# Online Appendix for the article
**"Voting over Selfishly Optimal Tax Schedules: Can Pigouvian Tax Redistribute Income?"**

Darong Dai*

June 16, 2020

# 1 Appendix B: Comparative Statics of Pigouvian Tax Rate with respect to Welfare Weight

To conduct a comparative static analysis about how the Pigouvian tax rate changes when welfare weights change, I follow Jacquet et al. (2013) and Jacquet and Lehmann (2020) to adopt the generalized Bergson-Samuelson social welfare function $\mathcal{W}(U(w), w)$ that also depends on individuals' type $w$. In particular, this social objective encompasses weighted utilitarian preferences (see, Weymark, 1987) with type-dependent welfare weights. Replacing by this one the social goal adopted in the text, the optimal tax formula,

$$\begin{aligned}\frac{\tau(w)}{1-\tau(w)} &= \frac{\gamma_1}{\gamma_2} \\ &+ \underbrace{[1+\varepsilon_{h_1,l}(w,\bar{c})]}_{\mathcal{A}(w)} \cdot \underbrace{\left[\frac{1-F(w)}{wf(w)}\right]}_{\mathcal{B}(w)} \cdot \underbrace{\frac{(1/\gamma_2)\int_w^{\overline{w}}[\gamma_1+\gamma_2-\mathcal{W}'(U(t))]f(t)dt}{1-F(w)}}_{\mathcal{C}(w)},\end{aligned} \tag{1}$$

shown in Lemma 3.1 of the text is slightly modified as follows:

$$\begin{aligned}\frac{\tau(w)}{1-\tau(w)} &= \frac{\gamma_1}{\gamma_2} \\ &+ [1+\varepsilon_{h_1,l}(w,\bar{c})] \cdot \left[\frac{1-F(w)}{wf(w)}\right] \cdot \frac{(1/\gamma_2)\int_w^{\overline{w}}[\gamma_1+\gamma_2-\mathcal{W}_U(U(t),t)]f(t)dt}{1-F(w)}.\end{aligned}$$

Using this formula, I immediately get the Pigouvian tax rate as $\tau^{\mathrm{P}} = \gamma_1/(\gamma_1+\gamma_2)$, which is actually the optimal marginal tax rate (MTR) facing the bottom and top skill types. Without loss of generality, I assume that $\mathcal{W}(U(w), w) \equiv \Upsilon(w)\mathcal{W}(U(w))$ with the welfare weight $\Upsilon(w)$ assigned to type $w$. Then, I will derive and sign the expression of $\partial\tau^{\mathrm{P}}/\partial\Upsilon(w)$ in what follows.

*Institute for Advanced Research, Key Laboratory of Mathematical Economics (SUFE), Ministry of Education of China, Shanghai University of Finance and Economics, Shanghai 200433, People's Republic of China. E-mail: dai.darong@sufe.edu.cn; darongdai@hotmail.com

Rearranging the following optimality condition that appears in the proof of Lemma 3.1:

$$\gamma_1 - (\gamma_1 + \gamma_2) \int_{\underline{w}}^{\overline{w}} h_2\left(l(w), \bar{c}\right) f(w) dw + \int_{\underline{w}}^{\overline{w}} q(w) h_{12}\left(l(w), \bar{c}\right) \frac{l(w)}{w} dw = 0,$$

I get that

$$\tau^{\mathrm{P}} = -\int_{\underline{w}}^{\overline{w}} \left[\frac{q(w)}{\gamma_1 + \gamma_2}\right] h_{12}\left(l(w), \bar{c}\right) \frac{l(w)}{w} dw + \int_{\underline{w}}^{\overline{w}} h_2\left(l(w), \bar{c}\right) f(w) dw. \tag{2}$$

In addition, making use of the following optimality condition that appears in the proof of Lemma 3.1:

$$-q(w) = -\int_{w}^{\overline{w}} \mathcal{W}'(U(t)) f(t) dt + (\gamma_1 + \gamma_2)[1 - F(w)], \tag{3}$$

I get

$$-\frac{q(w)}{\gamma_1 + \gamma_2} = -\frac{1}{\gamma_1 + \gamma_2} \int_{w}^{\overline{w}} \mathcal{W}'(U(t)) \Upsilon(t) f(t) dt + [1 - F(w)],$$

differentiating both sides of which with respect to $\Upsilon(w)$ gives the following derivative:

$$\frac{\partial\left[-\frac{q(w)}{\gamma_1+\gamma_2}\right]}{\partial \Upsilon(w)} = \frac{\partial(\gamma_1 + \gamma_2)/\partial \Upsilon(w)}{(\gamma_1 + \gamma_2)^2} \int_{w}^{\overline{w}} \mathcal{W}'(U(t)) \Upsilon(t) f(t) dt - \frac{\mathcal{W}'(U(w)) f(w)}{\gamma_1 + \gamma_2}. \tag{4}$$

