The influence of income inequality aversion on redistribution in a democratic context[1]

Dooseok Jang

Korea Advanced Institute of Science and Technology (KAIST)
School of Humanities and Social Sciences #1310, 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea

Joel Atkinson

Graduate School of International and Area Studies (GSIAS), Hankuk University of Foreign Studies (HUFS), #815 Faculty Bldg, 107 Imun Rd, Seoul 02450, Republic of Korea

---

Abstract

Post-war declines in income inequality have reversed sharply since the 1970s. Democracy's role in mediating this trend remains controversial, with ambiguity in the empirical findings. In contrast, experiments have shown clearer aversion toward inequality. This paper examines one mechanism that could account in part for this failure of democracy to mitigate income inequality despite a widespread popular preference against it. We investigate redistribution among other-regarding worker-voters who hold non-material as well as material concerns. Workers can adjust their working hours to reduce inequality after voting on a flat tax rate to fund redistribution. Our stylized example shows that as the income distribution skews to the right, a lower tax rate can be more preferred if the degree of inequality aversion goes beyond a certain level. This derives from workers adjusting their working hours in consideration of income inequality and decrease in efficiency of an economy, which is strengthened by a right-skewed skill distribution.

Keywords: inequality aversion, redistribution game, tax, efficiency, labor supplied, income share

[1]We would like to extend a special thanks to Mark Stegeman, Price Fisherback, Stanley S. Reynolds, Derek Lemoine and Martin Dufwenberg for providing helpful feedback.

JEL codes: D91, H23, J20

1. Introduction

Declines in income inequality in developed market economies following World War II have reversed sharply since the 1970s, and in some cases income concentration has returned to levels not seen since the pre-World War I period (Cornia and Kiiski, 2001; Piketty and Saez, 2003, 2006; Atkinson et al., 2017). The reasons for this remain controversial, as does the role of democratic politics in mediating this trend. (Atkinson et al., 2011; Bonica et al., 2013). In theory Meltzer and Richard (1981)'s seminal paper shows that an expansion in voting franchise leads to a preference for an increase in taxes and subsequent decrease in inequality. Acemoglu and Robinson (2006) further specifies Meltzer and Richard (1981)'s model by contrasting preferences between the rich and the poor. In a redistribution model studying the rationale behind elite voters, who are almost always rich voters, agree to expand franchises, redistribution in a broad sense takes place via government spending instead of government transfer which is beneficial to a small range of elite voters (Lizzeri and Persico, 2004). On the other hand, if elite voters use de facto power to maintain their advantages, increase in de jure democracy which is represented by franchise expansion does not significantly reduce inequality (Acemoglu and Robinson, 2008). Moreover, if middle-class voters are the majority in a society that consists of the poor, the middle-class, and the elite-rich, they would vote for the tax policy favorable to themselves, i.e. transferring wealth from the rich class to the middle class (Acemoglu et al., 2015). Thereby, a high tax rate may be less associated with a reduction in inequality.

Some empirical country-level literature supports this finding that an expansion of democracy mitigates income inequality (Lindert, 1994; Justman and Gradstein, 1999; Shen and Yao, 2008). In Acemoglu and Robinson (2000)'s study of western countries during late 19 and early 20 centuries, political elites agreed to expand the voting franchise to mitigate the possibility of revolution, which reduces inequality. Rodrik (1999) argues democracy causes an increase of the labor wage in manufacturing. Others see no or limited correlation between democracy and redistribution (Bollen and Jackman, 1985; Sirowy and Inkeles, 1990; Gradstein and Milanovic, 2004). In Mulligan et al. (2004)'s study of public spendings among 142 countries from 1960 to 1990, political institution is not associated with government spending or a tax structure increasing redistribution.

This ambiguity in empirical findings at the country-level contrasts to some extent with experimental studies, which have generally shown significant aversion toward material output inequality. For endowments received without expending effort, subjects typically reveal other-regarding inequality concerns along with an interest in efficiency and maximizing the lowest income in a distributional setting (Fehr and Schmidt, 1999; Bolteon and Ockenfels, 2000; Tricomi et al., 2010; Charness and Rabin, 2002; Engelmann and Strobel, 2004). Moreover, in a public good experiment in which voters decided whether to contribute their endowments to a real charity, subjects expressed a moral non-material preference (Tyran, 2004; Feddersen et al., 2009). Even in an experiment where subjects are required to carry out a production task in order to receive a material payoff, one fourth of subjects were classified as an egalitarian who significantly considered inequality aversion (Cappelen et al., 2010). One experimental study of Höchtl et al. (2012) found that if the rich are a majority in an economy, a small number of inequality averse rich voters are likely to be pivotal in a majority voting mechanism. On the other hand, if the poor are a majority, a small number of inequality averse poor voters rarely influence the preferred tax level.

Moreover, several empirical findings are in line with the existence of inequality aversion among individuals. People's happiness decreases as inequality increases (Alesina et al., 2004; Schwarze and Härpfer, 2007). By studying 12 countries' survey data, Corneo and Grüner (2002) argue people are willing to sacrifice their self-benefit for a policy supporting public value. In Lambert et al. (2003)'s study finding the existence and the determinants of inequality aversion, inequality aversion is associated with a high level of GDP per capita. The authors provide the following explanation, "modern, mature economies have a stronger sense of ethical principles regarding the relatively poor."

With the possibility that inequality aversion is present among a significant proportion of the population and among pivotal voters are inequality averse, it follows that the influence of inequality averse voters on a preferred tax rate becomes of interest. Part of the answer may be found through investigating redistribution among inequality averse voters who hold non-material as well as material concerns. Some progress has been made in this direction, with Alesina and Angeletos (2005) notably studying the relationship between what workers consider a fair income and their income tax rate preferences. Dhami and Al-Nowaihi (2010) further modelled the influence of income inequality aversion on workers that compare own utility with the utility of others, finding that such a worker contributes the same amount of labor as a worker who does not consider inequality. Moreover, a worker's preference for a high income tax rate increases along with the degree of inequality aversion. However, as real world workers have no information on other's utility as this requires knowledge of the opportunity cost of labor, questions remain about how aversion to income inequality will affect labor supply and tax rate preferences.

In this paper, we extend and refine Meltzer and Richard (1981) and Dhami and Al-Nowaihi's (2010) model, comparing workers with the more realistic assumption of information on own and other's after-tax income. We do not study whether a pivotal voter has inequality aversion preference, but rather focus on if this is a case whether the existence of inequality aversion monotonously increases tax preference. As a main result, workers can adjust their labor as a means to mitigate overall income inequality. A high-skilled worker can decrease the amount of labor supplied to decrease disutility due to differences in his or her after income and that of the relatively poor. In the same way, a low-skilled worker can increase the amount of labor supplied. However, in a stylized example, in the case that the skill distribution is strongly right-skewed, workers can prefer a lower tax rate in accordance with the degree of income inequality.

This is mainly because a high degree of income inequality can lead to a strong adjustment of own labor supplied, which implies a decrease in the marginal tax benefit. This result sheds some light on why democracy can fail to mitigate income inequality. When the degree of inequality aversion is very high, if people expect that inequality is mitigated but efficiency is exacerbated, they may choose a lower tax rate.

The remainder of the paper is organized as follows. Section 2 introduces the formal model for the two-stage distribution game. Section 3 discusses the propositions we assume define worker behavior. Section 4 describes how the degree of income inequality affects the optimal tax rate of the median voter in a stylized example. Section 5 checks robustness of the main result of section 4. Section 6 concludes. All proofs are in the Appendix.

2. Model

This is a two-stage distribution game. Participants vote on a flat tax rate for income in the first stage, and then the government distributes a budget derived from having taxed voters evenly. The flat tax means that the tax ratio is constant across incomes. With the tax rate set, each voter, who is also a worker, then chooses her hours worked in the second stage. A direct voting mechanism is applied, which leads voters to express their optimal tax rates non-strategically, and then a tax rate corresponding to a Condorcet winner, typically a median voter, is chosen. No inefficiency in the process of income redistribution is assumed. Technically, the continuum of voters is distributed in the support $[0, 1]$ based on a c.d.f. function F. For simplicity, assume that F is twice continuously differentiable and that f is the probability density function of F; $f = F'$.

Each voter $i \in [0, 1]$ has a different skill level s_i such that $s_i = i$. The skill level provides a conversion ratio for working hours into income. After a flat tax rate t is chosen, each worker i chooses her working hours l_i simultaneously with all other workers. Each voter has a physical limitation of maximum hours worked that is normalized to 1; thus, $l_i \in [0, 1]$. The pre-tax income of worker

i is $y_i := s_i l_i$. Additionally, no worker possesses any assets or other income. As this is a one-shot game, workers have no savings or chance to build wealth from investment. Therefore, relative rich and poor in this model are determined by hours worked l_i and skill-determined per hour income. Based on the tax revenue, the government then redistributes the aggregated tax evenly to the workers by government transfer and government spending.[2] Subsequently, the consumption level of a worker i is $c_i = (1-t)y_i + g$, where $g = t\bar{y} = t\int_0^1 y_j dF(j)$. The income redistributed by government, g, is limited to the total revenue collected from individual income taxes.

The (opportunity) cost that a worker i necessarily produces to work as much as l_i is represented by the increasing, convex, continuously differentiable function $C(-)$. Also, it is zero for no work; $C(0) = 0$. What subtracts a cost from the consumption level c_i is the self-interest utility denoted by π_i. Therefore, the self-interest utility consists of $\pi_i := c_i - C(l_i)$.[3] Assume $C'(0) = 0$, implying that all workers, $s_i > 0$, can be satiated with leisure and choose $l_i > 0$. The marginal utility for leisure is positive $C'(l) > 0$ as long as $l \in (0, 1]$. Also, assume $C'(1) \geq 1$ to guarantee an interior solution except for $i = 0$ and $i = 1$ workers, meaning that a worker should increase her leisure from zero leisure. The convexity of the cost function is assumed for the uniqueness of the hours worked. For a notation, let us denote $L(x)$ as the inverse function of the marginal cost function. Hence, $L(x)$ exists, is positive-valued, and has a positive derivative.

The utility of a worker i based on inequality aversion is called an inequality-averse utility and denotes the function Ω_i. The utility function of the worker i is quasi-linear between π_i and Ω_i so that it consists of consumption, leisure,

[2]It is also possible to set up a government transfer and a tax imposition progressively by assuming that both are functions of income; however, for simplicity and to focus on the role of inequality aversion, a flat tax and an even redistribution are assumed.

