utilities remain constant at $u$ when facing $T^{\zeta,u}$. The taxes faced by type $B$ agents do not depend on $\zeta$. Hence, all agents are indifferent when facing $T^{\zeta,u}$ as $\zeta$ varies while $u$ is held fixed, and so by Pareto indifference along paths (Proposition 2), which, as explained above, continues to hold in the non-quasilinear case, we have:

$$T^{\zeta,u} \sim^g T^{\zeta',u}, \quad \forall \zeta, \zeta', \forall u. \tag{D.10}$$

Let $T_A(z,\zeta,u)$ be the taxes paid by type $A$ agents under $T^{\zeta,u}$ when earning income $z$. ($T_B(z,\zeta,u)$ is defined similarly for type $B$ agents.) Let $U_A(\zeta,u)$ be type $A$ agents' utility when facing tax policy $T^{\zeta,u}$ and let $c_A\left(T^{\zeta,u}\right)$ and $z_A\left(T^{\zeta,u}\right)$ be respectively the optimal consumption and income for type $A$ agents when facing tax policy $T^{\zeta,u}$. It follows form the envelope theorem that

$$\frac{\partial}{\partial u}U_A(\zeta,u) = -\frac{\partial}{\partial c}U_A\left(c_A\left(T^{\zeta,u}\right), z_A\left(T^{\zeta,u}\right)\right)\left.\frac{\partial}{\partial u'}\right|_{u'=u} T_A\left(z\left(T^{\zeta,u}\right),\zeta,u'\right). \tag{D.11}$$

On the other hand because, for all $u$ and $\zeta$, $U_A(\zeta,u) = u$, it follows that

$$\frac{\partial}{\partial u}U_A(\zeta,u) = 1, \quad \forall \zeta, \forall u. \tag{D.12}$$

Putting (D.11) and (D.12) together, we have

$$\left.\frac{\partial}{\partial u}\right|_{u=u'} T_A\left(z_A\left(T^{\zeta,u}\right),\zeta,u'\right) = -\frac{1}{\frac{\partial}{\partial c}U_A\left(c_A\left(T^{\zeta,u}\right), z_A\left(T^{\zeta,u}\right)\right)}.$$

By construction we have:

$$\left.\frac{\partial}{\partial u'}\right|_{u'=u} T_B\left(z_B\left(T^{\zeta,u}\right),\zeta,u'\right) = \frac{b}{g_B(\hat{c},\hat{z})}$$

Note that when $\zeta = z_0$ and $u = u^*$, $z_A\left(T^{\zeta,u}\right) = z_0$ and $c_A\left(T^{\zeta,u}\right) = \tilde{c}_A(u^*,z_0) = c_0$. Also, when $u = u^*$, type $B$ agents face no taxes under $T^{\zeta,u}$, and hence $c_B\left(T^{\zeta,u}\right) = \hat{c}$ and $z_B\left(T^{\zeta,u}\right) = \hat{z}$. It follows that

$$\begin{aligned}
&\int g_i\left(T^{z_0,u^*}\right)\left.\frac{\partial}{\partial u}\right|_{u=u^*} T_i\left(z_i\left(T^{z_0,u^*}\right),z_0,u\right)\mathrm{d}i\\
=&\underbrace{\int_0^{\frac{1}{2}} g_A\left(c_0,z_0\right)\left(-\frac{1}{\frac{\partial}{\partial c}U_A\left(c_0,z_0\right)}\right)\mathrm{d}i}_{\text{type } A \text{ agents}}+\underbrace{\int_{\frac{1}{2}} g\left(\hat{c},\hat{z}\right)\frac{b}{g_B\left(\hat{c},\hat{z}\right)}\mathrm{d}i}_{\text{type } B \text{ agents}}\\
=&\frac{1}{2}\left(-\frac{g_A\left(c_0,z_0\right)}{\frac{\partial}{\partial c}U_A\left(c_0,z_0\right)}+b\right)>0,
\end{aligned}$$

where the inequality follows from (D.9). Similarly,

$$\int g_i\left(T^{z_1,u^*}\right)\left.\frac{\partial}{\partial u}\right|_{u=u^*} T_i\left(z_i\left(T^{z_1,u^*}\right),z_1,u\right)\mathrm{d}i=\frac{1}{2}\left(-\frac{g_A\left(c_1,z_1\right)}{\frac{\partial}{\partial c}U_A\left(c_1,z_1\right)}+b\right)<0.$$