In light of the transversality condition, namely $q(\underline{w}) = 0$, as well as the optimality condition (3) with replacing the social welfare function by $\Upsilon(w)\mathcal{W}(U(w))$, I have

$$\gamma_1 + \gamma_2 = \int_{\underline{w}}^{\overline{w}} \mathcal{W}'(U(t)) \Upsilon(t) f(t) dt, \tag{5}$$

by which $\partial(\gamma_1 + \gamma_2)/\partial \Upsilon(w) = \mathcal{W}'(U(w)) f(w)$ is immediate. In consequence, equation (4) can be rewritten as

$$\frac{\partial\left[-\frac{q(w)}{\gamma_1+\gamma_2}\right]}{\partial \Upsilon(w)} = \frac{\mathcal{W}'(U(w)) f(w)}{(\gamma_1 + \gamma_2)^2} \left[\int_{w}^{\overline{w}} \mathcal{W}'(U(t)) \Upsilon(t) f(t) dt - (\gamma_1 + \gamma_2)\right] \leq 0, \tag{6}$$

in which, given that $w \geq \underline{w}$, the inequality sign follows from using equation (5). Applying (6) to (2) shows that

$$\frac{\partial \tau^{\mathrm{P}}}{\partial \Upsilon(w)} = \frac{\mathcal{W}'(U(w)) f(w)}{(\gamma_1 + \gamma_2)^2} \left[\int_{w}^{\overline{w}} \mathcal{W}'(U(t)) \Upsilon(t) f(t) dt - (\gamma_1 + \gamma_2)\right] h_{12}\left(l(w), \bar{c}\right) \frac{l(w)}{w} \leq 0.$$

Therefore, for any given skill type $w$, the larger the welfare weight imposed on this type, the lower the Pigouvian tax rate is. This result critically depends on the status-seeking motive characterized by the assumption $h_{12} > 0$. In particular, I obtain, in light of condition (5), $\partial \tau^{\mathrm{P}}/\partial \Upsilon(w)|_{w=\underline{w}} = 0$, namely, the marginal effect of bottom type's welfare weight on the Pigouvian tax rate vanishes. Given that the bottom type earns the lowest income in the present context, the underlying intuition of this result seems straightforward.

# 2 Appendix C: Alternative Constraints on the Reference Consumption

In the text, I follow the common practice and use the average consumption in the population to represent the reference consumption level for all skill types. There are, of course, other specifications of the reference consumption that generates a negative consumption externality. I check how the redistributive tax schedules in the social optimum and selfish optimum change across alternative measures of status-seeking motive. For simplicity, I shall focus on the first-order approach in the following analysis.

## 2.1 Type-dependent Reference Consumption

It is reasonable to assume that individuals make local rather than global social comparisons, and hence the reference consumption might be specified as being type-dependent. Also, an emulation-driven individual is likely to become less happy when the average consumption (or earning) of the higher income class becomes larger, which suggests the following constraint on the reference consumption:

$$\bar{c}(w) \;\geq\; \int_{w}^{\overline{w}} c(t) f(t) dt \tag{7}$$

for any $w \in [\underline{w}, \overline{w}]$. As is obvious, the reference consumption $\bar{c}(w)$ is type-dependent as the right-hand side of (7) varies with skill type.

As before, I shall derive first the socially optimal income taxation schedule and then the selfishly optimal income schedule. The benevolent social planner chooses the bundle $\{y(w), c(w), U(w), \bar{c}(w)\}_{w\in[\underline{w},\overline{w}]}$ that solves the problem:

$$\max \int_{\underline{w}}^{\overline{w}} \mathcal{W}(U(w)) f(w) dw \tag{8}$$

subject to the reference consumption constraint (7), the resource constraint:

$$\int_{\underline{w}}^{\overline{w}} [y(w) - c(w)] f(w) dw \;\geq\; 0, \tag{9}$$

and the following first-order incentive compatibility constraint and the individual gross utility:

$$\begin{aligned} U'(w) \;&= h_1\left(\frac{y(w)}{w}, \bar{c}(w)\right) \frac{y(w)}{w^2}, \\ U(w) \;&= c(w) - h\left(\frac{y(w)}{w}, \bar{c}(w)\right), \end{aligned}$$

for all $w \in [\underline{w}, \overline{w}]$.

The socially optimal income taxation schedule is obtained from solving problem (8).

**Lemma 2.1** *Under the type-dependent status seeking, the tax formula for socially optimal MTRs reads as follows:*

$$
\begin{aligned}
\frac{\tau(w)}{1-\tau(w)} &= \frac{1}{\gamma_2}\int_{\underline{w}}^{w}\gamma_1(t)dt \\
&+ \underbrace{[1+\varepsilon_{h_1,l}(w)]}_{\mathcal{A}(w)}\cdot\underbrace{\left[\frac{1-F(w)}{wf(w)}\right]}_{\mathcal{B}(w)}\cdot\underbrace{\frac{(1/\gamma_2)\int_{w}^{\overline{w}}\left[\int_{\underline{w}}^{t}\gamma_1(z)dz+\gamma_2-\mathcal{W}'(U(t))\right]f(t)dt}{1-F(w)}}_{\mathcal{C}(w)},
\end{aligned}
\tag{10}
$$

*where $w \in [\underline{w}, \overline{w}]$, $\gamma_1(w), \gamma_2 > 0$ are Lagrangian multipliers on the constraints (7) and (9), respectively, and the labor supply elasticity $\varepsilon_{h_1,l}(w)$ is given by*

$$
\varepsilon_{h_1,l}(w) \equiv \frac{h_{11}(l(w),\bar{c}(w))l(w)}{h_1(l(w),\bar{c}(w))}. \tag{11}
$$

Comparing (10) to (1), the difference in Pigouvian tax arises from the fact that the shadow price, denoted by $\gamma_1(\cdot)$, of the constraint on reference consumption becomes type-dependent. Under constraint (7), the contribution to the negative externality increases as skill level increases, and hence the socially optimal Pigouvian tax rate is no longer type-independent. In fact, I can give the following proposition.

