

Heterogeneous preferences and in-kind redistribution*

Luna Bellani[§]

University of Konstanz, Germany

Francesco Scervini[‡]

University of Turin, Italy

Abstract

This paper examines the impact of social heterogeneity on in-kind redistribution. We contribute to the previous literature in two ways: we consider *i)* the provision of several public goods and *ii)* agents different not only in income, but also in their preferences over the various goods provided by the public sector. In this setting, both the distribution and size of goods provision depend on the heterogeneity of preferences. Our main result is that preference heterogeneity tends to decrease in-kind redistribution, while income inequality tends to increase it. An empirical investigation based on United States Census Bureau data confirms these theoretical findings.

Keywords: Heterogeneous preferences, in-kind redistribution, voting, social distance, local public budgets.

JEL classification codes: D31, D72, H42, H70

*The authors thank Vincent Hildebrand, Dirk Van de gaer, Bertrand Verheyden and participants to the EPCS 2013, APET 2014, 23rd Silvaplane Workshop in Political Economy, IIPF 2014 and to seminars at CEPS/INSTEAD, DEMM - University of Milan and ESOMAS - University of Turin for helpful suggestions and comments to this or earlier versions of the paper. The authors also thank the Società Italiana di Economia Pubblica for having awarded this paper with the SIEP prize at the 26th meeting in September 2014. This research is supported by an AFR grant (PDR 2011-1) from the Luxembourg 'Fonds National de la Recherche' co-funded under the Marie Curie Actions of the European Commission (FP7-COFUND). Usual disclaimers apply.

[§]University of Konstanz, Department of Economics, Box 132, 78457, Konstanz, Germany. E-mail: luna.bellani@uni-konstanz.de

[‡]Università di Torino, ESOMAS, C.so Unione Sovietica, 218 bis, 10134, Torino, Italy. E-mail: francesco.scervini@unito.it

1 Introduction

The amount and the types of public goods provided in modern societies depend on a collective choice attempting to aggregate individual preferences. The preferred level of redistribution is in general considered to vary across individuals due to mere differences in income, the poor desiring more progressive tax schemes than the rich. If one extends the notion of redistribution to the provision of various public goods,¹ preferred allocations may also vary due to differences in preferences. Such differences may be due to age (young couples are more in favor of public education, while the elderly are more in favor of the provision of public health services), but also to heterogeneity in education, ethnicity, employment status, etc. (Alesina and Giuliano, 2010).

The present contribution aims at analyzing public goods provision in the presence of both heterogeneous incomes and preferences. Individual preferences are formed over *i*) private consumption, i.e. net income; *ii*) quantity and *iii*) nature of public goods, whose provision is financed by taxation.

The basic intuition behind our setting is that, although general support for redistribution is higher for the individuals with income lower than the mean, all the individuals, independently of their income levels, are less inclined to support taxation if they anticipate that part of the public budget is going to be spent on goods and services which are not among their preferred ones. If one considers that fractionalization in a society implies a higher heterogeneity in preferences, then the support for taxation and public provision is likely to be lower in societies with higher levels of fractionalization.

In more unequal society, poorer individuals (usually assumed to be the majority of the population) desire a higher degree of redistribution (see Borck (2007) for a survey of models of voting on redistribution). Standard theories on cash redistribution state that support for redistribution decreases with income. If the median voter is able to determine the level of redistribution, then it is increasing with inequality, defined as the distance between mean and median income (Meltzer and Richard, 1981).² When redistribution is made in kind, an alternative result, the Director's law (Stigler, 1970), can be obtained. This result, reprised by Epple and Romano (1996), is that when the middle class benefits from the public provision of goods more than both richer and poorer classes, the rich and the poor may agree on a lower level of in-kind redistribution.

On the one hand, a vast, mostly empirical literature examines the role of income inequality or ethnic and linguistic differences in determining the extent of income redistribution or the provision of public

¹Following most of the literature on this topic, all along the paper we refer to public goods as both pure public goods and publicly provided private goods (see, among others, Alesina et al. (1999)).

²This assuming that the distribution of income is left-skewed, i.e. median income lower than the average income. The opposite would be true otherwise.

goods. The general findings common to those contributions are *i*) a negative correlation across countries, or municipalities, between ethno-linguistic (or/and religious) diversity and indicators of public goods provision and *ii*) a tendency for individuals to increase their support for welfare spending if a larger fraction of welfare recipients in their area belongs to their ethno-linguistic group (Alesina et al., 1999; Luttmer, 2001; Alesina and La Ferrara, 2005; Banerjee et al., 2005; Miguel and Gugerty, 2005; Habyarimana et al., 2007; Desmet et al., 2009; Alesina and Zhuravskaya, 2011).

On the other hand, the theoretical literature has extended the classical framework of majority voting over public goods provision by considering multiple dimensions in either the policy space (e.g. the set of public goods) or the space of voters' characteristics (e.g. income and preferences). These contributions consider either the choice of the quantity of a public good (Alesina and Spalore, 1997) or the "type" of this good (location and/or proportion of the budget to be allocated to a specific good, Bolton and Roland (1997)) or, as in this paper, both aspects (Alesina et al., 1999; Gregorini, 2009). In some settings, voters differ only in their income (Bolton and Roland, 1997), only in their preferences (Alesina et al., 1999) or, as in the present contribution, in both (Fernandez and Levy, 2008; Gregorini, 2009).³ More recently, De Donder et al. (2012) generalize this setting by considering *i*) both the type (i.e. location) and the size of a public good, with *ii*) voters differing both in their income and in their preferences. In particular, under lump sum taxation, they reproduce the benchmark results of Alesina et al. (1999), where individuals are located on a line and vote sequentially, first on the size, and second on the location of the public good. In both contributions, at equilibrium the median location is chosen to place the good and the equilibrium size of the good is determined by the agent whose distance from the equilibrium location is median. With proportional taxation, instead, they obtain a lower public good level than this benchmark, if they allow the correlation between income and location to be positive, although not perfect.

This paper departs from the approach of Alesina et al. (1999) and De Donder et al. (2012) by including several relevant refinements among those suggested by recent theoretical contributions reviewed above: *i*) income heterogeneity; *ii*) a proportional tax system; and *iii*) multiple public goods to be produced at the same time and whose structure does not impose strong assumptions such as an ordering on a uni-dimensional space. Finally, individuals loss functions, measured by the distance between individual preference and the type of goods actually provided by the government, are explicitly modeled allowing to derive a closed-form solution to the model. We show that these two types of heterogeneity affect the total provision of public goods through the level of taxation. At equilibrium, the total provision

³Both cited contribution are not directly comparable to our contribution as they only consider two classes of income (rich and poor) and they also differ in the political process they analyze, considering, instead of majority voting, the formation of parties and the existence of a social planner, respectively.

of public goods depends on the distance between individuals' preferences. In particular, we show that more fragmented societies generate lower public budgets, for any given income distribution, while income inequality tends to increase it, for any given level of fractionalization.

We test these results using United States Census Bureau data. The main advantage of this dataset is that it is possible to generate a panel of 750 observations (50 states over 15 years). Exploiting this panel structure with state fixed effect allows us to analyze units with a high level of heterogeneity, but within the same federal constitution. In order to measure the provision of public goods, we compute net expenditures of every state. The heterogeneity in the population is captured by distance in individuals' socio-economic characteristics and by income inequality measures.⁴ With these variables, we are able to test our main theoretical predictions. We provide several robustness checks of our results by considering different assumptions on the timing of the political process and on the indices used to measure inequality and social distance, and by addressing reverse causality and selective migration issues.

The remainder of this paper is organized as follows. We introduce the basic theoretical setting of our model of public good provision in a heterogeneous society in the next section and the main theoretical results are derived in Section 3. Section 4 presents the dataset and variables used in our empirical exercise. In Section 5 the methods and the results of the empirical part are presented. Section 6 concludes.

