Voting over Selfishly Optimal Tax Schedules: Can Pigouvian Tax Redistribute Income?

Darong Dai*

May 18, 2019

Abstract

This paper studies majority voting over selfishly optimal nonlinear income tax schedules proposed by a continuum of individuals who differ in skill and desire social status. It establishes three main results. First, all skills face the same Pigouvian tax in the social optimum, whereas in the voting equilibrium high skills face a uniformly higher Pigouvian tax than low skills face. Second, under Pareto, Champernowne, Weibull and lognormal skill distributions, the income tax schedule facing high skills tends to be more progressive when status concern of the lowest skilled becomes stronger, and that facing low skills tends to be more progressive when status concern of the highest skilled becomes stronger. Third, the median voter winning pairwise majority voting favors more redistribution than does the benevolent and inequality-averse social planner in the sense that high skills face higher marginal tax rates while low skills face lower ones than in the social optimum.

Keywords: Majority voting; positional externality; nonlinear income taxation; redistributive taxation.

JEL Codes: D62; D72; D82; H21.

1 Introduction

There is growing empirical evidence suggesting the interdependence among individuals in the form of status effects, showing that an individual’s utility depends on his or her absolute consumption (or income) as well as on how it compares to the consumption (or income) of others. Given the efficiency justification of using public policy remedies to attack this sort of positional externalities (e.g., Layard, 1980; Frank, 2008), some questions arise as follows.

How does status-seeking shape the structure of nonlinear income taxation? Does the externality induced by the desire for status leads to a more progressive or less progressive income tax system? How do the externality-correcting and income-redistributing roles of taxation policy interact?

---

*Institute for Advanced Research (IAR), Shanghai University of Finance and Economics, Shanghai 200433, P. R. China. E-mail addresses: dai.darong@mail.shufe.edu.cn; darongdai@hotmail.com

1See, Solnick and Hemenway (1998), Alpizar et al. (2005), Luttmer (2005), Clark et al. (2008), Heffetz and Frank (2011), Mujcic and Frijters (2015), and among others.
These questions have been addressed in the optimal tax framework either inspired by or à la Mirrlees (1971), such as Oswald (1983), Tuomala (1990), Ireland (2001), Aronsson and Johansson-Stenman (2008, 2011, 2013, 2014, 2015), Kanbur and Tuomala (2013), and Dai et al. (2018). Though they have provided some interesting insights, little is currently known about whether the introduction of realistic political-economy considerations would lead towards novel insights.

To address these questions with taking into account political-economic constraints, we incorporate relative consumption concern into the political economy environment considered by Brett and Weymark (2017). We establish first the socially optimal redistributive tax policy as a normative benchmark, and then derive the equilibrium taxation scheme under majority voting over selfishly optimal nonlinear income tax schedules. The following three main results are obtained.

First, all skills face the same Pigouvian tax in the social optimum, whereas high skills face a uniformly higher Pigouvian tax than do low skills in the voting equilibrium. Thus, the political economy consideration is indeed relevant in the sense that the externality-correcting tax generates an income-redistributing effect.

Second, under these well-known skill or ability distributions, namely Pareto, Champernowne (1952), Weibull and lognormal distributions, if status-seeking motive of the lowest skilled becomes stronger, then the income tax schedule facing high skills in the voting equilibrium tends to be more progressive in the sense that marginal tax rates rise faster with income; if status-seeking motive of the highest skilled becomes stronger, then the income tax schedule facing low skills tends to be more progressive as well.

Third, the self-interested voter of median skill level winning the majority voting turns out to favor more redistribution than does the benevolent and inequality-averse social planner in the sense that high skills face strictly higher marginal tax rates while low skills face strictly lower marginal tax rates than they face in the social optimum. The major reason behind this result is that the Pigouvian tax in the social optimum is larger than the Pigouvian tax facing low skills while is smaller than that facing high skills in the voting equilibrium.

This study makes three contributions to the literature. Firstly, the externality-correcting tax policy in the voting equilibrium exhibits a nontrivial redistributive effect, which is a novel departure from the related normative results of Kanbur and Tuomala (2013) and some other papers. Secondly, though Kanbur and Tuomala (2013) show that higher relative consumption concerns increase progressivity of taxation, a novel feature arises in the current voting equilibrium: the source and degree of the impact of status seeking on the progressivity of taxation are in general different between high and low skills. Thirdly, rather than being zero obtained by Brett and Weymark (2017) in a setting without status seeking, the marginal tax rate is positive for both the lowest skilled and the highest skilled, and it is larger for the highest skilled than for the lowest skilled.

The outline of the study is as follows. Section 2 describes the model. Section 3 establishes the social optimum as a benchmark result. Section 4 proceeds to the majority voting over selfishly optimal nonlinear income tax schedules. Section 5 provides some
concluding remarks. Proofs are relegated to Appendix.

2 The Model

The economy is populated by individuals that differ in skill (or labor productivity) denoted by $w$ that is continuously distributed with cumulative distribution function $F(w)$ and density function $f(w) = F'(w) > 0$ over the support $[w, \overline{w}]$, satisfying $0 < w < \overline{w} \leq \infty$. The measure of individuals is normalized to one. Assuming a perfectly competitive labor market, an individual with skill $w$ earns an income of $y = wl$, in which $l \geq 0$ is the measure of labor supply. Following Lehmann et al. (2014) and Brett and Weymark (2016, 2017), the utility function in terms of observable variables, namely before-tax income $y$, after-tax income $c$ (consumption) and average consumption level $\bar{c}$, is of the quasilinear-in-consumption form:

$$u(y, c, \bar{c}; w) = c - h\left(\frac{y}{w}, \bar{c}\right),$$

in which the reference consumption term driven by social comparison is given by

$$\bar{c} \equiv \int_{w}^{\overline{w}} c(w)f(w)dw.$$  

For the preferences specified in (1), we let $h_1$ and $h_2$ denote the partial derivatives with respect to the first and the second arguments, respectively. As usually assumed for the disutility function of labor supply, we have $h_1, h_{11} > 0$, for any positive amount of labor supply. To capture the motivation of status-seeking, we have $h_2 > 0$, also a property that denotes envy or jealousy. In particular, we follow Aronsson and Johansson-Stenman (2013) to let $h_{12} > 0$, a feature inspired by Veblen (1899) that states that leisure has a displaying role in making relative consumption more visible.

