B Proofs

B.1 Proofs of Section A

We start by evaluating the incidence effects of a general perturbation $(d\tau_1, d\tau_2)$ of the baseline tax system.

Proof of equations (58) to (61). The perturbed individual first-order condition of individual θ_i writes

$$(1 - \tau_i - d\tau_i)\, \tilde{w}_i = v'\left(l_i + dl_i\right),$$

where the perturbed wage rate $\tilde{w}_i$ satisfies, to a first-order in $(d\tau_1, d\tau_2)$,

$$\begin{aligned} &\tilde{w}_i\left(\left(l_1 + dl_1\right) F_1, \left(l_2 + dl_2\right) F_2\right) - w_i\left(l_1 F_1, l_2 F_2\right) \\ =&\mathscr{F}_i'\left(\left(l_1 + dl_1\right) F_1, \left(l_2 + dl_2\right) F_2\right) - \mathscr{F}_i'\left(l_1 F_1, l_2 F_2\right) \\ =&\sum_{n=1}^{2} \mathscr{F}_{i,n}''\left(l_1 F_1, l_2 F_2\right) F_n dl_n = \sum_{n=1}^{2} \frac{l_n F_n \mathscr{F}_{i,n}''}{\mathscr{F}_i'} \mathscr{F}_i' \frac{dl_n}{l_n} = \sum_{n=1}^{2} w_i \gamma_{i,n} d\hat{l}_n. \end{aligned}$$

Denoting by $d\hat{w}_i \equiv \frac{\tilde{w}_i - w_i}{w_i}$, this is a linear system of two equations (indexed by $i \in \{1, 2\}$) with two unknowns $(d\hat{w}_1, d\hat{w}_2)$, which can be rewritten as

$$d\hat{w}_i = \gamma_{i,1} d\hat{l}_1 + \gamma_{i,2} d\hat{l}_2,\ \forall i \in \{1, 2\},$$

which immediately leads to the matrix form (59).

We thus get, to a first order in $(d\tau_1, d\tau_2)$,

$$(1-\tau_i)\, w_i + (1-\tau_i)\,(\tilde{w}_i - w_i) - w_i d\tau_i = v'\,(l_i) + v''\,(l_i)\, dl_i,$$

i.e., using the first-order condition at the baseline tax system,

$$\begin{aligned} w_i d\tau_i &= (1-\tau_i)\,(\tilde{w}_i - w_i) - v''\,(l_i)\, dl_i = (1-\tau_i) \sum_{n=1}^{2} w_i \gamma_{i,n} d\hat{l}_n - l_i v''\,(l_i)\, d\hat{l}_i \\ &= \sum_{n=1}^{2} \left[(1-\tau_i)\, w_i \gamma_{i,n} - l_i v''\,(l_i)\, \mathbb{I}_{\{n=i\}}\right] d\hat{l}_n. \end{aligned}$$

Using the formula (57) for the labor supply elasticity, we can rewrite this equation as

$$\begin{aligned} \frac{d\tau_i}{1-\tau_i} &= \sum_{n=1}^{2} \left[\gamma_{i,n} - \frac{l_i v''\,(l_i)}{(1-\tau_i)\, w_i} \mathbb{I}_{\{n=i\}}\right] d\hat{l}_n = \sum_{n=1}^{2} \left[\gamma_{i,n} - \frac{l_i v''\,(l_i)}{v'\,(l_i)} \mathbb{I}_{\{n=i\}}\right] d\hat{l}_n \\ &= \sum_{n=1}^{2} \left[\gamma_{i,n} - \frac{1}{\varepsilon_i} \mathbb{I}_{\{n=i\}}\right] d\hat{l}_n. \end{aligned}$$

This is a linear system of two equations (indexed by $i \in \{1,2\}$) with two unknowns $(d\hat{l}_1, d\hat{l}_2)$, which can be rewritten as

$$\begin{aligned} \left(\gamma_{1,1} - \frac{1}{\varepsilon_1}\right) d\hat{l}_1 + \gamma_{1,2} d\hat{l}_2 &= \frac{d\tau_1}{1-\tau_1} \\ \gamma_{2,1} d\hat{l}_1 + \left(\gamma_{2,2} - \frac{1}{\varepsilon_2}\right) d\hat{l}_2 &= \frac{d\tau_2}{1-\tau_2}, \end{aligned}$$

and hence, in matrix form, as

$$\left[\begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} - \begin{pmatrix} \frac{1}{\varepsilon_1} & 0 \\ 0 & \frac{1}{\varepsilon_2} \end{pmatrix}\right] \begin{pmatrix} d\hat{l}_1 \\ d\hat{l}_2 \end{pmatrix} = \begin{pmatrix} \frac{d\tau_1}{1-\tau_1} \\ \frac{d\tau_2}{1-\tau_2} \end{pmatrix}.$$

Assuming that the matrix in square brackets on the left hand side is invertible, we immediately obtain (58).

The utility of individual θ_i changes, to a first-order in $(d\tau_1, d\tau_2)$, by

$$\begin{aligned} dU_i \equiv\, & U\left((1-\tau_i - d\tau_i)\,\tilde{w}_i\,(l_i + dl_i) - v\,(l_i + dl_i)\right) - U\left((1-\tau_i)\, w_i l_i - v\,(l_i)\right) \\ =\, & \left[-w_i l_i d\tau_i + (1-\tau_i)\, l_i\,(\tilde{w}_i - w_i) + (1-\tau_i)\, w_i dl_i - v'\,(l_i)\, dl_i\right] U'\left((1-\tau_i)\, w_i l_i - v\,(l_i)\right) \\ =\, & (1-\tau_i)\, w_i l_i \left[-\frac{d\tau_i}{1-\tau_i} + d\hat{w}_i + d\hat{l}_i - \frac{v'\,(l_i)}{(1-\tau_i)\, w_i} d\hat{l}_i\right] U'\left((1-\tau_i)\, w_i l_i - v\,(l_i)\right). \end{aligned}$$

But the individual's first order condition (55) implies $\frac{v'(l_i)}{(1-\tau_i)w_i} = 1$, so that we obtain (60).

Government revenue changes, to a first-order in $(d\tau_1, d\tau_2)$, by

$$d\mathscr{R} \equiv \sum_{i=1}^{2} \left[(\tau_i + d\tau_i)\,\tilde{w}_i\,(l_i + dl_i) - \tau_i w_i l_i\right] F_i = \sum_{i=1}^{2} \left[w_i l_i d\tau_i + \tau_i l_i\,(\tilde{w}_i - w_i) + \tau_i w_i dl_i\right] F_i$$
$$= \sum_{i=1}^{2} \tau_i w_i l_i \left(\frac{d\tau_i}{\tau_i} + d\hat{w}_i + d\hat{l}_i\right) F_i,$$

which implies (61).

□

Next we focus on the partial equilibrium case.

Proof of equations (62) and (63). With two types, we can easily invert explicitly the matrix in equation (58) to obtain

$$\begin{pmatrix} d\hat{l}_1 \\ d\hat{l}_2 \end{pmatrix} = \begin{pmatrix} \gamma_{11} - \frac{1}{\varepsilon_1} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} - \frac{1}{\varepsilon_2} \end{pmatrix}^{-1} \begin{pmatrix} \frac{d\tau_1}{1-\tau_1} \\ \frac{d\tau_2}{1-\tau_2} \end{pmatrix} = \frac{\begin{pmatrix} \gamma_{22} - \frac{1}{\varepsilon_2} & -\gamma_{12} \\ -\gamma_{21} & \gamma_{11} - \frac{1}{\varepsilon_1} \end{pmatrix} \begin{pmatrix} \frac{d\tau_1}{1-\tau_1} \\ \frac{d\tau_2}{1-\tau_2} \end{pmatrix}}{\left(\gamma_{11} - \frac{1}{\varepsilon_1}\right)\left(\gamma_{22} - \frac{1}{\varepsilon_2}\right) - \gamma_{12}\gamma_{21}}.$$

Consider a perturbation of τ_2 only, i.e., $d\tau_2 = 0$. Then the previous expression implies that the changes in labor supplies are given by

$$\begin{pmatrix} d\hat{l}_1 \\ d\hat{l}_2 \end{pmatrix} = \frac{\begin{pmatrix} \gamma_{22} - \frac{1}{\varepsilon_2} \\ -\gamma_{21} \end{pmatrix} \frac{d\tau_1}{1-\tau_1}}{\left(\gamma_{11} - \frac{1}{\varepsilon_1}\right)\left(\gamma_{22} - \frac{1}{\varepsilon_2}\right) - \gamma_{12}\gamma_{21}} = \frac{\begin{pmatrix} \varepsilon_2\gamma_{22} - 1 \\ -\varepsilon_2\gamma_{21} \end{pmatrix} \frac{d\tau_1}{1-\tau_1}}{\left(\gamma_{11} - \frac{1}{\varepsilon_1}\right)(\varepsilon_2\gamma_{22} - 1) - \varepsilon_2\gamma_{12}\gamma_{21}}.$$

Now take the limit as type-θ_2 labor supply becomes inelastic, i.e., $\varepsilon_2 \to 0$. We get

$$\begin{pmatrix} d\hat{l}_1 \\ d\hat{l}_2 \end{pmatrix} \to \frac{\frac{d\tau_1}{1-\tau_1}}{\frac{1}{\varepsilon_1} - \gamma_{11}} \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \text{ i.e., } \begin{cases} d\hat{l}_1 = & -\frac{1}{\frac{1}{\varepsilon_1} - \gamma_{11}} \frac{d\tau_1}{1-\tau_1}, \\ d\hat{l}_2 = & 0, \end{cases}$$

which shows (63). Finally, (59) implies that

$$d\hat{w}_1 = \gamma_{1,1} d\hat{l}_1 = -\frac{\gamma_{1,1}}{\frac{1}{\varepsilon_1} - \gamma_{11}} \frac{d\tau_1}{1-\tau_1}$$

and $d\hat{w}_2 = \gamma_{2,1} d\hat{l}_1 = \frac{\gamma_{2,1}}{\gamma_{1,1}} d\hat{w}_1$. The Euler's homogeneous equation theorem states that

$$w_1 L_1 + w_2 L_2 = F(L_1, L_2)$$

so that

$$w_1\gamma_{1,1} + w_2\gamma_{2,1} = L_1 \frac{\partial w_1}{\partial L_1} + L_2 \frac{\partial w_2}{\partial L_1} = 0,$$

and hence $d\hat{w}_2 = -\frac{w_1}{w_2} d\hat{w}_1$.

□

B.2 Proofs of Section 1

B.2.1 Proof of equations (6), (7), and (8)

We first derive the expressions for the labor supply elasticities.

Proof. We first derive the labor supply elasticity along the linear budget constraint (6). Rewrite the first-order condition (1) as

$$v'\left(l\left(\theta\right)\right)=\left(1-\tau\left(\theta\right)\right)w\left(\theta\right),$$

where $\tau\left(\theta\right)=T'\left(w\left(\theta\right)l\left(\theta\right)\right)$ is the marginal tax rate of agent θ. The first-order effect of perturbing the marginal tax rate $\tau\left(\theta\right)$ by $d\tau_{\theta}$ on the labor supply $l\left(\theta\right)$, in partial equilibrium (i.e., keeping $w\left(\theta\right)$ constant), is obtained by a Taylor approximation of the first-order condition characterizing the perturbed equilibrium,

$$v'\left(l\left(\theta\right)+dl_{\theta}\right)=\left(1-\tau\left(\theta\right)-d\tau_{\theta}\right)w\left(\theta\right),$$

around the baseline equilibrium. We obtain

$$v'\left(l\left(\theta\right)\right)+v''\left(l\left(\theta\right)\right)dl_{\theta}=\left(1-\tau\left(\theta\right)\right)w\left(\theta\right)-w\left(\theta\right)d\tau_{\theta}=v'\left(l\left(\theta\right)\right)-\frac{v'\left(l\left(\theta\right)\right)}{1-\tau\left(\theta\right)}d\tau_{\theta},$$

and thus

$$\frac{dl_{\theta}}{l\left(\theta\right)}=-\frac{v'\left(l\left(\theta\right)\right)}{l\left(\theta\right)v''\left(l\left(\theta\right)\right)}\frac{d\tau_{\theta}}{1-\tau\left(\theta\right)},$$

which immediately leads equation (6).

Next we derive the labor supply elasticity along the non-linear budget constraint (6), keeping the wage constant. The perturbed individual first-order condition writes

$$v'\left(l\left(\theta\right)+dl_{\theta}\right)=\left[1-T'\left(w\left(\theta\right)\left(l\left(\theta\right)+dl_{\theta}\right)\right)-d\tau_{\theta}\right]w\left(\theta\right).$$

A first-order Taylor expansion implies

$$v'\left(l\left(\theta\right)\right)+v''\left(l\left(\theta\right)\right)dl_{\theta}=\left(1-T'\left(w\left(\theta\right)l\left(\theta\right)\right)\right)w\left(\theta\right)-T''\left(w\left(\theta\right)l\left(\theta\right)\right)\left(w\left(\theta\right)\right)^{2}dl_{\theta}-w\left(\theta\right)d\tau_{\theta},$$

i.e.,

$$\begin{aligned}\frac{dl_{\theta}}{l\left(\theta\right)}&=-\frac{1}{l\left(\theta\right)}\frac{w\left(\theta\right)}{v''\left(l\left(\theta\right)\right)+T''\left(w\left(\theta\right)l\left(\theta\right)\right)w_{\theta}^{2}}d\tau_{\theta}\\&=-\frac{\frac{v'\left(l\left(\theta\right)\right)}{l\left(\theta\right)}}{\left(1-T'\left(w\left(\theta\right)l\left(\theta\right)\right)\right)v''\left(l\left(\theta\right)\right)+v'\left(l\left(\theta\right)\right)w\left(\theta\right)T''\left(w\left(\theta\right)l\left(\theta\right)\right)}d\tau_{\theta}\\&=\frac{\varepsilon_{l,1-\tau}\left(\theta\right)\left(1-T'\left(w\left(\theta\right)l\left(\theta\right)\right)\right)}{1-T'\left(w\left(\theta\right)l\left(\theta\right)\right)+\varepsilon_{l,1-\tau}\left(\theta\right)w\left(\theta\right)l\left(\theta\right)T''\left(w\left(\theta\right)l\left(\theta\right)\right)}\frac{d\left(1-\tau\left(\theta\right)\right)}{1-T'\left(w\left(\theta\right)l\left(\theta\right)\right)},\end{aligned}$$

which yields equation (7).

Finally, we derive the partial equilibrium labor supply elasticity with respect to the wage along the non-linear budget constraint (8). The perturbed individual first-order condition writes

$$v'\left(l\left(\theta\right)+dl_{\theta}\right)=\left[1-T'\left(\left(w\left(\theta\right)+dw_{\theta}\right)\left(l\left(\theta\right)+dl_{\theta}\right)\right)\right]\left(w\left(\theta\right)+dw_{\theta}\right).$$

A first-order Taylor expansion then implies

$$\begin{aligned}v'\left(l\left(\theta\right)\right)+v''\left(l\left(\theta\right)\right)dl_{\theta}=&\left(1-T'\left(w\left(\theta\right)l\left(\theta\right)\right)\right)w\left(\theta\right)-T''\left(w\left(\theta\right)l\left(\theta\right)\right)\left(w\left(\theta\right)\right)^{2}dl_{\theta}\\&-T''\left(w\left(\theta\right)l\left(\theta\right)\right)w\left(\theta\right)l\left(\theta\right)dw_{\theta}+\left(1-T'\left(w\left(\theta\right)l\left(\theta\right)\right)\right)dw_{\theta},\end{aligned}$$

i.e.,

$$\begin{aligned}\frac{dl_{\theta}}{l\left(\theta\right)}=&\frac{\left(1-T'\left(w\left(\theta\right)l\left(\theta\right)\right)-w\left(\theta\right)l\left(\theta\right)T''\left(w\left(\theta\right)l\left(\theta\right)\right)\right)w\left(\theta\right)}{v''\left(l\left(\theta\right)\right)+\left(w\left(\theta\right)\right)^{2}T''\left(w\left(\theta\right)l\left(\theta\right)\right)}\frac{dw_{\theta}}{w\left(\theta\right)l\left(\theta\right)}\\=&\frac{\left(1-T'\left(w\left(\theta\right)l\left(\theta\right)\right)-w\left(\theta\right)l\left(\theta\right)T''\left(w\left(\theta\right)l\left(\theta\right)\right)\right)\frac{v'\left(l\left(\theta\right)\right)}{l\left(\theta\right)v''\left(l\left(\theta\right)\right)}}{\left(1-T'\left(w\left(\theta\right)l\left(\theta\right)\right)\right)+w\left(\theta\right)l\left(\theta\right)\frac{v'\left(l\left(\theta\right)\right)}{l\left(\theta\right)v''\left(l\left(\theta\right)\right)}T''\left(w\left(\theta\right)l\left(\theta\right)\right)}\frac{dw_{\theta}}{w\left(\theta\right)},\end{aligned}$$

which implies (8).

□

B.2.2 Proofs for Examples 1 and 2

We first provide algebra details for the CES technology. It is straightforward to show that the cross-wage elasticities are given by (15). We have

$$\bar{\gamma}\left(\theta,\theta^{*}\right)=\left(1-\rho\right)\frac{a\left(\theta^{*}\right)L\left(\theta^{*}\right)^{\rho}}{\int_{\Theta}a\left(x\right)L\left(x\right)^{\rho}dx},$$

which implies, for all $\theta\in\Theta$,

$$\int_{\Theta}\bar{\gamma}\left(\theta,\theta^{*}\right)d\theta^{*}=1-\rho=\frac{1}{\sigma}.$$

Moreover, denoting interchangeably by $\bar{\gamma}\left(y,y^{*}\right)\equiv\bar{\gamma}\left(\theta,\theta^{*}\right)$ where $y=y\left(\theta\right)$ and $y^{*}=y\left(\theta^{*}\right)$, Euler's homogeneous function theorem writes

$$\int_{\Theta}\gamma\left(\theta,\theta^{*}\right)y\left(\theta\right)f_{\theta}\left(\theta\right)d\theta=0,$$

which can be rewritten as

$$\int_{\Theta}\bar{\gamma}\left(\theta,\theta^{*}\right)y\left(\theta\right)f_{\theta}\left(\theta\right)d\theta+\bar{\bar{\gamma}}\left(\theta^{*},\theta^{*}\right)y\left(\theta^{*}\right)f_{\theta}\left(\theta^{*}\right)=0,$$

or equivalently,

$$\int_{\mathbb{R}_{+}}\bar{\gamma}\left(y,y^{*}\right)yf_{y}\left(y\right)dy-\left(1-\rho\right)y^{*}f_{y}\left(y^{*}\right)\frac{dy\left(\theta^{*}\right)}{d\theta}=0.$$

Since $\bar{\gamma}(y, y^*) = \bar{\gamma}(\theta, \theta^*)$ does not depend on y (or θ), this implies

$$(1-\rho) \frac{a(\theta^*) L(\theta^*)^\rho}{\int_\Theta a(x) L(x)^\rho dx} \int_{\mathbb{R}_+} y f_y(y) dy - (1-\rho) y^* f_y(y^*) \frac{dy(\theta^*)}{d\theta} = 0,$$

i.e.,

$$\frac{a(\theta^*) L(\theta^*)^\rho}{\int_\Theta a(x) L(x)^\rho dx} = \frac{y^* f_y(y^*)}{\int_{\mathbb{R}_+} y f_y(y) dy} \frac{dy(\theta^*)}{d\theta}.$$

Substituting in expression (15), we hence get

$$\bar{\gamma}(y, y(\theta^*)) = (1-\rho) \frac{y(\theta^*) f_y(y(\theta^*))}{\int_{\mathbb{R}_+} y f_y(y) dy} \frac{dy(\theta^*)}{d\theta}. \tag{64}$$

We now provide the algebra details for the Translog production function.

Proof. First, we show that if the conditions θ, θ', $\int_\Theta a(\theta') d\theta' = 1$, $\bar{\beta}(\theta, \theta') = \bar{\beta}(\theta', \theta)$, and $\bar{\bar{\beta}}(\theta, \theta) = -\int_\Theta \bar{\beta}(\theta, \theta') d\theta'$ hold, then the production function (16) has constant returns to scale. Indeed, we can then write

$$\begin{aligned} \ln \mathscr{F}(\lambda \mathscr{L}) = & \ln \mathscr{F}(\mathscr{L}) + \left(\int_\Theta \alpha_\theta d\theta \right) \ln \lambda + \frac{1}{2} (\ln \lambda)^2 \left(\int_\Theta \bar{\bar{\beta}}_{\theta,\theta} d\theta + \int_{\Theta \times \Theta} \bar{\beta}_{\theta,\theta'} d\theta d\theta' \right) \\ & + \ln \lambda \left(\int_\Theta \alpha_\theta d\theta + \int_\Theta \bar{\bar{\beta}}_{\theta,\theta} \ln L(\theta) d\theta + \int_{\Theta \times \Theta} \bar{\beta}_{\theta,\theta'} \ln L(\theta) d\theta d\theta' \right) \end{aligned}$$

which is equal to $\ln \mathscr{F}(\mathscr{L}) + \ln \lambda$.

Second, we derive expression (17) for the wage. We have

$$\begin{aligned} & \ln \mathscr{F}(\mathscr{L} + \mu \delta_{\theta^*}) - \ln \mathscr{F}(\mathscr{L}) \\ = & \int_\Theta \alpha_\theta \left[\ln (L(\theta) + \mu \delta_{\theta^*}(\theta)) - \ln L(\theta) \right] d\theta + \frac{1}{2} \int_\Theta \bar{\bar{\beta}}_{\theta,\theta} \left[\ln^2 (L(\theta) + \mu \delta_{\theta^*}(\theta)) - \ln^2 L(\theta) \right] d\theta \\ & + \frac{1}{2} \int_{\Theta \times \Theta} \bar{\beta}_{\theta,\theta'} \left[\ln (L(\theta) + \mu \delta_{\theta^*}(\theta)) \ln (L(\theta') + \mu \delta_{\theta^*}(\theta')) - \ln (L(\theta)) \ln (L(\theta')) \right] d\theta d\theta'. \end{aligned}$$

A first-order Taylor approximation of the right-hand side as $\mu \to 0$ yields

$$\begin{aligned} & \ln \mathscr{F}(\mathscr{L} + \mu \delta_{\theta^*}) - \ln \mathscr{F}(\mathscr{L}) \\ \underset{\mu \to 0}{=} & \mu \int_\Theta \alpha_\theta \frac{1}{L(\theta)} \delta_{\theta^*}(\theta) d\theta + \mu \int_\Theta \bar{\bar{\beta}}_{\theta,\theta} \frac{\ln L(\theta)}{L(\theta)} \delta_{\theta^*}(\theta) d\theta \\ & + \frac{1}{2} \mu \int_{\Theta \times \Theta} \bar{\beta}_{\theta,\theta'} \left[\delta_{\theta^*}(\theta) \frac{\ln (L(\theta'))}{L(\theta)} + \delta_{\theta^*}(\theta') \frac{\ln (L(\theta))}{L(\theta')} \right] d\theta d\theta' + o(\mu) \\ = & \mu \left\{ \frac{\alpha_{\theta^*}}{L(\theta^*)} + \bar{\bar{\beta}}_{\theta^*,\theta^*} \frac{\ln L(\theta^*)}{L(\theta^*)} + \frac{1}{2} \int_\Theta \bar{\beta}_{\theta^*,\theta'} \frac{\ln (L(\theta'))}{L(\theta^*)} d\theta' + \frac{1}{2} \int_\Theta \bar{\beta}_{\theta,\theta^*} \frac{\ln (L(\theta))}{L(\theta^*)} d\theta \right\} + o(\mu), \end{aligned}$$

Thus the wage of type θ^* is equal to

$$
\begin{aligned}
w(\theta^*) &= \lim_{\mu\to 0}\frac{1}{\mu}\left[\mathscr{F}(\mathscr{L}+\mu\delta_{\theta^*})-\mathscr{F}(\mathscr{L})\right]\\
&=\frac{\mathscr{F}(\mathscr{L})}{L(\theta^*)}\left\{\alpha_{\theta^*}+\bar{\bar{\beta}}_{\theta^*,\theta^*}\ln L(\theta^*)+\int_\Theta\bar{\beta}_{\theta^*,\theta''}\ln\left(L(\theta'')\right)d\theta''\right\},
\end{aligned}
$$

where the second equality follows from the symmetry of $\left\{\bar{\beta}_{\theta,\theta'}\right\}_{(\theta,\theta')\in\Theta^2}$. Note that the same expression can be obtained heuristically by computing the derivative of $\mathscr{F}\left(\{L(\theta'')\}_{\theta''\in\Theta}\right)$ with respect to $L(\theta^*)$:

$$
w(\theta^*) = \frac{\mathscr{F}(\mathscr{L})}{L(\theta^*)}\frac{\partial\ln\mathscr{F}\left(\{L(\theta'')\}_{\theta''\in\Theta}\right)}{\partial\ln L(\theta^*)}.
$$

Third, we derive the cross-wage elasticities and own-wage elasticities (18). For simplicity we derive these formulas heuristically by directly evaluating the derivatives; they can be easily obtained rigorously following the same steps as above for the wages. For the cross-wage elasticities, suppose that $\theta\neq\theta'$. We have

$$
\begin{aligned}
\bar{\gamma}(\theta,\theta') &= \frac{\partial\ln\left(\mathscr{F}(\mathscr{L})/L(\theta)\right)}{\partial\ln L(\theta')}+\frac{\partial\ln\left\{\alpha_\theta+\bar{\bar{\beta}}_{\theta,\theta}\ln L(\theta)+\int_\Theta\bar{\beta}_{\theta,\theta''}\ln L(\theta'')\,d\theta''\right\}}{\partial\ln L(\theta')}\\
&=\left\{\alpha_{\theta'}+\bar{\bar{\beta}}_{\theta',\theta'}\ln L(\theta')+\int_\Theta\bar{\beta}_{\theta',\theta''}\ln L(\theta'')\,d\theta''\right\}+\frac{\bar{\beta}_{\theta,\theta'}}{\alpha_\theta+\bar{\bar{\beta}}_{\theta,\theta}\ln L(\theta)+\int_\Theta\bar{\beta}_{\theta,\theta''}\ln L(\theta'')\,d\theta''}\\
&=\left(\frac{w(\theta')L(\theta')}{\mathscr{F}(\mathscr{L})}\right)+\left(\frac{w(\theta)L(\theta)}{\mathscr{F}(\mathscr{L})}\right)^{-1}\bar{\beta}_{\theta,\theta'}.
\end{aligned}
$$

Similarly, the own wage elasticities are given by:

$$
\begin{aligned}
\bar{\bar{\gamma}}(\theta,\theta) =& \frac{\partial\ln\left(Y(\mathscr{L})/L(\theta)\right)}{\partial\ln L(\theta)}+\frac{\partial\ln\left\{\alpha_\theta+\bar{\bar{\beta}}_{\theta,\theta}\ln L(\theta)+\int_\Theta\bar{\beta}_{\theta,\theta''}\ln L(\theta'')\,d\theta''\right\}}{\partial\ln L(\theta)}-\bar{\gamma}(\theta,\theta)\\
=&-1+\frac{\bar{\bar{\beta}}_{\theta,\theta}}{\alpha_\theta+\bar{\bar{\beta}}_{\theta,\theta}\ln L(\theta)+\int_\Theta\bar{\beta}_{\theta,\theta''}\ln L(\theta'')\,d\theta''}=-1+\left(\frac{w(\theta)L(\theta)}{\mathscr{F}(\mathscr{L})}\right)^{-1}\bar{\bar{\beta}}_{\theta,\theta}.
\end{aligned}
$$

Finally, note that

$$
\begin{aligned}
\ln\left(\frac{w(\theta)}{w(\theta')}\right) &= \ln\left(\frac{L(\theta')}{L(\theta)}\right)+\ln\frac{\alpha_\theta+\bar{\bar{\beta}}_{\theta,\theta}\ln L(\theta)+\int_\Theta\bar{\beta}_{\theta,\theta''}\ln L(\theta'')\,d\theta''}{\alpha_{\theta^*}+\bar{\bar{\beta}}_{\theta',\theta'}\ln L(\theta')+\int_\Theta\bar{\beta}_{\theta',\theta''}\ln L(\theta'')\,d\theta''}\\
&=\ln\left(\frac{L(\theta')}{L(\theta)}\right)+\ln\frac{\alpha_\theta+\ln L(\theta)\left\{\bar{\bar{\beta}}_{\theta,\theta}+\int_\Theta\bar{\beta}_{\theta,\theta''}d\theta''\right\}+\int_\Theta\bar{\beta}_{\theta,\theta''}\left\{\ln L(\theta'')-\ln L(\theta)\right\}d\theta''}{\alpha_{\theta^*}+\ln L(\theta')\left\{\bar{\bar{\beta}}_{\theta',\theta'}+\int_\Theta\bar{\beta}_{\theta',\theta''}d\theta''\right\}+\int_\Theta\bar{\beta}_{\theta',\theta''}\left\{\ln L(\theta'')-\ln L(\theta')\right\}d\theta''}\\
&=\ln\left(\frac{L(\theta')}{L(\theta)}\right)+\ln\frac{\alpha_\theta+\int_\Theta\bar{\beta}_{\theta,\theta''}\ln\left(L(\theta'')/L(\theta)\right)d\theta''}{\alpha_{\theta'}+\int_\Theta\bar{\beta}_{\theta',\theta''}\ln\left(L(\theta'')/L(\theta')\right)d\theta''},
\end{aligned}
$$

so that the elasticities of substitution are given by:

$$
\begin{aligned}
&-\frac{1}{\sigma(\theta,\theta')}=\frac{\partial\ln\left(w(\theta)/w(\theta')\right)}{\partial\ln\left(L(\theta)/L(\theta')\right)}\\
=&-1-\frac{\bar{\beta}_{\theta,\theta'}}{\alpha_{\theta}+\int_{\Theta}\bar{\beta}_{\theta,\theta''}\ln\left(\frac{L(\theta'')}{L(\theta)}\right)d\theta''}-\frac{\bar{\beta}_{\theta',\theta}}{\alpha_{\theta'}+\int_{\Theta}\bar{\beta}_{\theta',\theta''}\ln\left(\frac{L(\theta'')}{L(\theta')}\right)d\theta''}\\
=&-1-\frac{\bar{\beta}_{\theta,\theta'}}{\alpha_{\theta}+\bar{\bar{\beta}}_{\theta,\theta}\ln L(\theta)+\int_{\Theta}\bar{\beta}_{\theta,\theta''}\ln L(\theta'')\,d\theta''}-\frac{\bar{\beta}_{\theta',\theta}}{\alpha_{\theta'}+\bar{\bar{\beta}}_{\theta',\theta'}\ln L(\theta')+\int_{\Theta}\bar{\beta}_{\theta',\theta''}\ln L(\theta'')\,d\theta''}\\
=&-1-\left[\left(\frac{w(\theta)L(\theta)}{\mathscr{F}(\mathscr{L})}\right)^{-1}+\left(\frac{w(\theta')L(\theta')}{\mathscr{F}(\mathscr{L})}\right)^{-1}\right]\bar{\beta}_{\theta,\theta'}.
\end{aligned}
$$

This concludes the derivations of the wage elasticities for a general Translog production function.