**Proposition 2.1** *Under the type-dependent status-seeking, the socially optimal tax schedule features three properties:*

*(a) The Pigouvian tax is type-dependent, and high skills face higher Pigouvian tax rates than low skills;*

*(b) Individuals of the top skill level face the highest Pigouvian tax rate, and they do not face any Mirrleesian tax;*

*(c) Individuals of the bottom skill level face a zero MTR.*

Since the consumption of higher skills imposes a negative externality to a larger number of individuals than the consumption of lower skills, the intuition of observation (a) follows. As two extremes, the consumption of top-skill individuals causes a negative externality to the largest number of individuals whereas the consumption of bottom-skill individuals does not cause an externality at all, the intuition of observations (b)-(c) is also immediate.

I now proceed to derive the selfishly optimal income taxation schedule. Similar to the following problem

$$
\begin{aligned}
\max_{\{y(\cdot),\bar{c}\}}\int_{\underline{w}}^{k}&\left\{\left[y(w)-h\left(\frac{y(w)}{w},\bar{c}\right)\right]f(w)+\frac{y(w)}{w^2}h_1\left(\frac{y(w)}{w},\bar{c}\right)F(w)\right\}dw \\
&+\int_{k}^{\overline{w}}\left\{\left[y(w)-h\left(\frac{y(w)}{w},\bar{c}\right)\right]f(w)-\frac{y(w)}{w^2}h_1\left(\frac{y(w)}{w},\bar{c}\right)[1-F(w)]\right\}dw
\end{aligned}
$$

shown in the text, for any proposer of skill type $k \in [\underline{w}, \overline{w}]$, his program can be written as follows:

$$
\begin{aligned}
\max_{\{y(\cdot),\bar{c}(\cdot)\}} & \int_{\underline{w}}^{k} \left[ y(w) - h\left(\frac{y(w)}{w}, \bar{c}(w)\right)\right] f(w)dw \\
& + \int_{\underline{w}}^{k} \frac{y(w)}{w^2} h_1 \left(\frac{y(w)}{w}, \bar{c}(w)\right) F(w)dw \\
& + \int_{k}^{\overline{w}} \left[ y(w) - h\left(\frac{y(w)}{w}, \bar{c}(w)\right)\right] f(w)dw \\
& - \int_{k}^{\overline{w}} \frac{y(w)}{w^2} h_1 \left(\frac{y(w)}{w}, \bar{c}(w)\right) [1 - F(w)]dw \\
& + \int_{\underline{w}}^{\overline{w}} \lambda(w)\bar{c}(w)dw - \int_{\underline{w}}^{\overline{w}} c(w)f(w) \left[\int_{\underline{w}}^{w} \lambda(t)dt\right] dw,
\end{aligned}
\tag{12}
$$

where $\lambda(w) > 0$ denotes the Lagrangian multiplier on constraint (7).

Solving this problem (12), the selfishly optimal income taxation schedules under type-dependent status seeking are given in the following lemma.

**Lemma 2.2** *For any proposer of type $k \in (\underline{w}, \overline{w})$ with a constant elasticity of labor supply, the selfishly optimal schedule of before-tax incomes $y(\cdot)$ is given by*

$$
y(w) = \begin{cases} y^{max}(w) & \text{for } w \in [\underline{w}, k), \\ y^{min}(w) & \text{for } w \in (k, \overline{w}] \end{cases}
$$

*in which there is a downward jump discontinuity at $w = k$, and the corresponding maxi-max and maxi-min MTRs are given, respectively, by the following tax formulas:*

$$
\frac{\tau^{max}(w)}{1-\tau^{max}(w)} = \int_{\underline{w}}^{w} \lambda^{max}(t)dt - [1 + \varepsilon_{h_1,l}(w)] \left[\frac{F(w)}{wf(w)}\right] \tag{13}
$$

*and*

$$
\frac{\tau^{min}(w)}{1-\tau^{min}(w)} = \int_{\underline{w}}^{w} \lambda^{min}(t)dt + [1 + \varepsilon_{h_1,l}(w)] \left[\frac{1-F(w)}{wf(w)}\right], \tag{14}
$$

*in which*

$$
\begin{aligned}
\lambda^{max}(w) &= h_2\left(\frac{y(w)}{w}, \bar{c}(w)\right) f(w) - \frac{y(w)}{w^2} h_{12}\left(\frac{y(w)}{w}, \bar{c}(w)\right) F(w), \\
\lambda^{min}(w) &= h_2\left(\frac{y(w)}{w}, \bar{c}(w)\right) f(w) + \frac{y(w)}{w^2} h_{12}\left(\frac{y(w)}{w}, \bar{c}(w)\right) [1 - F(w)].
\end{aligned}
\tag{15}
$$

By comparing the tax schedule established in Lemma 2.1 to that established in Lemma 2.2, the following observation immediately follows.

**Corollary 2.1** *Under the type-dependent status-seeking, the selfishly optimal tax schedule features as well the properties given in Proposition 2.1.*

We thus conclude that, given the present type-dependent reference consumption constraint, the Pigouvian tax in both the socially optimal and selfishly optimal tax schedules exhibits a novel income redistributive effect.

## 2.2 Median-skill Consumption as the Reference Consumption

In addition to the average level of consumption to which individuals make social comparisons, it is also reasonable to consider the following constraint on the reference consumption:

$$\bar{c} \;\geq\; c(w_{\text{median}}), \tag{16}$$

in which $c(w_{\text{median}})$ denotes the consumption of workers of the median skill level.