It follows from the local improvement principle (Proposition 1), which also continues to hold in the non-quasilinear case, that for sufficiently small $\epsilon>0$,

$$\begin{aligned}
T^{z_0,u^*} &\succ_g T^{z_0,u^*+\epsilon}, \text{ and}\\
T^{z_1,u^*} &\prec_g T^{z_1,u^*+\epsilon}.
\end{aligned} \tag{D.13}$$

Putting together (D.10) and (D.13), we have the social preference cycle:

$$T^{z_1,u^*}\prec_g T^{z_1,u^*+\epsilon}\sim_g T^{z_0,u^*+\epsilon}\prec_g T^{z_0,u^*}\sim^g T^{z_1,u^*}.$$

So, on the assumption that welfare weights are not structurally utilitarian, we have derived a cycle.

### D.4 Proof of Theorem 3 without quasilinearity

In this section, I explain how to modify the proof of Theorem 3 when quasilinearity is no longer assumed. (The statement of the theorem must be modified to appeal to the assumptions of Sections D.1 and D.4.1 rather than Section V.A.)

#### D.4.1 Additional structure for the non-quasilinear version of Theorem 3

I now assume, as in Section V.A, that there are no observable characteristics, but there is a single one-dimensional real valued unobservable characteristic $y$, so that we can write $U_i\left(c_i,z_i\right)=U\left(c_i,z_i;y_i\right)$, and that the function $i\mapsto y_i$ is smooth and that the derivative of $y_i$ with respect to $i$ is positive at all values of $i$ in $I=[0,1]$. Moreover, I assume the single-crossing condition that, for all $(c,z,y)\in\mathbb{R}\times Z\times Y$, $\frac{\mathrm{d}}{\mathrm{d}y}\frac{\frac{\partial}{\partial z}U(c,z,y)}{\frac{\partial}{\partial c}U(c,z,y)}>0$. This single crossing condition implies that for every regular tax policy $T$, $i\mapsto z_i\left(T\right)$ is strictly increasing in $i$. Note that in Section V.A, we assumed that $\frac{\partial^2}{\partial y\partial z}v\left(z,y\right)<0$, so that when $U\left(c,z;y\right)=u\left(c-v\left(z,y\right)\right)$, $\frac{\mathrm{d}}{\mathrm{d}y}\frac{\frac{\partial}{\partial z}U(c,z,y)}{\frac{\partial}{\partial c}U(c,z,y)}=-\frac{\partial^2}{\partial y\partial z}v\left(z,y\right)>0$. So the above single-crossing condition generalizes the assumption we made in the quasilinear case.

#### D.4.2 Modifications of the main lemmas

Theorem 3 is proven by means of a series of lemmas, and in this section I will discuss how these lemmas must be altered when we drop the assumption of quasilinearity and revert to the weaker assumptions of Sections D.1 and D.4.1 above.

##### D.4.2.1 Lemma 1

Lemma 1 is unaltered relative to the quasilinear case and the proof is identical.

##### D.4.2.2 Corollary 3

The following result is the non-quasilinear analog of Corollary 3.

**Corollary D.1** *Let $h$ be related to $g$ as specified by (D.3) and (D.5). If $g$ is not structurally utilitarian, then there exists a regular tax policy $T$ for which there exist agents $i_a, i_b \in (0,1)$ with $i_a < i_b$ such that either*

$$\forall i \in (i_a, i_b), \quad \frac{\partial}{\partial z_i} h_i (U_i(T), z_i(T)) < 0 \tag{D.14}$$

*or*

$$\forall i \in (i_a, i_b), \quad \frac{\partial}{\partial z_i} h_i (U_i(T), z_i(T)) > 0. \tag{D.15}$$

In the proof of Corollary D.1, Proposition D.1 plays the role that Proposition 3 plays in the proof of Corollary 3; moreover the proof of Corollary D.1 is a bit more involved than that of Corollary 3 because one cannot rely on the convenient properties of quasilinear preferences.