2 Basic Setting

Consider a population of individuals of measure 1. This population is perfectly partitioned in J groups $j = 1 \dots J$, with ν_j the proportion of individual belonging to group j . Every group is characterized by a type θ_j , which represents a certain preference over the types of public goods to be provided. The distribution of incomes y in group j is denoted by F_j , with density f_j over $[\underline{y}_j, \bar{y}_j]$. Individuals know their type and income, and know the distribution of these characteristics in the society.

A set of K goods and services are provided by the government, such as public schools, parks, homeland security, bicycle paths, health-care facilities, infrastructures and so on. In order to finance the provision of these K goods, the government levies taxes on individuals' incomes at a constant tax rate t .⁵ Total tax revenue is $G = t\bar{y}$, where \bar{y} is the mean income:

$$\bar{y} = \sum_{j=1}^J \nu_j \int_{\underline{y}_j}^{\bar{y}_j} y f_j(y) dy. \quad (1)$$

⁴Refer to Section 4 for a description of the variables and the methods used to proxy individuals' heterogeneity.

⁵In absence of flat cash transfers, it gives rise to a proportional tax system.

The production function is the same for all public goods: $g_k = q_k$, where g_k is the amount of k -th public good produced and q_k is public expenditure for the production of the k -th public good.

Government budget constraint must be balanced at every period, requiring equalization of tax revenues and public expenditures, taking the following form:

$$G = t\bar{y} = \sum_{k=1}^K q_k. \quad (2)$$

In order to distinguish the size of the budget (G) and its allocation across goods, let us call s_k the share of total budget that is used to produce the k -th public good, then we have:

$$g_k = q_k = s_k G, \quad (3)$$

where $\sum_{k=1}^K s_k = 1$.

If we call \mathbf{q} , \mathbf{g} , \mathbf{s} the k -dimensional vectors of q_k , g_k , s_k respectively, then we can restate eq.(2) as:

$$G = t\bar{y} = \mathbf{i}'\mathbf{q} \quad (4)$$

and eq.(3) as:

$$\mathbf{q} = \mathbf{g} = \mathbf{s}G. \quad (5)$$

Individuals' utility depends positively on private consumption and public goods. With respect to the latter, we assume that utility depends both on the quantity of public goods and on the distance between the quantity of the goods preferred by the individuals and those actually produced. As for the quantity, the axiom of non-satiation holds in our utility function: holding constant the types of public goods produced, the more public goods, the better off the individual is. Considering the types of public goods, we assume that, for any public budget, the more similar the produced goods to the individual preferences, the higher the individual's utility. We allow these two terms, quantity and types of public goods, to interact in the utility function. Summarizing:

$$u_{ij} = c_i + V(G, L(d_j(\mathbf{s}_j, \mathbf{s}^*))), \quad (6)$$

where u_{ij} represents utility of an individual i belonging to group j , c_i is his/her private consumption and V the utility associated to the provision of a quantity G of public goods, distributed according to the shares \mathbf{s}^* (and therefore $\mathbf{g}^* = \mathbf{s}^*G$ are the public goods actually provided). $\mathbf{s}_j = \{s_{1j}, \dots, s_{Kj}\}$ is the vector of the composition preferred by the group j , and $d_j(\mathbf{s}_j, \mathbf{s}^*)$ is a scalar representing the distance

⁶This budget constraint implicitly states that there is no cash redistribution *per se*, although it could be feasible to consider it as one of the K goods.

between the two vectors of composition and $L(\cdot)$ is a generic loss function associated to this distance. We assign the simplest concept of distance, that is the Euclidean distance, to the function $d_j(\mathbf{s}_j, \mathbf{s}^*)$:

$$d_j(\mathbf{s}_j, \mathbf{s}^*) = \|\mathbf{s}_j - \mathbf{s}^*\| = \sqrt{\sum_{k=1}^K (s_{jk} - s_k^*)^2}. \quad (7)$$

The loss function associated to the distance in eq.(7) represents the reduction of utility associated to the realization of a policy \mathbf{s}^* different from the preferred one, \mathbf{s}_j . In order to keep the model tractable, we need to assume an explicit loss function, which is defined as the square of the normalized Euclidean distance,⁷ as follows:

$$L_j = \left(\frac{d_j(\mathbf{s}_j, \mathbf{s}^*)}{\max d_j(\mathbf{s}_j, \mathbf{s}^*)} \right)^2 = \left(\frac{d_j(\mathbf{s}_j, \mathbf{s}^*)}{\sqrt{2}} \right)^2 = \sum_{k=1}^K \frac{(s_{jk} - s_k^*)^2}{2}. \quad (8)$$

This quadratic functional form has been frequently used as a loss function (see for instance Davis et al. (1970) on mathematics for electoral processes) and has appealing properties.⁸ Its economic intuition is the following. The loss function reaches its minimum (in zero) when the preferred and the produced bundles coincide. In contrast, it reaches its maximum (in one) for a group whose preferred bundle is composed of a single good, whereas the actual bundle is composed of a unique, but different, good.

Furthermore, let us assume that $V(G, L_j)$, the utility associated to size and composition of public expenditure (eq.6) is the following:

$$V(G, L_j) = G^{(1-\gamma)} (1 - L_j)^\gamma, \quad (9)$$

where the parameter $\gamma \in (0, 1)$ can be interpreted as the sensitivity to the difference in preferences. When $\gamma \rightarrow 1$ individuals do not care about the quantity of public goods, but only to their relative shares. Opposite, when $\gamma \rightarrow 0$, they only care about the amount of public goods. Given the utility function in eq.(6), $\gamma = 0$ is exactly the textbook case of a pure cash redistribution. In this case, the preferred tax rate is 1 for individuals poorer than the mean and 0 for individuals richer than the mean.

If we replace private consumption with net income and write the government budget in terms of total revenues, we can re-write utility in terms of the tax rate as:

$$u_{ij} = (1 - t) y_i + (t\bar{y})^{(1-\gamma)} (1 - L_j)^\gamma. \quad (10)$$

⁷It is convenient in this context to bound the distance between 0 and 1. Therefore, we normalize the distance for its maximum, that can be shown to be: $\max d_j(\mathbf{s}_j, \mathbf{s}^*) = \sqrt{2}$ (see appendix A.1 for the formal proof).

⁸ The advantage of the quadratic form consists in being smoother and differentiable in the whole domain. With respect to the normalized Euclidean distance, its squared form gives proportionally greater weight on greater distances. However, since it is a monotonic transformation, it does not alter the ranking of individual losses, so that there is no qualitative difference in using one or the other.

Regarding the political behavior, in this paper we focus on a representative democracy where we assume that all individuals (voters and politicians) act rationally. In particular, the voters seek to maximize their utility function given a combination of quantity and composition of public goods, and we assume politicians to be purely self-interested and to care only about winning elections (“office-seeking”). With respect to the process of policy choices, we study a simple electoral competition, in which two candidates/parties propose policies to maximize their chances of winning the elections and once elected they implement their preannounced policies. Voters thus evaluate the policy platforms and cast their vote for their preferred candidate. In this paper we are interested to study the effects of *heterogeneous preferences* on the provision of public goods, and not the effects of the *political framework* on this provision.⁹

3 Equilibrium Tax Rate

3.1 Political Setting

In this paper we focus on a setting in which individuals vote over the level of taxation without knowing which types of public goods will be provided. This uncertainty may be motivated by the fact that the decisions on the allocation of the budget over the different goods are often made in a second step, at a different level, e.g. county or municipality. The hypothesis that decisions over tax rates and in-kind redistribution are partly taken at different levels of government (and thus with different timing) finds supports in the data. According to the United States Census Bureau (Barnett and Vidal, 2013), about 33% of municipal revenues are intergovernmental transfers from higher levels; the National League of Cities estimates that 20-25% of total revenues comes from state transfers, while the remaining 5-10% comes directly from the Federal government. Moreover, both states and the Federal government have some discretion in the funds allocation and targeting.¹⁰ Despite this fact, one might still consider to address the voting problem over the amount and type of public goods in a simultaneous model. Unfortunately, a well-known result in the literature states that simultaneous voting on multidimensional issues when individuals’ type is bidimensional generically has no equilibrium (see for example, in the context of public good provision, Proposition 5 in De Donder et al. (2012)).