□

B.3 Proofs of Section 2

B.3.1 Proof of Lemma 1

Proof. Denote the perturbed tax function by $\tilde{T}(y)=T(y)+\mu h(y)$ (later we let $\mu\to 0$). Denote by $\mathrm{d}l(\theta,h)$ the Gateaux derivative of the labor supply of type θ in response to this perturbation, and let $\mathrm{d}L(\theta,h)\equiv\mathrm{d}l(\theta,h)\,f_{\theta}(\theta)$. The labor supply response $\mathrm{d}l(\theta,h)$ of type θ is given by the solution to the perturbed first-order condition:

$$
\begin{aligned}
0=&v'\left(l(\theta)+\mu\mathrm{d}l(\theta,h)\right)\\
&-\left\{1-T'\left[\tilde{w}(\theta)\times\left(l(\theta)+\mu\mathrm{d}l(\theta,h)\right)\right]-\mu h'\left[\tilde{w}(\theta)\times\left(l(\theta)+\mu\mathrm{d}l(\theta,h)\right)\right]\right\}\tilde{w}(\theta),
\end{aligned}\tag{65}
$$

where $\tilde{w}(\theta)$ is the perturbed wage schedule. Heuristically, $\tilde{w}(\theta)$ satisfies

$$
\begin{aligned}
&\tilde{w}(\theta)-w(\theta)=\mathscr{F}'_{L(\theta)}\left(\{(l(\theta')+\mu\mathrm{d}l(\theta',h))\,f_{\theta}(\theta')\}_{\theta'\in\Theta}\right)-\mathscr{F}'_{L(\theta)}\left(\{l(\theta')\,f_{\theta}(\theta')\}_{\theta'\in\Theta}\right)\\
=&\int_{\Theta}\mathscr{F}''_{L(\theta),L(\theta')}\mu\mathrm{d}l(\theta',h)\,f_{\theta}(\theta')\,d\theta'+o(\mu)=\mu\mathscr{F}'_{L(\theta)}\int_{\Theta}\frac{L(\theta')\,\mathscr{F}''_{L(\theta),L(\theta')}}{\mathscr{F}'_{L(\theta)}}\mathrm{d}\hat{l}(\theta',h)\,d\theta'+o(\mu),
\end{aligned}
$$

so that, to a first order as $\mu\to 0$,

$$
\tilde{w}(\theta)-w(\theta)=\mu w(\theta)\int_{\Theta}\gamma(\theta,\theta')\,\mathrm{d}\hat{l}(\theta',h)\,d\theta'.\tag{66}
$$

To derive this equation formally, denote by $\mathrm{d}\mathscr{L}(h)\in\mathcal{M}$ the measure on Θ defined by the Gateaux derivative of the measure $\mathscr{L}$ in the direction h. We have $\mathrm{d}\mathscr{L}(h)=\int_{\Theta}\mathrm{d}L(\theta',h)\,\delta_{\theta'}d\theta'$. We then have

$$
\tilde{w}(\theta)-w(\theta)=\omega\left\{\theta,L(\theta)+\mu\mathrm{d}L(\theta,h),\mathscr{L}+\mu\mathrm{d}\mathscr{L}(h)\right\}-\omega\left(\theta,L(\theta),\mathscr{L}\right).
$$

The right hand side of this equation is equal to

$$\begin{aligned}
&\omega_2\left(\theta, L\left(\theta\right), \mathscr{L}\right)\mu \mathrm{d}L\left(\theta, h\right) + \int_\Theta \left\{\omega\left(\theta, L\left(\theta\right), \mathscr{L} + \mu \mathrm{d}L\left(\theta', h\right)\delta_{\theta'}\right) - \omega\left(\theta, L\left(\theta\right), \mathscr{L}\right)\right\} d\theta' \\
=&\bar{\bar{\gamma}}\left(\theta,\theta\right)\frac{w\left(\theta\right)}{L\left(\theta\right)}\mu \mathrm{d}L\left(\theta, h\right) + \int_\Theta \bar{\gamma}\left(\theta,\theta'\right)\frac{w\left(\theta\right)}{L\left(\theta'\right)}\mu \mathrm{d}L\left(\theta', h\right) d\theta' + o\left(\mu\right) \\
=&\mu w\left(\theta\right)\left\{\bar{\bar{\gamma}}\left(\theta,\theta\right)\mathrm{d}\hat{l}\left(\theta,h\right) + \int_\Theta \bar{\gamma}\left(\theta,\theta'\right)\mathrm{d}\hat{l}\left(\theta',h\right)d\theta'\right\} + o\left(\mu\right),
\end{aligned}$$

which leads to expression (66).

Next, for any function g, we have, to a first order as $\mu \to 0$,

$$\begin{aligned}
&g\left[\tilde{w}\left(\theta\right)\left(l\left(\theta\right) + \mu \mathrm{d}l\left(\theta,h\right)\right)\right] - g\left[w\left(\theta\right)l\left(\theta\right)\right] \\
=&\left\{\left[\tilde{w}\left(\theta\right) - w\left(\theta\right)\right]l\left(\theta\right) + w\left(\theta\right)\mu \mathrm{d}l\left(\theta,h\right)\right\}g'\left(w\left(\theta\right)l\left(\theta\right)\right) \\
=&\mu\left\{\int_\Theta \gamma\left(\theta,\theta'\right)\mathrm{d}\hat{l}\left(\theta',h\right)d\theta' + \mathrm{d}\hat{l}\left(\theta,h\right)\right\}w\left(\theta\right)l\left(\theta\right)g'\left(w\left(\theta\right)l\left(\theta\right)\right).
\end{aligned}$$

A first-order Taylor expansion of the perturbed first-order conditions (65) around the baseline allocation then yields:

$$\begin{aligned}
0 =&v''\left(l\left(\theta\right)\right)\mu \mathrm{d}l\left(\theta,h\right) + \mu h'\left(w\left(\theta\right)l\left(\theta\right)\right)w\left(\theta\right) \\
&- \left\{\left[\tilde{w}\left(\theta\right) - w\left(\theta\right)\right]\left(1 - T'\left(w\left(\theta\right)l\left(\theta\right)\right)\right) - \left[T'\left(\tilde{w}\left(\theta\right)\left(l\left(\theta\right) + \mu \mathrm{d}l\left(\theta,h\right)\right)\right) - T'\left(w\left(\theta\right)l\left(\theta\right)\right)\right]w\left(\theta\right)\right\}.
\end{aligned}$$

Using (66), we obtain

$$\begin{aligned}
0 =&\frac{l\left(\theta\right)v''\left(l\left(\theta\right)\right)}{w\left(\theta\right)}\mu \mathrm{d}\hat{l}\left(\theta,h\right) + \mu h'\left(w\left(\theta\right)l\left(\theta\right)\right) - \mu\left[\int_\Theta \gamma\left(\theta,\theta'\right)\mathrm{d}\hat{l}\left(\theta',h\right)d\theta'\right]\left(1 - T'\left(w\left(\theta\right)l\left(\theta\right)\right)\right) \\
&+ \mu\left[\int_\Theta \gamma\left(\theta,\theta'\right)\mathrm{d}\hat{l}\left(\theta',h\right)d\theta' + \mathrm{d}\hat{l}\left(\theta,h\right)\right]w\left(\theta\right)l\left(\theta\right)T''\left(w\left(\theta\right)l\left(\theta\right)\right).
\end{aligned}$$

Solving for $\mathrm{d}\hat{l}\left(\theta,h\right)$,

$$\begin{aligned}
\mathrm{d}\hat{l}\left(\theta,h\right) =&\frac{\left[1 - T'\left(y\left(\theta\right)\right) - y\left(\theta\right)T''\left(y\left(\theta\right)\right)\right]\int_\Theta \bar{\gamma}\left(\theta,\theta'\right)\mathrm{d}\hat{l}\left(\theta',h\right)d\theta' - h'\left(y\left(\theta\right)\right)}{\left[\left(1 - T'\left(y\left(\theta\right)\right)\right)\frac{l(\theta)v''(l(\theta))}{v'(l(\theta))} + y\left(\theta\right)T''\left(y\left(\theta\right)\right)\right] - \left[1 - T'\left(y\left(\theta\right)\right) - y\left(\theta\right)T''\left(y\left(\theta\right)\right)\right]\bar{\bar{\gamma}}\left(\theta,\theta\right)} \\
=&\frac{\frac{1-T'(y(\theta))-y(\theta)T''(y(\theta))}{1-T'(y(\theta))+\varepsilon_{l,1-\tau}(\theta)y(\theta)T''(y(\theta))}\varepsilon_{l,1-\tau}\left(\theta\right)}{1 - \frac{1-T'(y(\theta))-y(\theta)T''(y(\theta))}{1-T'(y(\theta))+\varepsilon_{l,1-\tau}(\theta)y(\theta)T''(y(\theta))}\varepsilon_{l,1-\tau}\left(\theta\right)\bar{\bar{\gamma}}\left(\theta,\theta\right)}\int_\Theta \bar{\gamma}\left(\theta,\theta'\right)\mathrm{d}\hat{l}\left(\theta',h\right)d\theta' \\
&- \frac{\frac{\varepsilon_{l,1-\tau}(\theta)}{1-T'(y(\theta))+\varepsilon_{l,1-\tau}(\theta)y(\theta)T''(y(\theta))}}{1 - \frac{1-T'(y(\theta))-y(\theta)T''(y(\theta))}{1-T'(y(\theta))+\varepsilon_{l,1-\tau}(\theta)y(\theta)T''(y(\theta))}\varepsilon_{l,1-\tau}\left(\theta\right)\bar{\bar{\gamma}}\left(\theta,\theta\right)}h'\left(y\left(\theta\right)\right),
\end{aligned}$$

which leads to equation (21).

□

B.3.2 Proof of Proposition 1

We now derive the general solution (22) of the integral equation (21). Assume that the condition $\int_{\Theta^2} |K_1(\theta,\theta')|^2 \, d\theta d\theta' < 1$ holds.

Proof. For ease of exposition, denote by $\hat{h}'(\theta) = \frac{h'(y(\theta))}{1-T'(y(\theta))}$, and $g(\theta) \equiv \mathrm{d}\hat{l}(\theta,h)$. The integral equation writes

$$g(\theta) = -\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}(\theta) + \int_\Theta K_1(\theta,\theta')\, g(\theta')\, d\theta',$$

where $K_1(\theta,\theta') = \tilde{E}_{l,w}(\theta)\,\bar{\gamma}(\theta,\theta')$. Substituting for $g(\theta')$ in the integral using the r.h.s. of the integral equation yields:

$$\begin{aligned}
g(\theta) =& -\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(\theta) + \int_\Theta K_1(\theta,\theta')\left\{-\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(\theta') + \int_\Theta K_1(\theta',\theta'')\, g(\theta'')\, d\theta''\right\} d\theta' \\
=& -\left\{\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(\theta) + \int_\Theta K_1(\theta,\theta')\,\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(\theta')\, d\theta'\right\} \\
&+ \int_\Theta \left\{\int_\Theta K_1(\theta,\theta')\, K_1(\theta',\theta'')\, g(\theta'')\, d\theta''\right\} d\theta'.
\end{aligned}$$

Applying Fubini's theorem yields

$$\begin{aligned}
g(\theta) =& -\left\{\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(\theta) + \int_\Theta K_1(\theta,\theta')\,\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(\theta')\, d\theta'\right\} \\
&+ \int_\Theta \left\{\int_\Theta K_1(\theta,\theta')\, K_1(\theta',\theta'')\, d\theta'\right\} g(\theta'')\, d\theta'' \\
\equiv& -\left\{\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(\theta) + \int_\Theta K_1(\theta,\theta')\,\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(\theta')\, d\theta'\right\} + \int_\Theta K_2(\theta,\theta')\, g(\theta')\, d\theta',
\end{aligned}$$

where $K_2(\theta,\theta') = \int_\Theta K(\theta,\theta'')\, K(\theta'',\theta')\, d\theta''$. A second substitution yields:

$$\begin{aligned}
g(\theta) =& -\left\{\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(\theta) + \int_\Theta K_1(\theta,\theta')\,\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(\theta')\, d\theta'\right\} \\
&+ \int_\Theta K_2(\theta,\theta')\left\{-\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(\theta') + \int_\Theta K_1(\theta',\theta'')\, g(\theta'')\, d\theta''\right\} d\theta',
\end{aligned}$$

which can be rewritten, following the same steps as above, as

$$\begin{aligned}
g(\theta) =& -\left\{\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(\theta) + \int_\Theta K(\theta,\theta')\,\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(\theta')\, d\theta' + \int_\Theta K_2(\theta,\theta')\,\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(\theta')\, d\theta'\right\} \\
&+ \int_\Theta K_3(\theta,\theta')\, g(\theta')\, d\theta',
\end{aligned}$$

where $K_3(\theta,\theta') = \int_\Theta K_2(\theta,\theta'')\, K(\theta'',\theta')\, d\theta''$. By repeated substitution, we thus obtain: for all

$n \geq 1$,

$$g(\theta) = -\left\{ \tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(\theta) + \sum_{i=1}^{n} \int_{\Theta} K_i(\theta,\theta')\,\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(\theta')\,d\theta' \right\}$$
$$+ \int_{\Theta} K_{n+1}(\theta,\theta')\,g(\theta')\,d\theta',$$

where for all n, $K_{n+1}(\theta,\theta') = \int_{\Theta} K_n(\theta,\theta'')\,K_1(\theta'',\theta')\,d\theta''$. Now the last step is to show that that $\int_{\Theta} K_{n+1}(\theta,\theta')\,g(\theta')\,d\theta'$ converges to zero as $n \to \infty$.

Now, applying the Cauchy-Schwartz inequality to the interated kernel yields

$$|K_{n+1}(\theta,\theta')|^2 \leq \left(\int_{\Theta} |K_n(\theta,\theta'')|^2\,d\theta'' \right) \left(\int_{\Theta} |K_1(\theta'',\theta')|^2\,d\theta'' \right).$$

Integrating this inequality with respect to θ' implies

$$\int_{\Theta} |K_{n+1}(\theta,\theta')|^2\,d\theta' \leq \left(\int_{\Theta} |K_n(\theta,\theta'')|^2\,d\theta'' \right) \left(\int_{\Theta}\int_{\Theta} |K_1(\theta'',\theta')|^2\,d\theta''d\theta' \right)$$
$$= \|K_1\|_2^2 \times \int_{\Theta} |K_n(\theta,\theta'')|^2\,d\theta''.$$

By induction, we obtain

$$\int_{\Theta} |K_{n+1}(\theta,\theta')|^2\,d\theta' \leq \|K_1\|_2^{2n} \times \int_{\Theta} |K_1(\theta,\theta'')|^2\,d\theta''.$$

We thus have, using the Cauchy-Schwartz inequality again,

$$\left| \int_{\Theta} K_{n+1}(\theta,\theta')\,\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(\theta')\,d\theta' \right|^2$$
$$\leq \left(\int_{\Theta} |K_{n+1}(\theta,\theta'')|^2\,d\theta'' \right) \left(\int_{\Theta} \left| \tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(\theta') \right|^2 d\theta'' \right)$$
$$\leq \|K_1\|_2^{2n} \times \left(\int_{\Theta} |K_1(\theta,\theta'')|^2\,d\theta'' \right) \times \left\| \tilde{E}_{l,1-\tau}\hat{h}' \right\|_2^2$$

Thus, for all $\theta \in \Theta$,

$$\left| \int_{\Theta} K_i(\theta,\theta')\,\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(\theta')\,d\theta' \right| \leq \left[\left\| \tilde{E}_{l,1-\tau}\hat{h}' \right\|_2 \sqrt{\int_{\Theta} |K_1(\theta,\theta'')|^2\,d\theta''} \right] \times \|K_1\|_2^i.$$

Denote by

$$\kappa_n(\theta) \equiv \sum_{i=1}^{n} \int_{\Theta} K_i(\theta,\theta')\,\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(\theta')\,d\theta'.$$

Since $\|K_1\|_2 < 1$, the previous arguments imply that the sequence $\{\kappa_n(\theta)\}_{n\geq 1}$ is dominated by a convergent geometric series of positive terms, and therefore it converges absolutely and uniformly to a unique limit $\kappa(\theta)$ on Θ. Similarly, we have

$$\lim_{n\to\infty}\left|\int_{\Theta} K_{n+1}(\theta,\theta') g(\theta') d\theta'\right| = 0.$$

Therefore, we can write

$$g(\theta) = -\tilde{E}_{l,1-\tau}(\theta)\hat{h}'(\theta) - \sum_{i=1}^{\infty}\int_{\Theta} K_i(\theta,\theta')\tilde{E}_{l,1-\tau}(\theta')\hat{h}'(\theta') d\theta',$$

which proves equation (22).

To show the uniqueness of the solution, suppose that $g_1(\theta)$ and $g_2(\theta)$ are two solutions to (22). Then $\Delta(\theta) \equiv g_2(\theta) - g_1(\theta)$ satisfies the homogeneous integral equation

$$\Delta(\theta) = \int_{\Theta} K_1(\theta,\theta')\Delta(\theta') d\theta'.$$

The Cauchy-Schwartz inequality reads

$$|\Delta(\theta)|^2 \leq \left(\int_{\Theta}|K_1(\theta,\theta')|^2 d\theta'\right)\left(\int_{\Theta}|\Delta(\theta')|^2 d\theta'\right).$$

Integrating with respect to θ yields

$$\int_{\Theta}|\Delta(\theta)|^2 d\theta \leq \|K_1\|^2\int_{\Theta}|\Delta(\theta')|^2 d\theta',$$

Assumption $\|K_1\|^2 < 1$ then implies $\int_{\Theta}|\Delta(\theta)|^2 d\theta$, i.e., $\Delta(\theta) = 0$ for all $\theta \in \Theta$.

□

In case the condition $\|K_1\|_2^2 < 1$ does not hold, there exist methods to express the solution to the integral equation (21). Schmidt's method consists of showing that the kernel of the integral equation (21) can be written as the sum of two kernels, $K_1(\theta,\theta') = \bar{K}(\theta,\theta') + K_{\varepsilon}(\theta,\theta')$, where $\bar{K}(\theta,\theta')$ is separable and $K_{\varepsilon}(\theta,\theta')$ satisfies the condition $\|K_{\varepsilon}\|_2^2 < 1$. This decomposition is allowed by the Weierstrass theorem: by appropriately choosing a separable polynomial in θ and θ' for $\bar{K}(\theta,\theta')$, we can make the norm of $K_{\varepsilon}(\theta,\theta')$ arbitrarily small. It is then easy to see that the solution to (21) satisfies an integral equation with kernel $K_{\varepsilon}(\theta,\theta')$, where the exogenous function on the right hand side (outside of the integral) is itself the solution to an integral equation with separable kernel $\bar{K}(\theta,\theta')$. The former integral equation can be analyzed using the same arguments as in the proof of Proposition 1. The latter integral equation can be analyzed using the arguments of the proof of Proposition 3. We can then derive a solution to (21) of the same form as (22), with a more general resolvent. This technique is detailed in Section 2.4 of Zemyan (2012).

B.3.3 Proof of Corollaries 1 and 2

We now derive the incidence of tax reforms on wages, utilities, government revenue, and social welfare.

Proof. First note that we can also write this equation as

$$
\begin{aligned}
\mathrm{d}\hat{l}(\theta,h) =& \frac{\left[1-T'(w(\theta)l(\theta))-y(\theta)T''(w(\theta)l(\theta))\right]\int_{\Theta}\gamma(\theta,\theta')\mathrm{d}\hat{l}(\theta',h)\,d\theta'-h'(y(\theta))}{(1-T'(y(\theta)))\frac{l(\theta)v''(l(\theta))}{v'(l(\theta))}+y(\theta)T''(w(\theta)l(\theta))} \\
=& \frac{\left[1-T'(y(\theta))-y(\theta)T''(y(\theta))\right]\varepsilon_{l,1-\tau}(\theta)}{1-T'(y(\theta))+\varepsilon_{l,1-\tau}(\theta)y(\theta)T''(y(\theta))}\int_{\Theta}\gamma(\theta,\theta')\mathrm{d}\hat{l}(\theta',h)\,d\theta' \\
&-\frac{\varepsilon_{l,1-\tau}(\theta)}{1-T'(y(\theta))+\varepsilon_{l,1-\tau}(\theta)y(\theta)T''(y(\theta))}h'(y(\theta)) \\
=& \tilde{\varepsilon}_{l,w}(\theta)\int_{\Theta}\gamma(\theta,\theta')\mathrm{d}\hat{l}(\theta',h)\,d\theta'-\tilde{\varepsilon}_{l,1-\tau}(\theta)\frac{h'(y(\theta))}{1-T'(y(\theta))},
\end{aligned}
$$

which, along with (66), proves equation (26).

Next, the first-order effects of the tax reform h on individual θ's tax liability are given by:

$$
\begin{aligned}
&\mathrm{d}_h T(w(\theta)l(\theta)) \\
\equiv& \lim_{\mu\to 0}\left\{\frac{1}{\mu}\left[T(\tilde{w}(\theta)(l(\theta)+\mu\mathrm{d}l(\theta,h)))-T(w(\theta)l(\theta))\right]+h(\tilde{w}(\theta)(l(\theta)+\mu\mathrm{d}l(\theta,h)))\right\} \\
=& T'(y(\theta))y(\theta)\left[\mathrm{d}\hat{l}(\theta,h)+\int_{\Theta}\gamma(\theta,\theta')\mathrm{d}\hat{l}(\theta',h)\,d\theta'\right]+h(y(\theta)) \\
=& T'(y(\theta))y(\theta)\left[\left(1+\frac{1}{\tilde{\varepsilon}_{l,w}(\theta)}\right)\mathrm{d}\hat{l}(\theta,h)+\frac{\tilde{\varepsilon}_{l,1-\tau}(\theta)}{\tilde{\varepsilon}_{l,w}(\theta)}\frac{h'(y(\theta))}{1-T'(y(\theta))}\right]+h(y(\theta)).
\end{aligned}
$$

The first-order effects of the tax reform h on individual welfare are given by

$$
\begin{aligned}
&\mathrm{d}_h u(y(\theta)-T(y(\theta))-v(l(\theta)))=\left[\mathrm{d}_h y(\theta)-\mathrm{d}_h T(y(\theta))-h(y(\theta))-\mathrm{d}_h v(l(\theta))\right]u'(\theta) \\
=&\left[(1-T'(y(\theta)))y(\theta)\mathrm{d}\hat{y}(\theta,h)-l(\theta)v'(l(\theta))\mathrm{d}\hat{l}(\theta,h)-h(y(\theta))\right]u'(\theta) \\
=&(1-T'(y(\theta)))y(\theta)\left[\left(1+\frac{1}{\tilde{\varepsilon}_{l,w}(\theta)}\right)\mathrm{d}\hat{l}(\theta,h)+\frac{\tilde{\varepsilon}_{l,1-\tau}(\theta)}{\tilde{\varepsilon}_{l,w}(\theta)}\frac{h'(y(\theta))}{1-T'(y(\theta))}\right]u'(\theta) \\
&-l(\theta)v'(l(\theta))\mathrm{d}\hat{l}(\theta,h)u'(\theta)-h(y(\theta))u'(\theta) \\
=&(1-T'(y(\theta)))y(\theta)\left[\frac{1}{\tilde{\varepsilon}_{l,w}(\theta)}\mathrm{d}\hat{l}(\theta,h)+\frac{\tilde{\varepsilon}_{l,1-\tau}(\theta)}{\tilde{\varepsilon}_{l,w}(\theta)}\frac{h'(y(\theta))}{1-T'(y(\theta))}\right]u'(\theta)-h(y(\theta))u'(\theta).
\end{aligned}
$$

The first-order effects of the tax reform h on government revenue are given by

$$
\begin{aligned}
&\mathrm{d}\mathscr{R}(T,h)=\mathrm{d}_h\left[\int_{\Theta}T(y(\theta))f_\theta(\theta)\,d\theta\right] \\
=&\int_{\Theta}h(y(\theta))f_\theta(\theta)\,d\theta \\
&+\int_{\Theta}T'(y(\theta))\left[\frac{\tilde{\varepsilon}_{l,1-\tau}(y(\theta))}{\tilde{\varepsilon}_{l,w}(y(\theta))}\frac{h'(y(\theta))}{1-T'(y(\theta))}+\left(1+\frac{1}{\tilde{\varepsilon}_{l,w}(y(\theta))}\right)\mathrm{d}\hat{l}(\theta,h)\right]y(\theta)f_\theta(\theta)\,d\theta \\
=&\int_{\mathbb{R}_+}h(y)f_y(y)\,dy+\int_{\mathbb{R}_+}T'(y)\left[\frac{\tilde{\varepsilon}_{l,1-\tau}(y)}{\tilde{\varepsilon}_{l,w}(y)}\frac{h'(y)}{1-T'(y)}+\left(1+\frac{1}{\tilde{\varepsilon}_{l,w}(y)}\right)\mathrm{d}\hat{l}(y,h)\right]yf_y(y)\,dy.
\end{aligned}
$$

The first-order effects of the tax reform h on the social objective are given by

$$\lambda^{-1}\mathrm{d}\mathscr{G}(T,h) = \lambda^{-1}\mathrm{d}_h \int_{\Theta} u\left(y(\theta) - T(y(\theta)) - v(l(\theta))\right)\tilde{f}_\theta(\theta)\,d\theta$$
$$= \int_{\Theta}\left[\frac{1-T'(y(\theta))}{\tilde{\varepsilon}_{l,w}(\theta)} y(\theta)\,\mathrm{d}\hat{l}(\theta,h) + \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta)}{\tilde{\varepsilon}_{l,w}(\theta)} y(\theta)\,h'(y(\theta)) - h(y(\theta))\right]\lambda^{-1}u'(\theta)\,\tilde{f}_\theta(\theta)\,d\theta$$
$$= \int_{\Theta}\left[\frac{1-T'(y(\theta))}{\tilde{\varepsilon}_{l,w}(\theta)} y(\theta)\,\mathrm{d}\hat{l}(\theta,h) + \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta)}{\tilde{\varepsilon}_{l,w}(\theta)} y(\theta)\,h'(y(\theta)) - h(y(\theta))\right] g_\theta(\theta)\,f_\theta(\theta)\,d\theta$$
$$= \int_{\mathbb{R}_+}\left[\frac{1-T'(y)}{\tilde{\varepsilon}_{l,w}(y)}\mathrm{d}\hat{l}(y,h) + \frac{\tilde{\varepsilon}_{l,1-\tau}(y)}{\tilde{\varepsilon}_{l,w}(y)} h'(y)\right] y g_y(y)\,f_y(y)\,dy - \int_{\mathbb{R}_+} h(y)\,g_y(y)\,f_y(y)\,dy.$$

The first-order effects of the tax reform h on social welfare are finally given by

$$\mathrm{d}\mathscr{W}(T,h) = \mathrm{d}\mathscr{R}(T,h) + \lambda^{-1}\mathrm{d}\mathscr{G}(T,h)$$
$$= \int_{\mathbb{R}_+}(1-g_y(y))\,h(y)\,f_y(y)\,dy$$
$$+ \int_{\mathbb{R}_+} T'(y)\left[\frac{\tilde{\varepsilon}_{l,1-\tau}(y)}{\tilde{\varepsilon}_{l,w}(y)}\frac{h'(y)}{1-T'(y)} + \left(1+\frac{1}{\tilde{\varepsilon}_{l,w}(y)}\right)\mathrm{d}\hat{l}(y,h)\right] y f_y(y)\,dy$$
$$+ \int_{\mathbb{R}_+}\left[\frac{1-T'(y)}{\tilde{\varepsilon}_{l,w}(y)}\mathrm{d}\hat{l}(y,h) + \frac{\tilde{\varepsilon}_{l,1-\tau}(y)}{\tilde{\varepsilon}_{l,w}(y)} h'(y)\right] y g_y(y)\,f_y(y)\,dy.$$

This concludes the proof of equations (28) and (29).