##### D.4.2.3 Lemma 2

We also have the following lemma, which is an analog of Lemma 2.

**Lemma D.1** *Assume that $\left(T^{\theta,\epsilon}\right)$ is well-behaved and satisfies (15). Then (17) holds if and only if*

$$\begin{aligned} \int m_i(\theta_0, \epsilon_0) \Bigg[ & \left.\frac{\partial^2}{\partial\theta\partial z_i}\right|_{\theta=\theta_0, z_i=z_i(\theta_0,\epsilon_0)} T(z_i, \theta, \epsilon_0) \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} T(z_i(\theta_0,\epsilon_0), \theta_0, \epsilon) \\ & - \left.\frac{\partial^2}{\partial\epsilon\partial z_i}\right|_{\epsilon=\epsilon_0, z_i=z_i(\theta_0,\epsilon_0)} T(z_i, \theta_0, \epsilon) \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} T(z_i(\theta_0,\epsilon_0), \theta, \epsilon_0) \Bigg] \mathrm{d}i < 0. \end{aligned} \tag{D.16}$$

*where*

$$m_i(\theta_0, \epsilon_0) = \frac{\left[\frac{\partial}{\partial c_i} U_i(c_i(\theta_0,\epsilon_0), z_i(\theta_0,\epsilon_0))\right]^2 \frac{\partial}{\partial z_i} h_i(U_i(\theta_0,\epsilon_0), z_i(\theta_0,\epsilon_0))}{\left.\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}\right|_{z_i=z_i(\theta_0,\epsilon_0)} U_i(z_i - T(z_i), z_i)}.$$

The structure of the proof is similar to the structure of the proof of Lemma 2, and features terms $A$, $B$, and $C$, which play the same role as the terms $A$, $B$, and $C$ in Lemma 2. However precise details of these terms differ in the two lemmas. In Lemma D.1,

$$
\begin{aligned}
A &= \int m_i(\theta_0,\epsilon_0) \left.\frac{\partial^2}{\partial\epsilon\partial z_i}\right|_{\epsilon=\epsilon_0, z_i=z_i(\theta_0,\epsilon_0)} T(z_i,\theta_0,\epsilon) \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} T(z_i(\theta_0,\epsilon_0),\theta,\epsilon_0)\,\mathrm{d}i \\
B &= \int m_i(\theta_0,\epsilon_0) \left.\frac{\partial^2}{\partial\theta\partial z_i}\right|_{\theta=\theta_0, z_i=z_i(\theta_0,\epsilon_0)} T(z_i,\theta,\epsilon_0) \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} T(z_i(\theta_0,\epsilon_0),\theta_0,\epsilon)\,\mathrm{d}i \\
C &= \int g_i \left( \frac{\partial^2 T_i}{\partial\theta\partial\epsilon} + \frac{\left[\frac{\partial^2 T_i}{\partial\theta\partial z_i}\frac{\partial T_i}{\partial\epsilon} + \frac{\partial^2 T_i}{\partial\epsilon\partial z_i}\frac{\partial T_i}{\partial\theta}\right]\left[\frac{\partial^2 U_i}{\partial c_i^2}\left(1-\frac{\partial T_i}{\partial z_i}\right) + \frac{\partial^2 U_i}{\partial c_i\partial z_i}\right] + \frac{\partial U_i}{\partial c_i}\frac{\partial^2 T_i}{\partial\theta\partial z_i}\frac{\partial^2 T_i}{\partial\epsilon\partial z_i}}{\frac{\mathrm{d}^2 U_i}{\mathrm{d}z_i^2}} \right)\mathrm{d}i \\
&\quad - \int \frac{\partial\tilde{g}_i}{\partial u_i}\frac{\partial U_i}{\partial c_i}\frac{\partial T_i}{\partial\theta}\frac{\partial T_i}{\partial\epsilon}\mathrm{d}i + \int \frac{\partial\tilde{g}_i}{\partial z_i}\frac{\left[\frac{\partial^2 U_i}{\partial c_i^2}\left(1-\frac{\partial T_i}{\partial z_i}\right)+\frac{\partial^2 U_i}{\partial c_i\partial z_i}\right]\frac{\partial T_i}{\partial\theta}\frac{\partial T_i}{\partial\epsilon}}{\frac{\mathrm{d}^2 U_i}{\mathrm{d}z_i^2}}\mathrm{d}i
\end{aligned}
\tag{D.17}
$$