When deciding over the size of the government, voters only know the distribution of the possible types

⁹Of course we are aware of the possible interactions between fractionalization and political institutions, but the study of such interactions goes beyond the scope of the present paper. Related to this, there are some theories showing that social fragmentation can actually *increase* the size of the redistribution whenever politicians have incentives to satisfy individual needs (opposite to general ones) to maximize the probability of winning elections (for instance, refer to Persson and Tabellini (1999) for a comparison between majoritarian and proportional voting).

¹⁰<http://www.nlc.org/build-skills-and-networks/resources/cities-101/city-finances/local-revenue-structures/revenue-from-intergovernmental-transfers>.

of goods (that is, they know all the possible realizations and the probability associated to each realization), but not their actual realization. We assume that the local government will either implement the policy preferred by the most influential group, or will opt for an intermediate solution by considering all groups' preferences. Formally, the local government restricts the set of possible vectors to those preferred by the J groups, and the average of these J vectors, $\mathbf{s}_{\bar{j}} = \frac{1}{J} \sum_{j=1}^J \mathbf{s}_j$:

$$\mathbf{s}^* \in \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_J, \mathbf{s}_{\bar{j}}\}. \quad (11)$$

Ex ante, voters assign the same probability to each of the vectors of this set. This probability is noted $p = 1/(J + 1)$.

The timing of the model is the following: *i*) the candidates ($c = A, B$) simultaneously and non-cooperatively announce the platforms t_p ; *ii*) elections are held where individuals choose between candidates by majority voting; *iii*) the winner implements the announced policy; *iv*) \mathbf{s}^* is realized.

3.2 Individual Preferred Tax Rate

Given the uncertainty on the actual set of public goods that will be provided, the utility of the individual in the first period is random, and thus it takes the following expected form:

$$E_{\mathbf{s}^*} [u_{ij}] = E_{\mathbf{s}^*} \left[(1 - t) y_i + (t\bar{y})^{(1-\gamma)} (1 - L(d_j(\mathbf{s}_j, \mathbf{s}^*)))^\gamma \right]. \quad (12)$$

Given that the uncertainty originates from the actual policy \mathbf{s}^* , and that \mathbf{s}^* only affects utility through L_j , it will prove useful to analyze the expected loss $E_{\mathbf{s}^*} [L_j]$.

Lemma 1. *The expected loss is:*

$$E_{\mathbf{s}^*} [L_j] = \frac{K}{\sqrt{2}} \left(\sigma_k^2 [s_{jk}] - 2\sigma [s_{jk}, \mu_j [s_{jk}]] + \mu_j [\sigma_k^2 [s_{jk}]] \right), \quad (13)$$

where $\sigma_k^2 [s_{jk}]$ is the variance of the shares of public goods preferred by individuals in group j ; $\mu_j [\sigma_k^2 [s_{jk}]]$ is the expected value of these variances across all groups, constant for the whole population; $\sigma [s_{jk}, \mu_j [s_{jk}]]$ is the covariance between individual preferences and the expected "social" preferences, that is the average of all individuals' preferences on public goods.

Proof. See proof in appendix A.2 ■

Since the composition of public expenditure is randomly chosen among all individual's \mathbf{s}_j , every individual chooses the tax rate that maximizes her own expected utility, given the value of the expected loss.

Proposition 1. *The preferred tax rate of individual i in group j is :*

$$t_{ij}^* = \frac{(1-\gamma)^{\frac{1}{\gamma}}}{\bar{y}} \left(\frac{\bar{y}}{y_i} \right)^{\frac{1}{\gamma}} ((1-\delta) - E_{\mathbf{s}^*} [L_j]) \quad (14)$$

where $\delta \in (0, 1)$ is a non-negative scalar defined in the proof.

This tax rate is:

1. decreasing in the individual income;
2. decreasing with the variance of his/her preferences over types of public goods;
3. decreasing with the mean variance of individuals' preferences over types of public goods; and
4. increasing with the covariance between his/her preferences and average individuals' preferences.

Proof. See proof in appendix A.3 ■

This proposition tells us that the lower the individual income related to the mean income the higher his/her optimal tax rate. Moreover, if the difference between the composition of public goods preferred and the one which is expected to be provided by the government increases, the tax rate declines. This result is quite intuitive if we consider voters as risk-averse agents. Therefore, the voter does not want high taxes if it turns out that *i*) her preferences are very particular, i.e. she has a strong preference only for a limited subset of goods; *ii*) there is a lot of variance on what the implemented shares can be; and *iii*) she is very different from the average citizen. Points *i*) and *ii*) increase the variance of the share vector, and *iii*) makes the expected outcome far away from the voter's ideal vector.

Corollary 1. *The preferred tax rate t_{ij}^* is decreasing in γ if the voter's income is lower than or equal to the average income in the society. Otherwise, it is increasing in γ only as long as $\frac{\bar{y}}{y_i} < e^{-\frac{\gamma}{1-\gamma} - \ln(1-\gamma)}$.*

Proof. See appendix A.4 ■

Corollary 1 states that – introducing a positive loss function, that is considering the effect of the type of goods produced together with their quantity – if the voter is poor the optimal tax rate will always be higher, the higher the importance given to the quantity of redistribution. While, in case of a richer voter the optimal tax rate could in principle also increase if the importance given to the composition of public expenditure with respect to the amount was too high.

In particular, it is worth noticing that, as shown in Figure 1, a voter poorer than the mean may demand the same redistribution as if she were richer than the mean in a more standard framework, if the loss function is positive. These results become more and more similar with the increase of the loss function.

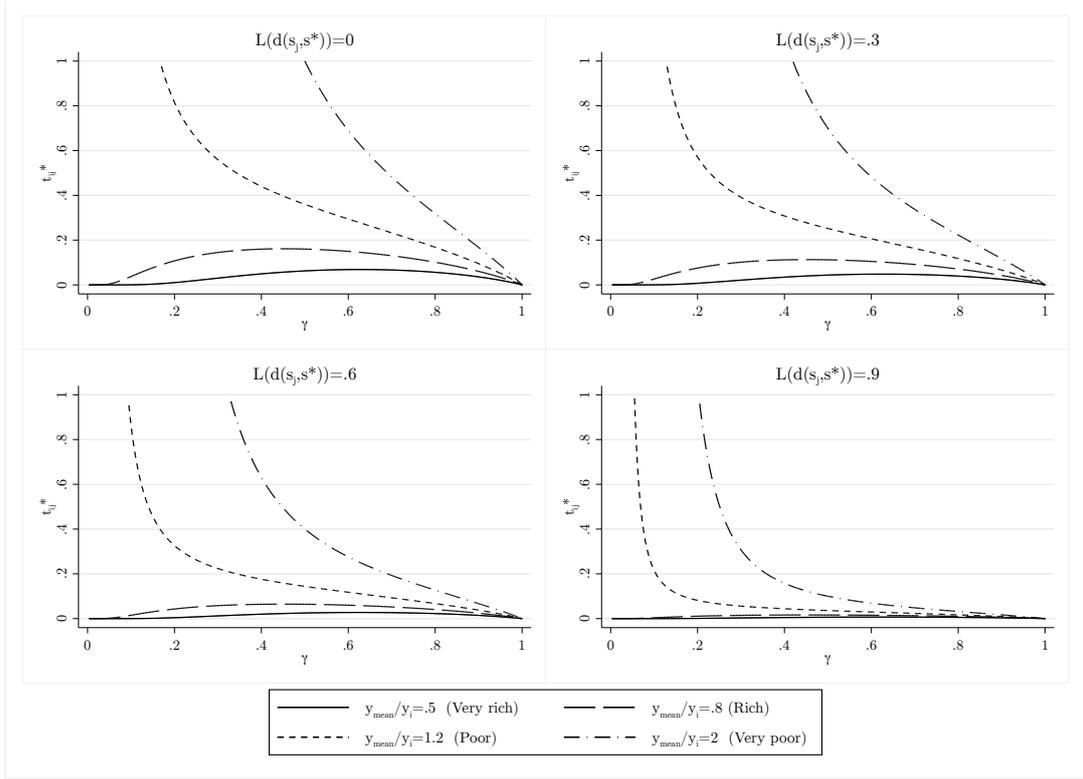


Figure 1: Optimal tax rate for different levels of income and different loss function

3.3 Voting Equilibrium

We know that in this political setting, the candidates' unique sub-game perfect equilibrium is to propose the same platform (Persson and Tabellini, 2002).