□

B.3.4 Proof of Proposition 2

We now specialize the production function to be CES. In this case the integral equation (21) has a simple solution.

Proof. If the production function is CES, the cross wage elasticities $\bar{\gamma}(\theta,\theta')$ do not depend on θ (see equation (15)). Hence the kernel of the integral equation, $K_1(\theta,\theta') = \tilde{E}_{l,w}(\theta)\,\bar{\gamma}(\theta,\theta')$, is multiplicatively separable, i.e., the product of a function of θ only and a function of θ' only:

$$K_1(\theta,\theta') = \frac{\tilde{\varepsilon}_{l,w}(\theta)}{1-\bar{\bar{\gamma}}(\theta,\theta)\,\tilde{\varepsilon}_{l,w}(\theta)}\bar{\gamma}(\theta,\theta')$$
$$= \left[\frac{\tilde{\varepsilon}_{l,w}(\theta)}{1-\bar{\bar{\gamma}}(\theta,\theta)\,\tilde{\varepsilon}_{l,w}(\theta)}\right] \times \left[\frac{(1-\rho)\,a(\theta')\,L(\theta')^{\rho}}{\int_{\Theta} a(x)\,L(x)^{\rho}\,dx}\right] \equiv \kappa_1(\theta)\times\kappa_2(\theta').$$

The integral equation (21) then writes, letting $\hat{h}'(y) = \frac{h'(y)}{1-T'(y)}$,

$$\mathrm{d}\hat{l}(\theta,h) = -\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(y(\theta)) + \kappa_1(\theta)\int_{\Theta}\kappa_2(\theta')\,\mathrm{d}\hat{l}(\theta',h)\,d\theta'$$

and can be solved as follows. Multiplying by $\kappa_2(\theta')$ both sides of the integral equation evaluated at

θ' yields

$$\kappa_2(\theta')\,\mathrm{d}\hat{l}(\theta',h) = -\,\kappa_2(\theta')\,\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(y(\theta')) + \kappa_1(\theta')\,\kappa_2(\theta')\int_{\Theta}\kappa_2(\theta'')\,\mathrm{d}\hat{l}(\theta'',h)\,d\theta''.$$

Integrating with respect to θ' gives

$$\int_{\Theta}\kappa_2(\theta')\,\mathrm{d}\hat{l}(\theta',h)\,d\theta' = -\int_{\Theta}\kappa_2(\theta')\,\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(y(\theta'))\,d\theta' + \left(\int_{\Theta}\kappa_1(\theta')\,\kappa_2(\theta')\,d\theta'\right)\left(\int_{\Theta}\kappa_2(\theta'')\,\mathrm{d}\hat{l}(\theta'',h)\,d\theta''\right),$$

i.e.,

$$\int_{\Theta}\kappa_2(\theta')\,\mathrm{d}\hat{l}(\theta',h)\,d\theta' = -\,\frac{\int_{\Theta}\kappa_2(\theta')\,\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(y(\theta'))\,d\theta'}{1-\int_{\Theta}\kappa_1(\theta')\,\kappa_2(\theta')\,d\theta'}.$$

Substituting into the right hand side of the integral equation (21) leads to

$$\mathrm{d}\hat{l}(\theta,h) = -\,\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(y(\theta)) - \kappa_1(\theta)\,\frac{\int_{\Theta}\kappa_2(\theta')\,\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(y(\theta'))\,d\theta'}{1-\int_{\Theta}\kappa_1(\theta')\,\kappa_2(\theta')\,d\theta'},$$

which implies equation (30).

Suppose in particular, as in Saez (2001), that the tax reform h is the step function $h(y) = \mathbb{I}_{\{y\geq y^*\}}$, so that $h'(y) = \delta_{y^*}(y)$ is the Dirac delta function (i.e., marginal tax rates are perturbed at income y^* only). To apply this equality to this non-differentiable perturbation, construct a sequence of smooth funtions $\varphi_{y^*,\varepsilon}(y)$ such that

$$\delta_{y^*}(y) = \lim_{\varepsilon\to 0}\varphi_{y^*,\varepsilon}(y),$$

in the sense that for all continuous functions ψ with compact support,

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}}\varphi_{y^*,\varepsilon}(y)\,\psi(y)\,dy = \psi(y^*),$$

i.e., changing variables in the integral,

$$\lim_{\varepsilon\to 0}\int_{\Theta}\varphi_{y^*,\varepsilon}(y(\theta'))\left\{\psi(y(\theta'))\,\frac{dy(\theta')}{d\theta}\right\}d\theta' = \psi(y^*).$$

This can be obtained by defining an absolutely integrable and smooth function $\varphi_{y^*}(y)$ with compact support and $\int_{\mathbb{R}}\varphi_{y^*}(y)\,dy = 1$, and letting $\varphi_{y^*,\varepsilon}(y) = \varepsilon^{-1}\varphi_{y^*}\left(\frac{y}{\varepsilon}\right)$. We then have, for all $\varepsilon > 0$,

$$\mathrm{d}\hat{l}(\theta,\Phi_{y^*,\varepsilon}) = -\,\tilde{E}_{l,1-\tau}(\theta)\,\frac{\varphi_{y^*,\varepsilon}(y(\theta))}{1-T'(y(\theta^*))} - \kappa_1(\theta)\,\frac{\int_{\Theta}\kappa_2(\theta')\,\tilde{E}_{l,1-\tau}(\theta')\,\frac{\varphi_{y^*,\varepsilon}(y(\theta'))}{1-T'(y(\theta'))}d\theta'}{1-\int_{\Theta}\kappa_1(\theta')\,\kappa_2(\theta')\,d\theta'}.$$

Letting $\varepsilon \to 0$, we get

$$\begin{aligned} d\hat{l}(\theta,h) &= -\tilde{E}_{l,1-\tau}(\theta)\frac{\delta_{y^*}(y(\theta))}{1-T'(y(\theta^*))} - \kappa_1(\theta)\frac{\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\kappa_2(\theta^*)\tilde{E}_{l,1-\tau}(\theta^*)\frac{1}{1-T'(y(\theta^*))}}{1-\int_\Theta \kappa_1(\theta')\kappa_2(\theta')d\theta'}, \\ &= -\frac{\tilde{E}_{l,1-\tau}(y(\theta^*))}{1-T'(y(\theta^*))}\left[\delta_{y^*}(y(\theta)) + \frac{1}{y'(\theta^*)}\frac{\tilde{E}_{l,w}(y(\theta))\bar{\gamma}(\theta,\theta^*)}{1-\int_\Theta \tilde{E}_{l,w}(y(\theta'))\bar{\gamma}(\theta',\theta')d\theta'}\right]. \end{aligned} \tag{67}$$

This formula will be useful for the proof of Corollary 3.

□

B.3.5 Proof of Corollary 3

Next, assume that the production function is CES and the baseline tax schedule is CRP, given by (31) for $p \in (-\infty, 1)$, so that in particular, $1 - T'(y) = (1-\tau)y^{-p}$ and $T''(y) = p(1-\tau)y^{-p-1}$. We start by deriving prelimilary properties of the labor supply and wage elasticities.

The labor supply elasticities (7), (7), and (12) are given by

$$\begin{aligned} \tilde{\varepsilon}_{l,1-\tau}(y) =& \frac{1-T'(y)}{1-T'(y)+\varepsilon y T''(y)}\varepsilon = \frac{(1-\tau)y^{-p}}{(1-\tau)y^{-p}+\varepsilon y p(1-\tau)y^{-p-1}}\varepsilon = \frac{\varepsilon}{1+p\varepsilon}, \\ \tilde{\varepsilon}_{l,w}(y) =& \frac{1-T'(y)-yT''(y)}{1-T'(y)+\varepsilon y T''(y)}\varepsilon = \frac{(1-\tau)y^{-p}-yp(1-\tau)y^{-p-1}}{(1-\tau)y^{-p}+\varepsilon y p(1-\tau)y^{-p-1}}\varepsilon = \frac{(1-p)\varepsilon}{1+p\varepsilon}, \end{aligned}$$

and

$$\begin{aligned} \tilde{E}_{l,1-\tau}(y) =& \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta)}{1-\bar{\bar{\gamma}}(\theta,\theta)\tilde{\varepsilon}_{l,w}(\theta)} = \frac{\frac{\varepsilon}{1+p\varepsilon}}{1+\frac{(1-p)\varepsilon}{(1+p\varepsilon)\sigma}} = \frac{\varepsilon}{1+p\varepsilon+(1-p)\frac{\varepsilon}{\sigma}}, \\ \tilde{E}_{l,w}(y) =& \frac{\tilde{\varepsilon}_{l,w}(\theta)}{1-\bar{\bar{\gamma}}(\theta,\theta)\tilde{\varepsilon}_{l,w}(\theta)} = \frac{\frac{(1-p)\varepsilon}{1+p\varepsilon}}{1+\frac{(1-p)\varepsilon}{(1+p\varepsilon)\sigma}} = \frac{(1-p)\varepsilon}{1+p\varepsilon+(1-p)\frac{\varepsilon}{\sigma}}. \end{aligned}$$

Note that all of these elasticities are constant. This is because the curvature of the CRP tax schedule (captured by the parameter p) is constant. This feature allows us to simplify further equation (29) to obtain (32), respectively.

We can now prove Corollary 3.

Proof. Suppose first that the baseline tax schedule is linear, i.e., $p = 0$ in (31). In this case we have

$$\int_\Theta \tilde{E}_{l,w}(y(\theta'))\bar{\gamma}(\theta',\theta')d\theta' = \frac{\varepsilon}{1+\frac{\varepsilon}{\sigma}}\int_\Theta \bar{\gamma}(\theta',\theta')d\theta' = \frac{\varepsilon}{1+\frac{\varepsilon}{\sigma}} \times \frac{1}{\sigma},$$

where the last equality follows from expression (15) for the cross-wage elasticities $\bar{\gamma}(\theta',\theta')$. Thus the integral equation (30) becomes

$$d\hat{l}(\theta,h) = -\frac{1}{1-\tau}\frac{\varepsilon}{1+\frac{\varepsilon}{\sigma}}h'(y(\theta)) - \frac{1}{1-\tau}\left(\frac{\varepsilon}{1+\frac{\varepsilon}{\sigma}}\right)^2\left(\frac{\sigma+\varepsilon}{\sigma}\right)\int_\Theta \bar{\gamma}(\theta,\theta')h'(y(\theta'))d\theta'.$$

Note that the first term in this expression for individual θ is proportional to the exogenous marginal tax rate perturbation at income $y(\theta)$, and the second term is a constant independent of θ.

Now integrate this equation to get the effect of the perturbation on government revenue. Formula (28) writes

$$
\begin{aligned}
& \mathrm{d}\mathscr{R}(T,h) \\
&= \int_{\mathbb{R}_+} h(y) f_y(y)\, dy + \int_{\mathbb{R}_+} T'(y) \left[\frac{\tilde{\varepsilon}_{l,1-\tau}(y)}{\tilde{\varepsilon}_{l,w}(y)} \frac{h'(y)}{1-T'(y)} + \left(1 + \frac{1}{\tilde{\varepsilon}_{l,w}(y)}\right) \mathrm{d}\hat{l}(y,h) \right] y f_y(y)\, dy \\
&= \int_{\mathbb{R}_+} h(y) f_y(y)\, dy + \frac{\tau}{1-\tau} \int_{\mathbb{R}_+} h'(y) y f_y(y)\, dy + \left(1 + \frac{1}{\varepsilon}\right) \tau \int_{\mathbb{R}_+} \mathrm{d}\hat{l}(y,h) y f_y(y)\, dy \\
&= \int_{\mathbb{R}_+} h(y) f_y(y)\, dy - \frac{(\sigma-1)\varepsilon}{\sigma+\varepsilon} \frac{\tau}{1-\tau} \int_{\mathbb{R}_+} h'(y) y f_y(y)\, dy \\
&\quad - \frac{\sigma(1+\varepsilon)\varepsilon}{\sigma+\varepsilon} \frac{\tau}{1-\tau} \int_{\mathbb{R}_+} \left[\int_{\mathbb{R}_+} \bar{\gamma}(y,y') h'(y') \left(\frac{dy(\theta')}{d\theta}\right)^{-1} dy' \right] y f_y(y)\, dy,
\end{aligned}
$$

where the last equality uses a change of variables to rewrite the integral $\int_\Theta \bar{\gamma}(\theta,\theta') h'(y(\theta'))\, d\theta'$. But using expression (64) to substitute for $\bar{\gamma}(y,y')$, we obtain

$$
\begin{aligned}
& \int_{\mathbb{R}_+} \left[\int_{\mathbb{R}_+} \bar{\gamma}(y,y') h'(y') \left(\frac{dy(\theta')}{d\theta}\right)^{-1} dy' \right] y f_y(y)\, dy \\
&= \left[\int_{\mathbb{R}_+} (1-\rho) \frac{y' f_y(y')}{\int_{\mathbb{R}_+} y f_y(y)\, dy} h'(y')\, dy' \right] \mathbb{E}[y] = (1-\rho) \int_{\mathbb{R}_+} h'(y) y f_y(y)\, dy.
\end{aligned}
$$

Hence we obtain

$$
\begin{aligned}
& \mathrm{d}\mathscr{R}(T,h) \\
&= \int_{\mathbb{R}} h(y) f_y(y)\, dy - \frac{(\sigma-1)\varepsilon}{\sigma+\varepsilon} \frac{\tau}{1-\tau} \int_{\mathbb{R}} h'(y) y f_y(y)\, dy - \frac{(1+\varepsilon)\varepsilon}{\sigma+\varepsilon} \frac{\tau}{1-\tau} \int_{\mathbb{R}} h'(y) y f_y(y)\, dy \\
&= \int_{\mathbb{R}} h(y) f_y(y)\, dy - \frac{\tau}{1-\tau} \varepsilon \int_{\mathbb{R}} h'(y) y f_y(y)\, dy,
\end{aligned}
$$

which is exactly the partial equilibrium formula.

Now suppose more generally that the baseline tax schedule is CRP with $p \in (-\infty, 1)$. Consider the Saez perturbation at income y^*, i.e. $h(y) = \mathbb{I}_{\{y \geq y^*\}}$ and $h'(y) = \delta_{y^*}(y)$. Substituting for the labor supply elasticities and using the fact that the elasticities $\tilde{E}_{l,1-\tau}(\theta), \tilde{E}_{l,w}(\theta)$ are constant, equation (67) can be rewritten as

$$
\begin{aligned}
\mathrm{d}\hat{l}(y(\theta),h) &= -\frac{\tilde{E}_{l,1-\tau}(y(\theta^*))}{1-T'(y(\theta^*))} \left[\delta_{y^*}(y) + \left(\frac{dy(\theta^*)}{d\theta}\right)^{-1} \frac{\tilde{E}_{l,w}(y(\theta^*)) \bar{\gamma}(\theta,\theta^*)}{1 - \int_\Theta \tilde{E}_{l,w}(y(\theta')) \bar{\gamma}(\theta',\theta')\, d\theta'} \right] \\
&= -\frac{\tilde{E}_{l,1-\tau}}{1-T'(y^*)} \left[\delta_{y^*}(y) + \left(\frac{dy(\theta^*)}{d\theta}\right)^{-1} \frac{\tilde{E}_{l,w}}{1-\tilde{E}_{l,w}/\sigma} \bar{\gamma}(y,y^*) \right].
\end{aligned}
$$

Substituting in the expression (28) that gives the revenue effects of the tax reform, we obtain:

$$\begin{aligned}
& \mathrm{d}\mathscr{R}\left(T,h\right) \\
=& \int_{\mathbb{R}_+} \mathbb{I}_{\{y\geq y^*\}} f_y\left(y\right)dy + \int_{\mathbb{R}_+} T'\left(y\right)\left[\frac{\tilde{\varepsilon}_{l,1-\tau}}{\tilde{\varepsilon}_{l,w}}\frac{\delta_{y^*}\left(y\right)}{1-T'\left(y\right)} + \left(1+\frac{1}{\tilde{\varepsilon}_{l,w}}\right)\mathrm{d}\hat{l}\left(y,h\right)\right] y f_y\left(y\right)dy \\
=& 1-F_y\left(y^*\right) + \int_{\mathbb{R}_+} \frac{T'\left(y\right)}{1-T'\left(y^*\right)}\left[\frac{\tilde{\varepsilon}_{l,1-\tau}}{\tilde{\varepsilon}_{l,w}} - \left(1+\frac{1}{\tilde{\varepsilon}_{l,w}}\right)\tilde{E}_{l,1-\tau}\right] y f_y\left(y\right)\delta_{y^*}\left(y\right)dy \\
& - \int_{\mathbb{R}_+} \frac{T'\left(y\right)}{1-T'\left(y^*\right)}\left(1+\frac{1}{\tilde{\varepsilon}_{l,w}}\right)\frac{\tilde{E}_{l,1-\tau}\tilde{E}_{l,w}}{1-\tilde{E}_{l,w}/\sigma}\bar{\gamma}\left(y,y^*\right)\left(\frac{dy\left(\theta^*\right)}{d\theta}\right)^{-1} y f_y\left(y\right)dy,
\end{aligned}$$

i.e.,

$$\begin{aligned}
\mathrm{d}\mathscr{R}\left(T,h\right) =& 1-F_y\left(y^*\right) - \tilde{\varepsilon}_{l,1-\tau}\frac{T'\left(y^*\right)}{1-T'\left(y^*\right)} y^* f_y\left(y^*\right) \\
& + \left(1+\frac{1}{\tilde{\varepsilon}_{l,w}}\right)\left(\tilde{\varepsilon}_{l,1-\tau} - \tilde{E}_{l,1-\tau}\right)\frac{T'\left(y^*\right)}{1-T'\left(y^*\right)} y^* f_y\left(y^*\right) \\
& - \left(1+\frac{1}{\tilde{\varepsilon}_{l,w}}\right)\tilde{E}_{l,1-\tau}\frac{\frac{\tilde{\varepsilon}_{l,w}}{1+\frac{\tilde{\varepsilon}_{l,w}}{\sigma}}}{1-\frac{\tilde{\varepsilon}_{l,w}}{\sigma+\tilde{\varepsilon}_{l,w}}}\int_{\mathbb{R}_+}\frac{T'\left(y\right)}{1-T'\left(y^*\right)}\bar{\gamma}\left(y,y^*\right)\left(\frac{dy\left(\theta^*\right)}{d\theta}\right)^{-1} y f_y\left(y\right)dy \\
=& 1-F_y\left(y^*\right) - \tilde{\varepsilon}_{l,1-\tau}\frac{T'\left(y^*\right)}{1-T'\left(y^*\right)} y^* f_y\left(y^*\right) - \tilde{E}_{l,1-\tau}\left(1+\tilde{\varepsilon}_{l,w}\right)\times\ldots \\
& \left\{-\frac{1}{\sigma}\frac{T'\left(y^*\right)}{1-T'\left(y^*\right)} y^* f_y\left(y^*\right) + \left(\frac{dy\left(\theta^*\right)}{d\theta}\right)^{-1}\int_{\mathbb{R}_+}\frac{T'\left(y\right)}{1-T'\left(y^*\right)}\bar{\gamma}\left(y,y^*\right) y f_y\left(y\right)dy\right\}.
\end{aligned}$$

The terms in curly brackets are equal to

$$\begin{aligned}
& \left(\frac{dy\left(\theta^*\right)}{d\theta}\right)^{-1}\left\{\int_{\mathbb{R}_+}\frac{T'\left(y\right)}{1-T'\left(y^*\right)}\bar{\gamma}\left(y,y^*\right) y f_y\left(y\right)dy + \frac{T'\left(y^*\right)}{1-T'\left(y^*\right)}\bar{\bar{\gamma}}\left(y^*,y^*\right) y^* f_y\left(y^*\right)\frac{dy\left(\theta^*\right)}{d\theta}\right\} \\
=& \left(\frac{dy\left(\theta^*\right)}{d\theta}\right)^{-1}\left\{\int_{\Theta}\frac{T'\left(y\left(\theta\right)\right)}{1-T'\left(y\left(\theta^*\right)\right)}\bar{\gamma}\left(y\left(\theta\right),y\left(\theta^*\right)\right) y\left(\theta\right) f_\theta\left(\theta\right)d\theta\right. \\
& \qquad \left. + \frac{T'\left(y\left(\theta^*\right)\right)}{1-T'\left(y\left(\theta^*\right)\right)}\bar{\bar{\gamma}}\left(y\left(\theta^*\right),y\left(\theta^*\right)\right) y\left(\theta^*\right) f_\theta\left(\theta^*\right)\right\} \\
=& \left(\frac{dy\left(\theta^*\right)}{d\theta}\right)^{-1}\left\{\int_{\Theta}\frac{T'\left(y\left(\theta\right)\right)}{1-T'\left(y\left(\theta^*\right)\right)}\gamma\left(y\left(\theta\right),y\left(\theta^*\right)\right) y\left(\theta\right) f_\theta\left(\theta\right)d\theta\right\} \\
=& \left(\frac{dy\left(\theta^*\right)}{d\theta}\right)^{-1}\left\{\int_{\mathbb{R}_+}\frac{T'\left(y\right)}{1-T'\left(y^*\right)}\gamma\left(y,y^*\right) y f_y\left(y\right)dy\right\}.
\end{aligned}$$

Thus we obtain

$$\begin{aligned}
\mathrm{d}\mathscr{R}\left(T,h\right) =& 1-F_y\left(y^*\right) - \tilde{\varepsilon}_{l,1-\tau}\frac{T'\left(y^*\right)}{1-T'\left(y^*\right)} y^* f_y\left(y^*\right) \\
& - \tilde{E}_{l,1-\tau}\left(1+\tilde{\varepsilon}_{l,w}\right)\left(\frac{dy\left(\theta^*\right)}{d\theta}\right)^{-1}\int_{\mathbb{R}_+}\frac{T'\left(y\right)}{1-T'\left(y^*\right)}\gamma\left(y,y^*\right) y f_y\left(y\right)dy.
\end{aligned}$$

Another way to write this formula is as follows: using the expression (64) derived above for the cross wage elasticities and the fact that $\bar{\gamma}(y, y^*)$ depends only on y^* (so that it can be taken out of the integral), we get

$$\begin{aligned}
&\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1} \bar{\gamma}(y, y^*) \int_{\mathbb{R}_+} \frac{T'(y)}{1-T'(y^*)} y f_y(y)\, dy + \frac{T'(y^*)}{1-T'(y^*)} \bar{\bar{\gamma}}(y^*, y^*)\, y^* f_y(y^*) \\
=&\frac{1}{\sigma} \frac{y^* f_y(y^*)}{\int_{\mathbb{R}_+} y f_y(y)\, dy} \int_{\mathbb{R}_+} \frac{T'(y)}{1-T'(y^*)} y f_y(y)\, dy - \frac{1}{\sigma} \frac{T'(y^*)}{1-T'(y^*)} y^* f_y(y^*) \\
=&\frac{1}{\sigma} \frac{y^* f_y(y^*)}{\int_{\mathbb{R}_+} y f_y(y)\, dy} \left\{ \int_{\mathbb{R}_+} \frac{T'(y) - T'(y^*)}{1-T'(y^*)} y f_y(y)\, dy \right\},
\end{aligned}$$

so that we can write

$$\begin{aligned}
\mathrm{d}\mathscr{R}(T,h) =& 1 - F_y(y^*) - \tilde{\varepsilon}_{l,1-\tau} \frac{T'(y^*)}{1-T'(y^*)} y^* f_y(y^*) \\
&- \frac{1}{\sigma} \tilde{E}_{l,1-\tau} (1 + \tilde{\varepsilon}_{l,w}) \frac{y^* f_y(y^*)}{\int_{\mathbb{R}_+} y f_y(y)\, dy} \int_{\mathbb{R}_+} \frac{T'(y) - T'(y^*)}{1-T'(y^*)} y f_y(y)\, dy.
\end{aligned}$$

We can rewrite this more concisely as

$$\begin{aligned}
\mathrm{d}\mathscr{R}(T,h) =& 1 - F_y(y^*) - \frac{1}{1-T'(y^*)} y^* f_y(y^*) \ldots \\
&\times \left[T'(y^*)\, \tilde{\varepsilon}_{l,1-\tau} + \frac{1}{\sigma} \tilde{E}_{l,1-\tau} (1 + \tilde{\varepsilon}_{l,w}) \int_{\mathbb{R}_+} (T'(y) - T'(y^*)) \frac{y f_y(y)}{\int_{\mathbb{R}_+} y' f_y(y')\, dy'} dy \right].
\end{aligned}$$