The following table explicitly defines the shorthand terms in the in the expression for $C$.

$$
\begin{array}{ll}
g_i = g_i(\theta_0,\epsilon_0) & \frac{\partial\tilde{g}_i}{\partial u_i} = \frac{\partial}{\partial u_i}\tilde{g}_i(U_i(\theta_0,\epsilon_0), z_i(\theta_0,\epsilon_0)) \\
\frac{\partial\tilde{g}_i}{\partial z_i} = \frac{\partial}{\partial z_i}\tilde{g}_i(U_i(\theta_0,\epsilon_0), z_i(\theta_0,\epsilon_0)) & \frac{\partial U_i}{\partial c_i} = \frac{\partial}{\partial c_i}U_i(c_i(\theta_0,\epsilon_0), z_i(\theta_0,\epsilon_0)) \\
\frac{\partial^2 U_i}{\partial c_i^2} = \frac{\partial^2}{\partial c_i^2}U_i(c_i(\theta_0,\epsilon_0), z_i(\theta_0,\epsilon_0)) & \frac{\partial^2 U_i}{\partial c_i\partial z_i} = \frac{\partial^2}{\partial c_i\partial z_i}U_i(c_i(\theta_0,\epsilon_0), z_i(\theta_0,\epsilon_0)) \\
\frac{\mathrm{d}^2 U_i}{\mathrm{d}z_i^2} = \left.\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}\right|_{z_i=z_i(\theta_0,\epsilon_0)} U_i(z_i - T(z_i), z_i) & \frac{\partial T_i}{\partial z_i} = \frac{\partial}{\partial z_i}T(z_i(\theta_0,\epsilon_0),\theta_0,\epsilon_0) \\
\frac{\partial T_i}{\partial\theta} = \left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} T(z_i(\theta_0,\epsilon_0),\theta,\epsilon_0) & \frac{\partial T_i}{\partial\epsilon} = \left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} T(z_i(\theta_0,\epsilon_0),\theta_0,\epsilon) \\
\frac{\partial^2 T_i}{\partial\theta\partial z_i} = \left.\frac{\partial^2}{\partial\theta\partial z_i}\right|_{\theta=\theta_0, z_i=z_i(\theta_0,\epsilon_0)} T(z_i,\theta,\epsilon_0) & \frac{\partial^2 T_i}{\partial\epsilon\partial z_i} = \left.\frac{\partial^2}{\partial\epsilon\partial z_i}\right|_{\epsilon=\epsilon_0, z_i=z_i(\theta_0,\epsilon_0)} T(z_i,\theta_0,\epsilon) \\
\frac{\partial^2 T_i}{\partial\theta\partial\epsilon} = \left.\frac{\partial^2}{\partial\theta\partial\epsilon}\right|_{\theta=\theta_0,\epsilon=\epsilon_0} T(z_i(\theta_0,\epsilon_0),\theta,\epsilon) &
\end{array}
\tag{D.18}
$$

The proof of Lemma D.1 relies on several facts. First observe that, by (D.4),

$$
\frac{\partial}{\partial z_i}k_i(u_i,z_i) = \frac{\partial^2}{\partial c_i^2}U_i(\tilde{c}_i(u_i,z_i),z_i)\frac{\partial}{\partial z_i}\tilde{c}_i(u_i,z_i) + \frac{\partial^2}{\partial z_i\partial c_i}U_i(\tilde{c}_i(u_i,z_i),z_i).
$$

It follows from agent $i$'s first order condition that $1-T'(z_i(\theta_0,\epsilon_0)) = -\frac{\frac{\partial}{\partial z_i}U_i(c_i(\theta_0,\epsilon_0),z_i(\theta_0,\epsilon_0))}{\frac{\partial}{\partial c_i}U_i(c_i(\theta_0,\epsilon_0),z_i(\theta_0,\epsilon_0))}$. Using (D.2) and the fact that $\tilde{c}_i(U_i(\theta_0,\epsilon_0), z_i(\theta_0,\epsilon_0)) = c_i(\theta_0,\epsilon_0)$ and the abbreviations in (D.18), we have

$$
\frac{\partial}{\partial z_i}k_i(U_i(\theta_0,\epsilon_0), z_i(\theta_0,\epsilon_0)) = \frac{\partial^2 U_i}{\partial c_i^2}\left(1-\frac{\partial T_i}{\partial z_i}\right) + \frac{\partial^2 U_i}{\partial c_i\partial z_i}. \tag{D.19}
$$