[3]A simple linear utility function for consumption is assumed because of the form of the income inequality aversion function. It is ambiguous whether a worker gains negative utility from unequal incomes or unequal utility from incomes. In practice, this comparison also depends on whether a worker can be aware of the utilities of other workers. Although it is technically common knowledge that everyone knows each other's utility and cost function, it is easily conceivable that this is problematic in practice. In this paper, this ambiguity is set aside by assuming quasi-linearity and an identity utility function of consumption.

and utility for an inequality preference.

$$U_i(l;s,t,F) := \pi_i(l_i;s,t,F) + \Omega_i(l;s,t,F)$$

where $l := (l_j)_{\forall j \in [0,1]}$.

The inequality aversion utility function follows Fehr and Schmidt (1999)'s inequality-aversion function. A player measures an inequality-averse utility as the difference on average between her after-tax income and each player's after-tax income. When her income is worse than that of another player, she experiences *envy*, the degree of which is represented by a non-negative parameter α. When her income is superior to that of another player, she experiences *pity*, the degree of which is represented by a non-negative parameter β. It is assumed that α is greater than or equal to β, meaning that this preference has a loss-aversion flavor because the player suffers more from his disadvantage. Additionally, β is less than 1. If β is greater than 1, a player who has more than another player disposes of her money to make their amounts even, which is assumed to be absurd.[4] Correspondingly, Definition 1 defines the IIA utility.

Definition 1 (Income-inequality aversion; IIA)**.**

$$\Omega_i(l;s,t,F) := -\beta(1-t)\int_0^1 max\{y_i - y_j, 0\}dF(j) - \alpha(1-t)\int_0^1 max\{y_j - y_i, 0\}dF(j)$$

where $\alpha \geq \beta$*, and* $0 \leq \beta < 1$*.*

This model differs from that of Dhami and Al-Nowaihi (2010). Dhami and Al-Nowaihi (2010) integrated Meltzer and Richard (1981) with Fehr and Schmidt (1999)' inequality aversion by comparing a player's self-interest utility, $(1-t)y_i + g - C(l_i)$, with another's, $(1-t)y_j + g - C(l_j)$. However, we compare

[4]For example, say that a player A has more material outcome m_1 than that of another player B m_2. If $\beta = 1$, A's inequality averse utility equals to his utility after A disposes of the difference between A' s and B' s outcomes.

$$U_A(m_1) = m_1 - 1 \cdot (m_1 - m_2) = m_2$$

Therefore, if $\beta \geq 1$, because the player is so painful due to pity led by the difference between his and her outcomes, he can burn his material outcome off to reduce this pain.

a player's after-tax income, $(1-t)y_i$, with another's, $(1-t)y_j$ to extend the ways that inequality aversion is integrated. Technically, it is also justified by the way that a worker compares her consumption with another's consumption to perceive inequality aversion. In practice, the after-tax income is more readily comparable between workers that belong to a group. On the other hand, the self-interest utility of others is rarely observable. Therefore, it is reasonable to assume that workers compare after-tax incomes objectively to consider inequality.

3. Workers' Behavior

3.1. Labor Supply of Workers

Proposition 1 characterizes the unique equilibrium of the labor supply in the second stage. Workers supply labor as long as they can earn more income than those less skilled than themselves.[5]

Proposition 1. *Given t, the unique equilibrium income of IIA Worker i is characterized by*

$$y_i^* = \max_{j \in [0,i]} y^*(s_j)$$

where $y^*(s) := s \cdot l^*(s)$ and $l^*(s) := L[(1-t)s\{1-\beta F(s) + \alpha(1-F(s))\}]$.

The next two paragraphs provide a verbal sketch of the argument in Proposition 1. Let $\delta(s) := 1 - \beta F(s) + \alpha(1 - F(s))$ for a notation. $\delta(s)$ represents the degree of labor supplied adjusted in association with inequality aversion. $\delta(s)$ decreases as the skill level of a worker increases because high-skilled workers are expected to have relatively more low-income workers available in their environment inducing *pity*. Also, for simplicity, let $\Delta := s\delta(s)$

[5] Anonymous reviewers correctly pointed out that workers adjusting their working hour based on the degree of inequality aversion is a strong assumption. However, we assume that inequality aversion is prevalent in society, and as a result, workers' behavior can be influenced by inequality aversion. Also, pity can be alternatively understood as a social pressure instead of an internal preference. This is similar to Fehr and Schmidt (1999) model, that stipulated that a player feels pity if his outcome is greater than another's, based on aggregated experiment results. However, they did not distinguish between an internal and external locus for this. Moreover, if workers refer to a reference point to decide their working hour (Camerer et al., 1997), workers who have a reference point consistent with the average income of the group they identify with, may work as much as the number of hours that others in the group work.

and $d\Delta = \partial\Delta/\partial s = \delta(s) - s(\alpha + \beta)f(s)$. $l^*(s)$ represents the labor-supply maximizing utility under the assumption that $y^*(s)$ strictly increases. However, suppose that there exists an interval such that $y^*(s)$ is decreasing, as illustrated by an interval $[s_{lb}, s_3]$ in Figure 1. This takes place if β and α are sufficiently high given F.[6][7] If F is a uniform distribution, then $y^*(s)$ decreases only for high-skilled workers. However, if F has a spike of high density in a certain skill level, an interval can occur of decreasing $y^*(s)$ for middle-skilled workers. A middle-skilled worker might experience a decrease in $\delta(s_i)$ over skill that is so fast that she has an incentive to decrease her income as the worker index i increases. This interval is illustrated by $[s_{lb}, s_3]$ in Figure 1.

For example, Worker i chooses to earn income $y^*(s_i)$ only if the incomes of workers in $[0, s_i]$ are less than that of Worker s_i and if those of workers in $[s_i, 1]$ are higher. However, Worker i has more workers earning greater incomes in $[s_1, s_i]$ than if $y^*(s_i)$ strictly increases. Also, Worker i has less workers earning greater incomes in $[s_i, s_2]$. Then $\delta(s_i)$ becomes greater because *envy* assumes to be greater that *pity*. It follows that Worker s_i provides more labor, which is shown by Proposition 1. Then all workers in $[s_i, s_{ub}]$ will recursively have at least as much income as $y_{s'}$ because $\delta(s_i)$ increases. These interactions eventually have all workers in $[s_{lb}, s_{ub}]$ at an income of y_b. However, if Worker s_i attempts to earn more than y_b, as depicted by a dot in Figure 1, then many workers in $[s_{lb}, s_{ub}]$ are discontinuously included in the group who earn less than Worker s_i. At the same time, for a worker to optimally earn more than y_b, the skill

[6] The derivative of $y^*(s)$ is as follows.

$$\frac{\partial y^*(s)}{\partial s} = L + L' \cdot s(1-t) \cdot d\Delta$$

In the derivative of $y^*(s)$, every term other than $d\Delta$ is positive if $t < 1$ and $s > 0$. α and β associated with $f(s)$ are necessarily large if there is an interval in which $y^*(s)$ decreases.

[7] Fehr and Schmidt (1999)'s model implements inequality aversion by perceiving unequally distributed outcome. This model does not consider whether people have equal opportunity to perceive inequality. As an anonymous referee indicated, even those who have inequality aversion may take unequally distributed outcome as a fair result if chances to get income are equally distributed. However, a worker who weighs the equality of chance has a totally different preference from inequality aversion preference which weights the equality of outcome in this paper. This seems to be out of the scope of this paper. Therefore, it needs to be investigated in a future study.

level s should exceed s_{ub} so that $\delta(s_i)$ is smaller. However, because s_i is less than s_{ub}, $\delta(s_i)$ must be greater, which is a contradiction. Therefore, when y_b begins to decrease from s_{lb}, all workers in $[s_{lb}, s_{ub}]$ should earn as much as y_b but not beyond y_b, which is proven in Proposition 1. Let us call the interval of skills such as $[s_{lb}, s_{ub}]$ the flat interval. flat intervals are completely determined by $y^*(s)$.

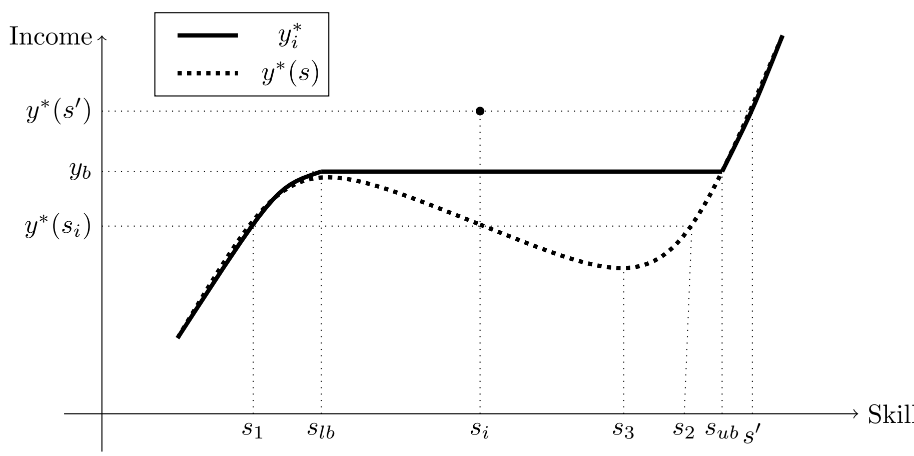

Figure 1: The income curve in a second stage

Note that the labor supply of an IIA worker is a function of her position in the skill distribution along with her skill itself. By doing so, she perceives inequality ex post.

If $y^*(s_i)$ strictly increases, $y^*(s_i)$ fully characterizes an equilibrium.

Corollary 1.1. *Given t, the unique equilibrium of the labor supply is $l_i^* = l^*(s_i)$ if $d\Delta \geq 0$ for all workers.*

The condition for strictly increasing equilibrium illustrates the degree of labor supply adjustment in association with the degree of inequality aversion. Because Δ represents the degree of the adjustment of labor supplied of the IIA worker i, if the degree of inequality aversion is weak, the workers slightly adjust their labor supply such that their ranks based on income are well aligned with their skill (Corollary 1.1). However, if the degree of inequality aversion is high, they adjust their labor supply to the extend that workers remain at levels where they earn similar incomes (Proposition 1).