Lemma 2. *In each group $j = 1 \dots J$ the preferences over t are single crossing.*

Proof. See proof in appendix A.5 ■

Lemma 2 states that within a given group, for a given distance, an individual with a lower income always prefers a higher tax rate than a richer individual. This result allows the existence of a Condorcet winner t_{CW} as stated in the next lemma.

Lemma 3. *There exists a tax rate $t_{CW}^* \in (0, 1)$ which is preferred by at least half of the voters when faced against any other alternative $t^* \in (0, 1)$.*

Proof. See appendix A.6 ■

Lemma 3 ensures that an equilibrium platform exists and thus the quantity of public goods provided in equilibrium is given by $G^* = t_{CW}^* \bar{y}$.

In our context, where voters are characterized by a bidimensional type, it is not possible to have a closed form solution for the equilibrium tax rate. The complexity in determining the voting equilibrium arises

from the fact that once individuals of different groups are pooled in order to determine the median tax rate of the whole, high income individuals from a given group may have a higher preferred tax rate than poorer individuals belonging to a different group (see proposition 1). Therefore, it is impossible to find out the characteristics of the agent(s) with the median t^* . As a result, studying the impacts of *i*) income inequality and *ii*) diversity in preferences on the equilibrium tax rate (t_{CW}^*) is a difficult task. Nonetheless, we are able to provide some insights about these two effects.

In order to analyze the effect of income inequality, we study the impact on in-kind redistribution of a reduction in inequality induced by a Pigou-Dalton transfer (x). This transfer between two individuals belonging to any of the J groups is mean preserving but median increasing. The next proposition states the impact of such a transfer (x) on the equilibrium in-kind redistribution (G^*).

Proposition 2. *If inequality decreases by means of a Pigou-Dalton transfer (x), in-kind redistribution (G^*) does not increase, unless the transfer is made by an individual whose pre-transfer preferred tax rate $t^* \in [t_{CW^a}^* - \frac{\Delta t^*}{\Delta x}, t_{CW^a}^*]$, where $t_{CW^a}^*$ denote the pre-transfer equilibrium tax rate.*

Proof. See appendix A.7 ■

Proposition 2 describes how inequality affects the provision of public goods in this specific framework. Whenever the recipient of the Pigou-Dalton transfer crosses the Condorcet Winner position in the distribution of preferred t^* , then the Condorcet Winner changes: in particular, if inequality declines, in-kind redistribution also declines. However, the individual whose preferences over t are the median of the distribution, who is the politically relevant individual in this context, might be unaffected by a Pigou-Dalton transfer (Dalton, 1920). Therefore, it is possible that a Pigou-Dalton transfer reducing inequality leaves unaffected the equilibrium amount of in-kind redistribution.¹¹ Finally, in our framework, it is not possible to exclude the case in which $t^* \in [t_{CW^a}^* - \frac{\Delta t^*}{\Delta x}, t_{CW^a}^*]$, that leads to a puzzling result: if the one who crosses the Condorcet Winner position is the giver of the Pigou-Dalton transfer, then a reduction of inequality may lead to an increase of redistribution. Intuitively, this might happen because, in our framework, preferences over the tax rate do not depend only on income, but also on the expected distance between preferred and actual public goods, and this reshuffles rich and poor in the distribution of t^* . If her expected distance is very low, it is possible that a rich individual prefers a tax rate close to the median. After a Pigou-Dalton transfer that makes her poorer, she might cross the median, thus increasing the t_{CW}^* . Even if theoretically possible, such a situation is unlikely in the real world: indeed, since income distributions are typically right skewed, in terms of relative position a

¹¹This is in line with the classic result by Meltzer and Richard (1981). Also in their framework the key voter is the median individual in the distribution of incomes, and any Pigou-Dalton transfer that does not affect the median income leaves the amount of redistribution unchanged.

Pigou-Dalton transfer affects much more the recipient than the giver.¹²

Finally, let us analyze the impact of preferences diversity on the equilibrium in-kind redistribution (G^*). In order to do so, we study the effect of a decrease of the population average covariance between individual preferences and the expected “social” preferences, that is the average of all individuals’ preferences on public goods ($\bar{\omega} = \sum_{j=1}^J \nu_j (\sigma [s_{jk}, \mu_j [s_{jk}]])$), induced by a polarization of those preferences.¹³ This covariance can be seen as a proxy for diversity in preference in the population as it expresses how on average any individual differs from the average citizen.

Proposition 3. *If diversity increases by mean of a decrease in $\bar{\omega}$, in-kind redistribution (G^*) decreases.*

Proof. See appendix A.8 ■

Proposition 3 states that the amount of in-kind redistribution decreases whenever the heterogeneity within a society increases. Intuitively, the more concentrated is the distribution of preferences over public goods, the lower is the expected distance between the preferred bundle of public goods of each individual and the one produced, the more willing individuals are to increase the amount of public goods produced. Therefore, an increase of heterogeneity given by a move of preferences towards the extreme of the distribution reduces the expected utility from the public goods provided by the government, and this, in turn, makes the individuals less inclined to finance these public goods, for any income level.

The relationship between our theoretical results and what happens in the real world will be tested in the next sections, where we empirically analyze our main predictions using US data.

4 Dataset and variables definitions

The principal features of our theoretical setting are heterogeneity in incomes and preferences towards public goods. As Propositions 2 and 3 above state, a measure of income inequality and a measure of diversity in preferences affect the level of public expenditure. Our model has thus two interesting implications that can be tested: *i)* the lower the median income related to the mean income the more redistribution there is in the society, and *ii)* the higher the social distance, the lower redistribution we expect in the society.

The present section illustrates the data and the variables used in the next section to empirically test our theoretical model. The empirical investigation focuses on the United States, because of both theoretical and practical reasons. Firstly, we want to analyze a set of public administrations that share a common

¹²The *relative* position of an individual whose income increases from 20,000\$ to 25,000\$ changes much more than the *relative* position of an individual whose income decreases from 200,000\$ to 195,000\$.

¹³Recall proposition 1.

institutional framework, that enjoy a certain degree of autonomy in deciding the budget policies, and that are different enough with respect to the core issues of the paper: income inequality, social distance, and provision of public goods. Secondly, we need detailed information both at individual level – in order to generate measures of social distance and income inequality – and at public level – in order to analyze the budget of the public sector. US states meet both of these requirements: on the one side, they belong to the same Federal institutional framework, but they have great autonomy in deciding the amount of taxes and public expenditures. Other possible groups of countries, such as European Union, have the drawback of higher institutional heterogeneity among countries, introducing much more disturbance in the empirical analysis. On the other side, the United States Census Bureau provides several information both at individual and at state level, including individual income and public expenditures and revenues, disaggregated in very detailed items. Finally, among western democracies, US is the most populated country and experiences one of the highest levels of inequality and heterogeneity, making it suitable for the present analysis.