Finally, the integral in the last expression can be easily calculated since the tax schedule is CRP. We have

$$\begin{aligned}
&\frac{1}{\int_{\mathbb{R}_+} y f_y(y)\, dy} \int_{\mathbb{R}_+} \frac{T'(y) - T'(y^*)}{1-T'(y^*)} y f_y(y)\, dy = \frac{1}{\int_{\mathbb{R}_+} y f_y(y)\, dy} \int_{\mathbb{R}_+} \frac{(y^*)^{-p} - y^{-p}}{(y^*)^{-p}} y f_y(y)\, dy \\
=&\frac{1}{\int_{\mathbb{R}_+} y f_y(y)\, dy} \left[\int_{\mathbb{R}_+} y f_y(y)\, dy - (y^*)^p \int_{\mathbb{R}_+} y^{1-p} f_y(y)\, dy \right] = 1 - \frac{\int_{\mathbb{R}_+} \left(\frac{y}{y^*}\right)^{1-p} dF_y(y)}{\int_{\mathbb{R}_+} \left(\frac{y}{y^*}\right) dF_y(y)} \\
=&1 - \frac{\int_{\mathbb{R}_+} \frac{\frac{1-\tau}{1-p} y^{1-p}}{\frac{1-\tau}{1-p} (y^*)^{1-p}} f_y(y)\, dy}{\int_{\mathbb{R}_+} \frac{y}{y^*} f_y(y)\, dy} = 1 - \frac{\int_{\mathbb{R}_+} (c(y)/c(y^*))\, f_y(y)\, dy}{\int_{\mathbb{R}_+} (y/y^*)\, f_y(y)\, dy},
\end{aligned}$$

where $c \equiv c(y)$ and $c^* \equiv c(y^*)$ are the consumptions (disposable incomes) of types θ and θ^* respectively. Therefore

$$\begin{aligned}
\mathrm{d}\mathscr{R}(T,h) =& 1 - F_y(y^*) - \frac{T'(y^*)}{1-T'(y^*)} \tilde{\varepsilon}_{l,1-\tau} y^* f_y(y^*) \\
&- \frac{1}{\sigma} \tilde{E}_{l,1-\tau} (1 + \tilde{\varepsilon}_{l,w})\, y^* f_y(y^*) \left(1 - \frac{\mathbb{E}[c/c^*]}{\mathbb{E}[y/y^*]}\right).
\end{aligned}$$

Finally, we derive the effects of the perturbation on social welfare. We have

$$\begin{aligned}
&\lambda^{-1}\mathrm{d}\mathscr{G}(T,h) \\
=&-\int_{\mathbb{R}_+} g_y(y)\,\mathbb{I}_{\{y\geq y^*\}} f_y(y)\,dy+\int_{\mathbb{R}_+} g_y(y)\left[\frac{\tilde{\varepsilon}_{l,1-\tau}}{\tilde{\varepsilon}_{l,w}}\delta_{y^*}(y)+\frac{1-T'(y)}{\tilde{\varepsilon}_{l,w}}\mathrm{d}\hat{l}(y,h)\right] yf_y(y)\,dy \\
=&-\int_{y^*}^{\infty} g_y(y)\,f_y(y)\,dy+\int_{\mathbb{R}_+} g_y(y)\left[\frac{\tilde{\varepsilon}_{l,1-\tau}}{\tilde{\varepsilon}_{l,w}}-\frac{1-T'(y)}{1-T'(y^*)}\frac{\tilde{E}_{l,1-\tau}}{\tilde{\varepsilon}_{l,w}}\right] yf_y(y)\,\delta_{y^*}(y)\,dy \\
&-\int_{\mathbb{R}_+} g_y(y)\,\frac{1-T'(y)}{1-T'(y^*)}\frac{\tilde{E}_{l,1-\tau}}{1-\tilde{E}_{l,w}/\sigma}\frac{\tilde{E}_{l,w}}{\tilde{\varepsilon}_{l,w}}\bar{\gamma}(y,y^*)\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1} yf_y(y)\,dy.
\end{aligned}$$

The second and third integrals of this expression are equal to

$$\begin{aligned}
&\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\Bigg\{\frac{\tilde{\varepsilon}_{l,1-\tau}-\tilde{E}_{l,1-\tau}}{\tilde{\varepsilon}_{l,w}}g_y(y^*)\,y^* f_\theta(\theta^*) \\
&\qquad\qquad -\frac{\tilde{E}_{l,1-\tau}}{1-\tilde{E}_{l,w}/\sigma}\frac{\tilde{E}_{l,w}}{\tilde{\varepsilon}_{l,w}}\int_{\mathbb{R}_+} g_y(y)\,\frac{1-T'(y)}{1-T'(y^*)}\bar{\gamma}(y,y^*)\,yf_y(y)\,dy\Bigg\} \\
=&\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\left\{\tilde{E}_{l,1-\tau}g_y(y^*)\,\frac{1}{\sigma}y^* f_\theta(\theta^*)-\tilde{E}_{l,1-\tau}\int_{\mathbb{R}_+} g_y(y)\,\frac{1-T'(y)}{1-T'(y^*)}\bar{\gamma}(y,y^*)\,yf_y(y)\,dy\right\} \\
=&-\tilde{E}_{l,1-\tau}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\int_{\mathbb{R}_+} g_y(y)\,\frac{1-T'(y)}{1-T'(y^*)}\gamma(y,y^*)\,yf_y(y)\,dy.
\end{aligned}$$

Using the fact that $\bar{\gamma}(y,y^*)$ depends only on y^* and equation (64), we obtain another way of writing these terms:

$$\begin{aligned}
&\tilde{E}_{l,1-\tau}\frac{1}{\sigma}g_y(y^*)\,y^* f_y(y^*)-\tilde{E}_{l,1-\tau}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\bar{\gamma}(y,y^*)\int_{\mathbb{R}_+} g_y(y)\,\frac{1-T'(y)}{1-T'(y^*)}yf_y(y)\,dy \\
=&\tilde{E}_{l,1-\tau}\frac{1}{\sigma}g_y(y^*)\,y^* f_y(y^*)-\tilde{E}_{l,1-\tau}\frac{1}{\sigma}\frac{y^* f_y(y^*)}{\int_{\mathbb{R}_+} yf_y(y)\,dy}\int_{\mathbb{R}_+} g_y(y)\,\frac{1-T'(y)}{1-T'(y^*)}yf_y(y)\,dy \\
=&\frac{\tilde{E}_{l,1-\tau}}{\sigma}\frac{y^* f_y(y^*)}{\int_{\mathbb{R}_+} yf_y(y)\,dy}g_y(y^*)\int_{\mathbb{R}_+}\left(1-\frac{(1-T'(y))\,g_y(y)}{(1-T'(y^*))\,g_y(y^*)}\right) yf_y(y)\,dy.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\lambda^{-1}\mathrm{d}\mathscr{G}(T,h) \\
=&-\int_{y^*}^{\infty} g_y(y)\,f_y(y)\,dy-\tilde{E}_{l,1-\tau}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\int_{\mathbb{R}_+} g_y(y)\,\frac{1-T'(y)}{1-T'(y^*)}\gamma(y,y^*)\,yf_y(y)\,dy \\
=&-\int_{y^*}^{\infty} g_y(y)\,f_y(y)\,dy+\frac{\tilde{E}_{l,1-\tau}}{\sigma}\frac{y^* f_y(y^*)}{\int_{\mathbb{R}_+} yf_y(y)\,dy}g_y(y^*)\int_{\mathbb{R}_+}\left(1-\frac{(1-T'(y))\,g_y(y)}{(1-T'(y^*))\,g_y(y^*)}\right) yf_y(y)\,dy.
\end{aligned}$$

Finally, the effect of the perturbation on social welfare is given by

$$\mathrm{d}\mathscr{W}(T,h)=\mathrm{d}\mathscr{R}(T,h)+\lambda^{-1}\mathrm{d}\mathscr{G}(T,h).$$

We thus obtain

$$\begin{aligned} d\mathscr{W}(T,h) =& 1 - F_y(y^*) - \int_{y^*}^{\infty} g(y) f_y(y)\, dy - \frac{T'(y^*)}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau} y^* f_y(y^*) - \frac{1}{1-T'(y^*)} y^* f_y(y^*) \dots \\ & \times \left[\frac{1}{\sigma}\tilde{E}_{l,1-\tau}(1+\tilde{\varepsilon}_{l,w}) \int_{\mathbb{R}_+} (T'(y) - T'(y^*)) \frac{y f_y(y)}{\int_{\mathbb{R}_+} y' f_y(y')\, dy'} dy \dots \right. \\ & \left. + \frac{1}{\sigma}\tilde{E}_{l,1-\tau} \int_{\mathbb{R}_+} ((1-T'(y)) g_y(y) - (1-T'(y^*)) g_y(y^*)) \frac{y f_y(y)}{\int_{\mathbb{R}_+} y' f_y(y')\, dy'} dy \right]. \end{aligned}$$

Reorganizing the terms easily leads to formula (32).

□

B.3.6 Proof of Proposition 3

Suppose that the production function is Translog, as defined in Example 2, with the functional form specification (19).

Proof. A Taylor expansion of the right hand side of (19) as $\theta \to \theta'$ writes

$$\bar{\beta}(\theta,\theta') = \alpha\left[1 - \exp\left(-\frac{1}{2s^2}(\theta-\theta')^2\right)\right] = -\alpha \sum_{n=1}^{N} \frac{1}{n!}\left(-\frac{1}{2s^2}(\theta-\theta')^2\right)^n + o(\theta-\theta')^{2N+1}.$$

Ignoring for now the error term $o(\theta-\theta')^{2N+1}$ and denoting by $\bar{\beta}_N(\theta,\theta')$ the first N terms of the Taylor expansion, we get

$$\begin{aligned} \bar{\gamma}_N(\theta,\theta') =& \chi(\theta') + \frac{1}{\chi(\theta)}\bar{\beta}_N(\theta,\theta') = \chi(\theta') + \frac{\alpha}{\chi(\theta)} \sum_{n=1}^{N} \frac{(-1)^{n+1}}{2^n s^{2n} n!} ((\theta-\mu_\theta) - (\theta'-\mu_\theta))^{2n} \\ =& \chi(\theta') + \frac{\alpha}{\chi(\theta)} \sum_{n=1}^{N} \frac{(-1)^{n+1}}{2^n s^{2n} n!} \left\{ \sum_{k=0}^{2n} \begin{pmatrix} 2n \\ k \end{pmatrix} (\theta-\mu_\theta)^k (\theta'-\mu_\theta)^{2n-k} \right\} \\ =& \chi(\theta') + \frac{\alpha}{\chi(\theta)} \sum_{k=0}^{2N} \sum_{n=\max\{1,\lceil \frac{k}{2} \rceil\}}^{N} \frac{(-1)^{n+1}}{2^n s^{2n} n!} \begin{pmatrix} 2n \\ k \end{pmatrix} (\theta-\mu_\theta)^k (\theta'-\mu_\theta)^{2n-k} \\ =& \chi(\theta') + \frac{\alpha}{\chi(\theta)} \sum_{k=0}^{2N} A_k(\theta')(\theta-\mu_\theta)^k, \end{aligned}$$

where $\mu_\theta = \int_0^\infty \theta f_\theta(\theta)\, d\theta$ and

$$A_k(\theta') \equiv \sum_{n=\max\{1,\lceil \frac{k}{2} \rceil\}}^{N} \frac{(-1)^{n+1}}{2^n s^{2n} n!} \begin{pmatrix} 2n \\ k \end{pmatrix} (\theta'-\mu_\theta)^{2n-k}.$$

We therefore obtain that $\bar{\gamma}_N(\theta,\theta')$, and hence the approximate kernel $K_1(\theta,\theta') = \tilde{E}_{l,w}(\theta)\bar{\gamma}_N(\theta,\theta')$ of the integral equation (21), can be written as the sum of multiplicatively separable functions, which allows us to derive a simple closed-form solution for $\mathrm{d}\hat{l}\left(\theta,h\right)$ at an arbitrary degree of precision. Letting $\hat{h}'\left(y\right)\equiv\frac{h'(y)}{1-T'(y)}$, we can write the approximate solution to the integral equation (21) as

$$\mathrm{d}\hat{l}_{N}\left(\theta,h\right)=-\tilde{E}_{l,1-\tau}\left(\theta\right)\hat{h}'\left(y\left(\theta\right)\right)+\int_{\Theta}\left\{\sum_{i=0}^{2N+1}\kappa_{i,1}\left(\theta\right)\kappa_{i,2}\left(\theta'\right)\right\}\mathrm{d}\hat{l}_{N}\left(\theta',h\right)d\theta', \tag{68}$$

where

$$\kappa_{0,1}\left(\theta\right)=\tilde{E}_{l,w}\left(\theta\right),\ \kappa_{0,2}\left(\theta'\right)=\chi\left(\theta'\right),$$

and for all $i\in\{1,\ldots,2N+1\}$,

$$\kappa_{i,1}\left(\theta\right)=\frac{\tilde{E}_{l,w}\left(\theta\right)}{\chi\left(\theta\right)}\left(\theta-\mu_{\theta}\right)^{i-1},\ \kappa_{i,2}\left(\theta'\right)=A_{i-1}\left(\theta'\right).$$

This equation can thus be rewritten as

$$\mathrm{d}\hat{l}_{N}\left(\theta,h\right)=-\tilde{E}_{l,1-\tau}\left(\theta\right)\hat{h}'\left(y\left(\theta\right)\right)+\sum_{i=0}^{2N+1}a_{i}\kappa_{i,1}\left(\theta\right),$$

where the constants $\left(a_{i}\right)_{i\geq0}$ are given by

$$a_{i}=\int_{\Theta}\kappa_{i,2}\left(\theta'\right)\mathrm{d}\hat{l}\left(\theta',h\right)d\theta'.$$

We thus obtain immediately that the solution to (21) is of the form

$$\mathrm{d}\hat{l}_{N}\left(\theta,h\right)=-\tilde{E}_{l,1-\tau}\left(\theta\right)\hat{h}'\left(y\left(\theta\right)\right)+\tilde{E}_{l,w}\left(\theta\right)\left[a_{0}+\frac{1}{\chi\left(\theta\right)}\sum_{n=0}^{N}a_{n+1}\left(\theta-\mu_{\theta}\right)^{n}\right].$$

To characterize the constants $\left(a_{i}\right)_{i\geq0}$ in closed form, integrate over Θ both sides of (68) evaluated at θ' and multiplied by $\kappa_{i,2}\left(\theta'\right)$:

$$\begin{aligned}a_{i}=\int_{\Theta}\kappa_{i,2}\left(\theta'\right)\mathrm{d}\hat{l}_{N}\left(\theta',h\right)d\theta'=&-\int_{\Theta}\kappa_{i,2}\left(\theta'\right)\tilde{E}_{l,1-\tau}\left(\theta'\right)\hat{h}'\left(y\left(\theta'\right)\right)d\theta'\\&+\sum_{j=0}^{2N+1}a_{j}\int_{\Theta}\kappa_{j,1}\left(\theta'\right)\kappa_{i,2}\left(\theta'\right)d\theta',\end{aligned}$$

so that the vector $\mathbf{a}=\left(a_{i}\right)_{0\leq i\leq2N+1}$ is the solution to the linear system

$$\left[\mathbf{I}_{2N+2}-\mathbf{A}\right]\mathbf{a}=\mathbf{h},$$

where $\mathbf{I}_{2N+2}$ is the $(2N+2)\times(2N+2)$-identity matrix, and the matrix $\mathbf{A}=\left(\mathbf{A}_{i,j}\right)_{0\leq i,j\leq2N+1}$ and the vector $\mathbf{h} = (\mathbf{h}_i)_{0\leq i\leq 2N+1}$ are given by:

$$\mathbf{A}_{i,j} = \int_\Theta \kappa_{j,1}(\theta)\,\kappa_{i,2}(\theta)\,d\theta,$$
$$\mathbf{h}_i = \int_\Theta \kappa_{i,2}(\theta)\,\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(y(\theta))\,d\theta.$$

We assume that the determinant $\det(\mathbf{I}_{2N+2} - \mathbf{A}) \neq 0$, so that this system can be inverted. The inverse matrix $[\mathbf{I}_{2N+2} - \mathbf{A}]^{-1}$ can be expressed as the transpose of the matrix of cofactors from $[\mathbf{I}_{2N+2} - \mathbf{A}]$, which we denote by $\mathbf{C}$, normalized by the Fredholm determinant $\det(\mathbf{I}_{2N+2} - \mathbf{A})$. Thus we have

$$a_i = \frac{1}{\det(\mathbf{I}_{2N+2} - \mathbf{A})}\sum_{j=0}^{2N+1} \mathbf{C}_{j,i}\mathbf{h}_j,$$

so that the solution to the integral equation writes

$$\begin{aligned}
\mathrm{d}\hat{l}_N(\theta,h) &= -\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(y(\theta)) + \sum_{i=0}^{2N+1}\left[\frac{1}{\det(\mathbf{I}_{2N+2} - \mathbf{A})}\sum_{j=0}^{2N+1}\mathbf{C}_{j,i}\mathbf{h}_j\right]\kappa_{i,1}(\theta)\\
&= -\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(y(\theta))\\
&\quad + \int_\Theta\left[\frac{1}{\det(\mathbf{I}_{2N+2} - \mathbf{A})}\sum_{i=0}^{2N+1}\sum_{j=0}^{2N+1}\mathbf{C}_{j,i}\kappa_{i,1}(\theta)\,\kappa_{j,2}(\theta')\right]\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(y(\theta'))\,d\theta',
\end{aligned}$$

or, denoting by $\mathcal{R}_N(\theta,\theta')$ the term in square brackets in the previous equation, or the resolvent kernel of the integral equation,

$$\mathrm{d}\hat{l}_N(\theta,h) = -\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(y(\theta)) + \int_\Theta \mathcal{R}_N(\theta,\theta')\,\tilde{E}_{l,1-\tau}(\theta')\,\hat{h}'(y(\theta'))\,d\theta'.$$

Finally, we can show that

$$\sum_{i=0}^{2N+1}\sum_{j=0}^{2N+1}\mathbf{C}_{j,i}\kappa_{i,1}(\theta)\,\kappa_{j,2}(\theta') = -\det(\mathbf{D}(\theta,\theta'))$$

with $\mathbf{D}_{1,1}(\theta,\theta') = 0$, $\mathbf{D}_{i,1}(\theta,\theta') = \kappa_{i-2,2}(\theta')$ for $i \geq 2$, $\mathbf{D}_{1,j}(\theta,\theta') = \kappa_{j-2,1}(\theta')$ for $j \geq 2$, and $(\mathbf{D}_{i,j}(\theta,\theta'))_{2\leq i,j\leq 2N+3} = \mathbf{I}_{2N+2} - \mathbf{A}$, so that

$$\mathcal{R}_N(\theta,\theta') = -\frac{\det(\mathbf{D}(\theta,\theta'))}{\det(\mathbf{I}_2 - \mathbf{A})}.$$

This proves the first Fredholm theorem with a separable kernel.

We now compute a bound on the error in the approximation of the true solution $\mathrm{d}\hat{l}(\theta,h)$ by $\mathrm{d}\hat{l}_N(\theta,h)$. We have

$$|\bar{\gamma}(\theta,\theta') - \bar{\gamma}_N(\theta,\theta')| \leq \frac{\alpha\,|\theta-\theta'|^{2N+1}}{2^{N+1}s^{2(N+1)}\,(N+1)!},$$

which implies, denoting $K_{1,N}(\theta,\theta') = \tilde{E}_{l,w}(\theta)\,\bar{\gamma}_N(\theta,\theta')$,

$$\int_\Theta |K_1(\theta,\theta') - K_{1,N}(\theta,\theta')|\, d\theta' \leq \int_\Theta \left|\tilde{E}_{l,w}(\theta)\right| \frac{\alpha\,|\theta-\theta'|^{2N+1}}{2^{N+1}s^{2(N+1)}(N+1)!} d\theta'$$
$$\leq \frac{\alpha\left(\bar{\theta}-\underline{\theta}\right)^{2(N+1)}\sup_\Theta\left|\tilde{E}_{l,w}(\theta)\right|}{2^{N+1}s^{2(N+1)}(N+1)!} \equiv \varepsilon_N,$$

which converges to zero as $N\to\infty$. Assume that $\varepsilon_N(1+M_N)<1$, where

$$M_N = \sup_{\theta\in\Theta}\int_\Theta |\mathcal{R}_N(\theta,\theta')|\, d\theta'.$$

We then have (see Theorem 2.6.1 in Zemyan (2012) and Section 13.14 in Polyanin and Manzhirov (2008)):

$$\left|\mathrm{d}\hat{l}(\theta,h) - \mathrm{d}\hat{l}_N(\theta,h)\right| \leq \varepsilon_N \frac{(1+M_N)^2\sup_\Theta\left|\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(y(\theta))\right|}{1-\varepsilon_N(1+M_N)}.$$

Now consider the approximation of $\mathrm{d}\hat{l}(\theta,h)$ obtained from the third-order Taylor expansion of $\bar{\beta}(\theta,\theta')$:

$$\begin{aligned}
\mathrm{d}\hat{l}(\theta,h) &\approx -\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(y(\theta)) + \tilde{E}_{l,w}(\theta)\left[a_0 + \frac{1}{\chi(\theta)}\sum_{n=0}^{2} a_{n+1}(\theta-\mu_\theta)^n\right]\\
&= -\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(y(\theta)) + \tilde{E}_{l,w}(\theta)\left[a_0 + \frac{a_1}{\chi(\theta)} + \frac{a_3}{\chi(\theta)}\left((\theta-\mu_\theta)^2 + \frac{a_2}{a_3}(\theta-\mu_\theta)\right)\right]\\
&= -\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(y(\theta)) + \tilde{E}_{l,w}(\theta)\left[a_0 + \frac{1}{\chi(\theta)}\left(a_1 - \frac{a_2^2}{4a_3}\right) + \frac{a_3}{\chi(\theta)}\left(\theta-\mu_\theta+\frac{a_2}{2a_3}\right)^2\right].
\end{aligned}$$

Letting $\tilde{\theta} = \mu_\theta - \frac{a_2}{2a_3}$ and recalling the approximation $\bar{\gamma}(\theta,\theta') = \chi(\theta') + \frac{\alpha}{2s^2}\frac{(\theta-\theta')^2}{\chi(\theta)}$, we get

$$\begin{aligned}
\mathrm{d}\hat{l}(\theta,h) &\approx -\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(y(\theta)) + \tilde{E}_{l,w}(\theta)\times\ldots\\
&\quad\left[a_0 - \frac{2s^2a_3}{\alpha}\chi\left(\tilde{\theta}\right) + \frac{1}{\chi(\theta)}\left(a_1-\frac{a_2^2}{4a_3}\right) + \frac{2s^2}{\alpha}a_3\left(\chi\left(\tilde{\theta}\right) + \frac{\alpha}{2s^2}\frac{1}{\chi(\theta)}\left(\theta-\tilde{\theta}\right)^2\right)\right]\\
&= -\tilde{E}_{l,1-\tau}(\theta)\,\hat{h}'(y(\theta)) + \tilde{E}_{l,w}(\theta)\left[c_1 + \frac{c_2}{\chi(\theta)} + c_3\bar{\gamma}\left(\theta,\tilde{\theta}\right)\right],
\end{aligned}$$

which proves equation (33).

We can obtain (33) by computing the second-order approximation directly. We find

$$\begin{aligned}
\bar{\beta}(\theta,\theta') \approx & \frac{1}{4s^2}\alpha(\theta)(\theta-\mu_\theta)^2\alpha(\theta') + \frac{1}{4s^2}\alpha(\theta)\alpha(\theta')(\theta'-\mu_\theta)^2\\
&- \frac{1}{2s^2}\alpha(\theta)(\theta-\mu_\theta)\alpha(\theta')(\theta'-\mu_\theta),
\end{aligned}$$

so that

$$\mathrm{d}\hat{l}\left(\theta,h\right)\approx-\tilde{E}_{l,1-\tau}\left(\theta\right)\hat{h}'\left(y\left(\theta\right)\right)+\int_{\Theta}\left\{\sum_{i=1}^{4}\kappa_{i,1}\left(\theta\right)\kappa_{i,2}\left(\theta'\right)\right\}\mathrm{d}\hat{l}\left(\theta',h\right)d\theta',$$

with

$$\begin{aligned}\kappa_{1,1}\left(\theta\right)=&\tilde{E}_{l,w}\left(\theta\right)\\ \kappa_{2,1}\left(\theta\right)=&\tilde{E}_{l,w}\left(\theta\right)\frac{\alpha\left(\theta\right)}{\chi\left(\theta\right)}\\ \kappa_{3,1}\left(\theta\right)=&\tilde{E}_{l,w}\left(\theta\right)\frac{\alpha\left(\theta\right)}{\chi\left(\theta\right)}\left(\theta-\mu_{\theta}\right)\\ \kappa_{4,1}\left(\theta\right)=&\tilde{E}_{l,w}\left(\theta\right)\frac{\alpha\left(\theta\right)}{\chi\left(\theta\right)}\left(\theta-\mu_{\theta}\right)^{2},\end{aligned}$$

and

$$\begin{aligned}\kappa_{1,2}\left(\theta'\right)=&\chi\left(\theta'\right)\\ \kappa_{2,2}\left(\theta'\right)=&\frac{\alpha\left(\theta'\right)}{4s^{2}}\left(\theta'-\mu_{\theta}\right)^{2}\\ \kappa_{3,2}\left(\theta'\right)=&-\frac{\alpha\left(\theta'\right)}{2s^{2}}\left(\theta'-\mu_{\theta}\right)\\ \kappa_{4,2}\left(\theta'\right)=&\frac{\alpha\left(\theta'\right)}{4s^{2}}.\end{aligned}$$

Thus

$$\mathrm{d}\hat{l}\left(\theta,h\right)=-\tilde{E}_{l,1-\tau}\left(\theta\right)\hat{h}'\left(y\left(\theta\right)\right)+\sum_{i=1}^{4}a_{i}\kappa_{i,1}\left(\theta\right),$$

with

$$\begin{pmatrix}a_{1}\\a_{2}\\a_{3}\\a_{4}\end{pmatrix}=\left[\mathbf{I}_{4}-\mathbf{A}\right]^{-1}\begin{pmatrix}\int_{\Theta}\kappa_{1,2}\left(\theta\right)\tilde{E}_{l,1-\tau}\left(\theta\right)\hat{h}'\left(y\left(\theta\right)\right)d\theta\\ \int_{\Theta}\kappa_{2,2}\left(\theta\right)\tilde{E}_{l,1-\tau}\left(\theta\right)\hat{h}'\left(y\left(\theta\right)\right)d\theta\\ \int_{\Theta}\kappa_{3,2}\left(\theta\right)\tilde{E}_{l,1-\tau}\left(\theta\right)\hat{h}'\left(y\left(\theta\right)\right)d\theta\\ \int_{\Theta}\kappa_{4,2}\left(\theta\right)\tilde{E}_{l,1-\tau}\left(\theta\right)\hat{h}'\left(y\left(\theta\right)\right)d\theta\end{pmatrix},$$

where the 4×4 matrix $\mathbf{A}$ is defined by

$$\mathbf{A}_{i,j}=\int_{\Theta}\kappa_{j,1}\left(\theta\right)\kappa_{i,2}\left(\theta\right)d\theta.$$

We can finally write

$$\begin{aligned}\mathrm{d}\hat{l}\left(\theta,h\right)\approx&-\tilde{E}_{l,1-\tau}\left(\theta\right)\hat{h}'\left(y\left(\theta\right)\right)+\tilde{E}_{l,w}\left(\theta\right)\left[a_{1}+\frac{\alpha\left(\theta\right)}{\chi\left(\theta\right)}\left(a_{2}+a_{3}\left(\theta-\mu_{\theta}\right)+a_{4}\left(\theta-\mu_{\theta}\right)^{2}\right)\right]\\ =&-\tilde{E}_{l,1-\tau}\left(\theta\right)\hat{h}'\left(y\left(\theta\right)\right)+\tilde{E}_{l,w}\left(\theta\right)\left[a_{1}+\frac{\alpha\left(\theta\right)}{\chi\left(\theta\right)}a_{2}+a_{4}\frac{\alpha\left(\theta\right)}{\chi\left(\theta\right)}\left(\left(\theta-\mu_{\theta}\right)^{2}+\frac{a_{3}}{a_{4}}\left(\theta-\mu_{\theta}\right)\right)\right].\end{aligned}$$

The term in square brackets can be rewritten as

$$
\begin{aligned}
&a_1 + \frac{\alpha(\theta)}{\chi(\theta)} a_2 + a_4 \frac{\alpha(\theta)}{\chi(\theta)} \left(\left(\theta - \mu_\theta + \frac{a_3}{2a_4} \right)^2 - \left(\frac{a_3}{2a_4} \right)^2 \right) \\
=&a_1 + \left(a_2 - \frac{a_3^2}{4a_4} \right) \frac{\alpha(\theta)}{\chi(\theta)} + a_4 \frac{\alpha(\theta)}{\chi(\theta)} \left(\theta - \tilde{\theta} \right)^2 \\
=&\left(a_1 - 2s^2 a_4 \chi\left(\tilde{\theta}\right) \right) + \left(a_2 - \frac{a_3^2}{4a_4} \right) \frac{\alpha(\theta)}{\chi(\theta)} + 2s^2 a_4 \left(\chi\left(\tilde{\theta}\right) + \frac{\alpha(\theta)}{2s^2 \chi(\theta)} \left(\theta - \tilde{\theta} \right)^2 \right),
\end{aligned}
$$

so that

$$
\mathrm{d}\hat{l}(\theta, h) \approx -\tilde{E}_{l,1-\tau}(\theta)\, \hat{h}'(y(\theta)) + \tilde{E}_{l,w}(\theta) \left[c_1 + c_2 \frac{\alpha(\theta)}{\chi(\theta)} + c_3 \bar{\gamma}\left(\theta, \tilde{\theta}\right) \right],
$$

where

$$
\begin{aligned}
\tilde{\theta} =& \mu_\theta - \frac{a_3}{2a_4}, \\
c_1 =& a_1 - 2s^2 a_4 \chi\left(\tilde{\theta}\right), \\
c_2 =& a_2 - \frac{a_3^2}{4a_4}, \\
c_3 =& 2s^2 a_4.
\end{aligned}
$$

This concludes the proof.