Replacing the reference consumption constraint by (16) in problem (8), I establish the following socially optimal income taxation schedule.

**Lemma 2.3** *The tax formula for socially optimal MTRs reads as follows:*

$$\begin{aligned}\frac{\tau(w)}{1-\tau(w)} \;&=\; \frac{\gamma_1}{\gamma_2}\cdot \mathbb{I}_{w=w_{median}}\\ &+ \underbrace{[1+\varepsilon_{h_1,l}(w,\bar{c})]}_{\mathcal{A}(w)}\cdot \underbrace{\left[\frac{1-F(w)}{wf(w)}\right]}_{\mathcal{B}(w)}\cdot \underbrace{\frac{(1/\gamma_2)\int_w^{\overline{w}}[\gamma_1\cdot \mathbb{I}_{w=w_{median}}+\gamma_2-\mathcal{W}'(U(t))]f(t)dt}{1-F(w)}}_{\mathcal{C}(w)},\end{aligned} \tag{17}$$

*where $w \in [\underline{w}, \overline{w}]$, $\mathbb{I}_{w=w_{median}}$ denotes the indicator function, and $\gamma_1, \gamma_2 > 0$ are Lagrangian multipliers on constraints (16) and (9), respectively.*

Using the tax formula (17), the following proposition is immediate.

**Proposition 2.2** *Imposing the median-skill consumption as the reference consumption for all skills, the socially optimal tax schedule features these two properties:*

*(a) Only individuals of the median skill level face a positive Pigouvian tax;*

*(b) The MTR is zero for both the highest skilled and the lowest skilled individuals.*

The intuition of observation (a) is easy to follow because it is the consumption of median-skill individuals that causes a negative externality to individuals of other skill types.

I now proceed to derive the tax formulas for selfishly optimal MTRs. Noting that $\partial U/\partial \bar{c} < 0$ and $c = y - T(y)$, the problem for a proposer of type-$k$ can be written as follows:

$$\begin{aligned}\max_{y(\cdot)} \int_{\underline{w}}^{k} &\left[y(w) - h\left(\frac{y(w)}{w}, y(w_{\text{median}}) - T(y(w_{\text{median}}))\right)\right] f(w)dw\\ &+ \int_{\underline{w}}^{k} \frac{y(w)}{w^2} h_1\left(\frac{y(w)}{w}, y(w_{\text{median}}) - T(y(w_{\text{median}}))\right) F(w)dw\\ &+ \int_{k}^{\overline{w}} \left[y(w) - h\left(\frac{y(w)}{w}, y(w_{\text{median}}) - T(y(w_{\text{median}}))\right)\right] f(w)dw\\ &- \int_{k}^{\overline{w}} \frac{y(w)}{w^2} h_1\left(\frac{y(w)}{w}, y(w_{\text{median}}) - T(y(w_{\text{median}}))\right) [1-F(w)]dw.\end{aligned} \tag{18}$$

Solving problem (18), the selfishly optimal income taxation schedules under the constraint (16) on status seeking are given in the following lemma.

**Lemma 2.4** *For any proposer of type $k \in (\underline{w}, \overline{w})$ with a constant elasticity of labor supply, the following statements are true.*

*(i) The selfishly optimal schedule of before-tax incomes $y(\cdot)$ is given by*

$$y(w) = \begin{cases} y^{max}(w) & \text{for } w \in [\underline{w}, k), \\ y^{min}(w) & \text{for } w \in (k, \overline{w}] \end{cases}$$

*in which there is a downward jump discontinuity at $w = k$, and the corresponding maxi-max and maxi-min MTRs for $w \neq w_{median}$ are given, respectively, by the following tax formulas:*

$$\frac{\tau^{max}(w)}{1 - \tau^{max}(w)} = -[1 + \varepsilon_{h_1,l}(w, \bar{c})] \left[ \frac{F(w)}{wf(w)} \right] \tag{19}$$

*and*

$$\frac{\tau^{min}(w)}{1 - \tau^{min}(w)} = [1 + \varepsilon_{h_1,l}(w, \bar{c})] \left[ \frac{1 - F(w)}{wf(w)} \right]. \tag{20}$$

*(ii) For $w = w_{median}$, the tax formulas are given as follows. If $w_{median} > k$, then workers of this skill type face maxi-min MTRs:*

$$\frac{\tau^{min}(w)}{1 - \tau^{min}(w)} = \lambda^{min}(w) + [1 + \varepsilon_{h_1,l}(w, \bar{c})] \left[ \frac{1 - F(w)}{wf(w)} \right] \tag{21}$$

*with*

$$\lambda^{min}(w) = h_2 \left( \frac{y(w)}{w}, \bar{c} \right) + \frac{y(w)}{w} h_{12} \left( \frac{y(w)}{w}, \bar{c} \right) \frac{1 - F(w)}{wf(w)}; \tag{22}$$

*if $w_{median} < k$, then they face maxi-max MTRs:*

$$\frac{\tau^{max}(w)}{1 - \tau^{max}(w)} = \lambda^{max}(w) - [1 + \varepsilon_{h_1,l}(w, \bar{c})] \left[ \frac{F(w)}{wf(w)} \right] \tag{23}$$

*with*

$$\lambda^{max}(w) = h_2 \left( \frac{y(w)}{w}, \bar{c} \right) - \frac{y(w)}{w} h_{12} \left( \frac{y(w)}{w}, \bar{c} \right) \frac{F(w)}{wf(w)}. \tag{24}$$

By comparing the tax schedule established in Lemma 2.3 to that established in Lemma 2.4, the following observation immediately follows.