In this paper we exploit two datasets, both of them issued by the United States Census Bureau: the March supplement of the Current Population Survey (March-CPS from now on), which includes personal information and is used to measure income inequality and social distance, and the State & Local Government Finances (SLGF), which refers to state public budgets and is used to compute state revenues and expenditures in public goods.

March-CPS is a supplemental inquiry on income added to the CPS, a monthly US household survey conducted jointly by the US Census Bureau and the Bureau of Labor Statistics since 1940s.¹⁴ It is the most widely used by social scientists and policy makers and includes detailed information on earnings and incomes at individual and household level. Moreover, being part of a population survey, it includes detailed information also on individual and household characteristics.

The US Census Bureau conducts also an Annual Survey of Government Finances. One of its component is the State & Local Government Finances (SLGF) that collects information on revenue and expenditure in “considerable functional detail,”¹⁵ debt and financial assets. Drawing on these data, we are able to compute the amount of government revenue and expenditure for the provision of public goods and the amount of total taxes collected by the state authorities.

Online data availability goes back to 1992 for SLGF and to 1962 for March-CPS.¹⁶ However, we restrict our analysis to the period 1996 - 2010 because of data consistency. During the years, several items have

¹⁴For more information see <http://www.census.gov/cps/> and <http://www.bls.gov/cps/>, respectively.

¹⁵See <http://www.census.gov/govs/www/financegen.html> for more details.

¹⁶www.nber.org/data/current-population-survey-data.html.

been changed and a full comparability of all the variables we are interested in – both at individual and at public finance level – is achieved only in this 15 years time span. Moreover, we limit the analysis to 50 states, excluding District of Columbia, for which several items in the SLGF are missing. Therefore, we end up with a perfectly balanced panel of 750 observations, with a cross-section of 50 states observed over 15 years.

The next three subsections describe in detail the procedures to generate our variables of interest: income inequality, social distance, and public finance budget deficit, respectively.

4.1 Income level and income inequality

Data on personal income are derived from the March-CPS. The module includes 20 different definitions of income, whose details can be found in the codebooks.¹⁷ Information on income are collected at individual level and aggregated at family and household levels. All incomes are gross, therefore excluding any personal tax, social security, union dues, Medicare deductions, and so on. Moreover, they do not include non-cash benefits, such as food stamps, health benefits, subsidized housing, and goods produced and consumed on the farm. March-CPS data are top-coded and this may lead to underestimation of income inequality. In order to overcome this issue, we applied the procedure suggested by Larrimore et al. (2008) and a further refinement proposed by Burkhauser et al. (2012). These procedures replicate very closely the aggregate indicators obtained by uncensored internal CPS data, published in yearly reports (see DeNavas-Walt et al. (2012) for the most recent version).

Moving from the household total income, we retrieve equivalent household income by dividing the household income by the square root of the household size.¹⁸ We use equivalent household income under the assumption that labor market and consumption decisions are agreed within the household and that individuals vote on redistribution according to the household's needs.¹⁹ On the basis of equivalent incomes, we compute average and median income and inequality indices (mean/median ratio and Gini index) for every state and every year. Table 1 below lists the summary statistics for these variables.

¹⁷Sources of incomes are: wage and salary; self-employment (non-farm); self-employment (farm); unemployment compensation; worker's compensation; social security payments; supplemental security; public assistance or welfare; veterans' benefits; survivor's benefits; disability benefits; retirement income; interest; dividends; rents; educational assistance benefits; child support; alimony; financial assistance; other incomes. The complete set of codebooks and detailed definitions can be found at: <http://cps.ipums.org/cps/codebooks.shtml>.

¹⁸This is usually called the OECD equivalence scale. Other scales give different weights to household members, depending on the age and on the role in the household, but results are usually very similar. We decided to use the OECD equivalence scale because it is less sensitive to household composition (see below for the measurement problems on this variable).

¹⁹Imagine a household of two individuals, one with a top-income and the other who chooses to be out of the labour market. Our assumption is that they share the household's consumption and they vote for the level of redistribution that maximizes household members' utility.

4.2 Social distance

Society's social fragmentation and economic inequality are two different concepts. While the latter is related to the gap in economic resources among individuals, the former is a much broader concept that includes also differences due to social, cultural, and personal characteristics (Smeeding, 2002; Weeden et al., 2007).

Measures of social distance are computed using the personal information included in the March-CPS dataset. Among several variables available,²⁰ we select information on age, gender, educational level, ethnicity and labor force status. This choice follows not only from the belief that these characteristics are those most significantly affecting the preferences over the amount and composition of public expenditure (see the review by Alesina and Giuliano (2010) on this point), but also from the presence of two issues that makes it impossible to exploit the full set of information to compute the social distance. One is the quality of the dataset: some of the variables – namely migration status, employment, household composition and size of the home town – are missing for several state/year cells or are not fully comparable across cells. The other is the computational requirements of the social distance indices: given the complexity of the index, as we will show below, and the relatively low sample size, we had to limit the number of variables used.

The procedure to generate social distance indicators requires a disaggregation of the sample in a number of cells that is the product of all possible categories of the main variables. In general, it is $\prod_{c=1}^C n_c$, where c are categories and n the number of types per category. The cells need to satisfy two criteria: internal homogeneity, since we assume that preferences over public goods are similar for all individuals in the same cell; and parsimony, since the minimum sample size is 1137 individuals. Therefore we aggregated age in four categories,²¹ gender in two, educational level in four,²² ethnicity in five,²³ labor force status in four,²⁴ so that the population of every state/year is partitioned in 640 groups.

In order to obtain a synthetic indicator of social distance, we first implement a polychoric principal component analysis over the socioeconomic characteristics and then define the type for each group as the resulting normalized first component. Starting from this measure, we compute the weighted social distance (WSD) between *all* the types, a measure that is based on the Euclidean distance (and so closer in spirit to our theoretical setting).²⁵

²⁰For instance, gender, age, formal education, migration status, ethnicity, employment, labor force status, household composition and size of the home town.

²¹18-25, 26-45, 46-65, 66 and over.

²²No formal education, elementary completed, high school completed, college completed.

²³White non-hispanic, hispanic, black, asian, others.

²⁴Not in the labor force, full-time employed, part-time employed, unemployed.

²⁵See D'Agostino and Dardanoni (2009) for properties and applications of the Euclidean distance in similar contexts.

More in details, let us consider J groups, each formed by N_j individuals with type θ_j , defined in our case by the index given by the first component of the polychoric principal component analysis. We define the social distance of group j , SD_j , with respect to the rest of the society as:

$$SD_j = \sum_{i=1}^J \|\theta_j - \theta_i\| = \sqrt{\sum_{i=1}^J (\theta_j - \theta_i)^2} \quad (15)$$

with $i, j \in (1, J)$ and $j \neq i$. Group j social distance is thus given by the sum of the Euclidean distances with every other group.

As a measure of diversity in the whole society under consideration, we use the sum of these distances weighted by the population shares, as follows:

$$WSD = \sum_{j=1}^J \frac{N_j}{N} SD_j \quad (16)$$

As a robustness check, in the next section we also use another indicator of diversity, recently proposed in the literature, the generalized ethno-linguistic fractionalization index (GELF), introduced by Bossert et al. (2011). This measure is a generalization of the well-know ELF index,²⁶ and it is here defined as:

$$GELF = 1 - \frac{1}{J^2} \sum_{i=1}^J \sum_{j=1}^J s_{ij} \quad (17)$$

where $s_{ij} \in (0, 1)$ is the similarity between group i and group j and $s_{ij} = 1 - \frac{|\theta_j - \theta_i|}{(\max\theta - \min\theta)}$.²⁷

Although differing in the specific definition, both the indicators measure heterogeneity in the society taking into consideration not only the presence of different socioeconomic groups, but also how distant each group is from the others.²⁸ Table 1 includes summary statistics for both indicators.