3.2. Does IIA influence income efficiency and the income share of the richest?

Proposition 2. *The average income increases with an increase in envy and decreases with more pity. Moreover, if a c.d.f. $F(i)$ first-order stochastic dominates over a c.d.f. $G(i)$, the average income associated with $G(i)$ is smaller than that associated with $F(i)$.*

The average income, or economic efficiency, decreases with more *pity* and increases when there is more *envy.* Also, if the skill distribution is more right skewed, the high-skilled workers who take more part of the average income decrease their working hours due to *pity.* Thereby, the average income decreases.

Proposition 3 (The income share of the 1-h richest)**.** *Let $\rho(\alpha, \beta)$ be the income share of the $1-h$ richest:*

$$\rho(\alpha, \beta) := \frac{\int_h^1 y_i^* dF(i)}{\int_0^1 y_j^* dF(j)}$$

For α and β such that $y_i^*(s)$ strictly increases, if $C(l) = \frac{l^2}{2}$ and $F(s)$ is a quadratic function, $\rho(\alpha, \beta)$ decreases as α or β increases.

The income share of the $1-h$ richest decreases as the degree of inequality aversion increases. As α increases, the lower skilled-workers increase their labour worked faster than the higher skilled-workers. Also, as β increases, the higher skilled-workers decrease their labour worked faster than the lower skilled-workers. Proposition 3 supports Piketty and Saez (2003)' discussion that there is a bridge between personal and social norms and the trend in income inequality, saying that "changing social norms regarding inequality and the acceptability of very high wages might partly explain the rise in U.S. top wage shares observed since the 1970s."

4. Voters' Behavior

4.1. The Optimal Tax Rate of the Decisive Voter

In the first stage, Proposition 4 says that the optimal tax rate for each worker strictly decreases as the worker index i. Therefore, the preferred tax rate of the median voter is decisive.

Proposition 4. *The optimal tax rate t_i^* for each voter i exists, is continuous, and strictly decreasing in s_i. As a result, the median voter is decisive.*

4.2. An Example of Income-Inequality Aversion

Hereafter, assume $C(l) = l^2/2$ and $\alpha = \beta = \gamma \in [0,1)$ for simplicity. It also means for a given α we take the highest β to derive a counter-example showing that strong IIA workers do not necessarily prefer a higher tax rate. Denote the value function of the median voter, $U_m(l_m^*)$, in the 1st stage by V_m. If we consider a γ such that $y^*(s)$ strictly increases, the derivative of V_m with regard to the tax rate t is

$$\frac{\partial V_m}{\partial t} = -y^*(s_m) + g' + \Omega \tag{1}$$

where $g' = \dfrac{\partial g}{\partial t}$ and $\Omega = \gamma \left\{ \displaystyle\int_0^m \frac{\partial (1-t) y^*(s_j)}{\partial t} dF(j) - \int_m^1 \frac{\partial (1-t) y^*(s_j)}{\partial t} dF(j) \right\}$.

Three terms represents three different channels to influence the tax preference. The first term $-y^*(s_m)$ represents the marginal tax cost and g' does the marginal tax benefit. Moreover, Ω is the discrepancy between the mean changes in after-tax income of the workers having higher skill than the median voter's skill and that of the workers having lower. According to Lemma 3 in the appendix, in a case that $y^*(s)$ strictly increases as the skill s, $y^*(s_m)$ is unchanged as γ. Also, g' decreases as γ if $t \leq 0.5$ while otherwise g' increases. It follows because increase in γ leads to decrease in income inequality, change in g' across t is reduced as γ increases. Moreover, Ω increases as γ when γ is sufficiently low. Eventually, because γ mitigates the difference of the mean of the after-tax incomes between those having higher skill than the median voter and those having lower, the change in Ω decreases as γ.

As a result, as γ increases, if increase in Ω is bigger than decrease in g' at the low tax rate, the optimal tax rate increases. However, increase in Ω decreases as γ while change in g' is constant (Lemma 3). As Proposition 5 shown the second derivative of $\frac{\partial V_m}{\partial t}$ decreases as γ increases. It implies that the optimal tax rate does not necessarily increase as γ increases.

Proposition 5. *For a given $t < 1$ and a γ in which $y^*(s)$ strictly increases, the second derivative of $\frac{\partial V_m}{\partial t}$ with regard to γ is negative.*

To confirm Proposition 5, we provide an example that the optimal tax rate decreases as γ in which two c.d.f. functions $F(i) = i$ and $G(i) = 2i - i^2$ are compared. When a distribution $F(i)$ is assumed, we explicitly articulate $y^*(s)$ by $y^*(s; F)$. Moreover, so do g' and Ω. The median workers for $F(i)$ and $G(i)$ are $m = 0.5$ and $m = 1 - \sqrt{2}/2 \simeq 0.29$ respectively. Moreover, $y^*(s; F) = (1-t)s^2\{1+\gamma-2\gamma s\}$ and $y^*(s; G) = (1-t)s^2\{1+\gamma-4\gamma s+2\gamma s^2\}$ respectively. $y^*(s; F)$ strictly increases across all skills s if $\gamma < 0.5$ while otherwise there is an interval in which $y^*(s; F)$ decreases. $y^*(s; G)$ strictly increase if $\gamma < 0.8$. Denote the optimal tax rate of the median voter for a distribution $F(s)$ by $t^*_M(\gamma; F)$.

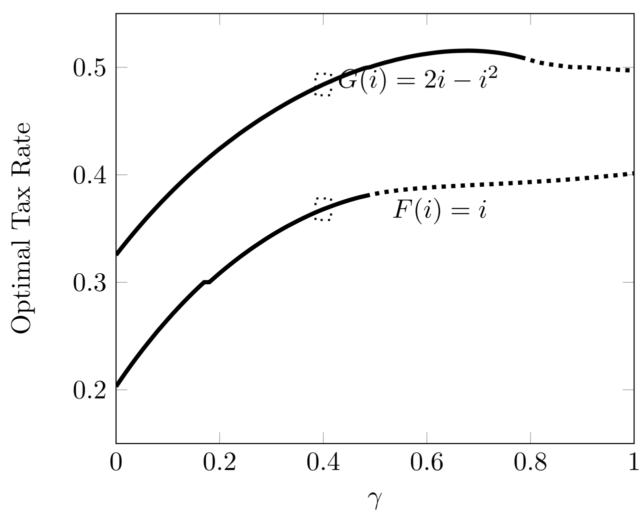

Figure 2: The optimal tax rate of the median voter

Note: The solid line represents that $y^*(s)$ strictly increases at a given γ while $y^*(s)$ weakly increases at a corresponding γ in the dotted line.

Figure 2 illustrates the optimal tax rate across γ and two distributions $F(s)$ and $G(s)$. At the corresponding γ in the solid line, $y^*(s)$ strictly increases as the skill s while at the γ in the dotted line $y^*(s)$ weakly increases. The optimal tax rate is derived from a closed form solution in the interval of $y^*(s)$ increases.

Otherwise because of difficulties of finding the closed form solution due to the interval that $y^*(s)$ does not change, the optimal tax rate is derived from a numerical analysis in which we look for the optimal tax rate which maximizes the value function V_m.[8]

Observation 1. *For all γ, a higher tax rate is preferred when $G(i)$ is assumed than when $F(i)$ is assumed.*

When the distribution is $G(i)$ compared with $F(i)$, decrease in the marginal tax cost, $y^*(s_m)$, outweighs changes in the marginal tax benefit and Ω. For example, compare two points represented by two square boxes in the figure 2, at $\gamma = 0.4$. Given t, $y^*(s_m = 0.5; F) = 0.25(1-t)$ and $y^*(s_m = 0.29; G) = 0.0858(1-t)$. Also, $g'(F) = 0.2667 - 0.5333t$, $g'(G) = 0.1267 - 0.2533t$, $\Omega(F) = 0.14 - 0.14t$, and $\Omega(G) = 0.0784 - 0.0784t$. At the optimal $t_{m=0.5}(F) = 0.37$, $y^*(s_m; F) = 0.1575$ is changed to be $y^*(s_m; G) = 0.0541$ by 0.1034 while $g'(F) + \Omega(F) = 0.1575$ is changed to be $g'(F) + \Omega(F) = 0.0824$ by 0.0751. A higher tax is preferred by the median voter because his income $y^*_{s_m}$ significantly decreases mainly due to the right-skewed distribution.

Observation 2. *A lower tax rate can be increasingly preferred as γ when $G(i)$ is assumed.*

In the example, $\Omega(G)$ is decreased after it peaks, which is consistent with Lemma 3 in the appendix. Moreover, the marginal tax revenue g' decreases as γ. Along with that $y^*(s_M)$ is unchanged, those decreases create an interval in which the optimal tax rate decreases.

This is ironic in that it is contrary to the intuition that when a distortion becomes severe in a way that a highest pity and a more right skewed distribution are assumed, voters are more likely to prefer that the rich be sacrificed. At first, increased concern on who pays more burden covers decrease in efficiency as γ. However, because a higher tax along with preference on income inequality

[8]We mainly used the internal optimization functions of Matlab, such as fminsearch or fminbnd, to find the optimal tax rate from the value function V_m. The numerical analysis also shows the same solution with the closed-form solution in the interval that $y^*(s)$ increases. The Matlab code is available upon request.

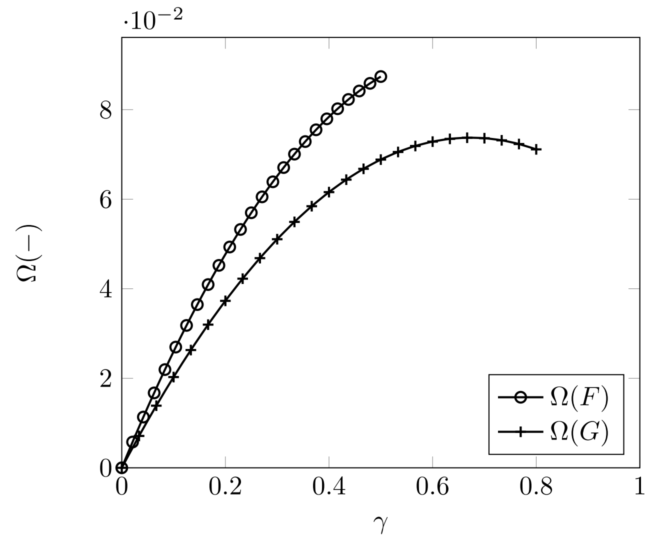

Figure 3: The value of $\Omega(-)$ across $F(i)$ and $G(i)$

Note: The above plot represents $\Omega(-)$ at a given tax rate chosen in the square box of Figure 2. We only represent in the interval that income strictly increases.

mitigates income inequality and decreases efficiency of this economy, voters can increasingly prefer a lower tax rate.