The government cannot observe individual types, $(w, l)$, and can only condition transfers on earnings $y$ through an integrable income tax function $T(y)$. By exploiting the Taxation Principle (see, Hammond, 1979; Guesnerie, 1995; Bierbrauer, 2011), here the attention is restricted to the simple direct mechanism. Thus, the bundle allocated to individuals of type $w$ is $(y(w), c(w), \bar{c})$, resulting in the gross utility:

$$U(w) = c(w) - h\left(\frac{y(w)}{w}, \bar{c}\right), \quad \forall w \in [w, \overline{w}].$$

The standard requirement of incentive compatibility calls for that

$$c(w) - h\left(\frac{y(w)}{w}, \bar{c}\right) \geq c(w') - h\left(\frac{y(w')}{w}, \bar{c}\right), \quad \forall w, w' \in [w, \overline{w}],$$

by which we have the first-order necessary condition for truth-telling:

$$U'(w) = h_1\left(\frac{y(w)}{w}, \bar{c}\right)\frac{y(w)}{w^2};$$

---

3As shown by Brett and Weymark (2017), the assumption of quasi-linearity in consumption helps establish the single-peakedness of individual preferences that is crucial for applying the median voter theorem. In addition, as shown by Broadway and Jacquet (2008), the characterization of the optimal tax structure with quasi-linear preferences can be facilitated.

4This follows from the common practice in the literature, such as Dupor and Liu (2003), Abel (2005), Liu and Turnovsky (2005), Clark et al. (2008), Kanbur and Thuomala (2013), and Dai et al. (2018).
as well as the second-order sufficient condition,
\[ y'(w) \geq 0, \] 
(5)
for any \( w \in [\underline{w}, \overline{w}] \).

In addition, the government budget constraint is written as
\[ \int_{\underline{w}}^{\overline{w}} [y(w) - c(w)]f(w)dw \geq 0, \] 
(6)
in which the exogenous revenue requirement is normalized to zero for expositional convenience. This is reasonable as we assume away public good provision and focus on income redistribution.

In case that \( T(y) \) is not differentiable, we, as in Brett and Weymark (2017), define the implicit form of marginal tax rates (MTRs) (or labor wedge) as
\[ \tau(w) = 1 - h_1 \left( \frac{y(w)}{w}, \bar{c} \right) \frac{1}{w}, \forall w \in [\underline{w}, \overline{w}]. \] 
(7)
That is, the MTR is equal to one minus the marginal rate of substitution between before-tax income and after-tax income.

3 The Social Optimum

As a benchmark to refer to, we establish first the socially optimal allocation or the optimum income taxation in the sense of Mirrlees (1971). We adopt a welfarist criterion that sums over all types of individuals a transformation, \( W(U(w)) \), of individuals’ gross utility \( U \), with \( W \) twice-continuously differentiable and \( W''(\cdot) > 0 \). Also, inequality aversion or the preferences for redistribution would be induced by the concavity of \( W \), \( W''(\cdot) \leq 0 \), which encompasses the special case with \( W(U(w)) = U(w) \).

Thus, the benevolent social planner chooses the bundle \( \{y(w), c(w), U(w), \bar{c}\}_{w \in [\underline{w}, \overline{w}]} \) that solves the problem:
\[ \max \int_{\underline{w}}^{\overline{w}} W(U(w))f(w)dw \] 
(8)
subject to constraints (2)-(6).

The characterization of constrained efficiency of the economy is, therefore, obtained from solving problem (8).

**Lemma 3.1** At the second-best optimum, there are Lagrangian multipliers \( \gamma_1, \gamma_2 > 0 \) on the constraints (2) and (6), respectively, and a differentiable function \( q \) such that:

(i) The optimal choice of individual utility levels \( \{U(w)\}_{w \in [\underline{w}, \overline{w}]} \) satisfies
\[ \int_{\underline{w}}^{\overline{w}} W''(U(w))f(w)dw = \gamma_1 + \gamma_2; \]

(ii) The optimal choice of reference consumption level \( \bar{c} \) satisfies
\[ \gamma_1 f(w) = (\gamma_1 + \gamma_2) f(w) h_2 \left( \frac{y(w)}{w}, \bar{c} \right) - q(w) h_{12} \left( \frac{y(w)}{w}, \bar{c} \right) \frac{y(w)}{w^2}; \]
(iii) If there is no bunching at skill $w$:

$$h_1 \left( \frac{y(w)}{w}, \tilde{c} \right) \frac{1}{w} - \frac{\gamma_2}{\gamma_1 + \gamma_2} \left( \frac{q(w)}{(\gamma_1 + \gamma_2)f(w)} \left[ h_{11} \left( \frac{y(w)}{w}, \tilde{c} \right) \frac{y(w)}{w^3} + h_1 \left( \frac{y(w)}{w}, \tilde{c} \right) \frac{1}{w^2} \right] \right);$$