□

B.4 Proofs of Sections 3 and 4

B.4.1 Proof of Proposition 4

We start by deriving the formula (42) for optimal taxes using mechanism design tools, i.e., by solving the optimal control problem (37, 38, 40, 41).

Proof. The Lagrangian writes:

$$
\begin{aligned}
\mathcal{L} =& \int_\Theta u(V(\theta))\, \tilde{f}_\theta(\theta) d\theta + \lambda \left\{ \mathscr{F}(\mathscr{L}) - \int_\Theta \left[V(\theta) + v(l(\theta)) \right] f_\theta(\theta) d\theta \right\} \\
&- \int_\Theta \mu(\theta) v'(l(\theta))\, l(\theta) \frac{\omega_1[\theta, l(\theta) f_\theta(\theta), \mathscr{L}] + [l(\theta) f_\theta'(\theta) + b(\theta) f_\theta(\theta)]\, \omega_2[\theta, l(\theta) f_\theta(\theta), \mathscr{L}]}{\omega[\theta, l(\theta) f_\theta(\theta), \mathscr{L}]} d\theta \\
&- \int_\Theta \mu'(\theta) V(\theta) d\theta - \int_\Theta \eta(\theta) b(\theta) d\theta - \int_\Theta \eta'(\theta)\, l(\theta) d\theta.
\end{aligned}
$$

For simplicity of notation we denote

$$
\begin{aligned}
\hat{w}(\theta) \equiv& \hat{\omega}[\theta, l(\theta) f_\theta(\theta), \mathscr{L}] \\
\equiv& \frac{\omega_1[\theta, l(\theta) f_\theta(\theta), \mathscr{L}] + [l(\theta) f_\theta'(\theta) + b(\theta) f_\theta(\theta)]\, \omega_2[\theta, l(\theta) f_\theta(\theta), \mathscr{L}]}{\omega[\theta, l(\theta) f_\theta(\theta), \mathscr{L}]}.
\end{aligned}
\tag{69}
$$

The first-order condition for $V(\theta)$ writes:

$$u'\left(V(\theta)\right)\tilde{f}_{\theta}(\theta)-\lambda f_{\theta}(\theta)-\mu'(\theta)=0. \tag{70}$$

The first-order conditions for $b(\theta)$ writes:

$$-\mu(\theta)v'\left(l(\theta)\right)l(\theta)\frac{\partial\hat{w}(\theta)}{\partial b(\theta)}-\eta(\theta)=0. \tag{71}$$

The first-order conditions for $l(\theta)$ is obtained by perturbing $\mathscr{L}$ in the Dirac direction δ_{θ} and evaluating the Gateaux derivative of $\mathcal{L}$ (i.e., heuristically, "$\frac{\partial\mathcal{L}}{\partial l(\theta)}$"):

$$\begin{aligned}\mathrm{d}\mathcal{L}\left(\mathscr{L},\delta_{\theta}\right)=&\lambda w(\theta)f_{\theta}(\theta)-\lambda v'(l(\theta))f_{\theta}(\theta)-\mu(\theta)v''\left(l(\theta)\right)l(\theta)\hat{w}(\theta)-\mu(\theta)v'(l(\theta))\hat{w}(\theta)\\&-\int_{\Theta}\mu(\theta')v'(l(\theta'))l(\theta')\mathrm{d}\hat{\omega}\left(\theta',\delta_{\theta}\right)d\theta'-\eta'(\theta)=0,\end{aligned} \tag{72}$$

where $\mathrm{d}\hat{\omega}\left(\theta',\delta_{\theta}\right)$ (or, heuristically, "$\frac{\partial\hat{w}(\theta')}{\partial l(\theta)}$") is defined as:

$$\mathrm{d}\hat{\omega}\left(\theta',\delta_{\theta}\right)\equiv\lim_{\mu\to 0}\frac{1}{\mu}\left\{\hat{\omega}\left[\theta',\left(l\left(\theta'\right)+\mu\delta_{\theta}\left(\theta'\right)\right)f_{\theta}\left(\theta'\right),\mathscr{L}+\mu\delta_{\theta}\right]-\hat{\omega}\left[\theta',l\left(\theta'\right)f_{\theta}\left(\theta'\right),\mathscr{L}\right]\right\}. \tag{73}$$

Now, note that

$$\frac{\partial\hat{w}(\theta)}{\partial b(\theta)}=\frac{\omega_{2}\left[\theta,l(\theta)f_{\theta}(\theta),\mathscr{L}\right]f_{\theta}(\theta)}{\omega(\theta,l(\theta)f_{\theta}(\theta),\mathscr{L})}=\frac{\bar{\bar{\gamma}}(\theta,\theta)}{l(\theta)},$$

where the second equality comes from the fact that, by definition of the own-wage elasticity,

$$\bar{\bar{\gamma}}\left(\theta,\theta\right)=\frac{\omega_{2}\left[\theta,l(\theta)f_{\theta}(\theta),\mathscr{L}\right]f_{\theta}(\theta)}{w\left(\theta\right)}l\left(\theta\right) \tag{74}$$

(intuitively, $\bar{\bar{\gamma}}=\frac{l}{w}\frac{\partial w}{\partial l}$, keeping $\mathscr{L}$ constant). Thus, we can rewrite (71) as:

$$\eta(\theta)=-\mu(\theta)v'\left(l(\theta)\right)l(\theta)\frac{\bar{\bar{\gamma}}\left(\theta,\theta\right)}{l(\theta)}=-\mu(\theta)v'\left(l(\theta)\right)\bar{\bar{\gamma}}\left(\theta,\theta\right), \tag{75}$$

which implies

$$\eta'(\theta)=-\mu'(\theta)v'\left(l(\theta)\right)\bar{\bar{\gamma}}\left(\theta,\theta\right)-\mu(\theta)v''\left(l(\theta)\right)l'(\theta)\bar{\bar{\gamma}}\left(\theta,\theta\right)-\mu(\theta)v'\left(l(\theta)\right)\bar{\bar{\gamma}}'\left(\theta,\theta\right).$$

Using this expression to substitute for $\eta'(\theta)$ into (72) yields

$$\begin{aligned}0=&\lambda w(\theta)f_{\theta}(\theta)-\lambda v'(l(\theta))f_{\theta}(\theta)-\mu(\theta)v''\left(l(\theta)\right)l(\theta)\hat{w}(\theta)-\mu(\theta)v'(l(\theta))\hat{w}(\theta)\\&+\mu'(\theta)v'\left(l(\theta)\right)\bar{\bar{\gamma}}\left(\theta,\theta\right)+\mu(\theta)v''\left(l(\theta)\right)b\left(\theta\right)\bar{\bar{\gamma}}\left(\theta,\theta\right)+\mu(\theta)v'\left(l(\theta)\right)\bar{\bar{\gamma}}'\left(\theta,\theta\right)\\&-\int_{\Theta}\mu(\theta')v'(l(\theta'))l(\theta')\frac{\partial\hat{w}(\theta')}{\partial l(\theta)}d\theta'.\end{aligned} \tag{76}$$

We now analyze the last line of this equation. From (73), we have

$$\begin{aligned} d\hat{\omega}\left(\theta', \delta_\theta\right) =& \hat{\omega}_2\left(\theta', l\left(\theta'\right) f_\theta\left(\theta'\right), \mathscr{L}\right) f_\theta\left(\theta'\right) \delta_\theta\left(\theta'\right) \\ &+ \lim_{\mu\to 0} \frac{1}{\mu} \left\{\hat{\omega}\left[\theta', l\left(\theta'\right) f_\theta\left(\theta'\right), \mathscr{L}+\mu\delta_\theta\right] - \hat{\omega}\left[\theta', l\left(\theta'\right) f_\theta\left(\theta'\right), \mathscr{L}\right]\right\} \\ \equiv& \hat{\omega}_2\left(\theta', l\left(\theta'\right) f_\theta\left(\theta'\right), \mathscr{L}\right) f_\theta\left(\theta'\right) \delta_\theta\left(\theta'\right) + \hat{\omega}_{3,\theta}\left[\theta', l\left(\theta'\right) f_\theta\left(\theta'\right), \mathscr{L}\right], \end{aligned}$$

where we introduce the short-hand notation $\hat{\omega}_{3,\theta}$ in the last line for simplicity of exposition. Denote by $\hat{\omega}_{13,\theta}$ and $\hat{\omega}_{23,\theta}$ the derivatives of $\hat{\omega}_{3,\theta}$ with respect to its first and second variables, respectively.

Now recall the notation (69) and note that

$$\begin{aligned} \hat{\omega}_2(\theta, l(\theta) f_\theta\left(\theta\right), \mathscr{L}) f_\theta\left(\theta\right) =& \frac{\left[f_\theta\left(\theta\right)\omega_{12} + f_\theta'(\theta)\omega_2 + \left(l(\theta) f_\theta'(\theta) + b(\theta) f_\theta(\theta)\right) f_\theta\left(\theta\right)\omega_{22}\right] w(\theta)}{w^2(\theta)} \\ &- \frac{\left[\omega_1 + \left(l(\theta) f_\theta'(\theta) + b(\theta) f_\theta(\theta)\right)\omega_2\right] f_\theta\left(\theta\right)\omega_2}{w^2(\theta)}, \end{aligned}$$

so that, using the definition of $\hat{w}(\theta)$ and equation (74),

$$\hat{\omega}_2\left[\theta, l(\theta) f_\theta\left(\theta\right), \mathscr{L}\right] f_\theta\left(\theta\right) = \frac{\omega_{12} + \frac{f_\theta'(\theta)}{f_\theta(\theta)}\omega_2 + \left[l(\theta) f_\theta'(\theta) + b(\theta) f_\theta(\theta)\right]\omega_{22}}{w(\theta)} f_\theta\left(\theta\right) - \frac{\hat{w}(\theta)}{l\left(\theta\right)} \bar{\bar{\gamma}}\left(\theta, \theta\right).$$

Now, we have

$$\begin{aligned} \left(\frac{\bar{\bar{\gamma}}\left(\theta,\theta\right)}{l\left(\theta\right)}\right)' =& \frac{\left\{\left[\omega_{21} + \left(l(\theta) f_\theta'(\theta) + b(\theta) f_\theta(\theta)\right)\omega_{22}\right] f_\theta\left(\theta\right) + \omega_2 f_\theta'\left(\theta\right)\right\} w\left(\theta\right)}{w^2\left(\theta\right)} \\ &- \frac{\omega_2 f_\theta\left(\theta\right)\left[\omega_1 + \left(l(\theta) f_\theta'(\theta) + b(\theta) f_\theta(\theta)\right)\omega_2\right]}{w^2\left(\theta\right)} \\ =& \frac{\omega_{12} + \frac{f_\theta'(\theta)}{f_\theta(\theta)}\omega_2 + \left[l(\theta) f_\theta'(\theta) + b(\theta) f_\theta(\theta)\right]\omega_{22}}{w\left(\theta\right)} f_\theta\left(\theta\right) - \frac{\hat{w}(\theta)}{l\left(\theta\right)} \bar{\bar{\gamma}}\left(\theta, \theta\right). \end{aligned}$$

Therefore, the previous two equalities imply

$$\hat{\omega}_2\left[\theta, l(\theta) f_\theta\left(\theta\right), \mathscr{L}\right] f_\theta\left(\theta\right) = \left(\frac{\bar{\bar{\gamma}}\left(\theta,\theta\right)}{l\left(\theta\right)}\right)' = \frac{\bar{\bar{\gamma}}'\left(\theta,\theta\right) l\left(\theta\right) - \bar{\bar{\gamma}}\left(\theta,\theta\right) b\left(\theta\right)}{l^2\left(\theta\right)}. \tag{77}$$

Next, from definition (69) we have (omitting the arguments $(\theta', L(\theta'), \mathscr{L})$ on the right hand side)

$$\begin{aligned} &\hat{\omega}_{3,\theta}\left[\theta', l\left(\theta'\right) f_\theta\left(\theta'\right), \mathscr{L}\right] \\ =& \frac{\omega_{13,\theta} + \left(l(\theta') f_\theta'(\theta') + b(\theta') f_\theta(\theta')\right)\omega_{23,\theta}}{w\left(\theta'\right)} - \frac{\left[\omega_1 + \left(l(\theta') f_\theta'(\theta') + b(\theta') f_\theta(\theta')\right)\omega_2\right]\omega_{3,\theta}}{w^2\left(\theta'\right)} \\ =& \frac{\omega_{13,\theta} + \left(l(\theta') f_\theta'(\theta') + b(\theta') f_\theta(\theta')\right)\omega_{23,\theta}}{w\left(\theta'\right)} - \frac{\hat{w}\left(\theta'\right)}{l\left(\theta\right)} \bar{\gamma}\left(\theta', \theta\right), \end{aligned} \tag{78}$$

where the second equality follows from the definition of $\hat{w}(\theta')$ and of the cross-wage elasticities

$$
\begin{aligned}
\bar{\gamma}(\theta',\theta) =& \frac{l(\theta)}{w(\theta')} \times \lim_{\mu\to 0} \frac{1}{\mu} \left\{ \omega\left[\theta', l(\theta') f_\theta(\theta'), \mathscr{L} + \mu\delta_\theta\right] - \omega\left[\theta', l(\theta') f_\theta(\theta'), \mathscr{L}\right]\right\} \\
=& \frac{l(\theta)}{w(\theta')} \omega_{3,\theta}\left(\theta', l(\theta') f_\theta(\theta'), \mathscr{L}\right).
\end{aligned}
$$

Note moreover that this equality implies

$$
\begin{aligned}
\frac{\partial \bar{\gamma}(\theta',\theta)}{\partial \theta'} =& l(\theta) \frac{\omega_{13,\theta}\left(\theta', l(\theta') f_\theta(\theta'), \mathscr{L}\right) + \left(l(\theta') f_\theta'(\theta') + b(\theta') f_\theta(\theta')\right) \omega_{23,\theta}\left(\theta', l(\theta') f_\theta(\theta'), \mathscr{L}\right)}{w(\theta')} \\
& - l(\theta) \frac{\omega_{3,\theta}\left(\theta', l(\theta') f_\theta(\theta'), \mathscr{L}\right)}{w^2(\theta')} \times \dots \\
& \left[\omega_1\left(\theta', l(\theta') f_\theta(\theta'), \mathscr{L}\right) + \left(l(\theta') f_\theta'(\theta') + b(\theta') f_\theta(\theta')\right) \omega_2\left(\theta', l(\theta') f_\theta(\theta'), \mathscr{L}\right)\right] \\
=& l(\theta) \frac{\omega_{13,\theta} + \left(l(\theta') f_\theta'(\theta') + b(\theta') f_\theta(\theta')\right) \omega_{23,\theta}}{w(\theta')} - \hat{w}(\theta') \bar{\gamma}(\theta',\theta),
\end{aligned}
$$

and thus, from (78)

$$
\hat{\omega}_{3,\theta}\left[\theta', l(\theta') f_\theta(\theta'), \mathscr{L}\right] = \frac{1}{l(\theta)} \frac{\partial \bar{\gamma}(\theta',\theta)}{\partial \theta'}. \tag{79}
$$

Substitute equations (77) and (79) in (76) to get:

$$
\begin{aligned}
0 =& \lambda w(\theta) f_\theta(\theta) - \lambda v'(l(\theta)) f_\theta(\theta) - \mu(\theta) v''(l(\theta)) l(\theta) \hat{w}(\theta) - \mu(\theta) v'(l(\theta)) \hat{w}(\theta) \\
& + \mu'(\theta) v'(l(\theta)) \bar{\bar{\gamma}}(\theta,\theta) + \mu(\theta) v''(l(\theta)) l'(\theta) \bar{\bar{\gamma}}(\theta,\theta) + \mu(\theta) v'(l(\theta)) \bar{\bar{\gamma}}'(\theta,\theta) \\
& - \int_\Theta \mu(\theta') v'(l(\theta')) l(\theta') \times \dots \\
& \left\{\hat{\omega}_2\left(\theta', l(\theta') f_\theta(\theta'), \mathscr{L}\right) f_\theta(\theta') \delta_\theta(\theta') + \hat{\omega}_{3,\theta}\left[\theta', l(\theta') f_\theta(\theta'), \mathscr{L}\right]\right\} d\theta' \\
=& \lambda w(\theta) f_\theta(\theta) - \lambda v'(l(\theta)) f_\theta(\theta) - \mu(\theta) v''(l(\theta)) l(\theta) \hat{w}(\theta) - \mu(\theta) v'(l(\theta)) \hat{w}(\theta) \\
& + \mu'(\theta) v'(l(\theta)) \bar{\bar{\gamma}}(\theta,\theta) + \mu(\theta) v''(l(\theta)) b(\theta) \bar{\bar{\gamma}}(\theta,\theta) \\
& + \left\{\mu(\theta) v'(l(\theta)) \frac{b(\theta)}{l(\theta)} \bar{\bar{\gamma}}(\theta,\theta) + \mu(\theta) v'(l(\theta)) l(\theta) \hat{\omega}_2 f_\theta(\theta)\right\} \\
& - \left\{\mu(\theta) v'(l(\theta)) l(\theta) \hat{\omega}_2\left(\theta, l(\theta) f_\theta(\theta), \mathscr{L}\right) f_\theta(\theta) + \int_\Theta \mu(\theta') v'(l(\theta')) l(\theta') \frac{1}{l(\theta)} \frac{\partial \bar{\gamma}(\theta',\theta)}{\partial \theta'} d\theta'\right\},
\end{aligned}
$$

and hence,

$$
\begin{aligned}
0 =& \lambda w(\theta) f_\theta(\theta) - \lambda v'(l(\theta)) f_\theta(\theta) - \mu(\theta) v''(l(\theta)) l(\theta) \hat{w}(\theta) - \mu(\theta) v'(l(\theta)) \hat{w}(\theta) \\
& + \mu'(\theta) v'(l(\theta)) \bar{\bar{\gamma}}(\theta,\theta) + \mu(\theta) v''(l(\theta)) b(\theta) \bar{\bar{\gamma}}(\theta,\theta) + \mu(\theta) v'(l(\theta)) \frac{b(\theta)}{l(\theta)} \bar{\bar{\gamma}}(\theta,\theta) \\
& - \int_\Theta \mu(\theta') v'(l(\theta')) \frac{l(\theta')}{l(\theta)} \frac{\partial \bar{\gamma}(\theta',\theta)}{\partial \theta'} d\theta'.
\end{aligned} \tag{80}
$$

Using the definition of the labor supply elasticity (6), we finally obtain

$$
\begin{aligned}
0 =& \lambda w(\theta) f_\theta(\theta) - \lambda v'(l(\theta)) f_\theta(\theta) - \mu(\theta) v''(l(\theta)) l(\theta)\hat{w}(\theta) - \mu(\theta) v'(l(\theta))\hat{w}(\theta) \\
&+ \mu'(\theta) v'(l(\theta)) \bar{\bar{\gamma}}(\theta,\theta) + \mu(\theta)\left(1+\frac{1}{\varepsilon_{l,1-\tau}(\theta)}\right) v'(l(\theta)) \frac{b(\theta)}{l(\theta)} \bar{\bar{\gamma}}(\theta,\theta) \\
&- \int_\Theta \mu(\theta') v'(l(\theta')) \frac{l(\theta')}{l(\theta)} \frac{\partial \bar{\gamma}(\theta',\theta)}{\partial \theta'} d\theta'. \qquad (81)
\end{aligned}
$$

Moreover, defining the wedge $(1-\tau(\theta))w(\theta) = v'(l(\theta))$ and noting that

$$
v'(l(\theta)) + v''(l(\theta)) l(\theta) = (1-\tau(\theta))w(\theta)\left(1+\frac{1}{\varepsilon_{l,1-\tau}(\theta)}\right),
$$

we can rewrite (81) as

$$
\begin{aligned}
0 =& \lambda w(\theta) f_\theta(\theta) - \lambda(1-\tau(\theta))w(\theta) f_\theta(\theta) - \mu(\theta)(1-\tau(\theta))w(\theta)\left(1+\frac{1}{\varepsilon_{l,1-\tau}(\theta)}\right)\hat{w}(\theta) \\
&+ \mu'(\theta)(1-\tau(\theta))w(\theta)\bar{\bar{\gamma}}(\theta,\theta) + \mu(\theta)\left(1+\frac{1}{\varepsilon_{l,1-\tau}(\theta)}\right)(1-\tau(\theta))w(\theta)\frac{b(\theta)}{l(\theta)}\bar{\bar{\gamma}}(\theta,\theta) \\
&- \frac{1}{l(\theta)}\int_\Theta \mu(\theta')(1-\tau(\theta'))y(\theta')\frac{\partial \bar{\gamma}(\theta',\theta)}{\partial \theta'} d\theta' \\
=& \lambda\tau(\theta) w(\theta) f_\theta(\theta) + \mu(\theta)(1-\tau(\theta))w(\theta)\left(1+\frac{1}{\varepsilon_{l,1-\tau}(\theta)}\right)\left\{\frac{l'(\theta)}{l(\theta)}\bar{\bar{\gamma}}(\theta,\theta) - \hat{w}(\theta)\right\} \\
&+ \mu'(\theta)(1-\tau(\theta))w(\theta)\bar{\bar{\gamma}}(\theta,\theta) - \frac{1}{l(\theta)}\int_\Theta \mu(\theta')(1-\tau(\theta'))y(\theta')\frac{\partial \bar{\gamma}(\theta',\theta)}{\partial \theta'} d\theta'.
\end{aligned}
$$

Using the fact that $\hat{w}(\theta) = \frac{w'(\theta)}{w(\theta)}$ and dividing through by $\lambda(1-\tau(\theta))w(\theta)f_\theta(\theta)$ yields

$$
\begin{aligned}
0 =& \frac{\tau(\theta)}{1-\tau(\theta)} + \frac{\mu(\theta)}{\lambda f_\theta(\theta)}\left(1+\frac{1}{\varepsilon_{l,1-\tau}(\theta)}\right)\left\{\frac{l'(\theta)}{l(\theta)}\bar{\bar{\gamma}}(\theta,\theta) - \frac{w'(\theta)}{w(\theta)}\right\} \\
&+ \frac{\mu'(\theta)}{\lambda f_\theta(\theta)}\bar{\bar{\gamma}}(\theta,\theta) - \frac{1}{\lambda(1-\tau(\theta))y(\theta)f_\theta(\theta)}\int_\Theta \mu(\theta')(1-\tau(\theta'))y(\theta')\frac{\partial \bar{\gamma}(\theta',\theta)}{\partial \theta'} d\theta',
\end{aligned}
$$

and hence, using the relationship between the densities of productivities and wages:

$$
\begin{aligned}
\frac{\tau(\theta)}{1-\tau(\theta)} =& \left(1+\frac{1}{\varepsilon_{l,1-\tau}(\theta)}\right)\frac{\mu(\theta)}{\lambda w(\theta) f_w(w(\theta))}\left(1-\bar{\bar{\gamma}}(\theta,\theta)\frac{\frac{l'(\theta)}{l(\theta)}}{\frac{w'(\theta)}{w(\theta)}}\right) - \frac{\mu'(\theta)}{\lambda f_\theta(\theta)}\bar{\bar{\gamma}}(\theta,\theta) \\
&+ \frac{1}{\lambda(1-\tau(\theta))y(\theta)f_\theta(\theta)}\int_\Theta \mu(\theta')(1-\tau(\theta'))y(\theta')\frac{\partial \bar{\gamma}(\theta',\theta)}{\partial \theta'} d\theta'. \qquad (82)
\end{aligned}
$$

Note that for a CES production function, we have $\frac{\partial \bar{\gamma}(\theta',\theta)}{\partial \theta'} = 0$.