5. Robustness Check

5.1. The optimal tax rate with a progressive tax mechanism

Some audiences would wonder whether the main implication in the example would hold even if a progressive tax mechanism is imposed. We integrate a progressive tax mechanism such that a low tax $t_1 = t$ is imposed if one's income is smaller than an exogenuously given income threshold y_c, and otherwise a high tax $t_2 = \theta t$ is imposed. Accordinly, it assumes $0 \leq t \leq \frac{1}{\theta}$ and $\theta > 1$.

Definition 2. *Let* $Z_1(s) = L[(1-\theta t)s\delta(s)]$, $Z_2(s) = L[(1-t)s\delta(s)]$*. Moreover, for a given* y_c*, define an interval* $\Gamma(y_c) := [c, d]$ *where* $c := Z_2^{-1}(y_c)$ *and* $d := Z_1^{-1}(y_c)$*.*

Provided that $Z_1(s)$ and $Z_2(s)$ strictly increase, then $\Gamma(y_c)$ is well-defined if y_c is appropriately given.

Proposition 6. *If both $Z_1(s)$ and $Z_2(s)$ strictly increase, there is the unique equilibrium strategy of l_i^* such that*

$$y_i^* = \begin{cases} s_i Z_1(s_i) & , if\, i > Z_1^{-1}(y_c) \\ y_c & , if\, i \in \Gamma(y_c) \\ s_i Z_2(s_i) & , if\, i < Z_2^{-1}(y_c) \end{cases}$$

Proposition 7. *There are sufficiently small α and β such that the median voter is decisive.*

Proposition 7 says for a sufficiently small α and β, the median voter is decisive. For bigger α and β, Corollary 8.1 partly covers in the appendix. In the following example, the result that a lower tax rate is preferred as the degree of inequality increases survives in spite of introduction of a progressive tax mechanism.

Let $C(l) = l^2/2$. When $\theta = 1.5$ and $G(i) = 2i - i^2$ are assumed, Figure 4 contrasts the optimal tax rate across $y_c = 1, 0.4$ and 0.3. Therefore, the optimal tax rate is consistent with that of without the progress tax mechanism. For $y_c = 0.4$, when a degree of income inequality is low so that limited part of workers' behavior are changed due to y_c, the median voter prefers a little lower tax rate than at $y_c = 1$ as much. This pattern is also confirmed as y_c is lowered because more workers' behavior is also influenced by this criteria. However, there is an interval of γ such that the lower tax rate is still preferred. Introduction of a progressive tax mechanism does not necessarily remove such an interval of γ.

5.2. Heterogeneous preference toward income inequality

We also consider heterogeneous preference toward income inequality such that the low skill workers, $s \in [0, \bar{s})$, are inequality averse while the high skill workers including the median voter have self-interest preference: $\bar{s} \leq s_m$. We do not consider the case that their preference are varied based on their income mainly because Fehr and Schmidt (1999)'s inequality model assumes that degree of the inequality aversion is extraneously given.

In the 2nd stage, the high skilled workers $s \in [\bar{s}, 1]$ provide their labors as much as $L[(1-t)s]$ as a result of solely considering self-interest while the low

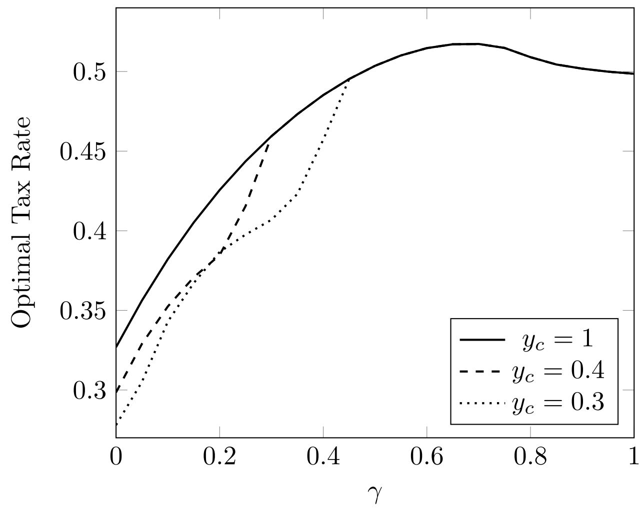

Figure 4: The optimal tax rate across y_c at $G(s) = 2s - s^2$ and $\theta = 1.5$

skilled workers provide their labor following $y^*(s)$. Then, in the first stage, the median voter chooses his optimal tax rate following

$$\frac{\partial V_m}{\partial t} = -y^*(s_m) + g'$$

Proposition 8 shows that a higher tax rate is preferred as γ increases. The rationale follows because the relative lower-skilled workers are influenced by envy more than pity, increase in γ is reflected to the lower-skilled workers to work harder by stimulating their envy. Eventually, g' increases while the median voter's income does not change. However, let the workers of which skill is in $[0, \bar{s})$ also including the median voter are self-interest and otherwise the workers have inequality averse preference. Then, as γ increases, the optimal tax rate of the median voter decreases.

Proposition 8. *Let $C(l) = \frac{l^2}{2}$. If the workers of which skill is in $[0, \bar{s})$ and $\bar{s} \leq s_m$ have inequality averse preference while otherwise the workers only consider self-interest, the optimal tax rate of the median voter increases as γ. However, if the workers of which skill is in $[0, \bar{s})$ and $\bar{s} > s_m$ only consider self-interest while otherwise the workers have inequality averse preference, the optimal tax rate of the median voter decreases as γ.*

If the morality of the rich is associated with inequality aversion (Lambert et al., 2003), even in the case that the median voter is fully self-interest, a degree of inequality aversion might lead to a lower tax preference.

6. Conclusion

This paper illuminates one potential mechanism by which democracy may fail to mitigate income inequality despite a widespread preference against it. When a strong preference for income inequality is prevalent in a society, workers could be expected to adjust their number of hours worked to decrease income inequality. In other words, the inequality averse can adjust their working time following the group to which they can refer, such as their own group. When they have an opportunity to choose a tax level to fund income distribution, they might be reluctant to choose a strong distribution if they expect a severe decrease in societal efficiency.

In line with the model presented here, a way of decreasing income inequality while maintaining high overall social income would be to stimulate envy. If worker' envy toward people who are ahead increases with all other conditions are unchanged, the willingness to work harder to match the income of those who are ahead increases, which eventually increases efficiency. Since asserting envy is the greater influence on relatively lower-skilled workers, overall inequality would be decrease regardless of the tax preference or with a high tax preference. Interestingly, this account is in line with the historical experience of many countries that experienced overall income increases coinciding with relatively high tax rates, and before income inequality had become severe.

Appendix A. The Proofs of Propositions and Lemmas

Lemma 1. *For a continuous function* $k : [0,1] \to [0,1]$*, let* $\mathbb{B}_i = \{j \in [0,1] | k(i) \geq k(j)\}$ *and* $\mathbb{A}_i = \{j \in [0,1] | k(i) \leq k(j)\}$*. Moreover, let* $\gamma(i) := 1 - \beta \int_{\mathbb{B}_i} dF(j) + \alpha \int_{\mathbb{A}_i} dF(j)$*. Then, i)* $\gamma(i) > 0$ *and ii)* $\gamma(i) \geq \gamma(i')$ *for* $i < i'$ *if* $k(i)$ *is increasing.*

Proof. For a continuous function k, $\mathbb{B}_i$ and $\mathbb{A}_i$ are closed sets. Then, $\int_{\mathbb{B}_i} dF(j)$ and $\int_{\mathbb{A}_i} dF(j)$ exist. i) $1 - \beta \int_{\mathbb{B}_i} dF(j) + \alpha \int_{\mathbb{A}_i} dF(j) \geq 1 - \beta \int dF(j) > 0$ because $\beta < 1$ and $\int dF(j) \leq 1$. ii) Assume $k(i)$ is increasing function. Because $k(i') > k(i)$ for $i' > i$, $\mathbb{B}_i \subset \mathbb{B}_{i'}$ and $\mathbb{A}_i \supset \mathbb{A}_{i'}$. Because $f(i)$ is

nonnegative, $\int_{\mathbb{B}_i} dF(j) \leq \int_{i'} dF(j)$ and $\int_{\mathbb{A}_i} dF(j) \geq \int_{\mathbb{A}_{i'}} dF(j)$. Therefore, $\gamma(i) = 1 - \beta \int_{\mathbb{B}_i} dF(j) + \alpha \int_{\mathbb{A}_i} dF(j) \geq 1 - \beta \int_{\mathbb{B}_{i'}} dF(j) + \alpha \int_{\mathbb{A}_{i'}} dF(j) = \gamma(i')$.
□

Lemma 2. *Given t and y_{-i}, let $b_i(l_i) = \{j \in [0,1] | y_j < s_i l_i\}$ and $a_i(l_i) = \{j \in [0,1] | y_j > s_i l_i\}$. Then, the unique optimal labor supply l_i^* necessarily holds $l_i^* = L[(1-t)s_i\{1 - \beta B_i(l_i^*) + \alpha A_i(l_i^*)\}]$ if y_{-i} is strictly monotone and differentiable or disconnected near $s_i l_i^*$ once y_{-i} are relabeled after sorting in ascending order, where $B_i(l_i) = \int_{b_i(l_i)} dF(j)$ and $A_i(l_i) = \int_{a_i(l_i)} dF(j)$.*

Proof. Suppose that y_{-i} is relabeled after sorting in ascending order. Let $\bar{b}_i(l_i) = \sup b_i(l_i)$ and $\underline{a}_i(l_i) = \inf a_i(l_i)$.

$$\begin{aligned} U_i(l_i) =& (1-t)s_i l_i + g - C(l_i) \\ & - \beta(1-t) \int_0^{\bar{b}_i(l_i)} y_i - y_j dF(j) - \alpha(1-t) \int_{\underline{a}_i(l_i)}^1 y_j - y_i dF(j) \end{aligned} \tag{A.1}$$