(iv) If there is bunching over $[w_0, w_1] \subseteq [w, \overline{w}]$:

$$\int_{w_0}^{w_1} h_1 \left( \frac{y(w)}{w}, \tilde{c} \right) \frac{f(w)}{w} dw = \frac{\gamma_2[F(w_1) - F(w_0)]}{\gamma_1 + \gamma_2} + \frac{1}{\gamma_1 + \gamma_2} \int_{w_0}^{w_1} q(w) \left[ h_{11} \left( \frac{y(w)}{w}, \tilde{c} \right) \frac{y(w)}{w^3} + h_1 \left( \frac{y(w)}{w}, \tilde{c} \right) \frac{1}{w^2} \right] dw.$$

Proof. See Appendix. ■

Making use of this lemma and (7) gives rise to the following tax formula for socially optimal MTRs:

$$\frac{\tau(w)}{1 - \tau(w)} = \frac{\gamma_1}{\gamma_2} + \left[ 1 + \varepsilon_{h_1,l}(w, \tilde{c}) \right] \frac{1 - F(w)}{w[f(w)] \frac{1}{B(w)}} \frac{(1/\gamma_2) \int_{w}^{\overline{w}} (\gamma_1 + \gamma_2 - \mathcal{W}'(U(t))) f(t) dt}{1 - F(w)} \frac{1}{c(w)},$$

where the elasticity is defined as

$$\varepsilon_{h_1,l}(w, \tilde{c}) \equiv \frac{h_{11}[l(w), \tilde{c}]l(w)}{h_1[l(w), \tilde{c}]}. \quad (10)$$

As in Kanbur and Tuomala (2013), the first term $\gamma_1/\gamma_2$ is a kind of Pigouvian tax correcting for an externality. The Mirrleesian-tax component consists of the efficiency term $\mathcal{A}(w)$ that reflects the elasticity of labor supply on the intensive margin, the skill distribution term $\mathcal{B}(w)$, and the average redistributional gain of imposing a marginal tax rate on skill levels above $w$, denoted by $\mathcal{C}(w)$.

In particular, if the disutility function $h$ is of the following explicit form:

$$h(l, \tilde{c}) = \frac{l^{1+(1/\epsilon)}}{1+(1/\epsilon)} \tilde{c}^\eta \quad \text{for } \epsilon > 0 \text{ and } \eta \in (0, 1), \quad (11)$$

then we have from (10) that $\mathcal{A}(w) \equiv 1 + \varepsilon_{h_1,l}(w, \tilde{c}) = \varepsilon_{h,l} = 1 + (1/\epsilon)$, a constant.

### 4 The Political-Economic Equilibrium

We are in line with Brett and Weymark (2016, 2017) in terms of the underlying political economy: each individual proposes an income tax schedule that is his or her selfishly optimal choice and then pairwise majority voting rule is used to select the one that shall be actually implemented. Naturally, all proposals must be incentive-feasible and meet the government budget constraint (6). Formally, an individual of skill type $k \in [w, \overline{w}]$ determines his selfishly optimal allocation schedule for all types of individuals by solving the problem:

$$\max_{\{c(w), y(w), \tilde{c}\} w \in [w, \overline{w}]} U(k) \text{ subject to constraints } (2) - (6). \quad (12)$$
4.1 The First-Order Approach

We begin our analysis by ignoring the monotonicity constraint (5) in problem (12), generating a relaxed problem for the proposer of type \( k \). Using the technique developed by Brett and Weymark (2017) for proving their Proposition 1, we have:

**Lemma 4.1** For any \( \lambda_k \) by Brett and Weymark (2017) for proving their Proposition 1, we have:

where \( \lambda_k > 0 \) denotes the Lagrangian multiplier on constraint (2).

When \( k = w \), the solution to the problem in Lemma 4.1 is usually called the maxi-min schedule, which we denote by \( \{y_{\min}(w), \bar{c}_{\min}\} \), and when \( k = \bar{w} \), the solution is usually called the maxi-max schedule, which we denote by \( \{y_{\max}(w), \bar{c}_{\max}\} \). Applying the simple differentiation, the first-order conditions of the problem are given as follows:

\[
\lambda_k = \int_{w}^{\bar{w}} \left[ f(w)h_2 \left( \frac{y(w)}{w}, \bar{c} \right) - F(w)h_{12} \left( \frac{y(w)}{w}, \bar{c} \right) \frac{y(w)}{w^2} \right] dw \tag{13}
\]

and

\[
\phi_{\min}(w, y(w), \bar{c}) = 0, \quad \forall w \in [w, k];
\]

\[
\phi_{\max}(w, y(w), \bar{c}) = 0, \quad \forall w \in [k, \bar{w}],
\tag{14}
\]

in which

\[
\phi_{\max}(w, y(w), \bar{c}) = \left[ 1 - \lambda_k - h_1 \left( \frac{y(w)}{w}, \bar{c} \right) \frac{1}{w} \right] f(w)
\]

\[+ h_{11} \left( \frac{y(w)}{w}, \bar{c} \right) \frac{y(w)}{w^3} + h_1 \left( \frac{y(w)}{w}, \bar{c} \right) \frac{1}{w^2} \right] F(w); \tag{15}
\]

and

\[
\phi_{\min}(w, y(w), \bar{c}) = \left[ 1 - \lambda_k - h_1 \left( \frac{y(w)}{w}, \bar{c} \right) \frac{1}{w} \right] f(w)
\]

\[ - h_{11} \left( \frac{y(w)}{w}, \bar{c} \right) \frac{y(w)}{w^3} + h_1 \left( \frac{y(w)}{w}, \bar{c} \right) \frac{1}{w^2} \right] [1 - F(w)]. \tag{16}\]

Using (7) and (13)-(16), the optimal maxi-min MTRs are given by

\[
\tau_{\min}(w) = \frac{\lambda_w + \frac{1 - F(w)}{w f(w)} \left[ 1 + \varepsilon_{h_{11} f}(w, \bar{c}) \right]}{1 + \frac{1 - F(w)}{w f(w)} \left[ 1 + \varepsilon_{h_1 f}(w, \bar{c}) \right]}, \quad \forall w \in [w, \bar{w}] \tag{17}\]
where $\varepsilon_{h_1,l}(w, \bar{c})$ is defined by (10) and
\[
\lambda_w = \int_w^\infty \left\{ f(w)h_2\left(\frac{y(w)}{w}, \bar{c}\right) + [1 - F(w)]h_{12}\left(\frac{y(w)}{w}, \bar{c}\right) \frac{y(w)}{w^2} \right\} dw. \tag{18}
\]
Hence, the MTR is equal to $\lambda_w$ for the highest skilled, and is strictly greater than $\lambda_w$ for all other types under the maxi-min objective. Here $\lambda_w$ is the corresponding optimal Pigouvian tax rate.