Finally, an alternative optimal tax formula is given by integrating the previous equation by parts (with the appropriate boundary conditions on $\mu(\theta')$):

$$
\int_\Theta \mu(\theta')(1-\tau(\theta'))y(\theta')\frac{\partial \bar{\gamma}(\theta',\theta)}{\partial \theta'} d\theta' = -\int_\Theta \bar{\gamma}(\theta',\theta)\frac{d}{d\theta'}\left[\mu(\theta')(1-\tau(\theta'))y(\theta')\right] d\theta'.
$$

We therefore obtain

$$
\begin{aligned}
\frac{\tau(\theta)}{1-\tau(\theta)} =& \left(1+\frac{1}{\varepsilon_{l,1-\tau}(\theta)}\right)\frac{\mu(\theta)}{\lambda w(\theta)f_w(w(\theta))}\left(1-\bar{\bar{\gamma}}\left(\theta,\theta\right)\frac{\frac{l'(\theta)}{l(\theta)}}{\frac{w'(\theta)}{w(\theta)}}\right) \\
&-\frac{\mu'(\theta)}{\lambda f_\theta(\theta)}\bar{\bar{\gamma}}\left(\theta,\theta\right)-\frac{\int_\Theta\left[\mu(x)v'(l(x))l(x)\right]'\bar{\gamma}\left(x,\theta\right)dx}{\lambda(1-\tau(\theta))y(\theta)f_\theta(\theta)} \\
=& \left(1+\frac{1}{\varepsilon_{l,1-\tau}(\theta)}\right)\frac{\mu(\theta)}{\lambda w(\theta)f_w(w(\theta))}\left(1-\bar{\bar{\gamma}}\left(\theta,\theta\right)\frac{\frac{l'(\theta)}{l(\theta)}}{\frac{w'(\theta)}{w(\theta)}}\right) \\
&-\frac{\int_\Theta\left[\mu(x)v'(l(x))l(x)\right]'\gamma\left(x,\theta\right)dx}{\lambda(1-\tau(\theta))y(\theta)f_\theta(\theta)}-\frac{\mu'(\theta)}{\lambda f_\theta(\theta)}\bar{\bar{\gamma}}\left(\theta,\theta\right) \\
&+\frac{\left[\mu'(\theta)v'(l(\theta))l(\theta)+\mu(\theta)v''(l(\theta))l(\theta)+\mu(\theta)v'(l(\theta))l'(\theta)\right]\bar{\bar{\gamma}}\left(\theta,\theta\right)}{\lambda(1-\tau(\theta))y(\theta)f_\theta(\theta)},
\end{aligned}
$$

where the second equality uses the definition (11) of $\gamma\left(x,\theta\right)$. This implies

$$
\begin{aligned}
\frac{\tau(\theta)}{1-\tau(\theta)} =& \left(1+\frac{1}{\varepsilon_{l,1-\tau}(\theta)}\right)\frac{\mu(\theta)}{\lambda w(\theta)f_w(w(\theta))}-\frac{\int_\Theta\left[\mu(x)v'(l(x))l(x)\right]'\gamma\left(x,\theta\right)dx}{\lambda(1-\tau(\theta))y(\theta)f_\theta(\theta)} \\
&-\bar{\bar{\gamma}}\left(\theta,\theta\right)\left\{\left(1+\frac{v''(l(\theta))l(\theta)}{v'(l(\theta))}\right)\frac{\mu(\theta)}{\lambda f_\theta(\theta)}\frac{l'\left(\theta\right)}{l\left(\theta\right)}+\frac{\mu'(\theta)}{\lambda f_\theta(\theta)}-\dots\right. \\
&\left.\frac{\mu'(\theta)v'(l(\theta))l(\theta)+\mu(\theta)v''(l(\theta))l'(\theta)l(\theta)+\mu(\theta)v'(l(\theta))l'(\theta)}{\lambda v'(l(\theta))l(\theta)f_\theta(\theta)}\right\}.
\end{aligned}
$$

The terms in the curly brackets in the second and third lines cancel each other out, therefore

$$
\frac{\tau(\theta)}{1-\tau(\theta)}=\left(1+\frac{1}{\varepsilon_{l,1-\tau}(\theta)}\right)\frac{\mu(\theta)}{\lambda w(\theta)f_w(w(\theta))}-\frac{\int_\Theta\left[\mu(x)v'(l(x))l(x)\right]'\gamma\left(x,\theta\right)dx}{\lambda(1-\tau(\theta))y(\theta)f_\theta(\theta)}.
$$

Note finally that the first-order condition (70) implies

$$
\mu'(\theta)=u'\left(V(\theta)\right)\tilde{f}_\theta(\theta)-\lambda f_\theta(\theta),
$$

so that, using $\mu(\bar{\theta})=0$,

$$
\begin{aligned}
\mu(\theta)=&-\int_\theta^{\bar{\theta}}\left[u'\left(V(x)\right)\tilde{f}_\theta(x)-\lambda f_\theta(x)\right]dx \\
=&\lambda\int_\theta^{\bar{\theta}}\left[1-\frac{u'\left(V(x)\right)\tilde{f}_\theta(x)}{\lambda f_\theta(x)}\right]f_\theta(x)dx=\lambda\int_\theta^{\bar{\theta}}\left(1-g_\theta\left(x\right)\right)f_\theta(x)dx.
\end{aligned}
$$

This concludes the proof.

□

We now derive optimal taxes using the variational tools introduced in Section 2, i.e., we tackle problem (34, 35, 36) directly. We start by deriving the expression (48) for the counteracting perturbation h_2.

B.4.2 Proof of equation (48)

Proof. Consider the perturbation h_1 defined by $h_1(y) = \mathbb{I}_{\{y \geq y^*\}}$ and $h_1'(y) = \delta_{y^*}(y)$. As usual, denote by θ^* the type such that $y(\theta^*) = y^*$. We impose that $h_1 + h_2$ has the same effects on labor supply as those that h_1 induces in the partial equilibrium framework. The general equilibrium response to $h_1 + h_2$ is given by the solution to the following integral equation: for all $\theta \in \Theta$,

$$\begin{aligned} \mathrm{d}\hat{l}(\theta, h_1 + h_2) = &- \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta)}{1 - \bar{\bar{\gamma}}(\theta,\theta)\,\tilde{\varepsilon}_{l,w}(\theta)} \left\{ \frac{h_1'(y(\theta)) + h_2'(y(\theta))}{1 - T'(y(\theta))} \right\} \\ &+ \frac{\tilde{\varepsilon}_{l,w}(\theta)}{1 - \bar{\bar{\gamma}}(\theta,\theta)\,\tilde{\varepsilon}_{l,w}(\theta)} \int_\Theta \bar{\gamma}(\theta,\theta')\,\mathrm{d}\hat{l}(\theta', h_1 + h_2)\, d\theta'. \end{aligned} \tag{83}$$

The partial equilibrium effect of h_1, on the other hand, is given by: for all $\theta \in \Theta$,

$$\begin{aligned} \mathrm{d}\hat{l}_{PE}(\theta, h_1) &= - \tilde{\varepsilon}_{l,1-\tau}(\theta) \left\{ \frac{h_1'(y(\theta))}{1 - T'(y(\theta))} \right\} \\ &= - \frac{\varepsilon_{l,1-\tau}(\theta^*)}{1 - T'(y(\theta^*)) + \varepsilon_{l,1-\tau}(\theta^*)\, y(\theta^*)\, T''(y(\theta^*))} \delta_{y^*}(y(\theta)). \end{aligned} \tag{84}$$

In particular, note that in partial equilibrium, we have $\mathrm{d}\hat{l}_{PE}(\theta, h_1) = 0$ for all $\theta \neq \theta^*$, i.e., the only individuals who respond to a change in the marginal tax rate at income y^* are those whose type is θ^* (and hence whose income is y^*). Substituting for (84) in the left hand side and under the integral sign of (83) yields:

$$\begin{aligned} - \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta)}{1 - T'(y(\theta))} \delta_{y^*}(y(\theta)) = &- \frac{1}{1 - T'(y(\theta))} \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta)}{1 - \bar{\bar{\gamma}}(\theta,\theta)\,\tilde{\varepsilon}_{l,w}(\theta)} \left\{ \delta_{y^*}(y(\theta)) + h_2'(y(\theta)) \right\} \\ &- \frac{\tilde{\varepsilon}_{l,w}(\theta)}{1 - \bar{\bar{\gamma}}(\theta,\theta)\,\tilde{\varepsilon}_{l,w}(\theta)} \int_\Theta \bar{\gamma}(\theta,\theta') \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta')}{1 - T'(y(\theta'))} \delta_{y^*}(y(\theta'))\, d\theta', \end{aligned}$$

i.e., after changing variables in the integral in the second line of the previous expression,

$$\begin{aligned} \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta)}{1 - \bar{\bar{\gamma}}(\theta,\theta)\,\tilde{\varepsilon}_{l,w}(\theta)} \frac{h_2'(y(\theta))}{1 - T'(y(\theta))} = &\left(\tilde{\varepsilon}_{l,1-\tau}(\theta) - \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta)}{1 - \bar{\bar{\gamma}}(\theta,\theta)\,\tilde{\varepsilon}_{l,w}(\theta)} \right) \frac{\delta_{y^*}(y(\theta))}{1 - T'(y(\theta))} \\ &- \frac{\tilde{\varepsilon}_{l,w}(\theta)}{1 - \bar{\bar{\gamma}}(\theta,\theta)\,\tilde{\varepsilon}_{l,w}(\theta)} \left(\frac{dy(\theta^*)}{d\theta} \right)^{-1} \bar{\gamma}(\theta,\theta^*) \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1 - T'(y(\theta^*))}. \end{aligned}$$

This implies

$$\begin{aligned} \frac{h_2'(y(\theta))}{1 - T'(y(\theta))} = &- \bar{\bar{\gamma}}(\theta,\theta)\,\tilde{\varepsilon}_{l,w}(\theta) \frac{\delta_{y^*}(y(\theta))}{1 - T'(y(\theta))} \\ &- \frac{1}{1 - T'(y(\theta^*))} \frac{\tilde{\varepsilon}_{l,w}(\theta)\,\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{\tilde{\varepsilon}_{l,1-\tau}(\theta)} \left(\frac{dy(\theta^*)}{d\theta} \right)^{-1} \bar{\gamma}(\theta,\theta^*), \end{aligned}$$

which can be rewritten as

$$\begin{aligned}\frac{h_2'(y(\theta))}{1-T'(y(\theta))} &= -\frac{1}{1-T'(y(\theta^*))}\frac{\tilde{\varepsilon}_{l,w}(\theta)\,\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{\tilde{\varepsilon}_{l,1-\tau}(\theta)}\left[\bar{\bar{\gamma}}(\theta^*,\theta^*)\,\delta_{y^*}(y(\theta))+\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\bar{\gamma}(\theta,\theta^*)\right]\\ &= -\frac{1}{1-T'(y(\theta^*))}\frac{\tilde{\varepsilon}_{l,w}(\theta)\,\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{\tilde{\varepsilon}_{l,1-\tau}(\theta)}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\left[\bar{\bar{\gamma}}(\theta^*,\theta^*)\,\delta_{\theta^*}(\theta)+\bar{\gamma}(\theta,\theta^*)\right]\\ &= -\frac{1}{1-T'(y(\theta^*))}\frac{\tilde{\varepsilon}_{l,w}(\theta)}{\tilde{\varepsilon}_{l,1-\tau}(\theta)}\tilde{\varepsilon}_{l,1-\tau}(\theta^*)\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\gamma(\theta,\theta^*),\end{aligned}$$

or

$$\begin{aligned}h_2'(y(\theta)) &= -\left(1-T'(y(\theta))\right)\frac{\left(1-T'(y(\theta))-y(\theta)\,T''(y(\theta))\right)\varepsilon_{l,1-\tau}(\theta)}{1-T'(y(\theta))+\varepsilon_{l,1-\tau}(\theta)\,y(\theta)\,T''(y(\theta))}\\ &\quad\times\frac{1-T'(y(\theta))+\varepsilon_{l,1-\tau}(\theta)\,y(\theta)\,T''(y(\theta))}{\left(1-T'(y(\theta))\right)\varepsilon_{l,1-\tau}(\theta)}\frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\gamma(\theta,\theta^*)\\ &= -\frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\times\left(1-T'(y(\theta))-y(\theta)\,T''(y(\theta))\right)\gamma(\theta,\theta^*).\end{aligned}$$

This proves equation (48). Note that $h_2'(y(\theta))$ is a smooth function, except for a jump (formally, a Dirac term) at $\theta=\theta^*$, which adds to the jump in marginal tax rates defined by the tax reform h_1 at θ^* so that the total response of labor supply of individuals with income y^* is equal to their response to h_1 in the partial equilibrium environment.

Now, integrate this expression from 0 to y (letting $h_2(0)=0$) to get $h_2(y)$:

$$h_2(y(\theta)) = -\frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\int_0^y\left(1-T'(y')-y'T''(y')\right)\gamma(y',y^*)\,dy'.$$

The integral in this expression can be rewritten as

$$\begin{aligned}&\int_0^y\left(1-T'(y')-y'T''(y')\right)\bar{\bar{\gamma}}(y',y^*)\left(\frac{dy(\theta^*)}{d\theta}\right)\delta_{y^*}(y)\,dy'\\ &+\int_0^y\left(1-T'(y')-y'T''(y')\right)\bar{\gamma}(y',y^*)\,dy',\end{aligned}$$

and hence

$$\begin{aligned}h_2(y(\theta)) = -\frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\Bigg\{&\left(1-T'(y^*)-y^*T''(y^*)\right)\bar{\bar{\gamma}}(y^*,y^*)\,\mathbb{I}_{\{y\geq y^*\}}\\ &+\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\int_0^y\left(1-T'(y')-y'T''(y')\right)\bar{\gamma}(y',y^*)\,dy'\Bigg\}.\end{aligned}$$

(Note that we could have obtained this expression by integrating over $\theta\in\Theta$ rather than $y\in\mathbb{R}_+$, the expression

$$\frac{dh_2\left(y\left(\theta\right)\right)}{d\theta} = -\frac{\tilde{\varepsilon}_{l,1-\tau}\left(\theta^*\right)}{1-T'\left(y\left(\theta^*\right)\right)}\left(\frac{dy\left(\theta^*\right)}{d\theta}\right)^{-1}\left(1-T'\left(y\left(\theta\right)\right)-y\left(\theta\right)T''\left(y\left(\theta\right)\right)\right)$$
$$\times\left(\bar{\gamma}\left(\theta,\theta^*\right)+\bar{\bar{\gamma}}\left(\theta^*,\theta^*\right)\delta_{\theta^*}\left(\theta\right)\right)\times\left(\frac{dy\left(\theta\right)}{d\theta}\right),$$

and and using a change of variables in the integral.) This concludes the proof.

□

We now derive the effects of the combination of perturbations $h_1 + h_2$ on social welfare.

B.4.3 Proof of Proposition 5

Proof. We start by deriving the effect of the perturbation $h \equiv h_1 + h_2$ on the tax liability of any individual θ. We have, to a first order as $\mu \to 0$,

$$\mathrm{d}_h T\left(w\left(\theta\right)l\left(\theta\right)\right) \underset{\mu\to 0}{\sim} \mu^{-1}\left\{T\left(\tilde{w}\left(\theta\right)\left(l\left(\theta\right)+\mu \mathrm{d}l\left(\theta,h\right)\right)\right)-T\left(w\left(\theta\right)l\left(\theta\right)\right)\right\}+h_1\left(\tilde{w}\left(\theta\right)\left(l\left(\theta\right)+\mu dl_\theta\right)\right)$$
$$=\left\{\int_\Theta \gamma\left(\theta,\theta'\right)\mathrm{d}\hat{l}\left(\theta',h\right)d\theta'+\mathrm{d}\hat{l}\left(\theta,h\right)\right\}y\left(\theta\right)T'\left(y\left(\theta\right)\right)+h_1\left(y\left(\theta\right)\right)+\mu h_2\left(y\left(\theta\right)\right).$$

This can be rewritten as

$$\mathrm{d}_h T\left(y\left(\theta\right)\right) = \mathbb{I}_{\{y(\theta)\geq y^*\}} - \frac{\tilde{\varepsilon}_{l,1-\tau}\left(\theta^*\right)}{1-T'\left(y^*\right)}\left\{\int_\Theta \gamma\left(\theta,\theta'\right)\delta_{y^*}\left(y\left(\theta'\right)\right)d\theta'+\delta_{y^*}\left(y\left(\theta\right)\right)\right\}y\left(\theta\right)T'\left(y\left(\theta\right)\right)$$
$$-\frac{\tilde{\varepsilon}_{l,1-\tau}\left(\theta^*\right)}{1-T'\left(y\left(\theta^*\right)\right)}\left\{\left(1-T'\left(y^*\right)-y^*T''\left(y^*\right)\right)\bar{\bar{\gamma}}\left(y^*,y^*\right)\mathbb{I}_{\{y(\theta)\geq y^*\}}\right.$$
$$\left.+\left(\frac{dy\left(\theta^*\right)}{d\theta}\right)^{-1}\int_0^y\left(1-T'\left(y'\right)-y'T''\left(y'\right)\right)\bar{\gamma}\left(y',y^*\right)dy'\right\},$$

i.e.,

$$\mathrm{d}_h T\left(y\left(\theta\right)\right) = \mathbb{I}_{\{y(\theta)\geq y^*\}} - \frac{T'\left(y\left(\theta\right)\right)}{1-T'\left(y^*\right)}\tilde{\varepsilon}_{l,1-\tau}\left(\theta^*\right)y\left(\theta\right)\delta_{y^*}\left(y\left(\theta\right)\right)$$
$$-\frac{\tilde{\varepsilon}_{l,1-\tau}\left(\theta^*\right)}{1-T'\left(y\left(\theta^*\right)\right)}\left\{\gamma\left(\theta,\theta^*\right)\left(\frac{dy\left(\theta^*\right)}{d\theta}\right)^{-1}y\left(\theta\right)T'\left(y\left(\theta\right)\right)\ldots\right.$$
$$+\left(1-T'\left(y^*\right)-y^*T''\left(y^*\right)\right)\bar{\bar{\gamma}}\left(y^*,y^*\right)\mathbb{I}_{\{y(\theta)\geq y^*\}}\left(\frac{dy\left(\theta^*\right)}{d\theta}\right)$$
$$\left.+\int_0^y\left(1-T'\left(y'\right)-y'T''\left(y'\right)\right)\bar{\gamma}\left(y',y^*\right)dy'\right\},$$

i.e.,

$$\begin{aligned}\mathrm{d}_h T(y(\theta)) =& \mathbb{I}_{\{y(\theta)\geq y^*\}} - \frac{T'(y(\theta))}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau}(\theta^*)\, y(\theta)\,\delta_{y^*}(y(\theta)) \\ & - \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\left\{\gamma(\theta,\theta^*)\, y(\theta)\, T'(y(\theta))\ldots\right. \\ & \left. + \int_0^y \left(1-T'(y') - y'T''(y')\right)\gamma(y',y^*)\, dy'\right\}.\end{aligned}$$

Next, we derive the effect of the perturbation $h_1 + h_2$ on the aggregate tax liability. We have, to a first order as $\mu \to 0$,

$$\begin{aligned}& \mathrm{d}_h\left[\int_{\mathbb{R}_+} T(y)\, f_y(y)\, dy\right] \\ =& \int_{\mathbb{R}_+}\left\{\mathbb{I}_{\{y\geq y^*\}} - \frac{T'(y)}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau}(y^*)\, y\delta_{y^*}(y)\right\} f_y(y)\, dy - \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\ldots \\ & \times\int_{\mathbb{R}_+}\left\{T'(y)\,\gamma(y,y^*)\, y + \int_0^y \left(1-T'(y') - y'T''(y')\right)\gamma(y',y^*)\, dy'\right\} f_y(y)\, dy \\ =& 1 - F_y(y^*) - \frac{T'(y^*)}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau}(y^*)\, y^* f_y(y^*) \\ & - \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\int_{\mathbb{R}_+} T'(y)\,\gamma(y,y^*)\, y f_y(y)\, dy \\ & - \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\int_{y'=0}^{\infty}\int_{y=y'}^{\infty}\left(1-T'(y') - y'T''(y')\right)\gamma(y',y^*)\, f_y(y)\, dy dy',\end{aligned}$$

where we obtained the expression in the last line by switching the two integrals. Hence we get

$$\begin{aligned}& \mathrm{d}\mathscr{R}(T,h) \\ =& 1 - F_y(y^*) - \frac{T'(y^*)}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau}(y^*)\, y^* f_y(y^*) \\ & - \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\int_{\mathbb{R}_+} T'(y)\,\gamma(y,y^*)\, y f_y(y)\, dy \\ & - \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\int_{y'=0}^{\infty}\left(1-T'(y') - y'T''(y')\right)\gamma(y',y^*)\left(1-F_y(y')\right) dy' \\ =& 1 - F_y(y^*) - \frac{T'(y^*)}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau}(y^*)\, y^* f_y(y^*) - \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\ldots \\ & \times\int_0^{\infty}\left\{\left(1-T'(y) - yT''(y)\right)\left(1-F_y(y)\right) + T'(y)\, y f_y(y)\right\}\gamma(y,y^*)\, dy.\end{aligned}$$

But by Euler's homogeneous function theorem, we have

$$-\frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\int_0^{\infty}\gamma(y,y^*)\, y f_y(y)\, dy = 0,$$

so that substracting this (zero) term from the previous equation, we can write

$$\begin{aligned}
\mathrm{d}\mathscr{R}(T,h) =& 1-F_y(y^*)-\frac{T'(y^*)}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau}(y^*)\,y^* f_y(y^*)-\frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\ldots\\
&\times\int_0^\infty\left\{(1-T'(y)-yT''(y))(1-F_y(y))-(1-T'(y))\,y f_y(y)\right\}\gamma(y,y^*)\,dy.
\end{aligned}$$

Now note that

$$(1-T'(y)-yT''(y))(1-F_y(y))-(1-T'(y))\,y f_y(y)=\left[(1-T'(y))\,y\,(1-F_y(y))\right]',$$

so that we finally obtain

$$\begin{aligned}
\mathrm{d}\mathscr{R}(T,h) =& 1-F_y(y^*)-\frac{T'(y^*)}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau}(y^*)\,y^* f_y(y^*)\\
&-\frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\int_{\mathbb{R}_+}\left[(1-T'(y))\,y\,(1-F_y(y))\right]'\gamma(y,y^*)\,dy.
\end{aligned}$$

Note that using (11), this equation can also be expressed as

$$\begin{aligned}
\mathrm{d}\mathscr{R}(T,h) =& 1-F_y(y^*)-\frac{T'(y^*)}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau}(y^*)\,y^* f_y(y^*)\\
&-\frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left\{\psi'(y^*)\,\bar{\bar{\gamma}}(y^*,y^*)+\int_{\mathbb{R}_+}\psi'(y)\,\frac{\bar{\gamma}(y,y^*)}{y'(\theta^*)}dy\right\},
\end{aligned}$$

where we denote $\psi(y)\equiv(1-T'(y))\,y\,(1-F_y(y))$ and $y'(\theta)\equiv\frac{dy(\theta)}{d\theta}$.

Next, we derive the effect of the perturbation $h=h_1+h_2$ on the utility of any individual θ. First, the change in individual consumption due to the perturbation is

$$\begin{aligned}
&\mathrm{d}_h\left[y(\theta)-T(y(\theta))\right]=\mathrm{d}_h\left[y(\theta)\right]-\mathrm{d}_h\left[T(y(\theta))\right]\\
=&-\frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left\{\int_\Theta\gamma(\theta,\theta')\,\delta_{y^*}(y(\theta'))\,d\theta'+\delta_{y^*}(y(\theta))\right\}y(\theta)\,(1-T'(y(\theta)))\\
&-h_1(y(\theta))-h_2(y(\theta))\\
=&-\mathbb{I}_{\{y(\theta)\geq y^*\}}-\frac{1-T'(y(\theta))}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau}(\theta^*)\,y(\theta)\,\delta_{y^*}(y(\theta))-\frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\ldots\\
&\times\left\{(1-T'(y(\theta)))\,y(\theta)\,\gamma(\theta,\theta^*)-\int_0^y(1-T'(y')-y'T''(y'))\,\gamma(y',y^*)\,dy'\right\},
\end{aligned}$$

where the last equality follows from the same steps as those of the derivation of $\mathrm{d}_hT(y(\theta))$ above.

Thus the change in individual utility due to the perturbation is

$$\begin{aligned}
&\mathrm{d}_h \left[u\left(y\left(\theta\right)-T\left(y\left(\theta\right)\right)-v\left(l\left(\theta\right)\right)\right)\right] \\
=&\left(\mathrm{d}_h\left[y\left(\theta\right)\right]-\mathrm{d}_h\left[T\left(y\left(\theta\right)\right)\right]-v'\left(l\left(\theta\right)\right)\mathrm{d}_h l\left(\theta\right)\right)u'\left(y\left(\theta\right)-T\left(y\left(\theta\right)\right)-v\left(l\left(\theta\right)\right)\right) \\
=&\left\{-\mathbb{I}_{\{y(\theta)\geq y^*\}}-\frac{\tilde{\varepsilon}_{l,1-\tau}\left(\theta^*\right)}{1-T'\left(y^*\right)}\left[\left(1-T'\left(y\left(\theta\right)\right)\right)y\left(\theta\right)-l\left(\theta\right)v'\left(l\left(\theta\right)\right)\right]\delta_{y^*}\left(y\left(\theta\right)\right)\right\}u'\left(\theta\right) \\
&-\frac{\tilde{\varepsilon}_{l,1-\tau}\left(\theta^*\right)}{1-T'\left(y^*\right)}\left(\frac{dy\left(\theta^*\right)}{d\theta}\right)^{-1}u'\left(\theta\right)\left\{\left(1-T'\left(y\left(\theta\right)\right)\right)y\left(\theta\right)\gamma\left(\theta,\theta^*\right)\ldots\right. \\
&\left.\qquad-\int_0^y\left(1-T'\left(y'\right)-y'T''\left(y'\right)\right)\gamma\left(y',y^*\right)dy'\right\}.
\end{aligned}$$

Using the individual first-order condition (1), this can be rewritten as

$$\begin{aligned}
&\mathrm{d}_h \left[u\left(y\left(\theta\right)-T\left(y\left(\theta\right)\right)-v\left(l\left(\theta\right)\right)\right)\right] \\
=&-u'\left(\theta\right)\mathbb{I}_{\{y(\theta)\geq y^*\}}-\frac{\tilde{\varepsilon}_{l,1-\tau}\left(\theta^*\right)}{1-T'\left(y^*\right)}\left(\frac{dy\left(\theta^*\right)}{d\theta}\right)^{-1}\times\ldots \\
&\left\{\left(1-T'\left(y\left(\theta\right)\right)\right)y\left(\theta\right)\gamma\left(\theta,\theta^*\right)-\int_0^y\left(1-T'\left(y'\right)-y'T''\left(y'\right)\right)\gamma\left(y',y^*\right)dy'\right\}u'\left(\theta\right),
\end{aligned}$$

which is a manifestation of the envelope theorem. The first term in this expression is the PE term, the second is the GE welfare effect.