First, if y_{-i} strictly increases, $s_i l_i = y_{\bar{b}_i(l_i)} = y_{\underline{a}_i(l_i)}$. Suppose that y_j are differentiable near the optimal labor supply of Worker i. Then, $\underline{a}_i(l_i^*)$ and $\bar{b}_i(l_i^*)$ are differentiable near the optimal labor supply. The FOC is as follows.

$$\begin{aligned} FOC =& (1-t)s_i - C' - \beta(1-t) \int_0^{\bar{b}_i(l_i)} s_i dF(j) + \alpha(1-t) \int_{\underline{a}_i(l_i)}^1 s_i dF(j) \\ =& (1-t)s_i[1 - \beta F(\bar{b}_i(l_i)) + \alpha\{1 - F(\underline{a}_i(l_i))\}] - C' \\ & (\because \text{ By Leibniz's rule}) \end{aligned} \tag{A.2}$$

Moreover, the SOC is $-(1-t)s_i\{\beta f(\bar{b}_i(l_i))\bar{b}_i'(l_i) + \alpha f(\underline{a}_i(l_i)\underline{a}_i'(l_i)\} - C'' < 0$ because $\underline{a}_i(-)$ and $\bar{b}_i(-)$ strictly increase and $C(-)$ is a convex function. Therefore, at the optimal labor supply,

$$l_i^* = L[(1-t)s_i\{1 - \beta B_i(l_i^*) + \alpha A_i(l_i^*)\}] \tag{A.3}$$

($\because$ $a_i(l_i)$, $b_i(l_i)$, $A_i(l_i)$, and $B_i(l_i)$ are independent from sorting, $\int_0^{\bar{b}_i(l_i)} dF(j) = B_i(l_i)$, and $\int_{\underline{a}_i(l_i)}^1 dF(j) = A_i(l_i)$.)

Suppose that $y_i^* = s_i l_i^*$ is discontinuous at i. Without a loss of generosity, suppose that $\exists \delta > 0$ such that $\forall j > i$, $y_i^* + \delta < y_j$. I find a l_i^*'s interval that $\bar{b}_i(-)$ and $\underline{a}_i(-)$ are constant in such that $\Delta l_i^* < \frac{\delta}{s_i}$. This means that the shares of workers who are relatively rich or poor are fixed. Therefore, from (A.1), the necessary condition (A.3) is derived.
□

The Proof of Proposition 1

Proof. Denote the equilibrium income and hours worked by y_i^* and l_i^*. In the proof, first we show y_i^* is non-decreasing. Second, we characterize an equilibrium with l_i^* and corresponding y_i^*. Then, we check the uniqueness of the equilibrium.

First, let there is an equilibrium where y_i^* has an interval I_0 in which y_i^* is strictly decreasing. Then, there is i' such that $i' = i + \epsilon$ where $\epsilon > 0$ and $i' \in$ interior (I_0). Hence, $y_i^* > y_{i'}^*$ as well as $\gamma(s_i) \leq \gamma(s_i')$ due to Lemma 1. Note $s_i = i$ and $s_i' = i'$. Furthermore, from the first order condition, $y_i^* = s_i L[(1-t)s_i\gamma(s_i)]$. However, because $s_i < s_i'$, $\gamma(s_i) \leq \gamma(s_i')$, and $L[-]$ is an increasing function, it holds that $y_i^* < y_i'^*$. This contradicts. Therefore, y_i^* weakly increases at an equilibrium.

Second, call the interval of i in which y_i^* is constant a flat interval. For i that is not in a flat interval and its sufficiently close neighborhood, given l_{-i}^*, the optimal incomes of other players, y_{-i}^*, strictly and continuously increase. Due to Lemma 2, the following equality should hold.

$$l_i^* = L[(1-t)s_i\{1 - \beta B_i(l_i^*) + \alpha A_i(l_i^*)\}] \tag{A.4}$$

If $l_i^* = l^*(s_i)$, then the order of $y^*(s_i) > y^*(s_{i'})$ is equivalent to $s_i > s_{i'}$. It means $\bar{b}(l_i^*) = \underline{a}(l_i^*) = s_i$ and then $l_i^* = l^*(s_i)$ in (A.4). Therefore, this is a solution of the equality (A.4).

Let us check its uniqueness. First, if $s_i = 0$ or $t = 1$, $l_i^* = 0$. Suppose that $s_i > 0$ and $t < 1$. Let $h(l_i) := l_i - L[(1-t)s_i\{1 - \beta B_i(l_i) + \alpha A_i(l_i)\}]$. Because B_i and A_i are continuous, $h(l_i)$ is continuous. $h(l_i = 1) = 1 - L[(1-t)s_i\{1 - \beta B_i(1)\}] \geq 1 - L[(1-t)s_i\{1-\beta\}] > 0$. ($\because$ When $t = 0$ and $s_i = 1$, the value x satisfying $C'(x) = 1 - \beta$ is less than 1.) Likewise, $h(l_i = 0) = -L[(1-t)s_i\{1 + \alpha A_i(0)\}] < 0$. The derivative of $h(l_i)$ is as the below.

$$h'(l_i) = 1 - L' \cdot (1-t)s_i\{-\beta B_i'(l_i^*) + \alpha A_i'(l_i^*)\} > 0 \tag{A.5}$$

It holds because $B_i' = \bar{b}'(l_i)F(\bar{b}_i(l_i)) > 0$, $A_i' = -\underline{a}'(l_i)F(\underline{a}(l_i)) < 0$ meaning the equality (A.4) has the unique solution. Therefore, $l^*(s_i)$ is the unique solution of (A.4).

Now, suppose that there is a flat interval. Let s_{lb} and s_{ub} be a minimum skill level and a maximum skill level in the flat interval, respectively. Additionally, let $y_b = y^*(s_{lb}) = y^*(s_{ub})$. Therefore, Worker i in the flat interval provides a labor supply $l_i = y_b/s_i$, which leads to the income of worker i being constant. Note by definition, $y^*(s_i)$ decreases and recovers at the end of the flat interval.

First, given l_{-i}^*, let us check whether Worker i has an incentive to increase her labor supply. Without loss of generosity, let $l_i' = l_i + \epsilon$, where $\epsilon > 0$ and s' satisfying $s_i l_i' = y^*(s')$ and $s' > s_{ub}$. Let us compare the utilities deviating from the strategy.

$$U(l_i) = (1-t)y_b + g - C(l_i) - \beta(1-t)\int_0^{s_{lb}} y_b - y_j dF(j) - \alpha(1-t)\int_{s_{ub}}^1 y_j - y_b dF(j) \tag{A.6}$$

$$\begin{aligned} U(l_i + \epsilon) = &(1-t)(y_b + \epsilon s_i) + g - C(l_i + \epsilon) - \beta(1-t)\int_0^{s'} y_b + \epsilon s_i - y_j dF(j) \\ &- \alpha(1-t)\int_{s'}^1 y_j - y_b - \epsilon s_i dF(j) \end{aligned} \tag{A.7}$$

Hence,

$$U(l_i+\epsilon)-U(l_i) = -\epsilon s_i(1-t) - C(l_i) + C(l_i+\epsilon) + \epsilon s_i(1-t)$$
$$+\epsilon s_i(1-t)\left\{\beta\int_0^{s'} dF(j) - \alpha\int_{s'}^1 dF(j)\right\} + \beta(1-t)\int_{s_{lb}}^{s'} y_b - y_j dF(j) \geq 0 \tag{A.8}$$

After organizing,

$$\frac{C(l_i+\epsilon)-C(l_i)}{\epsilon} \geq (1-t)\left[s_i\left\{1-\beta\int_0^{s'} dF + \alpha\int_{s'}^1 dF(j)\right\} + \frac{\alpha+\beta}{\epsilon}\int_{s_{ub}}^{s'} y_j - y_b dF(j)\right] \tag{A.9}$$

For a small ϵ, the second term in right hand side turns to be zero. The inverse function of y^* exists in the neighbourhood of s' in which y_i^* strictly increases. Let $s' := q(l_i+\epsilon) := (y^*)^{-1}(s_i(l_i+\epsilon))$.

$$\lim_{\epsilon\to 0}(\alpha+\beta)\int_{s_{ub}}^{g(l_i+\epsilon)} \frac{y_j-y_b}{\epsilon} dF(j) = (\alpha+\beta)\lim_{\epsilon\to 0} g'(l_i+\epsilon)[y^*(g(l_i+\epsilon)) - y_b]_{s_{ub}}^{g'(l_i+\epsilon)} = 0 \tag{A.10}$$

The first equality holds due to L'hospital's rule and Leibniz's rule, and the second equality holds because $\lim_{\epsilon\to 0} y^*(g(l_i+\epsilon)) = y_b$. Then, the inequality (A.9) is simplified as the below.

$$C'(l_i) \geq (1-t)s_i\left\{1-\beta\int_0^{s_{ub}} dF(j) + \alpha\int_{s_{ub}}^1 dF(j)\right\} \tag{A.11}$$

Because $s_{ub} < s_i$, $(1-t)s_{ub}\{1-\beta\int_0^{s_{ub}} dF(j) + \alpha\int_{s_{ub}}^1 dF(j)\} > (1-t)s_i\{1-\beta\int_0^{s_{ub}} dF(j) + \alpha\int_{s_{ub}}^1 dF(j)\}$. Therefore, if the following inequality holds and the cost function is convex, the above inequality holds.

$$l^*(s_{lb}) = l_i \geq L\left[(1-t)s_{ub}\left\{1-\beta\int_0^{s_{ub}} dF(j) + \alpha\int_{s_{ub}}^1 dF(j)\right\}\right] = l^*(s_{ub}) \tag{A.12}$$

From the assumption, $y^*(s_{lb}) = y^*(s_{ub})$; because $s_{lb} < s_{ub}$, $l^*(s_{lb}) > l^*(s_{ub})$ should hold. Therefore, Worker i does not have an incentive to deviate.

Third, the equilibrium is unique because the income in the flat interval is constant and the income follows $y^*(s)$ for a non-flat interval along with its continuity.

□

The proof of Proposition 2

Proof. (i) Assume that y_i^* strictly increases with skill and $t < 1$, which leads to $y_i^* = y^*(s_i) = s_i l^*(s_i)$. Although $y^*(s)$ does not strictly increase with regard to skills, $y_i^*(s_i)$ in a flat interval has the same value as $y^*(s)$ that begins the flat interval. Therefore, changes of incomes in parameters other than skills are reflected in measuring efficiency in the same way.