Similarly, using (7) and (13)-(16), the optimal maxi-max MTRs are given by
\[
\tau_{\max}(w) = \lambda_{\overline{w}} - \frac{F(w)}{w f(w)} \left[ 1 + \varepsilon_{h_1,l}(w, \bar{c}) \right], \quad \forall w \in [w, \overline{w}] \tag{19}
\]
where
\[
\lambda_{\overline{w}} = \int_{w}^{\overline{w}} \left[ f(w)h_2\left(\frac{y(w)}{w}, \bar{c}\right) - F(w)h_{12}\left(\frac{y(w)}{w}, \bar{c}\right) \frac{y(w)}{w^2} \right] dw. \tag{20}
\]
As a result, the MTR is equal to $\lambda_{\overline{w}}$ for the lowest skilled, and is strictly smaller than $\lambda_{\overline{w}}$ for all other types under the maxi-max objective. Here $\lambda_{\overline{w}}$ is the corresponding optimal Pigouvian tax rate.

Since we see from (18) and (20) that $\lambda_w > \lambda_{\overline{w}}$, it is thus easy to verify from using (17) and (19) that $\tau_{\min}(w) > \tau_{\max}(w)$ for any $w \in [w, \overline{w}]$. The following result is, therefore, obvious.

**Lemma 4.2** For any proposer of type $k \in (w, \overline{w})$, the selfishly optimal schedule of before-tax incomes $y(\cdot)$ for his relaxed problem is given by
\[
y(w) = \begin{cases} 
y_{\max}(w) & \text{for } w \in [w, k); \\
y_{\min}(w) & \text{for } w \in (k, \overline{w}]. 
\end{cases}
\]
There is a downward discontinuity in this schedule at $w = k$.

Expressed in the usual form, we get from (17) that
\[
\frac{\tau_{\min}(w)}{1 - \tau_{\min}(w)} = \frac{\lambda_w}{1 - \lambda_w} + \left[ 1 + \varepsilon_{h_1,l}(w, \bar{c}) \right] \left[ \frac{1 - F(w)}{w f(w)} \right] \left( \frac{1}{1 - \lambda_{\overline{w}}} \right); \tag{21}
\]
and similarly we get from (19) that
\[
\frac{\tau_{\max}(w)}{1 - \tau_{\max}(w)} = \frac{\lambda_{\overline{w}}}{1 - \lambda_{\overline{w}}} - \left[ 1 + \varepsilon_{h_1,l}(w, \bar{c}) \right] \left[ \frac{F(w)}{w f(w)} \right] \left( \frac{1}{1 - \lambda_{\overline{w}}} \right) \tag{22}\]
for any $w \in [w, \overline{w}]$.

Using (21) and (22), the following proposition identifies the impact of introducing status-seeking motive into the political economy considered by Brett and Weymark (2017).

**Proposition 4.1** Under the selfishly optimal tax schedule proposed by a proposer of type $k \in (w, \overline{w})$, the following statements are true:
(i) Status-seeking generates two positive effects on the tax formula for MTRs: an additive effect and a multiplier effect. The higher skills with \( w > k \) face the same effects, denoted by \( \lambda_w/(1-\lambda_w) \) and \( 1/(1-\lambda_w) \), so do the lower skills with \( w < k \), denoted by \( \lambda_w/(1-\lambda_w) \) and \( 1/(1-\lambda_w) \). In particular, both effects are greater for high skills than for low skills.

(ii) Both the highest skilled and the lowest skilled face a strictly positive tax rate, rather than being zero established by Brett and Weymark (2017).

Using the tax formula (9) and Proposition 4.1, we arrive at the following corollary that shows an interesting difference between the social optimum and the selfish optimum in terms of externality-correcting taxation.

**Corollary 4.1** In the socially optimal tax schedule, all skills face the same Pigouvian tax. In contrast, in the selfishly optimal tax schedule, high skills face a higher Pigouvian tax than do low skills.

As maxi-min and maxi-max can be seen as extreme examples of social welfare functions, Corollary 4.1 implies that status effects on optimal tax rates vary with the choice of social welfare function. Importantly, the maxi-min and maxi-max considered here (and hence the choice of social welfare function), rather than being exogenously given, arise endogenously from the fundamental behaviour assumption of individuals pursuing selfish optimum when proposing tax schedules.

Moreover, with these tax formulas (17) and (19), we can calculate as Kanbur and Tuomala (2013) the effect of an increase in status-seeking concern on the degree of progressivity of the MTRs. Indeed, the following two results are obtained.