Finally we derive the effect of the perturbation $h_1 + h_2$ on the social welfare. Summing the previous expression over all individuals using the density function $\tilde{f}_\theta\left(\theta\right) = \tilde{f}_y\left(y\right)\frac{dy(\theta)}{d\theta}$, and defining the marginal social welfare weights as $g_y(y\left(\theta\right)) = \frac{u'(\theta)\tilde{f}_y(y(\theta))}{\lambda f_y(y(\theta))}$, we obtain that the change in the government objective due to the perturbation $h\left(\cdot\right) = h_1\left(\cdot\right) + h_2\left(\cdot\right)$ is

$$\begin{aligned}
&-\lambda^{-1}\mathrm{d}\mathscr{G}\left(T,h\right)=-\lambda^{-1}\mathrm{d}_h\left[\int_\Theta u\left(y\left(\theta\right)-T\left(y\left(\theta\right)\right)-v\left(l\left(\theta\right)\right)\right)\tilde{f}_\theta\left(\theta\right)d\theta\right] \\
=&\int_\Theta\left\{\mathbb{I}_{\{y(\theta)\geq y^*\}}+\frac{\tilde{\varepsilon}_{l,1-\tau}\left(\theta^*\right)}{1-T'\left(y^*\right)}\left(\frac{dy\left(\theta^*\right)}{d\theta}\right)^{-1}\left[\left(1-T'\left(y\left(\theta\right)\right)\right)y\left(\theta\right)\gamma\left(\theta,\theta^*\right)\ldots\right.\right. \\
&\left.\left.\qquad-\int_{y'=0}^y\left(1-T'\left(y'\right)-y'T''\left(y'\right)\right)\gamma\left(y',y^*\right)dy'\right]\right\}\lambda^{-1}u'\left(\theta\right)\tilde{f}_\theta\left(\theta\right)d\theta.
\end{aligned}$$

Switching the two integrals in the last line of this expression implies that $-\lambda^{-1}\mathrm{d}\mathscr{G}(T,h)$ is equal to

$$
\begin{aligned}
&\int_{\mathbb{R}_+} \mathbb{I}_{\{y\geq y^*\}}\lambda^{-1}u'(y)\,\tilde{f}_y(y)\,dy\\
&+\frac{\tilde{\varepsilon}_{l,1-\tau}(y^*)}{1-T'(y^*)}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\left\{\int_{\mathbb{R}_+}(1-T'(y))\,y\gamma(y,y^*)\,\lambda^{-1}u'(y)\,\tilde{f}_y(y)\,dy\ldots\right.\\
&\left.-\int_{y'=0}^{\infty}\int_{y=y'}^{\infty}(1-T'(y')-y'T''(y'))\,\gamma(y',y^*)\,\lambda^{-1}u'(y)\,\tilde{f}_y(y)\,dydy'\right\}\\
=&\int_{y^*}^{\infty} g_y(y)\,f_y(y)\,dy+\frac{\tilde{\varepsilon}_{l,1-\tau}(y^*)}{1-T'(y^*)}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\ldots\\
&\times\left\{\int_{\mathbb{R}_+}(1-T'(y))\,y\gamma(y,y^*)\,g_y(y)\,f_y(y)\,dy\right.\\
&\left.-\int_{y'=0}^{\infty}(1-T'(y')-y'T''(y'))\,\gamma(y',y^*)\left[\int_{y=y'}^{\infty}g_y(y)\,f_y(y)\,dy\right]dy'\right\},
\end{aligned}
$$

But note that

$$
\begin{aligned}
&(1-T'(y))\,yg_y(y)\,f_y(y)-(1-T'(y)-yT''(y))\left(\int_y^{\infty}g_y(y')\,f_y(y')\,dy'\right)\\
=&-\left[(1-T'(y))\,y\left(\int_y^{\infty}g_y(y')\,f_y(y')\,dy'\right)\right]'.
\end{aligned}
$$

Thus, we obtain

$$
\begin{aligned}
-\lambda^{-1}\mathrm{d}\mathscr{G}(T,h)=&\int_{y^*}^{\infty}g_y(y)\,f_y(y)\,dy-\frac{\tilde{\varepsilon}_{l,1-\tau}(y^*)}{1-T'(y^*)}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\ldots\\
&\times\int_{\mathbb{R}_+}\left[(1-T'(y))\,y\left(\int_y^{\infty}g_y(y')\,f_y(y')\,dy'\right)\right]'\gamma(y,y^*)\,dy.
\end{aligned}
$$

Therefore, the normalized effect of the perturbation on social welfare is finally given by:

$$
\begin{aligned}
&\mathrm{d}\mathscr{W}(T,h)=\mathrm{d}\mathscr{R}(T,h)+\lambda^{-1}\mathrm{d}\mathscr{G}(T,h)\\
=&1-F_y(y^*)-\int_{y^*}^{\infty}g_y(y)\,f_y(y)\,dy-\frac{T'(y^*)}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau}(y^*)\,y^*f_y(y^*)\\
&-\frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\int_{\mathbb{R}_+}[(1-T'(y))\,y\,(1-F_y(y))]'\,\gamma(y,y^*)\,dy\\
&+\frac{\tilde{\varepsilon}_{l,1-\tau}(y^*)}{1-T'(y^*)}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\int_{\mathbb{R}_+}\left[(1-T'(y))\,y\left(\int_y^{\infty}g_y(y')\,f_y(y')\,dy'\right)\right]'\gamma(y,y^*)\,dy.
\end{aligned}
$$

But note that

$$
\begin{aligned}
&\left[\left(1-T'(y)\right)y\left(1-F_y(y)\right)\right]' - \left[\left(1-T'(y)\right)y\left(\int_y^\infty g_y(y')\,f_y(y')\,dy'\right)\right]' \\
&= \left[\left(1-T'(y)\right)y\left(1-F_y(y)\right)\left(1-\int_y^\infty g_y(y')\,\frac{f_y(y')}{1-F_y(y)}dy'\right)\right]',
\end{aligned}
$$

so that, using the definition of $\bar{g}_y(y)$, we get

$$
\begin{aligned}
\mathrm{d}\mathscr{W}(T,h) &= 1-F_y(y^*) - \int_{y^*}^\infty g_y(y)\,f_y(y)\,dy - \frac{T'(y^*)}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau}(y^*)\,y^* f_y(y^*) \times \\
&- \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{1-T'(y(\theta^*))}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\int_{\mathbb{R}_+}\left[\left(1-T'(y)\right)y\left(1-F_y(y)\right)\left(1-\bar{g}_y(y)\right)\right]'\gamma(y,y^*)\,dy.
\end{aligned}
$$

Now, at the optimum we must have $\frac{\mathrm{d}\mathscr{W}(T,h)}{1-F(y^*)}=0$, therefore

$$
\begin{aligned}
0 =& 1-\bar{g}_y(y^*) - \frac{T'(y^*)}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau}(y^*)\,\frac{y^* f_y(y^*)}{1-F_y(y^*)} \\
&- \tilde{\varepsilon}_{l,1-\tau}(\theta^*)\int_{\mathbb{R}_+}\frac{\frac{d}{dy}\left[\left(1-\bar{g}_y(y)\right)\left(1-T'(y)\right)y\left(1-F_y(y)\right)\right]}{\left(1-T'(y(\theta^*))\right)\left(1-F_y(y^*)\right)y'(\theta^*)}\gamma(y,y^*)\,dy,
\end{aligned}
$$

which leads to the optimal tax formula (49).

□

We now prove that the optimal tax formula obtained by the variational approach coincides with the formula obtained by solving the mechanism design problem.

B.4.4 Proof of the equivalence of (42) and (49)

Proof. Substitute for $\mu(\theta)=\lambda\int_\theta^{\overline{\theta}}(1-g(x))dF_\theta(x)$ in the optimal tax formula (42) evaluated at θ^* to get:

$$
\begin{aligned}
&\frac{T'(y(\theta^*))}{1-T'(y(\theta^*))} \equiv \frac{\tau(\theta^*)}{1-\tau(\theta^*)} \\
&= \left(1+\frac{1}{\varepsilon_{l,1-\tau}(\theta^*)}\right)\frac{\int_{\theta^*}^{\overline{\theta}}(1-g_\theta(x))f_\theta(x)dx}{f_w(w(\theta^*))w(\theta^*)} - \frac{\int_\Theta\left[v'(l(x))l(x)\left(\int_x^{\overline{\theta}}(1-g_\theta(x'))f_\theta(x')dx'\right)\right]'\gamma(x,\theta^*)dx}{(1-\tau(\theta^*))\,y(\theta^*)f_\theta(\theta^*)} \\
&= \left(1+\frac{1}{\varepsilon_{l,1-\tau}(\theta^*)}\right)\frac{1-F_\theta(\theta^*)}{f_w(w(\theta^*))w(\theta^*)}\left(\int_{\theta^*}^{\overline{\theta}}(1-g_\theta(x))\frac{f_\theta(x)}{1-F_\theta(\theta^*)}dx\right) \\
&\quad - \frac{\int_\Theta\left[\left(1-T'(y(x))\right)y(x)\left(1-F_\theta(x)\right)\left(\int_x^{\overline{\theta}}(1-g_\theta(x'))\frac{f_\theta(x')}{1-F_\theta(x)}dx'\right)\right]'\gamma(x,\theta^*)dx}{\left(1-T'(y(\theta^*))\right)y(\theta^*)f_\theta(\theta^*)},
\end{aligned}
$$

where the last equality uses individual x's first order condition (1). Using the definition of the average marginal welfare weight $\bar{g}_\theta(\theta)=\int_\theta^{\overline{\theta}}g_\theta(x))\frac{f_\theta(x)}{1-F_\theta(\theta)}dx$, and multiplying and dividing the first term on the right hand side by $w'(\theta^*)/w(\theta^*)$, we can rewrite this expression as

$$\frac{T'\left(y\left(\theta^*\right)\right)}{1-T'\left(y\left(\theta^*\right)\right)}=\left(1+\frac{1}{\varepsilon_{l,1-\tau}\left(\theta^*\right)}\right)\frac{w'(\theta^*)}{w(\theta^*)}\left(1-\bar{g}_\theta\left(\theta^*\right)\right)\frac{1-F_\theta(\theta^*)}{f_w(w(\theta^*))w'(\theta^*)}$$
$$-\frac{\int_\Theta\left[\left(1-T'\left(y\left(\theta\right)\right)\right)y(\theta)\left(1-F_\theta(\theta)\right)\left(1-\bar{g}_\theta\left(\theta\right)\right)\right]'\gamma(\theta,\theta^*)d\theta}{\left(1-T'\left(y\left(\theta^*\right)\right)\right)y(\theta^*)f_\theta(\theta^*)}.$$

We now change variables from types and wages to incomes in each of the terms of this equation. First, recall that $F_\theta(\theta^*)=F_w\left(w\left(\theta^*\right)\right)=F_y\left(y\left(\theta^*\right)\right)$, and

$$f_\theta\left(\theta^*\right)=f_y\left(y\left(\theta^*\right)\right)\times\left.\frac{dy\left(\theta\right)}{d\theta}\right|_{\theta=\theta^*}.$$

Second, we can rewrite the integral as

$$\begin{aligned}
&\int_\Theta\frac{d}{d\theta}\left[\left(1-T'\left(y\left(\theta\right)\right)\right)y(\theta)\left(1-F_\theta(\theta)\right)\left(1-\bar{g}_\theta\left(\theta\right)\right)\right]\times\gamma\left(\theta,\theta^*\right)d\theta\\
=&\int_\Theta\Big[\left(1-T'\left(y\left(\theta\right)\right)-y\left(\theta\right)T''\left(y\left(\theta\right)\right)\right)\left(1-F_\theta(\theta)\right)\left(1-\bar{g}_\theta\left(\theta\right)\right)\frac{dy\left(\theta\right)}{d\theta}\\
&\quad-\left(1-T'\left(y\left(\theta\right)\right)\right)y\left(\theta\right)\left(1-g_\theta\left(\theta\right)\right)f_\theta\left(\theta\right)\Big]\gamma\left(\theta,\theta^*\right)d\theta\\
=&\int_\Theta\Big[\left(1-T'\left(y\left(\theta\right)\right)-y\left(\theta\right)T''\left(y\left(\theta\right)\right)\right)\left(1-F_\theta(\theta)\right)\left(1-\bar{g}_y\left(y\left(\theta\right)\right)\right)\\
&\quad-\left(1-T'\left(y\left(\theta\right)\right)\right)y\left(\theta\right)\left(1-g_y\left(y\left(\theta\right)\right)\right)f_y\left(y\left(\theta\right)\right)\Big]\frac{dy\left(\theta\right)}{d\theta}\gamma\left(y\left(\theta\right),y\left(\theta^*\right)\right)d\theta\\
=&\int_{\mathbb{R}_+}\Big[\left(1-T'\left(y\right)-yT''\left(y\right)\right)\left(1-F_y(y)\right)\left(1-\bar{g}_y\left(y\right)\right)\\
&\quad-\left(1-T'\left(y\right)\right)y\left(1-g_y\left(y\right)\right)f_y\left(y\right)\Big]\gamma\left(y,y^*\right)dy\\
=&\int_{\mathbb{R}_+}\frac{d}{dy}\left[\left(1-T'\left(y\right)\right)y\left(1-F_y(y)\right)\left(1-\bar{g}_y\left(y\right)\right)\right]\times\gamma\left(y,y^*\right)dy,
\end{aligned}$$

where the second equality uses

$$g_y\left(y\left(\theta\right)\right)=\frac{1}{\lambda}u'\left(\theta\right)\frac{\tilde{f}_y\left(y\left(\theta\right)\right)}{f_y\left(y\left(\theta\right)\right)}=\frac{1}{\lambda}u'\left(\theta\right)\frac{\left(y'\left(\theta\right)\right)^{-1}\times\tilde{f}_\theta\left(\theta\right)}{\left(y'\left(\theta\right)\right)^{-1}\times f_\theta\left(\theta\right)}=g_\theta\left(\theta\right),$$

and $\gamma\left(\theta,\theta^*\right)=\gamma\left(y\left(\theta\right),y\left(\theta^*\right)\right)$, and the third equality follows from a change of variables in the integral. Third, we have

$$\frac{w'(\theta^*)}{w(\theta^*)}\frac{1-F_\theta(\theta^*)}{f_w(w(\theta^*))w'(\theta^*)}=\frac{w'(\theta^*)}{w(\theta^*)}\frac{1-F_\theta(\theta^*)}{f_\theta(\theta^*)}=\frac{\frac{w'(\theta^*)}{w(\theta^*)}}{\frac{y'(\theta^*)}{y(\theta^*)}}\frac{1-F_y\left(y(\theta^*)\right)}{y(\theta^*)f_y\left(y(\theta^*)\right)}.$$

To compute $\frac{y'(\theta)}{y(\theta)}/\frac{w'(\theta)}{w(\theta)}$, note that the first order condition $l(\theta) = w(\theta)^{\varepsilon}(1-T'(y(\theta)))^{\varepsilon}$ implies

$$\frac{l'(\theta)}{l(\theta)} = \varepsilon\left(\frac{w'(\theta)}{w(\theta)} - \frac{y'(\theta)T''(y(\theta))}{1-T'(y(\theta))}\right).$$

Using (8), we can write

$$\tilde{\varepsilon}_{l,w}(\theta) = \frac{l(\theta)^{\frac{1}{\varepsilon}} - w(\theta)y(\theta)T''\left(y(\theta)\right)}{l(\theta)^{\frac{1}{\varepsilon}} + \varepsilon w(\theta)y(\theta)T''\left(y(\theta)\right)}\varepsilon.$$

These two equations imply that $\frac{l'(\theta)}{l(\theta)}/\frac{w'(\theta)}{w(\theta)}$ is equal to

$$\begin{aligned}\left(1-\frac{w(\theta)}{w'(\theta)}\frac{y'(\theta)T''(y(\theta))}{1-T'(y(\theta))}\right)\varepsilon = &\frac{\varepsilon}{l(\theta)^{\frac{1}{\varepsilon}} + \varepsilon w(\theta)y(\theta)T''\left(y(\theta)\right)}\left\{\left[l(\theta)^{\frac{1}{\varepsilon}} - w(\theta)y(\theta)T''\left(y(\theta)\right)\right]\dots\right.\\ &\left.+\left[\left(1+\varepsilon\right)w(\theta)y(\theta)T''\left(y(\theta)\right) - \frac{w(\theta)}{w'(\theta)}\frac{y'(\theta)T''(y(\theta))}{1-T'(y(\theta))}\left(l(\theta)^{\frac{1}{\varepsilon}} + \varepsilon w(\theta)y(\theta)T''\left(y(\theta)\right)\right)\right]\right\}\\ =&\tilde{\varepsilon}_{l,w}(\theta)\left\{1+\frac{\left(1+\varepsilon\right)w(\theta)y(\theta)T''\left(y(\theta)\right) - \frac{w(\theta)}{w'(\theta)}\frac{y'(\theta)T''(y(\theta))}{1-T'(y(\theta))}\left(l(\theta)^{\frac{1}{\varepsilon}} + \varepsilon w(\theta)y(\theta)T''\left(y(\theta)\right)\right)}{l(\theta)^{\frac{1}{\varepsilon}} - w(\theta)y(\theta)T''\left(y(\theta)\right)}\right\}.\end{aligned}$$

But the second term in the curly brackets is equal to zero; to see this, note that its numerator is proportional to

$$\begin{aligned}&\left(1+\varepsilon\right)y(\theta) - \frac{y'(\theta)}{\left(1-T'(y(\theta))\right)w'(\theta)}\left(l(\theta)^{\frac{1}{\varepsilon}} + \varepsilon w(\theta)y(\theta)T''\left(y(\theta)\right)\right)\\ &\propto\left(1+\varepsilon\right)w'(\theta)l(\theta) - \left(1+\varepsilon y(\theta)\frac{T''\left(y(\theta)\right)}{1-T'(y(\theta))}\right)y'(\theta) = 0,\end{aligned}$$

where the last equality follows from the fact that $y'(\theta) = w'(\theta)l(\theta) + w(\theta)l'(\theta)$ and the expression above for $\frac{l'(\theta)}{l(\theta)}$. Therefore, we finally obtain

$$\frac{\frac{y'(\theta)}{y(\theta)}}{\frac{w'(\theta)}{w(\theta)}} = 1 + \frac{\frac{l'(\theta)}{l(\theta)}}{\frac{w'(\theta)}{w(\theta)}} = 1 + \left(\frac{d\ln w(\theta)}{d\theta}\right)^{-1}\frac{d\ln l(\theta)}{d\theta} = 1 + \tilde{\varepsilon}_{l,w}(\theta).$$

This implies

$$\frac{1-F_w\left(w(\theta^*)\right)}{w(\theta^*)f_w\left(w(\theta^*)\right)} = \frac{1}{1+\tilde{\varepsilon}_{l,w}(\theta^*)}\frac{1-F_y\left(y(\theta^*)\right)}{y(\theta^*)f_y\left(y(\theta^*)\right)}.$$

Finally, note that

$$\frac{1}{1+\tilde{\varepsilon}_{l,w}\left(\theta\right)} = \frac{\frac{1-T'(\theta)}{1-T'(\theta)+\varepsilon_{l,1-\tau}(\theta)y(\theta)T''(\theta)}\varepsilon_{l,1-\tau}\left(\theta\right)}{1+\frac{1-T'(\theta)-y(\theta)T''(\theta)}{1-T'(\theta)+\varepsilon_{l,1-\tau}(\theta)y(\theta)T''(\theta)}\varepsilon_{l,1-\tau}\left(\theta\right)}\frac{1}{\tilde{\varepsilon}_{l,1-\tau}\left(\theta\right)} = \frac{\varepsilon_{l,1-\tau}\left(\theta\right)}{1+\varepsilon_{l,1-\tau}\left(\theta\right)}\frac{1}{\tilde{\varepsilon}_{l,1-\tau}\left(\theta\right)}.$$

Collecting all the terms, we obtain

$$
\begin{aligned}
\frac{T'(y(\theta^*))}{1-T'(y(\theta^*))} = & \left(1+\frac{1}{\varepsilon_{l,1-\tau}(\theta^*)}\right)(1-\bar{g}_\theta(\theta^*))\left(\frac{1}{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}\frac{\varepsilon_{l,1-\tau}(\theta^*)}{1+\varepsilon_{l,1-\tau}(\theta^*)}\right)\frac{1-F_y(y(\theta^*))}{y(\theta^*)f_y(y(\theta^*))} \\
& + \frac{\int_{\mathbb{R}_+}\frac{d}{dy}\left[(1-T'(y))\,y\,(1-F_y(y))\,(1-\bar{g}_y(y))\right]\times\gamma(y,y^*)\,dy}{(1-T'(y(\theta^*)))\,y(\theta^*)\,f_y(y(\theta^*))\times y'(\theta^*)} \\
= & \frac{1}{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}(1-\bar{g}_\theta(\theta^*))\frac{1-F_y(y(\theta^*))}{y(\theta^*)f_y(y(\theta^*))} \\
& - \frac{\int_{\mathbb{R}_+}\frac{d}{dy}\left[(1-T'(y))\,y\,(1-F_y(y))\,(1-\bar{g}_y(y))\right]\times\gamma(y,y^*)\,dy}{(1-T'(y(\theta^*)))\,y(\theta^*)\,f_y(y(\theta^*))\times y'(\theta^*)},
\end{aligned}
$$

which is exactly formula (49).

□

B.4.5 Optimum Characterization and Proof of Corollary 4

Here we show how to rewrite formula (49) as an integral equation.

Proof. The idea is to integrate by parts the last term in equation (49). Let

$$
\psi(y) = \frac{(1-\bar{g}_y(y))\,(1-T'(y))\,y\,(1-F_y(y))}{(1-T'(y^*))\,(1-F(y^*))},
$$

so that optimal taxes satisfy

$$
0 = 1-\bar{g}_y(y^*) - \frac{T'(y^*)}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau}(y^*)\frac{y^* f_y(y^*)}{1-F_y(y^*)} - \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{y'(\theta^*)}\int_{\mathbb{R}_+}\psi'(y)\,\gamma(y,y^*)\,dy.
$$

Disentangle the Dirac and the smooth part of $\gamma(y,y^*)$, i.e., the own- and cross-wage elasticities. We get:

$$
\begin{aligned}
0 = & 1-\bar{g}_y(y^*) - \frac{T'(y^*)}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau}(y^*)\frac{y^* f_y(y^*)}{1-F_y(y^*)} \\
& - \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{y'(\theta^*)}\int_{\mathbb{R}_+}\psi'(y)\left[\bar{\gamma}(y,y^*)+\bar{\bar{\gamma}}(y^*,y^*)\,y'(\theta^*)\,\delta_{y^*}(y)\right]dy \\
= & 1-\bar{g}_y(y^*) - \frac{T'(y^*)}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau}(y^*)\frac{y^* f_y(y^*)}{1-F_y(y^*)} \\
& - \tilde{\varepsilon}_{l,1-\tau}(\theta^*)\,\psi'(y^*)\,\bar{\bar{\gamma}}(y^*,y^*) - \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta^*)}{y'(\theta^*)}\int_{\mathbb{R}_+}\psi'(y)\,\bar{\gamma}(y,y^*)\,dy.
\end{aligned}
$$

Assuming that $y \mapsto \bar{\gamma}(y,y^*)$ is continuously differentiable for each y^*, we can integrate by parts the last term of this equation. Using $\psi(0)=0$ and $\psi(\bar{y})=0$ yields

$$
\begin{aligned}
0 = & 1-\bar{g}_y(y^*) - \frac{T'(y^*)}{1-T'(y^*)}\tilde{\varepsilon}_{l,1-\tau}(y^*)\frac{y^* f_y(y^*)}{1-F_y(y^*)} - \tilde{\varepsilon}_{l,1-\tau}(\theta^*)\,\psi'(y^*)\,\bar{\bar{\gamma}}(y^*,y^*) \\
& + \tilde{\varepsilon}_{l,1-\tau}(\theta^*)\int_{\mathbb{R}_+}\psi(y)\left[\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\frac{d\bar{\gamma}(y,y^*)}{dy}\right]dy.
\end{aligned}
\tag{85}
$$

Now, since the marginal social welfare weights $g_y(y)$ sum to 1, we have

$$
\begin{aligned}
(1-\bar{g}_y(y))(1-F_y(y)) =& 1-F_y(y)-\int_y^\infty g_y(y')\,f_y(y')\,dy' \\
=& \int_0^y g_y(y')\,f_y(y')\,dy' - F_y(y) \equiv G_y(y)-F_y(y),
\end{aligned}
$$

so that $[(1-\bar{g}_y(y))(1-F_y(y))]' = (g_y(y)-1)\,f_y(y)$ and thus

$$
\psi'(y) = \frac{(1-T'(y))\,y\,(g_y(y)-1)\,f_y(y) + (1-T'(y)-yT''(y))\,(1-\bar{g}_y(y))\,(1-F_y(y))}{(1-T'(y^*))\,(1-F(y^*))}.
$$

Hence (85) can be rewritten as

$$
\begin{aligned}
\frac{T'(y^*)}{1-T'(y^*)} =& \frac{1}{\tilde{\varepsilon}_{l,1-\tau}(y^*)}\frac{1-F_y(y^*)}{y^* f_y(y^*)}(1-\bar{g}_y(y^*))\ldots \\
&-\frac{1-F_y(y^*)}{y^* f_y(y^*)}\frac{(1-T'(y^*)-y^*T''(y^*))\,(1-\bar{g}_y(y^*))\,(1-F_y(y^*))}{(1-T'(y^*))\,(1-F(y^*))}\bar{\bar{\gamma}}(y^*,y^*) \\
&-\frac{1-F_y(y^*)}{y^* f_y(y^*)}\frac{(1-T'(y^*))\,y^*\,(g_y(y^*)-1)\,f_y(y^*)}{(1-T'(y^*))\,(1-F(y^*))}\bar{\bar{\gamma}}(y^*,y^*) \\
&+\frac{1-F_y(y^*)}{y^* f_y(y^*)}\int_{\mathbb{R}_+}\frac{(1-\bar{g}_y(y))\,(1-T'(y))\,y\,(1-F_y(y))}{(1-T'(y^*))\,(1-F(y^*))}\left[\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\frac{d\bar{\gamma}(y,y^*)}{dy}\right]dy.
\end{aligned}
$$

The right hand side of this expression is equal to

$$
\begin{aligned}
&\frac{1-F_y(y^*)}{y^* f_y(y^*)}(1-\bar{g}_y(y^*))\frac{1}{\tilde{\varepsilon}_{l,1-\tau}(y^*)}\left[1-\frac{(1-T'(y^*)-y^*T''(y^*))\,\tilde{\varepsilon}_{l,1-\tau}(y^*)}{1-T'(y^*)}\bar{\bar{\gamma}}(y^*,y^*)\right] \\
&-(g_y(y^*)-1)\,\bar{\bar{\gamma}}(y^*,y^*)+\int_{\mathbb{R}_+}\frac{(1-\bar{g}_y(y))\,(1-T'(y))\,y\,(1-F_y(y))}{(1-T'(y^*))\,y^* f_y(y^*)}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\frac{d\bar{\gamma}(y,y^*)}{dy}dy \\
=&\frac{1-F_y(y^*)}{y^* f_y(y^*)}(1-\bar{g}_y(y^*))\frac{1-\tilde{\varepsilon}_{l,w}(y^*)\,\bar{\bar{\gamma}}(y^*,y^*)}{\tilde{\varepsilon}_{l,1-\tau}(y^*)}+(1-g_y(y^*))\,\bar{\bar{\gamma}}(y^*,y^*) \\
&+\int_{\mathbb{R}_+}(1-\bar{g}_y(y))\frac{(1-T'(y))\,y}{(1-T'(y^*))\,y^*}\frac{1-F_y(y)}{f_y(y^*)}\left(\frac{dy(\theta^*)}{d\theta}\right)^{-1}\frac{d\bar{\gamma}(y,y^*)}{dy}dy.
\end{aligned}
$$

We thus get

$$
\begin{aligned}
\frac{T'(y^*)}{1-T'(y^*)} =& \frac{1}{\tilde{E}_{l,1-\tau}(y^*)}(1-\bar{g}_y(y^*))\left(\frac{1-F_y(y^*)}{y^* f_y(y^*)}\right)-(g(y^*)-1)\,\bar{\bar{\gamma}}(y^*,y^*) \\
&+\int_{\mathbb{R}_+}(1-\bar{g}_y(y))\left(\frac{1-F_y(y)}{y^* f_y(y^*)}\right)\left(\frac{1-T'(y)}{1-T'(y^*)}\right)y\frac{\bar{\gamma}'_y(y,y^*)}{y'(\theta^*)}dy.
\end{aligned}
$$

Denoting by $\frac{\bar{T}'(y^*)}{1-\bar{T}'(y^*)}$ the first line in the right hand side of the previous equation, and changing variables from incomes y to types θ (using the identities $f_y(y^*)\frac{dy(\theta^*)}{d\theta} = f_\theta(\theta^*)$, $F_y(y(\theta)) = F_\theta(\theta)$, and $\frac{d\bar{\gamma}(y,y^*)}{dy} = \frac{d\bar{\gamma}(\theta,\theta^*)}{d\theta}\left(\frac{dy(\theta)}{d\theta}\right)^{-1}$), we obtain

$$\frac{T'(y(\theta^*))}{1-T'(y(\theta^*))} = \frac{\bar{T}'(y(\theta^*))}{1-\bar{T}'(y(\theta^*))} + \int_\Theta (1-\bar{g}_\theta(\theta)) \frac{(1-T'(y(\theta)))\,y(\theta)}{(1-T'(y(\theta^*)))\,y(\theta^*)} \frac{1-F_\theta(\theta)}{f_\theta(\theta^*)} \frac{d\bar{\gamma}(\theta,\theta^*)}{d\theta} d\theta.$$

Equation (52) follows immediately from this expression since $\frac{d\bar{\gamma}(y,y^*)}{dy} = 0$ when the production function is CES.