Then, an increase in β reduces the income of Worker i for the other than $s_i = 0$ worker.

$$\frac{\partial y_i^*}{\partial \beta} = -s_i^2 L' \cdot (1-t)F(s_i) \leq 0 \tag{A.13}$$

Also, as α increases, the optimal income increases.

$$\frac{\partial y_i^*}{\partial \alpha} = L' \cdot (1-t)s_i^2(1-F(s_i)) \geq 0 \tag{A.14}$$

Therefore, $\bar{y} := E[y_i^*]$ weakly decreases as *pity* increases and weakly increases as *envy* increases.

(ii) Assume $\forall s \in [0,1]\ F(s) \leq G(s)$. Denote $\delta(s)$ and $y^*(s)$ associated with a c.d.f. $F(s)$ by $\delta(s, F)$ and $y^*(s, F)$ respectively. Then, $y^*(s, F) \geq y^*(s, G)$ holds due to $\delta(s, F) \geq \delta(s, G)$. Therefore, the following inequalities hold.

$$E_F[y^*(s, F)] \geq E_F[y^*(s, G)] \geq E_G[y^*(s, G)] \tag{A.15}$$

The second inequality holds because $y^*(s, G)$ weakly increases.

□

The proof of Proposition 3

Proof. The quadratic density function $F(s)$ is $(1-\eta)s + 2\eta s^2$, where $-1 \leq \eta \leq 1$. Then, the derivatives of $\rho(\alpha, \beta)$ with regard to α and β are

$$\frac{\partial \rho(\alpha, \beta)}{\partial \alpha} = -\frac{60(1-\beta)(1-h)h^3(5+\eta(-1+4h))}{(20+3\beta(-5+\eta)+\alpha(5+3\eta))^2} \leq 0$$

$$\frac{\partial \rho(\alpha, \beta)}{\partial \beta} = -\frac{60(1+\alpha)(1-h)h^3(5+\eta(-1+4h))}{(20+3\beta(-5+\eta)+\alpha(5+3\eta))^2} \leq 0$$

The inequality holds because $5 + \eta(4h-1)$ is positive and β and $h \leq 1$.

□

The Proof of Proposition 4

Proof. i) l_i^* is continuous in t such that y_i^* is continuous. $U_i^*(t, s) = (1-t)s_i l_i^* + G - C(l_i^*) - (1-t)\beta \int_0^{s_i} y_i^* - y_j^* dF - (1-t)\alpha \int_{s_i}^1 y_j^* - y_i^* dF$ Thus, U_i^* is also continuous and the domain of t is $[0, 1]$, which is compact. Therefore, there is a solution with the Weierstrass theorem.

ii) Suppose $y^*(s)$ strictly increases. Let U_i^* be the value function after substituting the optimal labor supply.

$$\begin{aligned}U_i^* =&(1-t)y^*(s_i) + g - C(l^*(s_i)) - \beta(1-t)\int_0^{s_i} y^*(s_i) - y^*(s_j)dF(j)\\&-\alpha(1-t)\int_{s_i}^1 y^*(s_j) - y^*(s_i)dF(j)\end{aligned}\tag{A.16}$$

From Lemma 6, U_i^* is continuously differentiable with regard to t such that its derivative exists.

$$\begin{aligned}\frac{\partial U_i^*}{\partial t} =& -y^*(s_i) + g' + \beta\int_0^{s_i} y^*(s_i) - y^*(s_j)dF(j) + \beta(1-t)\int_0^{s_i} y'(s_j)dF(j)\\&+\alpha\int_{s_i}^1 y^*(s_j) - y^*(s_i)dF(j) - \alpha(1-t)\int_{s_i}^1 y'(s_j)dF(j),\end{aligned}\tag{A.17}$$

where $g' = \frac{\partial g}{\partial t}$ and $y'(s) = \frac{\partial y^*(s)}{\partial t}$. Additionally,

$$\frac{\partial^2 U_i^*}{\partial s_i \partial t} = -\left.\frac{\partial y^*(s)}{\partial s}\right|_{s=s_i}\{1-\beta F(s_i) + \alpha(1-F(s_i))\} + (1-t)(\alpha+\beta)\left.\frac{\partial y^*(s)}{\partial t}\right|_{s=s_i} f(s_i) < 0 \tag{A.18}$$

($\because$ $\left.\frac{\partial y^*(s)}{\partial s}\right|_{s=s_i}$ is positive at $s_i \neq 0$ because of the assumption. $\left.\frac{\partial y^*(s)}{\partial t}\right|_{s=s_i} < 0$ at $s_i \neq 0$ and $t \neq 1$.) Therefore, the optimal tax rate is decreasing in s_i because of the Topkis theorem.

Suppose that there is an interval in which $y^*(s)$ weakly decreases calling a flat interval. For a voter that is in a flat interval, let s_{lb} and s_{ub} represent the minimum and maximum skill levels of the flat interval, respectively. The optimal tax rate of each voter outside the flat interval is continuous and strictly decreases. Let us consider continuity and decrease in the tax preference in the flat interval.

$$\begin{aligned}U_i^* =&(1-t)y_i^* + g - C(l_i^*) - \beta(1-t)\int_0^{s_{lb}} y_i^* - y^*(s_j)dF(j)\\&-\alpha(1-t)\int_{s_{ub}}^1 y^*(s_j) - y_i^* dF(j),\end{aligned}\tag{A.19}$$

where $l_i^* = \frac{s_{lb}l^*(s_{lb})}{s_i}$ and $y_i^* = y^*(s_{lb}) = s_{lb}l^*(s_{lb})$. From Lemma 6, l_i^* and y_i^* are continuously differentiable with regard to t.

$$
\begin{aligned}
\frac{\partial U_i^*}{\partial t} =& -y^*(s_{lb}) + (1-t)\frac{\partial y^*(s_{lb})}{\partial t} + g' - C'\frac{\partial y^*(s_{lb})}{\partial t}\frac{1}{s_i} \\
&+ \beta\int_0^{s_b} y^*(s_{lb}) - y^*(s_j)dF(j) - \beta(1-t)\int_0^{s_b}\frac{\partial y^*(s_{lb})}{\partial t} - \frac{\partial y^*(s_j)}{\partial t}dF(j) \\
&+ \alpha\int_{s_u}^{1} y^*(s_j) - y^*(s_{lb})dF(j) - \alpha(1-t)\int_{s_u}^{1}\frac{\partial y^*(s_j)}{\partial t} - \frac{\partial y^*(s_{lb})}{\partial t}dF(j)
\end{aligned} \tag{A.20}
$$

By taking a derivative with regard to s_i,

$$
\frac{\partial^2 U_i^*}{\partial s_i \partial t} = \left.\frac{\partial y^*(s)}{\partial t}\right|_{s=s_{lb}} \cdot \frac{1}{s_i^2}\left\{\frac{y^*(s_{lb})}{s_i}C'' + C'\right\} < 0 \tag{A.21}
$$

($\because \frac{\partial y^*(s)}{\partial t} < 0$ unless $t = 1$ or $s_i = 0$. C' and C'' are positive from Assumption 2.) Therefore, at least the optimal tax rate t strictly decreases in each flat interval.

Let us consider whether the optimal tax rate is continuous at $s_i = s_{lb}$ without a loss of generosity. $\lim_{s_i \to s_{lb}-} FOC_{(\mathrm{A.17})} = \lim_{s_i \to s_{lb}+} FOC_{(\mathrm{A.20})}$ because $\lim_{s_i \to s_{lb}} \int_{s_i}^1 y^*(s_j) - y^*(s_i)dF(j) = \lim_{s_i \to s_{lb}-} \int_{s_i}^{s_{lb}} y^*(s_{lb}) - y^*(s_j)dF(j) = 0$ and $\lim_{s_i \to s_{lb}+} C'(l_i^*) = C'(L[(1-t)s_{lb}R(s_{lb})]) = (1-t)s_{lb}R(s_{lb})$. The optimal tax rate strictly decreases in skill regardless of whether the skill is in a flat interval. Additionally, the optimal tax rate is continuous at the extreme points in a flat interval. Therefore, the optimal tax rate strictly and continuously decreases with skill.

iii) Based on the inequalities (A.18) and (A.21) associated with its continuity at the extreme points in a flat interval, $\forall s > s'$, $\left.\frac{\partial U_i^*}{\partial t}\right|_{i=s} < \left.\frac{\partial U_i^*}{\partial t}\right|_{i=s'}$ such that U_i^* is a decreasing difference. Therefore, based on Gans and Smart, the median voter is decisive.

□

Lemma 3. *Let $C(l) = l^2/2$. For a given t and γ such that $y^*(s)$ strictly increases,*

i. $y^*(s_m)$ does not change as γ.

ii. g' decreases as γ if $t \leq 0.5$ and otherwise increases. Moreover, $\frac{\partial g'}{\partial \gamma}$ does not change as γ.

iii. Ω is positive and increases at $\gamma = 0$. Moreover, $\frac{\partial \Omega}{\partial \gamma}$ decreases as γ.

Proof. For the part i: because $F(s_m) = 0.5$ that implies $\delta(s_m) = 1$, it turns out to be $y^*(s_m) = s_m L[(1-t)s_m]$.

Consider the part ii:

By integral by parts,

$$\int_0^1 s^2 f(s)ds = 1 - \int_0^1 2sF(s)ds,\ 2\int_0^1 s^2 F(s)f(s)ds = 1 - \int_0^1 2sF^2(s)ds$$

Along with the above terms,

$$\frac{\partial^2 g}{\partial\gamma\partial t} = (1-2t)\int_0^1 s^2(1-2F(s))dF(s) = (1-2t)\int_0^1 2sF(s)[F(s)-1]ds$$

If $t \leq 0.5$, it is negative while otherwise it is positive.