**Corollary 2.2** *Imposing the median-skill consumption as the reference consumption for all skill types, the selfishly optimal tax schedule features as well the properties given in Proposition 2.2.*

We accordingly conclude that, given the negative consumption externality caused only by the median-skill type, the Pigouvian tax in both the socially optimal and selfishly optimal tax schedules exhibits a novel income redistributive effect. In particular, the externality-correcting tax burden is totally placed on the median-skill type individuals, regardless of whether the social optimality or selfish optimality is adopted as the criterion of income taxation.

### 2.3 Proofs

**Proof of Lemma 2.1.** As in the proof of Lemma 3.1, let me treat the optimal choice of individual consumption $c(w)$ as an implicit function of $U(w)$, $l(w)$ and $\bar{c}(w)$, and equivalently rewrite it as $\varphi(\cdot)$. Applying the Implicit Function Theorem, I get

$$\frac{\partial\varphi}{\partial l(w)} = h_1(l(w),\bar{c}), \ \frac{\partial\varphi}{\partial U(w)} = 1 \ \text{ and } \ \frac{\partial\varphi}{\partial\bar{c}} = h_2(l(w),\bar{c}). \tag{25}$$

The Lagrangian of problem (8) is written as follows:

$$\begin{aligned} &\mathcal{L}\left(\{U(w),l(w),\bar{c}(w)\}_{w\in[\underline{w},\overline{w}]};\{\gamma_1(w),q(w)\}_{w\in[\underline{w},\overline{w}]},\gamma_2\right)\\ = &\int_{\underline{w}}^{\overline{w}}\mathcal{W}(U(w))f(w)dw+\int_{\underline{w}}^{\overline{w}}\gamma_1(w)\left[\bar{c}(w)-\int_{w}^{\overline{w}}\varphi(U(t),l(t),\bar{c}(t))f(t)dt\right]dw\\ &+\gamma_2\int_{\underline{w}}^{\overline{w}}\left[wl(w)-\varphi(U(w),l(w),\bar{c}(w))\right]f(w)dw\\ &+\int_{\underline{w}}^{\overline{w}}q(w)\left[h_1(l(w),\bar{c}(w))\frac{l(w)}{w}-U'(w)\right]dw \end{aligned}\tag{26}$$

in which $\gamma_1(w),\gamma_2$ are nonnegative Lagrangian multipliers, and $q(w)$ is a co-state variable. Reversing the order of integration in (26) gives rise to

$$\begin{aligned} &\int_{\underline{w}}^{\overline{w}}\gamma_1(w)\left[\int_{w}^{\overline{w}}\varphi(U(t),l(t),\bar{c}(t))f(t)dt\right]dw\\ =&\int_{\underline{w}}^{\overline{w}}\varphi(U(t),l(t),\bar{c}(t))f(t)\left[\int_{\underline{w}}^{t}\gamma_1(w)dw\right]dt. \end{aligned}\tag{27}$$

In light of the following optimality condition that appears in the proof of Lemma 3.1:

$$\int_{\underline{w}}^{\overline{w}}q(w)U'(w)dw = q(\overline{w})U(\overline{w})-q(\underline{w})U(\underline{w})-\int_{\underline{w}}^{\overline{w}}q'(w)U(w)dw,$$

I get by applying (27) to (26) that

$$\begin{aligned} &\mathcal{L}\left(\{U(w),l(w),\bar{c}(w)\}_{w\in[\underline{w},\overline{w}]};\{\gamma_1(w),q(w)\}_{w\in[\underline{w},\overline{w}]},\gamma_2\right)\\ = &\int_{\underline{w}}^{\overline{w}}\mathcal{W}(U(w))f(w)dw+\int_{\underline{w}}^{\overline{w}}\gamma_1(w)\bar{c}(w)dw\\ &-\int_{\underline{w}}^{\overline{w}}\varphi(U(t),l(t),\bar{c}(t))f(t)\left[\int_{\underline{w}}^{t}\gamma_1(w)dw\right]dt\\ &+\gamma_2\int_{\underline{w}}^{\overline{w}}\left[wl(w)-\varphi(U(w),l(w),\bar{c}(w))\right]f(w)dw\\ &+q(\underline{w})U(\underline{w})-q(\overline{w})U(\overline{w})+\int_{\underline{w}}^{\overline{w}}\left[q(w)h_1(l(w),\bar{c}(w))\frac{l(w)}{w}+q'(w)U(w)\right]dw. \end{aligned}$$

Assuming the existence of an interior solution and making use of (25), the optimality conditions read as:

$$\frac{\partial\mathcal{L}}{\partial U(w)} = \mathcal{W}'(U(w))f(w)-\left[\int_{\underline{w}}^{w}\gamma_1(t)dt\right]f(w)-\gamma_2 f(w)+q'(w)=0 \tag{28}$$

for $\forall w \in (\underline{w}, \overline{w})$ with the transversality conditions

$$q(\underline{w}) = q(\overline{w}) = 0; \tag{29}$$

and also

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \bar{c}(w)} &= \gamma_1(w) - \left[\int_{\underline{w}}^{w} \gamma_1(t)dt + \gamma_2\right] h_2\left(l(w), \bar{c}(w)\right) f(w) \\ &\quad + q(w) h_{12}\left(l(w), \bar{c}(w)\right) \frac{l(w)}{w} = 0 \ \forall w \in [\underline{w}, \overline{w}],\end{aligned} \tag{30}$$