4.3 Public finance

Data on public finance are based on SLGF dataset. The dataset, released by the Census Bureau, is very detailed and disaggregates public balances in 110 revenue items and 205 expenditure items, belonging to 33 different voices.²⁹ In order to compute variables meaningful for our purposes, we select items that *i)* are under the control of the policy maker; *ii)* are decided at state level. Therefore we exclude 10 of the 33 voices, some that are primarily or exclusively decided by municipalities, for instance local transit

²⁶The ethno-linguistic fractionalization index is given by one minus the Herfindahl index of concentration, and as so it does not take into consideration the “distance” among groups.

²⁷Refer to Bossert et al. (2011) for a formal characterization and for the properties of this index.

²⁸The correlation between the two indicators in our sample is equal to .63.

²⁹For instance, the voice “education” is split in 13 expenditure items and 6 revenue items, depending on the nature of the expenditure and of the revenue.

facilities, others because they are not under direct control of policy makers, for instance the service on debt. Finally, the dataset provides also the total amount of taxes collected by every state.

The provision of public good is computed taking the “net” expenditure, that is the difference between expenditure and revenue imputed to each category of the budget. The choice of subtracting the public revenue originating from the provision of goods or services is close in spirit to our definition of redistribution. If the public sector produces a good and sells it at its market price, it is not making any redistribution. Opposite, if the price is lower than the cost, or if it is charged only to a subgroup of the population, then redistribution occurs. We also compute the total amount of taxes. We then take the ratio of all variables with respect to state GDP, to make the figures comparable across states and years. It is important to highlight that revenues do *not* include Federal and intergovernmental grants, but they include only fees and tariffs directly related to the provision of goods and services.

From here on we refer to public deficit as the difference between total expenditures and total revenues for all the categories, that is the negative balance for all public goods provided by every state, in terms of GDP. Public deficit is positive whenever the expenditures are higher than the revenues, and negative otherwise. Last rows of table 1 show summary statistics of aggregated expenditures and taxes. Data on state GDP are taken from the Bureau of Economic Analysis dataset.

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
March supplement of Current Population Survey					
Mean / median ratio (in %)	126.491	6.09	110.251	150.171	750
Gini coefficient	0.408	0.026	0.331	0.478	750
WSD	0.514	0.026	0.429	0.648	750
GELF	0.247	0.015	0.204	0.29	750
State & Government Government Finances					
Deficit / GDP ratio (in %)	0.973	1.904	-4.102	18.518	750
Exp. / GDP ratio (in %)	6.584	2.256	3.467	23.67	750
Taxes / GDP ratio (in %)	5.611	1.276	2.422	17.532	750
Bureau of Economic Analysis					
GDP, billions \$	224.874	274.959	14.556	1900.463	750

Based on the indicators and variables illustrated here above, the next section describes the empirical analysis implemented to test our theoretical model and reports the results of the regression analysis.

5 Empirical strategy and results

Given the panel structure of the data, the most suitable model to test the effect of social distance on public goods provision is a time invariant unobserved heterogeneity model with state fixed effects, to capture all the time invariant unobservable characteristics of the states. Opposite to the frequently used “random effects” model, the assumptions of this model do not require the uncorrelation between the

independent variables and the time invariant unobserved heterogeneity term (Wooldridge, 2001).

The structure of our baseline model is the following:

$$Red_{i,t} = \alpha + \beta SD_{i,t} + \gamma Ineq_{i,t} + X_{i,t}\Lambda + u_i + \varepsilon_{i,t} \quad (18)$$

where, according to our theoretical model, $Red_{i,t}$ is redistribution, measured as the net public deficit in state i at time t , as defined in the previous section, SD is the social distance index, $Ineq$ is the mean-median ratio, X includes a small set of controls (the tax-GDP ratio and the nominal GDP), u is the time invariant unobserved heterogeneity term, and ε is the idiosyncratic error term. In the baseline specification time t refers to electoral terms. There are two main reasons for aggregating yearly variables by electoral term: first, the paper focuses on the link between heterogeneity of preferences and public expenditures. In all representative democracies, and US is not an exception, this link is mediated by a political process. Therefore, we think it is reasonable to consider the electoral mandate as the main time unit for the present analysis. Second, political processes need usually more time than private business decisions. Because of the public nature of the expenditures and of the timing of political decisions, balance items are somehow persistent and significant changes seldom take place in a short period of time. Therefore, aggregating variables in four years periods, the length of an electoral term in the US, raises more variability in the dataset and it is - in our opinion - the best way to isolate the effect of individual preferences on political decisions on provision of public goods. However, we include also calendar years regressions, and results are confirmed.

Before proceeding to the analysis of our results, we consider useful to highlight some differences between our empirical strategy and the contribution closest to ours in the literature, provided in the seminal paper by Alesina et al. (1999). First, our unit of analysis are the states, while they consider cities, metropolitan areas, and counties. Analyzing states is more coherent with our theoretical model, in which tax rate and budget decisions are taken at a higher level of government, while policies are implemented at a local level. Moreover, focusing on state level data, we rule out most of the Tiebout sorting issue: as Alesina et al. (1999, p.1255) state, “there is going to be far more Tiebout sorting between city and suburb of one metropolitan area than between different metropolitan areas”. As they do not find any evidence of a Tiebout sorting in their sample, *a fortiori* the issue should be even weaker at state level. A second, relevant difference is that we can exploit the panel dimension, while Alesina et al. (1999) were limited to cross sectional data, mainly due to data availability and comparability. This allows us to include state fixed effects to rule out possible time invariant unobserved heterogeneity. Thirdly, following the theory, we use a measure of redistribution as a dependent variable, instead of the level of expenditures: in our view, public expenditure for a service *per se* is only a partial indicator of

redistribution, since it does not consider the level of fees and tariffs charged on that service. Fourth, we do not disaggregate redistribution by categories, but we consider the aggregate of all discretionary items. Again, this choice is only driven by the willingness of being as close as possible to our theoretical predictions, that link social distance and income inequality to the size of in-kind redistribution, irrespectively of the public goods produced. Finally, for the same reason, we do not limit our measure of social distance to the racial heterogeneity, but we consider, as described in section 4.2, a much wider range of characteristics that might affect, in our view and according to Alesina and Giuliano (2010), the preferences for (different) public goods.

Results of our baseline model are reported in column 1 of Table 2 and they are in line with the theoretical predictions of the model: first, inequality has a positive impact on the public deficit,³⁰ meaning that – given the amount of taxes – public sector net expenditures are higher. Second, social distance has the opposite effect: the higher the social fragmentation,³¹ the lower the public deficit. Moreover, richer states seem to have more public deficit and there is an intuitive negative correlation between the amount of taxes collected by the state and the budget deficit.

Table 2: Baseline models

Time dimension:	Term		Year		2 years
Dependent variable:	Deficit / GDP ratio (in %)				
	(1)	(2)	(3)	(4)	(5)
Mean / median ratio	0.053***		0.025***		0.036***
	0.020		0.007		0.008
WSD Polych. PCA	-12.616***		-3.817***		-10.106***
	3.706		1.266		1.584
Mean / median ratio (lagged)		0.048**		0.022***	
		0.024		0.007	
WSD Polych. PCA (lagged)		-8.335*		-5.259***	
		4.566		1.285	
GDP, billions \$	0.003***	0.003**	0.003***	0.003***	0.003***
	0.001	0.001	0.000	0.000	0.000
Tax / GDP ratio	-1.024***	-1.045***	-1.106***	-1.140***	-1.135***
	0.084	0.090	0.043	0.043	0.043
Constant	9.822***	8.918**	9.233***	10.692***	11.524***
	3.236	3.856	1.146	1.129	1.394
State FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.870	0.892	0.847	0.860	0.890
Obs.	217	167	750	700	650

Note: * 10% confidence interval, ** 5% confidence interval, *** 1% confidence interval.