Moreover, applying the implicit function theorem to the agent's first order conditions, using the abbreviations in (D.18), we have

$$\begin{aligned}\frac{\partial}{\partial\theta}z\left(\theta_0,\epsilon_0\right)=&\frac{\left[\frac{\partial^2 U_i}{\partial c_i^2}\left(1-\frac{\partial T_i}{\partial z_i}\right)+\frac{\partial^2 U_i}{\partial c_i\partial z_i}\right]\frac{\partial T_i}{\partial\theta}+\frac{\partial U_i}{\partial c_i}\frac{\partial^2 T_i}{\partial\theta\partial z_i}}{\frac{\mathrm{d}^2 U_i}{\mathrm{d}z_i^2}},\\ \frac{\partial z_i}{\partial\epsilon}\left(\theta_0,\epsilon_0\right)=&\frac{\left[\frac{\partial^2 U_i}{\partial c_i^2}\left(1-\frac{\partial T_i}{\partial z_i}\right)+\frac{\partial^2 U_i}{\partial c_i\partial z_i}\right]\frac{\partial T_i}{\partial\epsilon}+\frac{\partial U_i}{\partial c_i}\frac{\partial^2 T_i}{\partial\epsilon\partial z_i}}{\frac{\mathrm{d}^2 U_i}{\mathrm{d}z_i^2}}.\end{aligned}\tag{D.20}$$

Using the envelope theorem, (D.6), (D.19), and (D.20), it is staightforward to show that when $A, B$, and $C$ are defined as in (D.17), then (21) and (22) hold, and then the argument for Lemma D.1 proceeds similarly to the argument for Lemma 2.

#### D.4.2.4 Lemma 3

Lemma 3 needs to be modified as follows for the non-quasilinear case:

**Lemma D.2** *Let $T$ be a regular tax policy and let $i_a, i_b \in (0,1)$ be such that $i_a < i_b$. Then there exists a well-behaved family $\left(T^{\theta,\epsilon}\right)$ with $T^{\theta_0,\epsilon_0} = T$ for some interior parameter values $\theta_0, \epsilon_0$ and that satisfies (15), (16), and*

$$\begin{aligned}&\left.\frac{\partial^2}{\partial\theta\partial z_i}\right|_{\theta=\theta_0,z_i=z_i(\theta_0,\epsilon_0)} T\left(z_i,\theta,\epsilon_0\right)\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} T\left(z_i\left(\theta_0,\epsilon_0\right),\theta_0,\epsilon\right)\\ &-\left.\frac{\partial^2}{\partial\epsilon\partial z_i}\right|_{\epsilon=\epsilon_0,z_i=z_i(\theta_0,\epsilon_0)} T\left(z_i,\theta_0,\epsilon\right)\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} T\left(z_i\left(\theta_0,\epsilon_0\right),\theta,\epsilon_0\right)\end{aligned}\begin{cases}>0, & \text{if } i\in\left(i_a,i_b\right),\\ =0, & \text{if } i\notin\left(i_a,i_b\right).\end{cases}\tag{D.21}$$

The reason that the inequality points in opposite directions in Lemmas 3 and D.2 is that, in the lemma preceding Lemma D.2, namely, Lemma D.1, the term $\frac{1}{\left.\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}\right|_{z_i=z_i(\theta_0,\epsilon_0)} U_i(z_i-T(z_i),z_i)}$, which is negative, has been absorbed into $m_i\left(\theta_0,\epsilon_0\right)$, whereas, in Lemma 2, the corresponding term was part of $\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} z_i\left(\theta,\epsilon_0\right)$ and $\left.\frac{\partial}{\partial\epsilon}\right|_{\epsilon=\epsilon_0} z_i\left(\theta_0,\epsilon\right)$.[21] In any event, just as in the quasilinear case, it was possible, with a slight modification in the construction to flip the inequality in (23) (see Lemma A.2), it is also possible to do the same for (D.21).