Consider the part iii:

$$\frac{\partial^2(1-t)y^*(s)}{\partial s\partial t} = -2(1-t)\frac{\partial s^2\delta(s)}{\partial s} < 0$$

Inequality holds because $y^*(s)$ strictly increases as s. Therefore, the second derivative of Ω is bigger than the first term. It implies Ω is non-negative.

$$\frac{\partial\Omega}{\partial\gamma} = 2(1-t)\left\{\int_m^1 s^2(1-2F(s))dF(s) - \int_0^m s^2(1-2F(s))dF(s)\right\} > 0$$

Inequality holds because $1-2F(s)$ is positive for s of which is less than the median voter while it is negative. □

Lemma 4. *Both c and d strictly increase as t. Also, if the cost function is a quadratic function and $Z_1(s)$ and $Z_2(s)$ are concave as s, then, $\frac{\partial d}{\partial t} > \frac{\partial c}{\partial t}$.*

Proof. Assume y_i^* strictly increases other than the flat interval. Without loss of generosity, consider the player i of which income is smaller than y_c. Let $c := Z_2^{-1}(y_c)$ and $d := Z_1^{-1}(y_c)$. Both c and d increase as t. Let $T = \theta t$. Then, denote the player satisfying its income is y_c by $x(T)$. Note both $Z_1(s)$ and $Z_2(s)$ are represented by the below term.

$$Z(x,T) := xL[(1-T)x\delta(x)] = y_c$$

Take a drivative with regard to T.

$$Z_x\frac{\partial x}{\partial T} + Z_T = 0 \tag{A.22}$$

where $Z_x := \partial Z(x,T)/\partial x$ and $Z_T := \partial Z(x,T)/\partial T$.

$$\frac{\partial x}{\partial T} = -\frac{Z_T}{Z_x} > 0, \forall x \in (0,1]$$

The last inequality holds because it assumes that y_x^* strictly increases. Consider the second derivative of the equality (A.22) with regard to T.

$$Z_{xx}\left(\frac{\partial x}{\partial T}\right)^2 + 2Z_{xT}\frac{\partial x}{\partial T} + Z_{TT} + Z_x\frac{\partial^2 x}{\partial T^2} = 0$$

Then,

$$\frac{\partial^2 x}{\partial T^2} = -\frac{Z_{xx}\left(\frac{\partial x}{\partial T}\right)^2 + 2Z_{xT}\frac{\partial x}{\partial T} + Z_{TT}}{Z_x}$$

If the cost function is a quadratic function, or $C(l) = \eta_1 l^2 + \eta_2 l + \eta_3$, then $Z(x,T) = \frac{(1-t)x^2\delta(x) - \eta_2 x}{2\eta_1}$ implying $Z_{TT} = 0$ and $Z_{xT} < 0$. Therefore, if $Z(x,T)$ is concave as x, $\frac{\partial^2 x}{\partial T^2}$ is positive. Also, it holds $\frac{\partial d}{\partial t} > \frac{\partial c}{\partial t}$. □

The proof of Proposition 6

Proof. Let $B_1(l_i, y_c) = \sup\{j|s_i l_i > y_j \, and \, y_j \geq y_c\}$, $B_2(l_i, y_c) = \sup\{j|s_i l_i > y_j \, and \, y_j < y_c\}$, $A_1(l_i, y_c) = \inf\{j|s_i l_i < y_j \, and \, y_j \geq y_c\}$, and $A_2(l_i, y_c) = \inf\{j|s_i l_i < y_j \, and \, y_j < y_c\}$. Consider the player i of which income is bigger than y_c. If y_{-i} is continuous almost everywhere after resorting in an ascending order, its utility is

$$U_i = (1-\theta t)s_i l_i + g - C(l_i) - \beta(1-\theta t)\int_{B_2(l_i,y_c)}^{B_1(l_i,y_c)} s_i l_i - y_j dF(j)$$

$$-\beta\int_0^{B_2(l_i,y_c)} (1-\theta t)s_i l_i - (1-t)y_j dF(j)$$

$$-\alpha(1-\theta t)\int_{A(l_i)}^{1} y_j - s_i l_i dF(j) \tag{A.23}$$

If y_{-i} is strictly increasing near $s_i l_i$, it holds $B_1(l_i, y_c) = A(l_i)$. Note $B_2(l_i, y_c)$ is unchanged as l_i. Take a derivative of U_i with regard to l_i.

$$\frac{\partial U_i}{\partial l_i} = (1-\theta t)s_i - C'(l_i) - \beta(1-\theta t)s_i F(B(l_i)) + \alpha(1-\theta t)s_i(1 - F(B(l_i)) = 0$$

Note $B(l_i) = B_1(l_i, y_c)$ and moreover $B_1(-)$, $B_2(-)$, and $A(-)$ are unchanged if y_{-i} is discontinuous near $s_i l_i$.

$$l_i^* = L\left[(1-\theta t)s_i\{1 - \beta F(B(l_i)) + \alpha(1 - F(B(l_i)))\}\right]$$

In the same way, a worker of who income is below than y_c is $l_i^* = L[(1-t)s_i\{1 - \beta F(B(l_i)) + \alpha(1 - F(B(l_i)))\}]$. Because $\{1 - \beta F(B(l_i)) + \alpha(1 - F(B(l_i)))\}$ decreases as $B(l_i)$, at the equilibrium y_i^* weakly increases. Therefore, at the equilibrium, it turns to be $B(l_i) = s_i$ if y_i strictly increases near i.

Let $Z_1(s) = L[(1-\theta t)s\delta(s)]$ and $Z_2 = L[(1-t)s\delta(s)]$. If both Z_1 and Z_2 strictly increase, because $Z_1(s) < Z_2(s)$, there is an unempty interval $\Gamma(y_c) = \{j \in [Z_2^{-1}(y_c), \, Z_1^{-1}(y_c)]\}$. Surely, $l_i^* = Z_1(s_i)$ if $i > Z_2^{-1}(y_c)$ and $l_i^* = Z_2(s_i)$ if $i < Z_1^{-1}(y_c)$.

Consider l_i^* such that $y_i^* = y_c$ if $i \in \Gamma(y_c)$. Without loss of generocity, let $l_i = l_i^* + \epsilon$. First, $Z_1(s_i) < y_c$ in the interval of $i \in \Gamma(y_c)$. Moreover, $U_i(l_i^*) > \lim_{l_i \to l_i^*} U(l_i)$ because the disutility derived from the workers of which income is below y_i^* decreases. Therefore, the player i does not have any incentive to deviate to be $l_i = l_i^* + \epsilon$.

To sum up, if both $Z_1(s)$ and $Z_2(s)$ strictly increase, y_i^* is continuous. The utility (A.23) is defined. Therefore, the below is an equilibrium.

$$l_i^* = \begin{cases} Z_1(s_i) & , if\, i > Z_1^{-1}(y_c) \\ \frac{y_c}{s_i} & , if\, i \in \Gamma(y_c) \\ Z_2(s_i) & , if\, i < Z_2^{-1}(y_c) \end{cases}$$

□

The proof of Proposition 7

Proof. Take a derivative the value function with regard to t. using $\frac{\partial C(l_i^*)}{\partial t} = (1-t)\delta(i)\frac{\partial y_j}{\partial t}$.

$$\frac{\partial U_i(l_i^*)}{\partial t} = -y_i^*\delta(i) + \frac{\partial g}{\partial t} - \beta\int_0^i y_j^* dF(j) + \beta(1-t)\int_0^i \frac{\partial y_j^*}{\partial t} dF(j)$$

$$+\alpha\int_i^c y_j^* dF(j) + \alpha\theta\int_c^1 y_j dF(j) - \alpha\int_i^c (1-t)\frac{\partial y_j^*}{\partial t} dF(j) - \alpha\int_c^1 (1-\theta t)\frac{\partial y_j^*}{\partial t} dF(j)$$

$$-\alpha(1-t)y_i^*\left\{\frac{\partial d}{\partial t} - \frac{\partial c}{\partial t}\right\} + \alpha y_c\left\{(1-\theta t)\frac{\partial d}{\partial t} - (1-t)\frac{\partial c}{\partial t}\right\}$$

where

$$\frac{\partial G}{\partial t} = \int_0^c y_j^* dF(j) + t\int_0^c \frac{\partial y_j^*}{\partial t} dF(j) + \theta(d-c)y_c + ty_c\frac{\partial c}{\partial t}(1-\theta)$$

$$+\theta\int_d^1 y_j^* dF(j) + \theta t\int_d^1 \frac{\partial y_j^*}{\partial t} dF(j).$$

Take a derivative with regard to i.

$$\frac{\partial^2 U_i(l_i^*)}{\partial i \partial t} = -\frac{\partial y_i^*}{\partial i}\delta(i) + y_i^*(\alpha+\beta)f(i) - \beta y_i^* f(i) - \alpha(1-t)\left\{\frac{\partial d}{\partial t} - \frac{\partial c}{\partial t}\right\}\frac{\partial y_i^*}{\partial i}$$

$$+\beta(1-t)\left.\frac{\partial y_j^*}{\partial t}\right|_{j=i} f(i) - \alpha y_i^* f(i) + \alpha(1-t)\left.\frac{\partial y_j^*}{\partial t}\right|_{j=i} f(i)$$

$$= -\frac{\partial y_i^*}{\partial i}\delta(i) + (1-t)(\alpha+\beta)\left.\frac{\partial y_j^*}{\partial t}\right|_{j=i} f(i) - \alpha(1-t)\left\{\frac{\partial d}{\partial t} - \frac{\partial c}{\partial t}\right\}\frac{\partial y_i^*}{\partial i}$$

Note y_i^* strictly decreases as i and y_i^* decreases as t. As α and β converge to zero, the above term turns to be negative because $\frac{\partial Z}{\partial t}$ and Z_x are bounded. In a same way, for $i > c$ worker, its $\frac{\partial U_i^2}{\partial i \partial t}$ is negative.