However, these results may be driven by the specific assumptions we made with respect to the timing of political and public choice or to the choice of a specific set of indicators. Luckily, drawing on the data available, it is possible to test our model under several hypotheses on the timing of the political choice

³⁰Recall that by public deficit we mean the negative difference between tariffs and expenditures in the provision of a set of publicly provided goods.

³¹The measure of social distance used in this section is the one presented in eq.(16).

and to test the robustness of our results to different sets of indicators.

With respect to the time dimension, many different assumptions can be made on how long the preferences of individuals take to be translated in political decisions and it is very difficult to order these hypotheses in terms of realism. In details, we decide to test the following assumptions:

- A one-term lag (column 2 of Table 2), meaning that preferences of individuals in the actual term will be translated in next term policies. Eq.(18) can be restated as:

$$Red_{i,t} = \alpha + \beta SD_{i,t-1} + \gamma Ineq_{i,t-1} + \Lambda X_{i,t} + u_i + \epsilon_{i,t} \quad (19)$$

where the variables have the same meaning as before and the only change is the time dimension. This represents the extreme case in which policies are unaffected by contemporaneous preferences, but they only depend on the political framework in the previous term, when elections took place.

- Contemporaneous effect, considering years instead of legislatures (column 3 of Table 2). This assumption is analogous to the eq.(18) where we replace electoral terms with calendar years. This represents the other extreme case: preferences are immediately translated into policies.
- A one-year lag (column 4 of Table 2), that combines the two previous assumptions: the preferences of individuals today affect the policies implemented next year, independently of the electoral cycle.
- A mixed effect, in which contemporaneous and one-year lags are considered jointly: the average public deficit in years $t - t + 1$ depends on individual preferences in years $t - 1 - t$ (column 5 of Table 2). This model allows to smooth the variables and to mediate between the two extreme assumptions of simultaneity and of term lags.

Results (models 2-5 in Table 2) confirm the baseline model (column 1) also under different hypotheses: the inequality term is always significant with the expected sign. Social distance is also always significant. The only model in which significance is lower than 1% is when we assume a one term lag between preference formation and policy implementation, the most extreme assumption we make. With respect to control variables, there are no relevant changes throughout the models, while the sample size varies according to the assumption on the time dimension: when it is the calendar year, sample size is 750 if the effects are simultaneous or 700 if they are lagged (intuitively, the first period is missing for each of the 50 states); when the time dimension is the term, since elections occur in different years, there is not a predetermined number of terms. We observe 217 terms (an average of 4.34 terms per state) that reduce

to 167 for lagged models (again, the first term for each state cannot be included). The regressions are robust to the reduced sample size and there seems not to be any fall of significance due to it.

5.1 Robustness checks

Following the literature on electoral cycle, we provide a first robustness check testing whether public deficit is subject to significant changes during the terms. Table 3, models (2) to (5) show the baseline model – in the version with calendar years, since we want to test exactly the difference among years within the same electoral term – with the addition of a dummy variable for every year of the term. Results suggest a possible effect of the political cycle, in particular an improvement of public deficit in the first year of the mandate. However, this does not influence the effect of inequality and social distance on the provision of public goods, confirming our main results.

Table 3: Electoral cycle

Time dimension: Dependent variable:	Year				
	Deficit / GDP ratio (in %)				
	(1)	(2)	(3)	(4)	(5)
Mean / median ratio	0.025*** 0.007	0.025*** 0.007	0.025*** 0.007	0.025*** 0.007	0.025*** 0.007
WSD Polych. PCA	-3.817*** 1.266	-3.964*** 1.265	-3.887*** 1.265	-3.824*** 1.266	-3.803*** 1.266
First year of term		-0.125** 0.064			
Second year of term			0.095 0.065		
Third year of term				0.069 0.064	
Fourth year of term					-0.041 0.068
GDP, billions \$	0.003*** 0.000	0.003*** 0.000	0.003*** 0.000	0.003*** 0.000	0.003*** 0.000
Tax / GDP ratio	-1.106*** 0.043	-1.103*** 0.043	-1.108*** 0.043	-1.104*** 0.043	-1.106*** 0.043
Constant	9.233*** 1.146	9.359*** 1.145	9.252*** 1.145	9.239*** 1.145	9.244*** 1.146
State FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.847	0.848	0.847	0.847	0.847
Obs.	750	750	750	750	750

Note: * 10% confidence interval, ** 5% confidence interval, *** 1% confidence interval.

The second type of robustness check regards the choice of indicators for the three main dimensions we focus on: public deficit, inequality, and social distance. Table 4 shows the baseline model (in column 1) and three variations of it. The first (column 2) includes GELF instead of WSD. The second considers a different inequality index, the Gini coefficient. This index differs from the mean-median ratio as it considers the whole distribution of income, instead of only two points of it. Finally, column 4 reports the expenditure-GDP ratio instead of the deficit-GDP ratio. The amount of taxes has been dropped, because of collinearity between public deficit, taxes and expenditures. All regressions in Table 4 are

consistent with the baseline model, therefore we can conclude that there are no reasons to suspect that our results are affected by the specific choice of indicators. For ease of exposition we do not show all the possible combinations of models and indicators, but they fully confirm our conclusion.³²

Table 4: Alternative indicators

Time dimension: Dependent variable:	Term			
	Deficit / GDP ratio (in %)			Exp./GDP
	(1)	(2)	(3)	(4)
Mean / median ratio	0.053*** 0.020	0.048** 0.020		0.054*** 0.020
Gini index			12.889** 5.153	
WSD Polych. PCA	-12.616*** 3.706		-12.556*** 3.713	-12.481*** 3.666
GELF		-20.395*** 6.378		
GDP, billions \$	0.003*** 0.001	0.003*** 0.001	0.003*** 0.001	0.003*** 0.001
Tax / GDP ratio	-1.024*** 0.084	-1.007*** 0.084	-1.029*** 0.084	
Constant	9.822*** 3.236	9.072*** 3.180	11.368*** 2.888	9.515*** 3.050
State FE	Yes	Yes	Yes	Yes
R-squared	0.870	0.869	0.870	0.918
Obs.	217	217	217	217

Note: * 10% confidence interval, ** 5% confidence interval, *** 1% confidence interval.

Finally, Table 5 shows the robustness checks on persistence, reverse causality and selective migration: first, one may think that the three variables of interest – in-kind redistribution, income inequality, and social distance – are very persistent, and this could affect the results. To control for this, we run the model in column 2, where public deficit in term $t - 1$ is regressed on inequality and social distance in term t . If the previous results were driven by intertemporal persistence, we should expect statistically significant correlations. Instead, coefficients are not statistically different from zero, suggesting that our results are not driven by the persistence of the economic process. Moreover, one may also think that there is an issue of reverse causality: it is the amount of redistribution that drives individual choices in terms of labor supply (that affects income inequality) or migration (that affects social distance, if one decides to move where the set of public goods provided is closer to her preferences). Columns 3 and 4 test whether redistribution and social distance depend significantly on the amount of in-kind redistribution in the previous term. Also in this case, the coefficients are not significant in both models, suggesting that our results are robust to this kind of criticism.³³ In addition, reverse causality with respect to migration behavior was already controlled for in column 2 and 4 of Table 2. Indeed, if we believe that there is a political process behind the redistribution policies, including lagged explanatory

³²These results are available from the authors upon request.

³³We find the same results if we consider as a baseline model the version with calendar years instead of electoral terms. Results are available upon request.

variable should partly solve this issue, since social distance in the previous term (year) is not affected by migration in the actual term (year). This is in line also with the findings in Alesina et al. (1999), who do not find any support for a Tiebout sorting of individuals.