The proof of Lemma D.2 is similar to the poof of Lemma 3. The construction of the family $\left(T^{\theta,\epsilon}\right)$ is the same as in Lemma 3; the fact that $i \mapsto z_i(T)$ is increasing, which is used in the construction, now follows from the single-crossing condition. Many other aspects of the argument

[21] In particular, in the quasilinear case, using the fact that, by construction, $T^{\theta_0,\epsilon_0} = T$, we have $\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} z_i\left(\theta,\epsilon_0\right) = -\frac{\left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=\theta_0}\left.\frac{\mathrm{d}}{\mathrm{d}z_i}\right|_{z_i=z_i(\theta_0,\epsilon_0)} U_i(z_i-T(z_i,\theta,\epsilon_0),z_i)}{\left.\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}\right|_{z_i=z_i(\theta_0,\epsilon_0)} U_i(z_i-T(z_i),z_i)}$ and $\left.\frac{\partial}{\partial\theta}\right|_{\epsilon=\epsilon_0} z_i\left(\theta,\epsilon_0\right)$ is similar. Note that, in the non-quasilinear case, the term $\left.\frac{\partial^2}{\partial\theta\partial z_i}\right|_{\theta=\theta_0,z_i=z_i(\theta_0,\epsilon_0)} T\left(z_i,\theta,\epsilon_0\right)$ differs from $\left.\frac{\partial}{\partial\theta}\right|_{\theta=\theta_0} z_i\left(\theta,\epsilon_0\right)$ in a number of ways, and not just in omitting the denominator $\left.\frac{\mathrm{d}^2}{\mathrm{d}z_i^2}\right|_{z_i=z_i(\theta_0,\epsilon_0)} U_i\left(z_i-T\left(z_i\right),z_i\right)$.

are unchanged. As (D.16) and (D.21), unlike (20) and (23), do not feature the terms $\frac{\partial}{\partial\theta}\big|_{\theta=\theta_0} z_i(\theta,\epsilon_0)$ and $\frac{\partial}{\partial\epsilon}\big|_{\epsilon=\epsilon_0} z_i(\theta_0,\epsilon)$, we no longer have to appeal to the conditions (A.37) and (A.38). In place of (A.40), we now derive the condition,

$$
\begin{aligned}
&\forall i \in (i_a, i_b),\\
&\frac{\partial^2}{\partial\theta\partial z_i}\bigg|_{\theta=\theta_0, z_i=z_i(\theta_0,\epsilon_0)} T(z_i,\theta,\epsilon_0)\, \frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=\epsilon_0} T(z_i(\theta_0,\epsilon_0),\theta_0,\epsilon)\\
&-\frac{\partial^2}{\partial\epsilon\partial z_i}\bigg|_{\epsilon=\epsilon_0, z_i=z_i(\theta_0,\epsilon_0)} T(z_i,\theta_0,\epsilon)\, \frac{\partial}{\partial\theta}\bigg|_{\theta=\theta_0} T(z_i(\theta_0,\epsilon_0),\theta,\epsilon_0)\\
&=\overbrace{\mu_1'(z_i(T))}^{+}\overbrace{\eta_1(z_i(T))}^{+}-\left(\overbrace{\eta_1'(z_i(T))}^{+\text{ on }(i_3,i_4),-\text{ on }(i_4,i_5)}\times\overbrace{[\mu_1(z_i(T))-1]}^{-\text{ on }(i_3,i_4),+\text{ on }(i_4,i_5)}\right)>0.
\end{aligned}
\tag{D.22}
$$

The equality (D.22) appeals to similar facts as (A.40) to derive and sign the relevant terms on the right hand side of the equality. The argument that the expression on the right hand side of (D.21) is equal to zero outside of $(i_a, i_b)$ is similar to the corresponding argument in Lemma 3. This completes the summary of how the proof of Lemma D.2 differs from that of Lemma 3.