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial l(w)} &= -\left[\int_{\underline{w}}^{w} \gamma_1(t)dt\right] h_1\left(l(w), \bar{c}(w)\right) f(w) + \gamma_2\left[w - h_1\left(l(w), \bar{c}(w)\right)\right] f(w) \\ &\quad + q(w)\left[h_{11}\left(l(w), \bar{c}(w)\right) \frac{l(w)}{w} + h_1\left(l(w), \bar{c}(w)\right) \frac{1}{w}\right] = 0 \ \forall w \in [\underline{w}, \overline{w}].\end{aligned} \tag{31}$$

Using the following equation given in the text:

$$\tau(w) \;=\; 1 - h_1\left(\frac{y(w)}{w}, \bar{c}\right)\frac{1}{w}, \ \ \forall w \in [\underline{w}, \overline{w}], \tag{32}$$

then applying (11) to (31) gives

$$\frac{\tau(w)}{1-\tau(w)} = \frac{1}{\gamma_2}\int_{\underline{w}}^{w} \gamma_1(t)dt - \frac{q(w)}{\gamma_2}\left[1 + \varepsilon_{h_1,l}(w)\right]\frac{1}{wf(w)}. \tag{33}$$

Integrating on both sides of equation (28) and using (29), I arrive at

$$-q(w) = -\int_{w}^{\overline{w}} \mathcal{W}'(U(t))f(t)dt + \int_{w}^{\overline{w}}\left[\int_{\underline{w}}^{t} \gamma_1(z)dz\right] f(t)dt + \gamma_2[1 - F(w)]. \tag{34}$$

Substituting (34) into (33) and rearranging the algebra, the tax formula (10) is, accordingly, established. ■

**Proof of Lemma 2.2.** The proof is quite similar to that of Lemma 4.1, so here I just show the main steps to economize on the space. Replacing $c(w)$ with $y(w) - T(y(w))$ and setting $k = \overline{w}$, problem (12) can be simplified as follows:

$$\begin{aligned}\max_{\{y(\cdot),\bar{c}(\cdot)\}} &\int_{\underline{w}}^{\overline{w}} \left\{\left[y(w) - h\left(\frac{y(w)}{w}, \bar{c}(w)\right)\right] f(w) + \frac{y(w)}{w^2} h_1\left(\frac{y(w)}{w}, \bar{c}(w)\right) F(w)\right\} dw \\ &+ \int_{\underline{w}}^{\overline{w}} \lambda(w)\bar{c}(w)dw - \int_{\underline{w}}^{\overline{w}} [y(w) - T(y(w))]f(w)\left[\int_{\underline{w}}^{w} \lambda(t)dt\right] dw.\end{aligned}$$

In view of equation (32), the first-order conditions with respect to $y(\cdot)$ and $\bar{c}(\cdot)$ are, respectively, given by

$$\begin{aligned}&\left[1 - h_1\left(\frac{y(w)}{w}, \bar{c}(w)\right)\frac{1}{w}\right] f(w) - [1 - \tau^{\max}(w)]f(w)\int_{\underline{w}}^{w} \lambda(t)dt \\ &+ \left[\frac{1}{w^2} h_1\left(\frac{y(w)}{w}, \bar{c}(w)\right) + \frac{y(w)}{w^3} h_{11}\left(\frac{y(w)}{w}, \bar{c}(w)\right)\right] F(w) = 0\end{aligned} \tag{35}$$

and

$$\lambda^{\max}(w) - h_2\left(\frac{y(w)}{w}, \bar{c}(w)\right) f(w) + \frac{y(w)}{w^2} h_{12}\left(\frac{y(w)}{w}, \bar{c}(w)\right) F(w) = 0 \tag{36}$$

for all $w \in [\underline{w}, \overline{w}]$. As a consequence, (36) implies the first equation of (15); applying equations (32) and (11) to (35) and rearranging the algebra gives the desired (13).

Similarly, setting $k = \underline{w}$ in problem (12) gives the program:

$$\max_{\{y(\cdot),\bar{c}(\cdot)\}} \int_{\underline{w}}^{\overline{w}} \left\{ \left[ y(w) - h\left(\frac{y(w)}{w}, \bar{c}(w)\right) \right] f(w) - \frac{y(w)}{w^2} h_1\left(\frac{y(w)}{w}, \bar{c}(w)\right) [1 - F(w)] \right\} dw$$
$$+ \int_{\underline{w}}^{\overline{w}} \lambda(w)\bar{c}(w) dw - \int_{\underline{w}}^{\overline{w}} [y(w) - T(y(w))] f(w) \left[ \int_{\underline{w}}^{w} \lambda(t) dt \right] dw.$$

Using equation (32) again, the first-order conditions with respect to $y(\cdot)$ and $\bar{c}(\cdot)$ are, respectively, given by

$$\begin{aligned} &\left[1 - h_1\left(\frac{y(w)}{w}, \bar{c}(w)\right)\frac{1}{w}\right] f(w) - [1 - \tau^{\min}(w)] f(w) \int_{\underline{w}}^{w} \lambda(t) dt \\ &- \left[\frac{1}{w^2} h_1\left(\frac{y(w)}{w}, \bar{c}(w)\right) + \frac{y(w)}{w^3} h_{11}\left(\frac{y(w)}{w}, \bar{c}(w)\right)\right] [1 - F(w)] = 0 \end{aligned} \tag{37}$$

and

$$\lambda^{\min}(w) - h_2\left(\frac{y(w)}{w}, \bar{c}(w)\right) f(w) - \frac{y(w)}{w^2} h_{12}\left(\frac{y(w)}{w}, \bar{c}(w)\right) [1 - F(w)] = 0 \tag{38}$$

for all $w \in [\underline{w}, \overline{w}]$. As a consequence, (38) implies the second equation of (15); applying equations (32) and (11) to (37) and rearranging the algebra gives the desired (14).