**Lemma 4.3** Suppose there is a type-independent elasticity of labor supply, then the maxi-min MTRs exhibit the following features:

(i) With a Pareto or Champernowne (1952) distribution, we have:

\[
\frac{\partial^2 \tau_{\min}(w)}{\partial \lambda_w \partial w} > 0 \quad \text{for } \bar{w} < \infty;
\]

\[
\frac{\partial^2 \tau_{\min}(w)}{\partial \lambda_w \partial w} = 0 \quad \text{for } \bar{w} = \infty.
\]

(ii) With a Weibull distribution, we have:

\[
\frac{\partial^2 \tau_{\min}(w)}{\partial \lambda_w \partial w} > 0 \quad \text{for } \bar{w} \leq \infty.
\]

(iii) With a lognormal distribution, we have:

\[
\frac{\partial^2 \tau_{\min}(w)}{\partial \lambda_w \partial w} > 0 \quad \text{for } \bar{w} \leq \infty
\]

if and only if

\[
(ln w - \mu) \exp \left[ \frac{(ln w - \mu)^2}{2\sigma^2} \right] [\Phi(\bar{w}) - \Phi(w)] < 0.8\sigma,
\]

in which \( \mu \) is the mean, \( \sigma^2 \) is the variance and

\[
\Phi(w) \equiv \frac{2}{\sqrt{\pi}} \int_{\frac{ln w - \mu}{\sqrt{2}\sigma}} e^{-t^2} dt \equiv \text{erf}\left( \frac{ln w - \mu}{\sqrt{2}\sigma} \right).
\]
Proof. See Appendix. ■

Hence, under these commonly used skill distributions, the maxi-min income tax schedule tend to become more progressive as a result of an increase in status-seeking concern of the lowest-skilled, measured by $\lambda_w$. As demonstrated by Boadway and Jacquet (2008), provided that the structure of optimal nonlinear income taxation is technically complex, the analysis must be limited to these empirically relevant distributions of skills/abilities so that clear-cut analytical results may be obtained.

**Lemma 4.4** Suppose there is a type-independent elasticity of labor supply, then the maxi-max MTRs exhibit the following features:

(i) With a Pareto or Champernowne distribution, we have:

$$\frac{\partial^2 \tau_{\text{max}}(w)}{\partial \lambda_w \partial w} > 0 \quad \text{for } w \leq \infty.$$

(ii) With a Weibull distribution, we have:

$$\frac{\partial^2 \tau_{\text{max}}(w)}{\partial \lambda_w \partial w} > 0 \quad \text{for } w \leq \infty \text{ and } w \geq \chi,$$

in which $\chi > 0$ is the scale parameter of the density function.

(iii) With a lognormal distribution, we have:

$$\frac{\partial^2 \tau_{\text{max}}(w)}{\partial \lambda_w \partial w} > 0 \quad \text{for } w \leq \infty$$

if and only if

$$(\mu - \ln w) \exp \left[ \frac{(\ln w - \mu)^2}{2\sigma^2} \right] \left[ 1 + \text{erf} \left( \frac{\ln w - \mu}{\sqrt{2}\sigma} \right) \right] < 0.8\sigma,$$

in which $\mu$ is the mean and $\sigma^2$ is the variance.

Proof. See Appendix. ■

Thus, under these commonly used skill distributions, the maxi-max income tax schedule tend to become more progressive as a result of an increase in status-seeking concern of the highest-skilled, measured by $\lambda_w$.

### 4.2 The Complete Solution

We now proceed to solve the problem (12). It is easy to verify that the optimal choice of $\bar{c}$ is independent of the monotonicity constraint (5), and hence is unchanged no matter whether there is bunching or not. We let $y_{\text{min}}^*(w)$ and $y_{\text{max}}^*(w)$ denote the optimal maxi-min and maxi-max income schedules when (5) is taken into account. Moreover, we let $B^\text{min}$ and $B^\text{max}$ denote the sets of types that are bunched with some other type in the complete solution to the maxi-min and maxi-max problems, respectively. Whenever $w$ is bunched, we let $[w_-, w_+]$ denote the set of types bunched with $w$. 

---

9
As in Brett and Weymark (2017), we rewrite the integrands of the problem in Lemma 4.1 as follows. If \( k = \bar{w} \), we have:

\[
G^{\max}(w, y(w), \bar{c}) = \begin{cases}
(1 - \lambda \pi) y(w) - h \left( \frac{y(w)}{w}, \bar{c} \right) f(w) + \frac{y(w)}{w^2} h_1 \left( \frac{y(w)}{w}, \bar{c} \right) F(w) & \text{for } \forall w \notin B^{\max}; \\
(1 - \lambda \pi) y(w) - h \left( \frac{y(w)}{w}, \bar{c} \right) \left[ f(w) + \frac{y(w)}{w^2} h_1 \left( \frac{y(w)}{w}, \bar{c} \right) F(w) \right] + \frac{y(w) - y(w-)}{w^2} h_1 \left( \frac{y(w)}{w}, \bar{c} \right) F(w-) & \text{for } \forall w \in B^{\max}.
\]

Similarly, when \( k = w \), we have:

\[
G^{\min}(w, y(w), \bar{c}) = \begin{cases}
(1 - \lambda w) y(w) - h \left( \frac{y(w)}{w}, \bar{c} \right) f(w) - \frac{y(w)}{w^2} h_1 \left( \frac{y(w)}{w}, \bar{c} \right) \left[ 1 - F(w) \right] & \text{for } \forall w \notin B^{\min}; \\
(1 - \lambda w) y(w) - h \left( \frac{y(w)}{w}, \bar{c} \right) \left[ f(w) - \frac{y(w)}{w^2} h_1 \left( \frac{y(w)}{w}, \bar{c} \right) \left[ 1 - F(w+) \right] \right] & \text{for } \forall w \in B^{\min}.
\]

The interpretation of (23) and (24) is similar to that of Brett and Weymark (2017), and hence is omitted here to economize on the space.

Following the same proof of Propositions 4-5 in Brett and Weymark (2017), we arrive at the following complete solution.