Multiplying both sides of the previous equation by $1-T'(y(\theta^*))$, we then get

$$\begin{aligned} T'(y(\theta^*)) = & \frac{\bar{T}'(y(\theta^*))}{1-\bar{T}'(y(\theta^*))}(1-T'(y(\theta^*))) \\ & + \int_\Theta (1-T'(y(\theta)))\,y(\theta) \frac{(1-\bar{g}_\theta(\theta))(1-F_\theta(\theta))}{y(\theta^*)\,f_\theta(\theta^*)} \frac{d\bar{\gamma}(\theta,\theta^*)}{d\theta} d\theta, \end{aligned}$$

which leads to the following formula for the optimal net-of-tax rate $1-\tau(\theta^*) \equiv 1-T'(y(\theta^*))$:

$$\begin{aligned} (1-\tau(\theta^*)) = & \frac{1}{1+\frac{\bar{\tau}(\theta^*)}{1-\bar{\tau}(\theta^*)}} \left[1 - \int_\Theta (1-\tau(\theta)) \frac{y(\theta)(1-\bar{g}_\theta(\theta))(1-F_\theta(\theta))\,\bar{\gamma}'(\theta,\theta^*)}{y(\theta^*)\,f_\theta(\theta^*)} d\theta\right] \\ = & (1-\bar{\tau}(\theta^*)) \left\{1 - \int_\Theta \left[(1-\bar{g}_\theta(\theta)) \left(\frac{1-F_\theta(\theta)}{y(\theta^*)\,f_\theta(\theta^*)}\right) y(\theta)\,\bar{\gamma}'(\theta,\theta^*)\right] (1-\tau(\theta))\,d\theta\right\}. \end{aligned}$$

This is now a well-defined integral equation in $(1-T'(y(\theta)))$, so that we can use the mathematical apparatus introduced in Section 2 to characterize its solution, i.e., the optimal tax schedule. Similar to the integral equation (21), it can be solved in closed form if $\bar{\gamma}'(\theta,\theta^*)$ is the sum of multiplicatively separable terms. We can use the same techniques as in Section xx in the Translog case to get a separable kernel.

□

B.4.6 Proof of Corollary 5

We now derive the formula for the optimal top tax rate when the production function is CES.

Proof. Assume that in the data (i.e., given the current tax schedule with a constant top tax rate, assuming that the aggregate production function is CES), the income distribution has a Pareto tail, so that the (observed) hazard rate $\frac{1-F_y^d(y^*)}{y f_y^d(y^*)}$ converges to a constant $1/\alpha$. We show that under these assumption, the income distribution at the optimum tax schedule is also Pareto distributed at the tail with the same Pareto coefficient $1/\alpha$. That is, the hazard rate of the income distribution at the top is independent of the level of the top tax rate. At the optimum, we have

$$\frac{1-F_y(y(\theta))}{y(\theta)f_y(\theta)} = \frac{1-F_\theta(\theta)}{\frac{y(\theta)}{y'(\theta)}f_\theta(\theta)} = \frac{1-F_\theta(\theta)}{\theta f_\theta(\theta)}\frac{\theta y'(\theta)}{y(\theta)} = \frac{1-F_\theta(\theta)}{\theta f_\theta(\theta)}\varepsilon_{y,\theta}, \tag{86}$$

where we define the income elasticity $\varepsilon_{y,\theta} \equiv d\ln y(\theta)/d\ln\theta$. To compute this elasticity, use the individual first order condition (1) with isoelastic disutility of labor to get $l(\theta) = (1-\tau(\theta))^{\varepsilon}\,w(\theta)^{\varepsilon}$.

Thus we have

$$\varepsilon_{l,\theta} \equiv \frac{d\ln l\left(\theta\right)}{d\ln\theta} = \varepsilon\frac{d\ln\left(1-\tau(\theta)\right)}{d\ln\theta} + \varepsilon\frac{d\ln w(\theta)}{d\ln\theta} = \varepsilon\left(\frac{\theta w'(\theta)}{w(\theta)} - \frac{\theta\tau'(\theta)}{1-\tau(\theta)}\right).$$

But since the production function is CES, we have, from equation (14),

$$\begin{aligned}\frac{d\ln w(\theta)}{d\ln\theta} =& \frac{d\ln a\left(\theta\right)}{d\ln\theta} + \left(\rho-1\right)\frac{d\ln L\left(\theta\right)}{d\ln\theta}\\ =& \frac{d\ln a\left(\theta\right)}{d\ln\theta} - \frac{1}{\sigma}\frac{d\ln l\left(\theta\right)}{d\ln\theta} - \frac{1}{\sigma}\frac{d\ln f_\theta\left(\theta\right)}{d\ln\theta} = \frac{\theta a'\left(\theta\right)}{a\left(\theta\right)} - \frac{1}{\sigma}\varepsilon_{l,\theta} - \frac{1}{\sigma}\frac{\theta f'_\theta\left(\theta\right)}{f_\theta\left(\theta\right)}.\end{aligned}$$

Thus, substituting this expression for $\frac{\theta w'(\theta)}{w(\theta)}$ in the previous equation, we obtain

$$\varepsilon_{l,\theta} = \varepsilon\left(\frac{\theta a'\left(\theta\right)}{a\left(\theta\right)} - \frac{1}{\sigma}\varepsilon_{l,\theta} - \frac{1}{\sigma}\frac{\theta f'_\theta\left(\theta\right)}{f_\theta\left(\theta\right)} - \frac{\theta\tau'(\theta)}{1-\tau(\theta)}\right).$$

Moreover, since we assume that the second derivative of the optimal marginal tax rate, $T''\left(y\right)$, converges to zero for high incomes, we have $\lim_{\theta\to\infty}\tau'(\theta)$. Therefore, the previous equation yields

$$\lim_{\theta\to\infty}\varepsilon_{l,\theta} = \frac{\varepsilon}{1+\frac{\varepsilon}{\sigma}}\left(\lim_{\theta\to\infty}\frac{\theta a'\left(\theta\right)}{a\left(\theta\right)} - \frac{1}{\sigma}\lim_{\theta\to\infty}\frac{\theta f'_\theta\left(\theta\right)}{f_\theta\left(\theta\right)}\right).$$

Note that the variables $\frac{\theta a'(\theta)}{a(\theta)}$ and $\frac{\theta f'_\theta(\theta)}{f_\theta(\theta)}$ are primitive parameters that do not depend on the tax rate. Assuming that they converge to constants as $\theta\to\infty$, we obtain that $\lim_{\theta\to\infty}\varepsilon_{l,\theta}$ is a constant independent of the tax rates, and hence

$$\varepsilon_{y,\theta} = \varepsilon_{l,\theta} + \varepsilon_{w,\theta} = \left(1+\frac{1}{\varepsilon}\right)\varepsilon_{l,\theta}$$

converges to a constant as $\theta\to\infty$. Therefore, the hazard rate of the income distribution at the optimum tax schedule, given by (86), converges to the same constant $1/\alpha$ as the hazard rate of incomes observed in the data.

Now let $y^*\to\infty$ in equation (52), to obtain an expression for the optimal top tax rate $\tau_{top} = \lim_{y^*\to\infty}T'\left(y^*\right)$. Since the production function is CES with parameter σ, and the disutility of labor is isoelastic with parameter ε, we have

$$\lim_{y^*\to\infty}\tilde{E}_{l,1-\tau}\left(y^*\right) = \frac{\varepsilon}{1+\varepsilon/\sigma}.$$

Furthermore assume that $\lim_{y^*\to\infty}g_y\left(y^*\right)=\bar{g}$, so that $\lim_{y^*\to\infty}\bar{g}_y\left(y^*\right)=\bar{g}$. Therefore (52) implies

$$\frac{\tau_{top}}{1-\tau_{top}} = \frac{1+\varepsilon/\sigma}{\varepsilon}\left(1-\bar{g}\right)\frac{1}{\alpha} + \frac{\bar{g}-1}{\sigma} = \frac{1-\bar{g}}{\alpha\varepsilon} + \frac{1-\bar{g}}{\alpha\sigma} + \frac{\bar{g}-1}{\sigma}.$$

This concludes the proof.

□

B.4.7 General equilibrium wedge accounting

We finally propose a decomposition into several components of the difference between the partial and general equilibrium optimal taxes, respectively given by formulas (54) and (52).

Proposition 6. *The optimal marginal tax rate of type θ in general equilibrium can be expressed as a function of $\tau_{PE}(\theta)$ and three additional terms:*

$$\begin{aligned}\frac{\tau(\theta)}{1-\tau(\theta)} =& \frac{\tau_{PE}(\theta)}{1-\tau_{PE}(\theta)} + \frac{g_\theta(\theta)-1}{\sigma} \\ &+ \frac{1-F_w(w(\theta))}{f_w(w(\theta))w(\theta)}\left(1+\frac{1}{\varepsilon}\right)(1-\bar{g}_\theta(\theta))\left(\frac{\tilde{\varepsilon}_{l,1-\tau}(\theta)}{\tilde{E}_{l,1-\tau}(\theta)}-1\right) \\ &+ \left(1+\frac{1}{\varepsilon}\right)(1-\bar{g}_\theta(\theta))\left(\frac{1-F_w(w(\theta))}{f_w(w(\theta))w(\theta)} - \frac{1-F_w^d(w_d(\theta))}{f_w^d(w_d(\theta))w_d(\theta)}\right). \end{aligned} \tag{87}$$

Proof. Adding and substracting from equation (52) the partial equilibrium tax $\frac{\tau_{PE}(\theta)}{1-\tau_{PE}(\theta)}$ constructed in (54), we find

$$\begin{aligned}\frac{\tau(\theta)}{1-\tau(\theta)} =& \frac{\tau_{PE}(\theta)}{1-\tau_{PE}(\theta)} + \frac{g_\theta(\theta)-1}{\sigma} \\ &+ \frac{1}{\tilde{E}_{l,1-\tau}(\theta^*)}(1-\bar{g}(\theta^*))\left(\frac{1-F_y(y(\theta^*))}{y(\theta^*))f_y(y(\theta^*))}\right) - \frac{\tau_{PE}(\theta)}{1-\tau_{PE}(\theta)}.\end{aligned}$$

Substituting for $\frac{\tau_{PE}(\theta)}{1-\tau_{PE}(\theta)}$ in the second line of this expression using the definition (54), we obtain

$$\begin{aligned}&\frac{\tau(\theta)}{1-\tau(\theta)} - \left[\frac{\tau_{PE}(\theta)}{1-\tau_{PE}(\theta)} + \frac{g_\theta(\theta)-1}{\sigma}\right] \\ =& \frac{1}{\tilde{E}_{l,1-\tau}(\theta^*)}(1-\bar{g}(\theta^*))\left(\frac{1-F_y(y(\theta^*))}{y(\theta^*))f_y(y(\theta^*))}\right) - \left(1+\frac{1}{\varepsilon}\right)\frac{1-F_w^d(w_d(\theta))}{f_w^d(w_d(\theta))w_d(\theta)}(1-\bar{g}(\theta)) \\ =& \frac{1}{\tilde{E}_{l,1-\tau}(\theta^*)}(1-\bar{g}(\theta^*))(1+\tilde{\varepsilon}_{l,w}(\theta^*))\left(\frac{1-F_w(w(\theta^*))}{w(\theta^*))f_w(w(\theta^*))}\right) - \left(1+\frac{1}{\varepsilon}\right)\frac{1-F_w^d(w_d(\theta))}{f_w^d(w_d(\theta))w_d(\theta)}(1-\bar{g}(\theta)),\end{aligned}$$

where the second equality follows from a change of variables from incomes to wages in the hazard rate, and the fact that

$$\begin{aligned}\frac{1-F_y(y(\theta^*))}{y(\theta^*))f_y(y(\theta^*))} =& \frac{y'(\theta^*))}{y(\theta^*))}\frac{1-F_y(y(\theta^*))}{y'(\theta^*))f_y(y(\theta^*))} = \frac{y'(\theta^*))}{y(\theta^*))}\frac{1-F_w(w(\theta^*))}{w'(\theta^*))f_w(w(\theta^*))} \\ =& \frac{\frac{y'(\theta^*))}{y(\theta^*))}}{\frac{w'(\theta^*))}{w(\theta^*))}}\frac{1-F_w(w(\theta^*))}{w(\theta^*))f_w(w(\theta^*))} = (1+\tilde{\varepsilon}_{l,w}(\theta))\frac{1-F_w(w(\theta^*))}{w(\theta^*))f_w(w(\theta^*))}\end{aligned}$$

(we showed the last equality above). Now, note that

$$\frac{1+\tilde{\varepsilon}_{l,w}(\theta)}{\tilde{\varepsilon}_{l,1-\tau}(\theta)} = \frac{1+\frac{1-\tau(\theta)-y(\theta)\tau'(\theta)}{1-\tau(\theta)+\varepsilon y(\theta)\tau'(\theta)}\varepsilon}{\frac{1-\tau(\theta)}{1-\tau(\theta)+\varepsilon y(\theta)\tau'(\theta)}\varepsilon} = \frac{1+\varepsilon}{\varepsilon},$$

so that

$$\frac{\tau(\theta)}{1-\tau(\theta)} - \frac{\tau_{PE}(\theta)}{1-\tau_{PE}(\theta)} + \frac{g_\theta(\theta)-1}{\sigma}$$
$$= \frac{\tilde{\varepsilon}_{l,1-\tau}(\theta)}{\tilde{E}_{l,1-\tau}(\theta^*)}(1-\bar{g}(\theta^*))\left(1+\frac{1}{\varepsilon}\right)\left(\frac{1-F_w(w(\theta^*))}{w(\theta^*))f_w(w(\theta^*))}\right) - \left(1+\frac{1}{\varepsilon}\right)\frac{1-F_w^d(w_d(\theta))}{f_w^d(w_d(\theta))w_d(\theta)}(1-\bar{g}(\theta)).$$

Adding and subtracting

$$\frac{1-F_w(w(\theta))}{f_w(w(\theta))w(\theta)}\left(1+\frac{1}{\varepsilon}\right)(1-\bar{g}(\theta))$$

yields the result.

□

The first correction in formula (87) is the cross-wage effect $(g(y^*)-1)/\sigma$ (see (52)). It is always negative for a Rawlsian planner, and hence pushes in the direction of lower tax rates. The second correction is due to the own-wage effect, and is captured by the adjusted elasticity $\tilde{E}_{l,1-\tau}$ vs. $\tilde{\varepsilon}_{l,1-\tau}$. It is always positive, and hence pushes in the direction of higher tax rates. Finally the third correction is due to the fact that (54) and (52) are evaluated at different wage distributions. In (87), this is accounted for by the difference between the hazard rate of the wage distribution at the general equilibrium optimum, $f_w(w(\theta))$, and that of the wage distribution inferred from the data, $f_w^d(w_d(\theta))$.[56] Figure 10 decomposes quantitatively the relative importance of each of these three forces.

C Numerical Simulations: Details and Robustness

C.1 Details on Calibration of Income Distribution

We assume that incomes are log-normally distributed apart from the top, where we append a Pareto distribution for incomes above \$150,000. To obtain a smooth resulting hazard ratio $\frac{1-F_y(y)}{yf_y(y)}$, we decrease the thinness parameter of the Pareto distribution linearly between \$150,000 and \$350,000 and let it be constant at 1.5 afterwards (Diamond and Saez, 2011). In the last step we use a standard kernel smoother to ensure differentiability of the hazard ratios at \$150,000 and \$350,000. We set the mean and variance of the lognormal distribution at 10 and 0.95, respectively. The mean parameter is chosen such that the resulting income distribution has a mean of \$64,000, i.e., approximately the average US yearly earnings. The variance parameter was chosen such that the hazard ratio at level \$150,000 is equal to that reported by Diamond and Saez (2011, Fig.2). The resulting hazard ratio is illustrated in Figure 6.

[56] With endogenous marginal social welfare weights, there would be an additional correction term to account for the fact that the welfare weights are endogenous to the tax schedule.

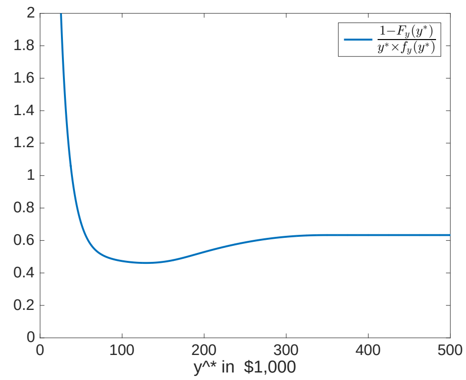


Figure 6: Calibrated Hazard Ratios of the Income Distribution

C.2 Additional Graphs for Benchmark Specification

Figure 7 illustrates optimal marginal tax rates as a function of income in the optimal allocation. Marginal tax rates in this graph reflect the policy recommendations of the optimal tax exercise which is to set marginal tax rates at each income (rather than unobservable productivity) level. A general pattern is that the marginal tax rate schedule is shifted to the left because individuals work less for optimal taxes than current taxes. This is visible most clearly for the top bracket and the bottom of the U that start earlier.

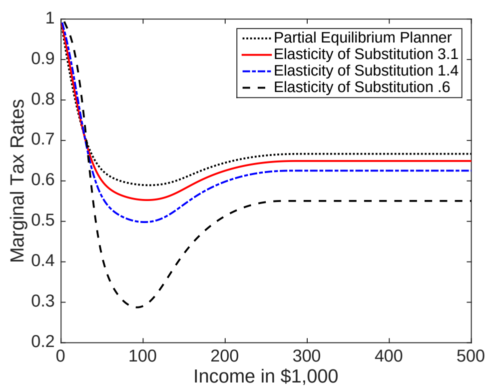
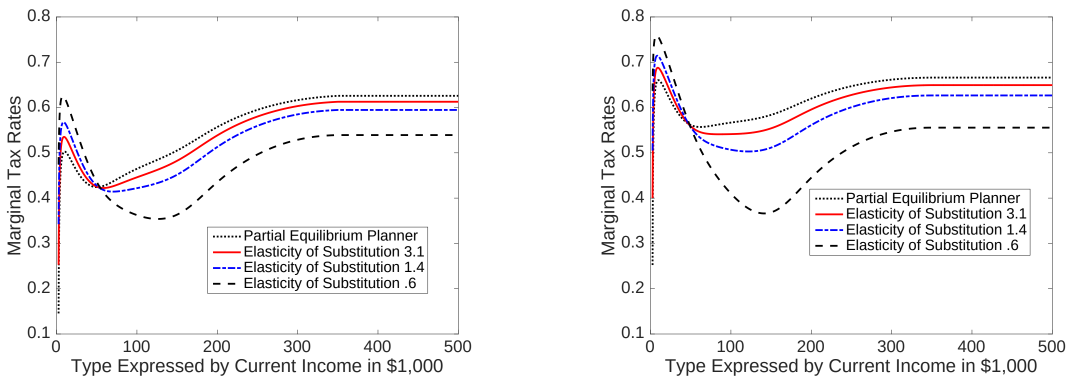


Figure 7: Optimal Marginal Tax Rates as a Function of Income

C.3 Utilitarian Welfare Function

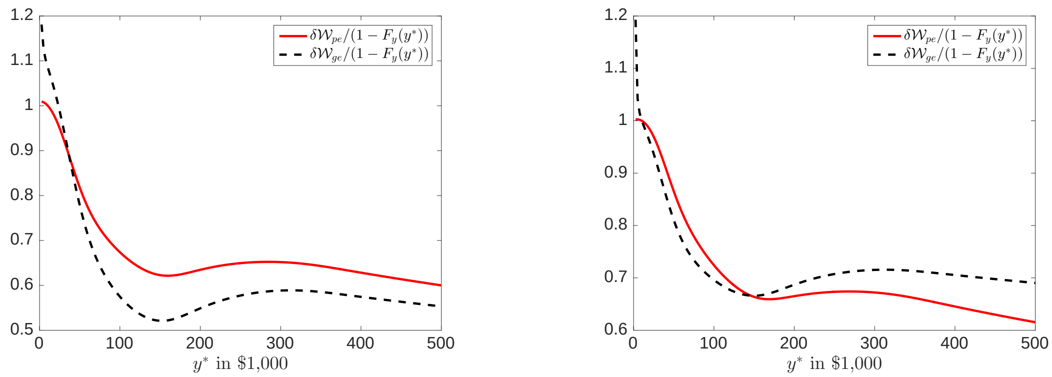
We here consider another often used social welfare function, namely the utilitarian welfare function. This implies we set $\tilde{f}(\theta) = f(\theta)$. To obtain a desire for redistribution, we assume the utility function to be $\frac{1}{1-\kappa}\left(c - l^{1+\frac{1}{\varepsilon}} / \left(1 + \frac{1}{\varepsilon}\right)\right)^{1-\kappa}$. Thus, κ determines the concavity of utility and therefore the desire for redistribution. Figure 8 illustrates optimal Utilitarian marginal tax rates for two values of κ (1 and 3). As in the Rawlsian case, the optimal U-shape of marginal tax rates is reinforced. Given that low income levels now also have positive welfare weights, the cross wage effect here is a force for higher marginal tax rates for low income levels. Thus, the result that marginal tax rates should be higher for low income levels is stronger than in the Rawlsian case in two ways: (i) the size of these effects is larger and (ii) it holds for a broader range (up to \$50,000).

Figure 8: Optimal Utilitarian Marginal Tax Rates ($CRRA = 1$ in left panel and $CRRA = 3$ in right panel)



Next we ask how the results about tax incidence are differ in the Utilitarian case. In contrast to the Rawlsian case, the policy implications of the optimal tax schedule are not necessarily overturned. For a relatively low desire to redistribute ($\kappa = 1$, see the left panel of Figure 9), the welfare gains of raising tax rates on high incomes are muted due to general equilibrium. For a stronger desire to redistribute ($\kappa = 3$, see the right panel of Figure 9), general equilibrium effects make raising top tax rates more desirable. How can that be explained? General equilibrium effects make rising top tax rates more desirable because the tax revenue increase is higher. At the same time the implied wage decreases for the working poor make them worse of. In case of very strong redistributive tastes, the tax revenue get a stronger weight (as they are used for lump-sum redistribution at the margin). In case where relatively richer workers (for whom the lump-transfer is less important relative to the very poor) still have significant welfare weights, the wage effects dominates.

Figure 9: Tax Incidence Utilitarian for $\sigma = 1.4$ ($CRRA = 1$ in left panel and $CRRA = 3$ in right panel)



C.4 General Equilibrium Wedge Accounting

Figure 10 decomposes quantitatively the relative importance of each of the three forces highlighted in Proposition 6, for $\varepsilon = 0.33$ and $\sigma = 1.4$. The partial (resp., general) equilibrium optimum is represented by the black dashed (resp., red bold) curve. The black dotted, blue dashed-dotted, and diamond-marked curves illustrate suboptimal tax schedules where each of the three elements of the decomposition (87) (respectively, the cross-wage term, the elasticity correction term, and the hazard rate correction term) are ignored. This graph shows that the hazard rate correction term (iii) has a minor quantitative importance.

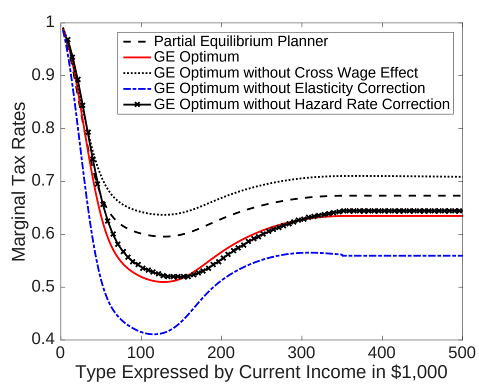


Figure 10: General Equilibrium Wedge Accounting

C.5 Translog Production Function

First we again look at tax incidence but change the parameterization of the Translog production function such that $\bar{\bar{\gamma}}(y^*, y^*) = -1$ for $y^* = \$80,000$ instead of \$250,000.

Figure 11: $\bar{\gamma}(\theta, \theta^*)$ with $y(\theta^*) = \$250,000$ (left panel) and the implied tax incidence (right panel)

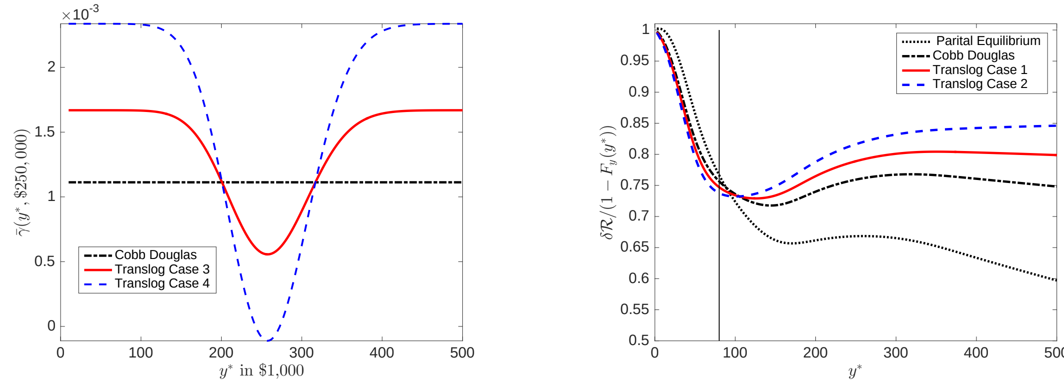


Figure 11 illustrates the distance dependent wage effects for these Cases 3 and 4. The tax incidence results are illustrated in the right panel. Here the best comparison level for y^* to understand the effects of distance dependence is \$80,000. In line with the results in the main body of the text, the general equilibrium effects on tax incidence are increased in magnitude through distance dependence.

Figure 12: Optimal Marginal Tax Rates for Translog Production Functions

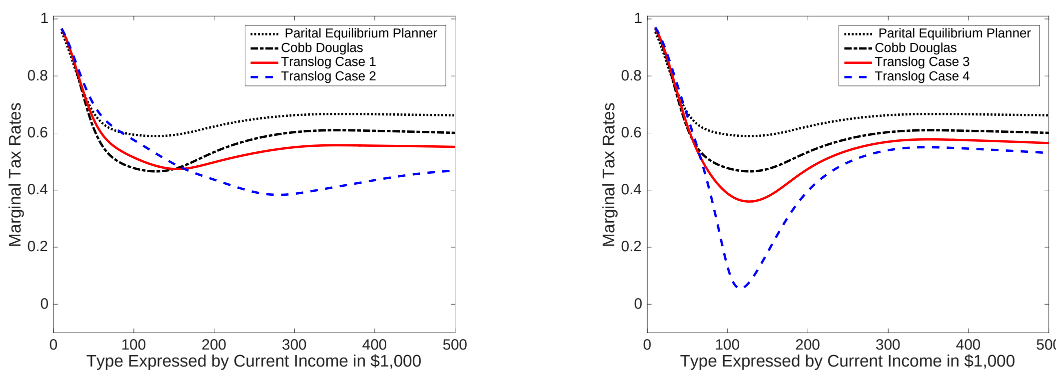


Figure 12 illustrates the resulting optimal marginal tax rates for all 4 cases. As cases 1 and 3 are more similar to Cobb Douglas than cases 2 and 4, not surprisingly, marginal tax rates are closer to the Cobb Douglas counterpart. These results can be best interpreted by looking at Figure 13 where the own-wage effects are illustrated.[57] For Cases 1 and 3, $\bar{\bar{\gamma}}(y, y)$ is relatively flat and close to -1 (as in Cobb Douglas). For Case 4, the wage effects are generally a bit larger in magnitude and still relatively flat (magnitude varies by a factor 2), which explains why the CES effects are mainly increased in magnitude. For Case 2, the magnitude of the wage effects strongly increases and varies

[57] They are illustrated for the current tax system but do look very similar for the respective optimal tax system.

by a factor of up to 6, which explains why the shape of marginal tax rates is more different from CES than in the other 3 cases. Finally, welfare gains in consumption equivalents are 1.09%, 0.31%, 1.85% and 2.53% respectively for these cases.

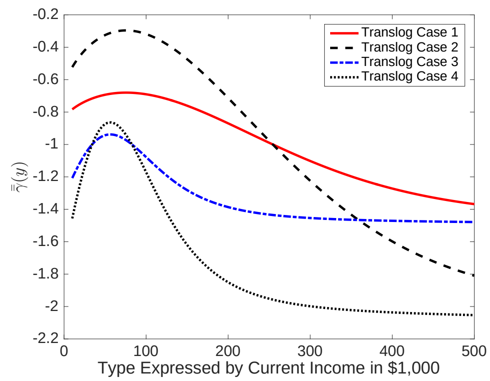


Figure 13: Illustration of the own-wage effect for current policies