Since by assumption I have $h_2, h_{12} > 0$, thus $\lambda^{\min}(w) > 0$ for all $w$, which then gives rise to $\tau^{\min}(w) > 0$ for all $w$ by using (14). If $\lambda^{\max}(w) < 0$ for all $w$, which then gives rise to $\tau^{\max}(w) < 0$ for all $w$ by using (13). The claim of the existence of a downward jump discontinuity of the selfishly optimal income schedule at $w = k$ thus follows under a constant elasticity of labor supply. Even if $\lambda^{\max}(w) \geq 0$ for some $w$, as I always have $\lambda^{\max}(w) < \lambda^{\min}(w)$ by (15), the downward discontinuity still emerges from the direct comparison of (13) and (14). ■

**Proof of Lemma 2.3.** As in the proof of Lemma 3.1, the Lagrangian reads as follows:

$$\begin{aligned} &\mathcal{L}\left(\{U(w), l(w)\}_{w\in[\underline{w},\overline{w}]}, \bar{c}; \gamma_1, \gamma_2, \{q(w)\}_{w\in[\underline{w},\overline{w}]}\right) \\ = &\int_{\underline{w}}^{\overline{w}} \mathcal{W}(U(w)) f(w) dw + \gamma_1 [\bar{c} - \varphi(U(w_{\text{median}}), l(w_{\text{median}}), \bar{c})] f(w_{\text{median}}) \\ &+ \gamma_2 \int_{\underline{w}}^{\overline{w}} [wl(w) - \varphi(U(w), l(w), \bar{c})] f(w) dw \\ &+ q(\underline{w})U(\underline{w}) - q(\overline{w})U(\overline{w}) + \int_{\underline{w}}^{\overline{w}} \left[q(w) h_1(l(w), \bar{c}) \frac{l(w)}{w} + q'(w)U(w)\right] dw. \end{aligned}$$

Assuming the existence of an interior solution and making use of equation (25), the optimality conditions read as:

$$\frac{\partial \mathcal{L}}{\partial U(w)} = \mathcal{W}'(U(w)) f(w) - \gamma_2 f(w) + q'(w) = 0 \tag{39}$$

for $\forall w \in (\underline{w}, w_{\text{median}}) \cup (w_{\text{median}}, \overline{w})$ with the transversality conditions

$$q(\underline{w}) = q(\overline{w}) = 0; \tag{40}$$

and

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial U(w_{\text{median}})} = {} & \mathcal{W}'(U(w_{\text{median}}))f(w_{\text{median}}) - \gamma_1 f(w_{\text{median}}) \\ & - \gamma_2 f(w_{\text{median}}) + q'(w_{\text{median}}) = 0, \end{aligned} \tag{41}$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \bar{c}} = {} & \gamma_1 \left[1 - h_2 \left(l(w_{\text{median}}), \bar{c}\right)\right] f(w_{\text{median}}) - \gamma_2 \int_{\underline{w}}^{\overline{w}} h_2 \left(l(w), \bar{c}\right) f(w) dw \\ & + \int_{\underline{w}}^{\overline{w}} q(w) h_{12} \left(l(w), \bar{c}\right) \frac{l(w)}{w} dw = 0, \end{aligned} \tag{42}$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial l(w)} = {} & \gamma_2 \left[w - h_1 \left(l(w), \bar{c}\right)\right] f(w) \\ & + q(w) \left[h_{11} \left(l(w), \bar{c}\right) \frac{l(w)}{w} + h_1 \left(l(w), \bar{c}\right) \frac{1}{w}\right] = 0 \end{aligned} \tag{43}$$

for $\forall w \in [\underline{w}, w_{\text{median}}) \cup (w_{\text{median}}, \overline{w}]$, and also

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial l(w)} = {} & -\gamma_1 h_1 \left(l(w), \bar{c}\right) f(w) + \gamma_2 \left[w - h_1 \left(l(w), \bar{c}\right)\right] f(w) \\ & + q(w) \left[h_{11} \left(l(w), \bar{c}\right) \frac{l(w)}{w} + h_1 \left(l(w), \bar{c}\right) \frac{1}{w}\right] = 0 \end{aligned} \tag{44}$$

for $w = w_{\text{median}}$.

Making use of equations (43)-(44) gives rise to

$$\frac{\tau(w)}{1 - \tau(w)} = \frac{\gamma_1}{\gamma_2} \cdot \mathbb{I}_{w = w_{\text{median}}} - \frac{q(w)}{\gamma_2} \left[1 + \varepsilon_{h_1, l}(w, \bar{c})\right] \frac{1}{w f(w)}. \tag{45}$$

Applying (40) to (39) and (41), I obtain

$$-q(w) = -\int_{w}^{\overline{w}} \mathcal{W}'(U(t)) f(t) dt + \int_{w}^{\overline{w}} \left(\gamma_1 \cdot \mathbb{I}_{t = w_{\text{median}}} + \gamma_2\right) f(t) dt. \tag{46}$$

Substituting (46) into (45) and rearranging the algebra, the tax formula (17) is thus established. ∎