**Lemma 4.5** For any proposer of type \( k \in (w, \bar{w}) \), the selfishly optimal schedule of before-tax incomes \( y^*(\cdot) \) for the problem (12) is given by

\[
y^*(w) = \begin{cases}
y^{\max}(w) & \text{for } w \in (w, \bar{w}), \\
y^{\max}(\bar{w}) & \text{for } w \in [\bar{w}, \hat{w}] \text{ if } \hat{w} > w, \\
y^{\min}(\hat{w}) & \text{for } w \in [\hat{w}, \bar{w}] \text{ if } \hat{w} < \bar{w}, \\
y^{\min}(\bar{w}) & \text{for } w \in (\hat{w}, \bar{w})
\end{cases}
\]

for some \( \hat{w}, \bar{w} \in (w, \bar{w}) \) for which \( \hat{w} < \bar{w} \) and \( k \in [\hat{w}, \bar{w}] \). In particular, the optimal values of the bridge endpoints \( \hat{w} \) and \( \bar{w} \) are determined by the equation

\[
\int_{\hat{w}}^{k} \phi^{\max}(w, y^{\max}(\bar{w}), \bar{c}) \, dw + \int_{k}^{\bar{w}} \phi^{\min}(w, y^{\max}(\bar{w}), \bar{c}) \, dw = 0
\]

if \( \bar{w} > w \) and by the equation

\[
\int_{\bar{w}}^{k} \phi^{\max}(w, y^{\min}(\hat{w}), \bar{c}) \, dw + \int_{k}^{\hat{w}} \phi^{\min}(w, y^{\min}(\hat{w}), \bar{c}) \, dw = 0
\]

if \( \bar{w} < \bar{w} \).

### 4.3 The Voting Equilibrium

Let \((c^*(w, k), y^*(w, k), \bar{c}^*(k))\) denote the optimal allocation assigned to an individual of type \( w \) by type \( k \)'s selfishly optimal tax schedule. The resulting utility of this individual is

\[
U(w, k) = c^*(w, k) - h \left( \frac{y^*(w, k)}{w}, \bar{c}^*(k) \right).
\]

Also, the bridge in type \( k \)'s selfishly optimal income schedule is now denoted by \([\hat{w}(k), \bar{w}(k)]\).

The following result can be established.
Lemma 4.6 The selfishly optimal income tax schedule for the median skill type is a Condorcet winner when pairwise majority voting is restricted to the income tax schedules that are selfishly optimal for some skill type.

Proof. See Appendix. ■

We now get by Lemmas 4.2-4.6 the following result.

**Proposition 4.2** Under these commonly used skill distributions, namely Pareto, Champernowne, Weibull and lognormal distributions, we have in the majority voting equilibrium that:

(i) The income tax schedule facing high skills tends to become more progressive as a result of an increase in status-seeking concern of the lowest-skilled, measured by $\lambda_w$.

(ii) The income tax schedule facing low skills tends to become more progressive as a result of an increase in status-seeking concern of the highest-skilled, measured by $\lambda_{\bar{w}}$.

This proposition thus shows how the desire for social status affects the progressivity of income taxation over the entire skill distribution in a majority voting equilibrium.

Moreover, a comparison of the political-economic equilibrium with the social optimum reveals the following result.

**Proposition 4.3** Suppose $\gamma_1 + \gamma_2 = 1$ and there is a constant elasticity of labor supply, then we have:

(i) $\tau^*(w) > \tau(w)$ for any $w > \hat{w}(w_{\text{median}})$;

(ii) $\tau^*(w) < \tau(w)$ for any $w < \hat{w}(w_{\text{median}})$

in which $\tau^*(w)$ is the MTR in the voting equilibrium, $\tau(w)$ is the MTR in the social optimum, and $[\hat{w}(w_{\text{median}}), \tilde{w}(w_{\text{median}})]$ is the bunching region of the median skill level $w_{\text{median}}$.

Proof. See Appendix. ■

That is, majority voting over selfishly optimal income tax schedules favors more redistribution than does the benevolent and inequality-averse social planner in the sense that high skills face higher MTRs while low skills face lower MTRs than they face in the social optimum. As shown in Corollary 4.1, all skills face the same Pigouvian tax in the social optimum; though the Pigouvian tax is different between high skills and low skills in the voting equilibrium, high skills face the same Pigouvian tax in the voting equilibrium, so do low skills. If there is a constant elasticity of labor supply, then the Pigouvian tax plays a determinant role in the MTRs comparison. In particular, if $h$ is of the usual functional form given by (11), then there is a constant elasticity of labor supply. The major driving force of Proposition 4.3 is thus that the Pigouvian tax in the social optimum is no smaller than the Pigouvian tax facing low skills while is strictly smaller than that facing high skills in the voting equilibrium.
5 Concluding Remarks

The human nature of the desire for status creates negative externalities as one’s gain in the relative position of society is a loss to someone else, à la a zero-sum game. Inspired by the seminal work of Mirrlees (1971) that focuses on the design of income taxation policy to tackle the efficiency-equity tradeoff, some papers have established optimal income tax schemes that also correct for negative positional externalities. The normative approach is definitely useful because it provides the benchmark to which we can refer when there is no such a benevolent dictator, which is more close to what we usually observe in real-world economies. Indeed, in addition to balance efficiency and fairness in resource allocation, a more practical issue concerning redistributive taxation policy lies in its political feasibility. This is why we are interested in analyzing the effect of social comparisons on the income taxation policy in a voting equilibrium.