Consider the i worker such that $i \in \Gamma$. Note y_c is given. Also, c and d is a function of t, but not a function of i. Because $Z_1(i)$ and $Z_2(i)$ strictly increase as i and also decrease as t, c and d increase as t.

$$U_i\left(l_i^* = \frac{y_c}{i}\right) = (1-t)y_c + g - C\left(\frac{y_c}{i}\right) - \beta(1-\theta t)\int_0^c y_c - y_j^* dF(j)$$

$$-\alpha \int_d^1 (1-\theta t)y_j^* - (1-t)y_c dF(j)$$

Take a derivative with regard to i and t.

$$\frac{\partial U_i}{\partial t} = -y_c + \frac{\partial g}{\partial t} + \theta \int_0^c y_c - y_j^* dF(j) + \beta(1-\theta t)\int_0^c \frac{\partial y_j^*}{\partial t} dF(j)$$

$$+\alpha t(1-\theta)y_c - \alpha \int_d^1 -\theta y_j^* + (1-\theta t)\frac{\partial y_j^*}{\partial t} + y_c dF(j) = 0$$

Then, by Topkis' theorem, the optimal tax rate decreases. Therefore, there are sufficiently small α and β such that the median voter is decisive. □

Lemma 5. *If the cost function and the c.d.f are a quadratic function, the x's interval in which $Z(x)$ is concave expands as α and β increase.*

Proof. If $C(l) = \eta_1 l^2 + \eta_2 l + \eta_3$, $Z_{xx} = \frac{1-t}{2\eta_1}\{x^2\delta"(x) + 4x\delta'(x) + 2\delta(x)\}$, where $\eta_1 > 0$ due to its convexity. If $F(x) = \xi_1 x^2 + \xi_2 x + \xi_3$, it satisfies $\xi_3 = 0$, $\xi_1 + \xi_2 = 1$, and $-1 \le \xi_1 \le 1$ because $F(x)$ is a c.d.f. Then, it simplies $f(x) = 2\eta_1 x + 1 - \eta_1$. Then, let $G(x) := -12\eta_1(\alpha+\beta)x^2 + 2(1-\eta_1)(\alpha+\beta)x + 2(1+\eta_1)$. Also, $G(0) \ge 0$. If $\eta_1 \le \frac{1}{7}$, $G(x) > 0$. If $n_1 > \frac{1}{7}$, $G(1)$ gets negative as α and β implying the interval of x in which $G(x)$ is negative expands as α and β. □

Corollary 8.1. *If the cost function is a quadratic function and $Z_1(s)$ and $Z_2(s)$ are concave as s, the median voter is decisive.*

Proof. Trivial along with Lemma 4 and the proof of Proposition 7. □

The proof of Proposition 8

Proof. If $C(l) = l^2/2$, then $y_m = (1-t)s_m^2$ and also

$$g = t\left[\int_0^{\bar{s}} (1-t)s^2\delta(s)dF(s) + \int_{\bar{s}}^1 (1-t)s^2 dF(s)\right]$$
$$= t(1-t)\left[\int_0^{\bar{s}} s^2\delta(s)dF(s) + \int_{\bar{s}}^1 s^2 dF(s)\right] \equiv t(1-t)\bar{y}$$

where $\bar{y}$ is the average income. If the SOC is positive, the optimal tax rate is zero because the utility at $t = 0$ is greater than that at $t = 1$. First,

$$\frac{\partial V_m}{\partial t} = -(1-t)s_m + (1-2t)\bar{y}$$

If $t \geq 0.5$, $\frac{\partial V_m}{\partial t}$ is negative. Therefore $t < 0.5$ at the optimal.

Moreover, $\frac{\partial^2 V_m}{\partial \gamma \partial t}$ fully depends on

$$\frac{\partial^2 g}{\partial \gamma \partial t} = (1-2t)\int_0^{\bar{s}} s^2(1-2F(s))dF(s) > 0$$

Note $\bar{s} < s_m$ so that $1 - 2F(s)$ is positive. Therefore, a higher tax rate is preferred as γ increases.

Consider the second part of the proposition. Assume the workers in $[0, \bar{s})$ are self-interest and $s_m < \bar{s}$ and otherwise the workers have inequality averse preference. Then, the above proof is the same except $\frac{\partial^2 V_m}{\partial \gamma \partial t}$.

$$\frac{\partial^2 g}{\partial \gamma \partial t} = (1-2t)\int_{\bar{s}}^{1} s^2(1-2F(s))dF(s) < 0$$

Therefore, a lower tax rate is preferred as γ increases. □

References

Acemoglu, D., Naidu, S., Restrepo, P. and Robinson, J. A. (2015), 'Democracy, redistribution, and inequality', *Handbook of Income Distribution* **2**, 1885–1966.

Acemoglu, D. and Robinson, J. A. (2000), 'Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective', *The Quarterly Journal of Economics* **115**, 1167–1199.

Acemoglu, D. and Robinson, J. a. (2006), *Economic Origins of Dictatorship and Democracy*, Vol. 136.

Acemoglu, D. and Robinson, J. A. (2008), 'Persistence of power, elites, and institutions', *American Economic Review* **98**(1), 267–293.

Alesina, A. and Angeletos, G.-M. (2005), 'Fairness and Redistribution', *American Economic Review* **95**(4), 960–980.

Alesina, A., Di Tella, R. and MacCulloch, R. (2004), 'Inequality and happiness: Are Europeans and Americans different?', *Journal of Public Economics* **88**(9-10), 2009–2042.

Atkinson, A. B., Hasell, J., Morelli, S. and Roser, M. (2017), *The Chartbook of Economic Inequality*, Institute for New Economic Thinking.
URL: *https://www.chartbookofeconomicinequality.com/*

Atkinson, A. B., Piketty, T. and Saez, E. (2011), 'Top iincome in the long run of history', *Journal of Economic Literature* **49**(1), 3–71.

Bollen, K. a. and Jackman, R. W. (1985), 'Political Democracy and the Size Distribution of Income', *American Sociological Review* **50**(4), 438.

Bolteon, G. E. and Ockenfels, A. (2000), 'ERC: A theory of equity, reciprocity, and competition', *American Economic Review* **90**(1), 166–193.

Bonica, A., McCarty, N., Poole, K. T. and Rosenthal, H. (2013), 'Why hasn't democracy slowed rising Inequality?', *Journal of Economic Perspectives* **27**(3), 103–124.

Camerer, C., Babcock, L., Loewenstein, G. and Thaler, R. (1997), 'Labor Supply of New York City Cabdrivers: One Day at a Time LABOR SUPPLY OF NEW YORK CITY CABDRIVERS: ONE DAY AT A TIME*', *Source: The Quarterly Journal of Economics* **112**(2), 407–441.
URL: *http://www.jstor.org/stable/2951241%5Cnhttp://www.jstor.org/%5Cnhttp://www.jstor.org/action/sh*

Cappelen, A. W., Sørensen, E., Tungodden, B., Sorensen, E. O., Tungodden, B., Sørensen, E. and Tungodden, B. (2010), 'Responsibility for what? Fairness and individual responsibility', *European Economic Review* **54**, 429–441.

Charness, G. and Rabin, M. (2002), 'Understanding Social Preferences with Simple Tests', *The Quarterly Journal of Economics* **117**(3), 817–869.
URL: *https://academic.oup.com/qje/article-lookup/doi/10.1162/003355302760193904*

Corneo, G. and Grüner, H. P. (2002), 'Individual preferences for political redistribution', *Journal of Public Economics* **83**(1), 83–107.

Cornia, G. A. and Kiiski, S. (2001), 'Trends In Income Distribution In The Post-World War Ii Period: Evidence And Interpretation', *Discussion Paper 2001/089 Helsinki: UNU-WIDER* .

Dhami, S. and Al-Nowaihi, A. (2010), 'Existence of a condorcet winner when voters have other-regarding preferences', *Journal of Public Economic Theory* **12**(5), 897–922.

Engelmann, D. and Strobel, M. (2004), 'Inequality aversion, efficiency, and maxmin preferences in simple distribution experiments', *American Economic Review* **94**(4), 857–869.

Feddersen, T., Gailmard, S. and Sandroni, A. (2009), 'Moral Bias in Large Elections: Theory and Experimental Evidence', *American Political Science Review* **103**(02), 175.
URL: *http://www.journals.cambridge.org/abstract_S0003055409090224*

Fehr, E. and Schmidt, K. M. (1999), 'A theory of Fairness, competition, and cooperation', *The Quarterly Journal of Economics* **114**(3), 817–868.

Gradstein, M. and Milanovic, B. (2004), 'Does liberte=egalite? A survey of the empirical links between democracy and inequality with some evidence on the transition economies', *Journal of Economic Surveys* .

Höchtl, W., Sausgruber, R., Tyran, J.-R. R., Hochtl, W., Sausgruber, R., Tyran, J.-R. R., Höchtl, W., Sausgruber, R. and Tyran, J.-R. R. (2012), 'Inequality aversion and voting on redistribution', *European Economic Review* **56**, 1406–1421.

Justman, M. and Gradstein, M. (1999), 'The Industrial Revolution, Political Transition, and the Subsequent Decline in Inequality in 19th-Century Britain', *Explorations in Economic History* **36**(2), 109–127.

Lambert, P. J., Millimet, D. L. and Slottje, D. (2003), 'Inequality aversion and the natural rate of subjective inequality', *Journal of Public Economics* **87**(5-6), 1061–1090.

Lindert, P. (1994), 'The Rise of Social Spending, 1880-1930', *Explorations in Economic History* **31**, 1–37.

Lizzeri, A. and Persico, N. (2004), 'Why did the elites extend the suffrage? Democracy and the scope of government, with an application to Britain's "age of reform"', *Quarterly Journal of Economics* **119**(2), 707–765.

Meltzer, A. H. and Richard, S. F. (1981), 'A rational theory of the size of government', *Journal of Political Economy* **89**(51), 914–927.

Mulligan, C. B., Gil, R. and Sala-i Martin, X. (2004), 'Do Democracies Have Different Public Policies than Nondemocracies?', *Journal of Economic Perspectives* **18**(1), 51–74.
URL: *http://pubs.aeaweb.org/doi/10.1257/089533004773563430*

Piketty, T. and Saez, E. (2003), 'Income Inequality in the United States, 1913-1998', *The Quarterly Journal of Economics* **118**, 1–39.

Piketty, T. and Saez, E. (2006), The evolution of top incomes: A historical and international perspective, *in* 'American Economic Review', Vol. 96, pp. 200–205.

Rodrik, D. (1999), 'Democracies pay higher wages', *Quarterly Journal of Economics* **114**(3), 707–738.

Schwarze, J. and Härpfer, M. (2007), 'Are people inequality averse, and do they prefer redistribution by the state?. Evidence from German longitudinal data on life satisfaction', *Journal of Socio-Economics* **36**(2), 233–249.

Shen, Y. and Yao, Y. (2008), 'Does grassroots democracy reduce income inequality in China?', *Journal of Public Economics* **92**(10-11), 2182–2198.

Sirowy, L. and Inkeles, A. (1990), 'The Effects of Democracy on Economic Growth and Inequality: A review'.

Tricomi, E., Rangel, A., Camerer, C. F. and O'Doherty, J. P. (2010), 'Neural evidence for inequality-averse social preferences', *Nature* **463**(7284), 1089–1091.
URL: *http://www.nature.com/doifinder/10.1038/nature08785*

Tyran, J. R. (2004), 'Voting when money and morals conflict: An experimental test of expressive voting', *Journal of Public Economics* **88**(7-8), 1645–1664.