**Proof of Lemma 2.4.** Since the proof is quite similar to those of Lemmas 4.1 and 2.2, here I just show the following main steps. Setting $k = \underline{w}$ in problem (18) gives

$$\begin{aligned} \max_{y(\cdot)} & \int_{\underline{w}}^{\overline{w}} \left[y(w) - h\left(\frac{y(w)}{w}, y(w_{\text{median}}) - T(y(w_{\text{median}}))\right)\right] f(w) dw \\ & - \int_{\underline{w}}^{\overline{w}} \frac{y(w)}{w^2} h_1 \left(\frac{y(w)}{w}, y(w_{\text{median}}) - T(y(w_{\text{median}}))\right) [1 - F(w)] dw. \end{aligned}$$

The first-order condition with respect to $y(w)$ for $w \neq w_{\text{median}}$ reads as follows

$$\begin{aligned}
&\left[1 - h_1\left(\frac{y(w)}{w}, \bar{c}\right)\frac{1}{w}\right] f(w) \\
-&\left[\frac{1}{w^2}h_1\left(\frac{y(w)}{w}, \bar{c}\right) + \frac{y(w)}{w^3}h_{11}\left(\frac{y(w)}{w}, \bar{c}\right)\right][1 - F(w)] = 0.
\end{aligned} \tag{47}$$

Applying the following equations of MTRs and labor supply elasticity,

$$\begin{aligned}
\tau(w) &= 1 - h_1\left(\frac{y(w)}{w}, \bar{c}\right)\frac{1}{w}, \ \forall w \in [\underline{w}, \overline{w}]; \\
\varepsilon_{h_1,l}(w, \bar{c}) &\equiv \frac{h_{11}(l(w), \bar{c})l(w)}{h_1(l(w), \bar{c})},
\end{aligned} \tag{48}$$

to (47) produces tax formula (20). Moreover, the first-order condition with respect to $y(w)$ for $w = w_{\text{median}}$ reads as follows

$$\begin{aligned}
&\left\{1 - h_1\left(\frac{y(w)}{w}, \bar{c}\right)\frac{1}{w} - h_2\left(\frac{y(w)}{w}, \bar{c}\right)[1 - \tau(w)]\right\} f(w) \\
-&\left[\frac{1}{w^2}h_1\left(\frac{y(w)}{w}, \bar{c}\right) + \frac{y(w)}{w^3}h_{11}\left(\frac{y(w)}{w}, \bar{c}\right)\right][1 - F(w)] \\
-&\frac{y(w)}{w^2}h_{12}\left(\frac{y(w)}{w}, \bar{c}\right)[1 - \tau(w)][1 - F(w)] = 0.
\end{aligned} \tag{49}$$

Applying equation (48) to (49) and rearranging the algebra gives the desired tax formula (21).

Similarly, setting $k = \overline{w}$ in problem (18) gives

$$\begin{aligned}
\max_{y(\cdot)} &\int_{\underline{w}}^{\overline{w}} \left[y(w) - h\left(\frac{y(w)}{w}, y(w_{\text{median}}) - T(y(w_{\text{median}}))\right)\right] f(w)dw \\
&+ \int_{\underline{w}}^{\overline{w}} \frac{y(w)}{w^2}h_1\left(\frac{y(w)}{w}, y(w_{\text{median}}) - T(y(w_{\text{median}}))\right) F(w)dw.
\end{aligned}$$

The first-order condition with respect to $y(w)$ for $w \neq w_{\text{median}}$ reads as follows

$$\begin{aligned}
&\left[1 - h_1\left(\frac{y(w)}{w}, \bar{c}\right)\frac{1}{w}\right] f(w) \\
+&\left[\frac{1}{w^2}h_1\left(\frac{y(w)}{w}, \bar{c}\right) + \frac{y(w)}{w^3}h_{11}\left(\frac{y(w)}{w}, \bar{c}\right)\right] F(w) = 0.
\end{aligned} \tag{50}$$

Applying equation (48) to (50) and rearranging the algebra gives tax formula (19). The first-order condition with respect to $y(w)$ for $w = w_{\text{median}}$ reads as follows

$$\begin{aligned}
&\left\{1 - h_1\left(\frac{y(w)}{w}, \bar{c}\right)\frac{1}{w} - h_2\left(\frac{y(w)}{w}, \bar{c}\right)[1 - \tau(w)]\right\} f(w) \\
+&\left[\frac{1}{w^2}h_1\left(\frac{y(w)}{w}, \bar{c}\right) + \frac{y(w)}{w^3}h_{11}\left(\frac{y(w)}{w}, \bar{c}\right)\right] F(w) \\
+&\frac{y(w)}{w^2}h_{12}\left(\frac{y(w)}{w}, \bar{c}\right)[1 - \tau(w)]F(w) = 0.
\end{aligned} \tag{51}$$

Applying equation (48) to (51) and rearranging the algebra gives the tax formula (23). ■

# References

[1] Jacquet, L. and E. Lehmann [2020]: "Optimal Tax Problems with Multidimensional Heterogeneity: A Mechanism Design Method." Unpublised Manuscript.

[2] Jacquet, L., E. Lehmann and B. V. d. Linden [2013]: "Optimal Redistributive Taxation with both Extensive and Intensive Responses." *Journal of Economic Theory*, 148(5): 1770-1805.

[3] Weymark, J. A. [1987]: "Comparative Static Properties of Optimal Nonlinear Income Taxes." *Econometrica*, 55(5): 1165-1185.