To the best of our knowledge, this paper represents the first study of majority voting over selfishly optimal nonlinear income tax schedules with status seeking. The underlying assumptions about human nature are reasonable for designing income redistribution policy, namely individuals exhibit self-interestedness as well as jealousy. How would the political institution featured by direct democracy respond to the human nature in terms of resource redistribution? This question may reflect the deeper motivation of this study. For future research, we may naturally investigate the case with the political institution featured by representative democracy in the probabilistic voting model developed by Lindbeck and Weibull (1987).
References


Appendix: Proofs

Proof of Lemma 3.1. As is standard, we equivalently rewrite the monotonicity constraint (5) as \( y'(w) = b(w) \) and \( b(w) \geq 0 \). Following Jacquet et al. (2013), the general Hamiltonian of problem (8) can be expressed as follows:

\[
H = W(U(w)) f(w) + \gamma_1[\bar{c} - c(w)] f(w) + \gamma_2[y(w) - c(w)] f(w) + \mu_1(w) \left( c(w) - h \left( \frac{y(w)}{w}, \bar{c} \right) - U(w) \right) f(w) + \mu_2(w) b(w) + \xi(w) b(w) + q(w) h_1 \left( \frac{y(w)}{w}, \bar{c} \right) \frac{y(w)}{w^2},
\]

in which \( \gamma_1, \gamma_2, \mu_1(w) \) and \( \mu_2(w) \) are nonnegative Lagrangian multipliers, and \( \xi(w) \) and \( q(w) \) are co-state variables. The optimality conditions thus read as:

\[
\begin{align*}
\mathcal{H}_c &= \mu_1(w) f(w) - \gamma_1 f(w) - \gamma_2 f(w) = 0, \\
\mathcal{H}_e &= \gamma_1 f(w) - \mu_1(w) f(w) h_2 \left( \frac{y(w)}{w}, \bar{c} \right) + q(w) h_{12} \left( \frac{y(w)}{w}, \bar{c} \right) \frac{y(w)}{w^2} = 0, \\
\mathcal{H}_b &= \mu_2(w) + \xi(w) = 0,
\end{align*}
\]

\( \dot{q}(w) = -\mathcal{H}_U = -W'(U(w)) f(w) + \mu_1(w) f(w), \quad (26) \)

and

\[
\begin{align*}
\dot{\xi}(w) &= -\mathcal{H}_y = -\gamma_2 f(w) + \mu_1(w) f(w) h_1 \left( \frac{y(w)}{w}, \bar{c} \right) \frac{1}{w} \\
&\quad - q(w) \left[ h_{11} \left( \frac{y(w)}{w}, \bar{c} \right) \frac{y(w)}{w^3} + h_1 \left( \frac{y(w)}{w}, \bar{c} \right) \frac{1}{w^2} \right].
\end{align*}
\]

(27)

Also, the transversality conditions read as:

\[
\xi(w) = \xi(\bar{w}) = 0 \quad \text{and} \quad q(w) = q(\bar{w}) = 0. \quad (28)
\]

Applying the first equation of (26) and (28) to the last equation of (26) immediately gives the result in part (i). Plugging the first equation of (26) in the second equation of (26) and rearranging the algebra, the result in part (ii) is immediate as well. If there is no bunching, then the complementary-slackness condition implies that \( \mu_2(w) = 0 \), by which and the third equation of (26) we get \( \xi(w) = 0 \) everywhere. As a consequence, \( \dot{\xi}(w) = 0 \) everywhere, applying which and the first equation of (26) to (27) gives the desired assertion in part (iii). Finally, if there is bunching over a maximal interval denoted by \([w_0, w_1]\), then there is no bunching at the boundaries \( w_0 \) and \( w_1 \). In consequence, \( \xi(w_0) = \xi(w_1) = 0 \) and \( \int_{w_0}^{w_1} \dot{\xi}(w) dw = 0 \). Applying this and the first equation of (26) to (27) gives the desired assertion in part (iv). \( \blacksquare \)

Proof of Lemma 4.3. If the elasticity of labor supply is type independent, then it follows from (17) that

\[
\frac{\partial^2 r_{\min}(w)}{\partial \lambda w \partial w} = -\frac{1 + \varepsilon_{h_1,l}(\bar{c})}{\left\{ 1 + \frac{1 - F(w)}{w f(w)} (1 + \varepsilon_{h_1,l}(\bar{c})) \right\}^2} \frac{d \left[ \frac{1 - F(w)}{w f(w)} \right]}{dw},
\]

16
which immediately leads to
\[
\frac{\partial^2 \tau_{\min}(w)}{\partial \lambda \partial w} \geq 0 \iff \frac{d\left[\frac{1-F(w)}{w f(w)}\right]}{d w} \leq 0.
\]

The results under Pareto, Weibull and lognormal distributions follow from Boadway and Jacquet (2008), and the result under Champernowne distribution follows from the derivation of Kanbur and Tuomala (2013).

**Proof of Lemma 4.4.** If the elasticity of labor supply is type independent, then it follows from (19) that
\[
\frac{\partial^2 \tau_{\max}(w)}{\partial \lambda \partial w} = \frac{1 + \varepsilon_{h1,l}(\bar{c})}{\left\{1 - \frac{F(w)}{w f(w)}[1 + \varepsilon_{h1,l}(\bar{c})]\right\}^2} \frac{d\left[\frac{F(w)}{w f(w)}\right]}{d w},
\]
which immediately leads to
\[
\frac{\partial^2 \tau_{\max}(w)}{\partial \lambda \partial w} > 0 \iff \frac{d\left[\frac{F(w)}{w f(w)}\right]}{d w} > 0.
\]

Under the Pareto distribution, we have:
\[
\frac{F(w)}{w f(w)} = \frac{w^a}{au^m} - \frac{1}{a}, \quad \text{for } \forall w \geq w_m
\]
in which \(a > 0\) is the Pareto index and \(w_m\) is the modal skill level. Under the Champernowne (1952) distribution, we have
\[
\frac{F(w)}{w f(w)} = \frac{w^\theta}{\theta^m \theta} + \frac{1}{\theta},
\]
in which \(\theta > 0\) is the shape parameter and \(m > 0\) is the scale parameter. The result in part (i) is thus immediate.