#### D.4.2.5 Lemma 4

Lemma 4 continues to hold in the non-quasilinear case, and its proof in the non-quasilinear case is very similar to its proof in the quasilinear case. In particular, note that, for each $i$, the terms $m_i(\theta_0,\epsilon_0)$ and $\frac{\partial}{\partial z_i} h_i(U_i(\theta_0,\epsilon_0), z_i(\theta_0,\epsilon_0))$ always have opposite signs when nonzero, and one term is equal to zero if and only if the other is equal to zero as well.

#### D.4.2.6 Lemma 5

The basic structure of the argument for Lemma 5, as explained in Section A.13 of the appendix, is unchanged. However, some of the lemmas supporting Lemma 5 must be modified. In the proof of Lemma A.3, the specific expressions in (B.21) must be modified because they depend on the assumption of quasilinearity, but the equality $\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\big|_{\varepsilon=0} z_i(T+\varepsilon(\Delta T_1 - s'(0)\Delta T_2)) = \frac{\mathrm{d}}{\mathrm{d}\xi}\Big|_{\xi=0} z_i(T+\Delta T^{\xi})$ continues to hold, so the proof can proceed as before. Similarly, in in the proof of Lemma B.1 the specific terms in (B.1) depend on quasilinearity but $\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\big|_{\varepsilon=0} z_i(T+\varepsilon(r_1\Delta T^{\gamma}+r_2\Delta T_2)) = r_1\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\big|_{\varepsilon=0} z_i(T+\varepsilon\Delta T_1) + r_2\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\big|_{\varepsilon=0} z_i(T+\varepsilon\Delta T_2)$ still holds, and so again the proof can proceed as before. In Lemma B.2, (B.2) becomes

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} R(T+\varepsilon\Delta T) = \int_{z_0}^{z_1} \Delta T(z)\,\ell_T(z)\,\mathrm{d}z - \int_{z_0}^{z_1} \Delta T'(z)\,k_T(z)\,\mathrm{d}z,$$

where

$$\ell_T(z) = \left[1 - \frac{\left[\frac{\partial^2}{\partial c^2} U_{\iota(z)}(z - T(z), z)(1 - T'(z)) + \frac{\partial^2}{\partial c \partial z} U_{\iota(z)}(z - T(z), z)\right] T'(z)}{\left.\frac{\mathrm{d}^2}{\mathrm{d}\tilde{z}^2}\right|_{\tilde{z}=z} U_{\iota(z)}(\tilde{z} - T(\tilde{z}), \tilde{z})}\right] h(z),$$

$$\forall z \in [z_0, z_1],$$

and $k_T(z)$ is modified to become:

$$k_T(z) = \frac{\frac{\partial}{\partial c} U_{\iota(z)}(z - T(z), z) T'(z)}{\left.\frac{\mathrm{d}^2}{\mathrm{d}\tilde{z}^2}\right|_{\tilde{z}=z} U_{\iota(z)}(\tilde{z} - T(\tilde{z}), \tilde{z})} h(z), \quad \forall z \in [z_0, z_1].$$

Accordingly, in the proof of Lemma B.3, (B.9) becomes

$$\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} R(T + \varepsilon \Delta T_1^\gamma) = \int_{z_0}^{z_*^\gamma} \Delta T_1^\gamma(z) \ell_T(z) \, \mathrm{d}z - \int_{z_0}^{z_*^\gamma} \frac{\mathrm{d}}{\mathrm{d}z} \Delta T_1^\gamma(z) k_T(z) \, \mathrm{d}z,$$

and (B.10) becomes

$$\lim_{\gamma \to \infty} \int_{z_0}^{z_*^\gamma} \Delta T_1^\gamma(z) \ell_T(z) \, \mathrm{d}z = 0.$$

Otherwise the proofs of Lemmas B.2 and B.3 remain the same. Some of the precise details of Lemma C.1 need to be changed, but the basic structure of the argument, which relies on the single-crossing property, remains the same. The proofs of Lemmas A.4, B.4, B.5 and C.2 are unchanged.

### References for Appendix


Hartman, P. (1982), *Ordinary Differential Equations*, second edn, Berkhauser.

Saez, E. and Stantcheva, S. (2016), 'Generalized social marginal welfare weights for optimal tax theory', *American Economic Review* **106**(1), 24–45.

Warner, F. W. (2013), *Foundations of Differentiable Manifolds and Lie Groups*, Springer.