Under the Weibull distribution, we have:
\[
\frac{F(w)}{w f(w)} = \frac{\chi^k}{k} - \frac{1}{w^k} \cdot \left[e^{(w/\chi)^k} - 1\right],
\]
in which \(k > 0\) is the shape parameter and \(\chi > 0\) is the scale parameter. Simple differentiation shows that
\[
\frac{d\left[\frac{F(w)}{w f(w)}\right]}{d w} = \frac{\chi^k}{w^{k+1}} + \frac{1}{w} e^{(w/\chi)^k} \left[1 - \frac{\chi^k}{w^k}\right],
\]
by which the result in part (ii) immediately follows.

Finally, under the lognormal distribution, we have:
\[
\frac{F(w)}{w f(w)} = \sigma \left[\int_{\ln w - \mu}^{\ln w - \mu} e^{-t^2/2} dt\right] \exp\left[\frac{(\ln w - \mu)^2}{2\sigma^2}\right],
\]

17
a differentiation of which gives the desired assertion in part (iii) after rearranging the algebra. ■

**Proof of Lemma 4.6.** Following the same proof of the Proposition 6 of Brett and Weymark (2017), we have that the bridge endpoints $\tilde{w}(k)$ and $\hat{w}(k)$ are nondecreasing in $k$ for all $k \in [w, \bar{w}]$. This finding helps to show that $y^*(w, k_1) \leq y^*(w, k_2)$ for all $w, k_1, k_2 \in [w, \bar{w}]$ with $k_1 < k_2$, i.e., any individual’s income is nondecreasing in the proposer’s type. We just need to show that individual preferences are (weakly) single-peaked on the set of skill types, namely $U(w, w) \geq U(w, k_1) \geq U(w, k_2)$ if $w < k_1 < k_2$ and also $U(w, w) \geq U(w, k_1) \geq U(w, k_2)$ if $w > k_1 > k_2$. The first inequalities are obvious under the selfishly optimal criterion. We just show that $U(w, k_1) \geq U(w, k_2)$ if $w < k_1 < k_2$, as the latter one is analogous. By using $U(w, w) \geq U(w, k_1)$ and (4), we have:

$$U(w, k_1) - U(w, k_2) = \int_{w}^{k_1} \left[ h_1 \left( \frac{y^*(t, k_2)}{t}, \int_{w}^{t} y^*(t, k_2)f(t)dt \right) \frac{y^*(t, k_2)}{t^2} \right] dt$$

$$- \int_{w}^{k_1} \left[ h_1 \left( \frac{y^*(t, k_1)}{t}, \int_{w}^{t} y^*(t, k_1)f(t)dt \right) \frac{y^*(t, k_1)}{t^2} \right] dt. \quad (29)$$

Since the differentiation of $h_1(y/t, \int_{w}^{t} f(t)dt)(y/t^2)$ with respect to $y$ is

$$h_{11}(\cdot, \cdot) \frac{y}{t^3} + h_{12}(\cdot, \cdot) \frac{y}{t^2} + h_1(\cdot, \cdot) \frac{1}{t^2} > 0,$$

we thus get by (29) that $U(w, k_1) \geq U(w, k_2)$ if $w < k_1 < k_2$, as desired. Finally, applying the Black (1948)’s Median Voter Theorem completes the proof. ■

**Proof of Proposition 4.3.** If $\gamma_1 + \gamma_2 = 1$, then it follows from the first and last equations of (26) that

$$q(w) = - \int_{w}^{\bar{w}} [1 - W'(U(t))] f(t)dt. \quad (30)$$

Integrating both sides of the equation in part (ii) of Lemma 3.1 over $[w, \bar{w}]$, replacing $q(w)$ by (30), and rearranging the algebra, we have:

$$\gamma_1 = \int_{w}^{\bar{w}} \left\{ f(w)h_2 \left( \frac{y(w)}{w}, \bar{c} \right) + [1 - F(w)]h_{12} \left( \frac{y(w)}{w}, \bar{c} \right) \frac{y(w)}{w^2} \right\} dw$$

$$- \int_{w}^{\bar{w}} \left\{ \left[ \int_{w}^{\bar{w}} W'(U(t))f(t)dt \right] h_{12} \left( \frac{y(w)}{w}, \bar{c} \right) \frac{y(w)}{w^2} \right\} dw$$

by which it is immediate that $\gamma_1 \leq \lambda_w$. By using part (i) of Lemma 3.1, we have

$$\int_{w}^{\bar{w}} W'(U(t))f(t)dt \leq \int_{w}^{\bar{w}} W'(U(t))f(t)dt = 1,$$

by which it is straightforward that $\gamma_1 \geq \lambda_{\bar{w}}$. 

18
Applying $\gamma_1 + \gamma_2 = 1$ to the tax formula (9) results in:

$$\frac{\tau(w)}{1 - \tau(w)} = \frac{\gamma_1}{1 - \gamma_1} + \left[1 + \varepsilon_{h,t}(w, \bar{c})\right] \left[\frac{1 - F(w)}{wf(w)}\right] \left(\frac{1}{1 - \gamma_1}\right) \frac{\int_w^\infty [1 - W'(U(t))] f(t) dt}{1 - F(w)}. \tag{31}$$

Since it is easy to see that

$$0 < \frac{\int_w^\infty [1 - W'(U(t))] f(t) dt}{1 - F(w)} < 1,$$

a comparison of (31) with (21) reveals that

$$\frac{\tau(w)}{1 - \tau(w)} < \frac{\tau_{\text{min}}(w)}{1 - \tau_{\text{min}}(w)}$$

under a constant elasticity of labor supply. Similarly, we can easily verify that $\tau(w) > \tau_{\text{max}}(w)$. Finally, an application of Lemmas 4.5 and 4.6 completes the proof. ■