Table 5: Persistence, reverse causality and sorting

Time dimension: Dep. var.	Term			
	Deficit / GDP ratio (in %) (1)	Deficit / GDP ratio (in %, lagged) (2)	WSD Polych. PCA (3)	Mean / median ratio (4)
Mean / median ratio	0.053*** 0.020	0.027 0.023	0.001 0.001	
WSD Polych. PCA	-12.616*** 3.706	3.155 3.772		22.774 15.565
GDP, billions \$	0.003*** 0.001			
Tax / GDP ratio	-1.024*** 0.084			
GDP, billions \$ (lagged)		0.003*** 0.001	0.000 0.000	-0.003 0.004
Tax / GDP ratio (lagged)		-0.642*** 0.131	-0.001 0.004	-1.036* 0.592
Deficit / GDP ratio (in %) (lagged)			0.002 0.002	0.474 0.388
Constant	9.822*** 3.236	1.763 3.518	0.417*** 0.078	118.136*** 9.509
State FE	Yes	Yes	Yes	Yes
R-squared	0.870	0.906	0.467	0.813
Obs.	217	167	167	167

Note: * 10% confidence interval, ** 5% confidence interval, *** 1% confidence interval.

6 Conclusion

The aim of the paper is to model the consequences of heterogeneous income and heterogeneous preferences on the provision of various public goods. We set up a simple framework in which individuals differ in both their income and their preferences over a bundle of public goods. We provide a theoretical framework that represents the variety of public goods in a more general way than the existing literature. Indeed, while goods are generally ordered in a unidimensional space, we do not impose such a restrictive structure. The total budget devoted to the production of public goods is decided through majority voting at the state level, while the composition of public expenditure is decided in a second step at a different level, e.g. county or municipality level. The composition of the public budget is thus uncertain when the size is voted upon, therefore the equilibrium quantity of in-kind redistribution depends on the dispersion of voters income and preferences over the type of good to be provided. In such a framework, the quantity of public goods provided ultimately depends negatively on the heterogeneity of individual preferences. *Ceteris paribus*, the higher diversity of preferences over public goods, the lower the provision.

The empirical analysis supports the theoretical predictions of our model. We use a panel of US survey and census data to test the linkage between social distance and public expenditure on provision of in-kind redistribution. We find that social distance is negatively associated to the quantity of redistribution. Moreover, in line with previous literature, we find that income inequality is positively associated to the quantity of redistribution. Our results hold under many different assumptions on the timing of the political process, on the indicators of income inequality and heterogeneity used, and are robust to several robustness checks on persistence, reverse causality and selective migration.

References

- Alesina, A., Baqir, R., and Easterly, W. (1999). Public goods and ethnic divisions. *Quarterly Journal of Economics*, 114:1243–1284.
- Alesina, A. and Giuliano, P. (2010). Preferences for redistribution. In Benhabib, J., Jackson, M. O., and Bisin, A., editors, *Handbook of Social Economics*, pages 93–131. The Netherlands: North Holland.
- Alesina, A. and La Ferrara, E. (2005). Ethnic diversity and economic performance. *Journal of Economic Literature*, 43:762–800.
- Alesina, A. and Spalatore, E. (1997). On the number and size of nation. *Quarterly Journal of Economics*, 113:1027–1056.
- Alesina, A. and Zhuravskaya, E. (2011). Segregation and the quality of government in a cross section of countries. *American Economic Review*, 101(5):1872–1911.
- Banerjee, A., Iyer, L., and Somanathan, R. (2005). History, social division, and public goods in rural India. *Journal of the European Economic Association*, 3(2-3):639–647.
- Barnett, J. L. and Vidal, P. M. (2013). State and local government finances summary: 2011. Governments division briefs, United States Census Bureau.
- Bolton, P. and Roland, G. (1997). The breakup of nations: A political economy analysis. *Quarterly Journal of Economics*, 112(4):1057–1090.
- Borck, R. (2007). Voting, inequality and redistribution. *Journal of Economic Surveys*, 21:90–109.
- Bossert, W., D’Ambrosio, C., and La Ferrara, E. (2011). A generalized index of fractionalization. *Economica*, 78:723–750.

- Burkhauser, R. V., Feng, S., Jenkins, S. P., and Larrimore, J. (2012). Recent trends in top income shares in the United States: Reconciling estimates from march CPS and IRS tax return data. *Review of Economics and Statistics*, 94(2):371–388.
- D’Agostino, M. and Dardanoni, V. (2009). What’s so special about Euclidean distance? A characterization with applications to mobility and spatial voting. *Social Choice and Welfare*, 33(2):211–233.
- Dalton, H. (1920). The measurement of the inequality of incomes. *Economic Journal*, 30(119):348–361.
- Davis, O. A., Hinich, M. J., and Ordeshook, P. C. (1970). An expository development of a mathematical model of the electoral process. *American political science review*, 64(2):426–448.
- De Donder, P. (2013). Majority voting and the single-crossing property when voters belong to separate groups. *Economics Letters*, 118(3):523–525.
- De Donder, P., Le Breton, M., and Peluso, E. (2012). On the (sequential) majority choice of public good size and location. *Social Choice and Welfare*, 39:457–489.
- DeNavas-Walt, C., Proctor, B. D., and Smith, J. C. (2012). Income, poverty, and health insurance coverage in the United States: 2011. Current population reports, United States Census Bureau.
- Desmet, K., Weber, S., and Ortuño Ortín, I. (2009). Linguistic diversity and redistribution. *Journal of the European Economic Association*, 7(6):1291–1318.
- Epple, D. and Romano, R. E. (1996). Ends against the middle: Determining public service provision when there are private alternatives. *Journal of Public Economics*, 62(3):297–325.
- Fernandez, R. and Levy, G. (2008). Diversity and redistribution. *Journal of Public Economics*, 92:925–943.
- Gregorini, F. (2009). Political geography and income inequalities. Working Papers 152, University of Milano-Bicocca, Department of Economics.
- Habyarimana, J., Humphreis, M., Posner, D. N., and Weinstein, J. M. (2007). Why does ethnic diversity undermine public goods provision? *American Political Science Review*, 101(04):709–725.
- Larrimore, J., Burkhauser, R. V., Feng, S., and Zayatz, L. (2008). Consistent cell means for topcoded incomes in the public use march CPS (1976-2007). *Journal of Economic and Social Measurement*, 33:89–128.

- Luttmer, E. F. P. (2001). Group loyalty and the taste for redistribution. *Journal of Political Economy*, 109(3):500–528.
- Meltzer, A. H. and Richard, S. F. (1981). A rational theory of the size of government. *Journal of Political Economy*, 89(5):914–27.
- Miguel, E. and Gugerty, M. K. (2005). Ethnic diversity, social sanctions, and public goods in Kenya. *Journal of Public Economics*, 89(11-12):2325–2368.
- Persson, T. and Tabellini, G. (1999). The size and scope of government: Comparative politics with rational politicians. *European Economic Review*, 43(4-6):699 – 735.
- Persson, T. and Tabellini, G. (2002). *Political Economics: Explaining Economic Policy*. The MIT Press.
- Smeeding, T. M. (2002). Globalization, inequality, and the rich countries of the G-20: Evidence from the Luxembourg Income Study (LIS). Working Paper 112, Center for Policy Research.
- Stigler, G. J. (1970). Director’s law of public income redistribution. *Journal of Law and Economics*, 13(1):1–10.
- Weeden, K. A., Kim, Y.-M., Carlo, M. D., and Grusky, D. B. (2007). Social class and earnings inequality. *American Behavioral Scientist*, 50:702–736.
- Wooldridge, J. M. (2001). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.

A Appendix

A.1 Maximum distance

In this appendix we formally prove that $\max d(\mathbf{s}_1, \mathbf{s}_2) = \sqrt{2}$, for any couple of vectors \mathbf{s}_j such that the sum of their elements is 1: $\sum_{k=1}^K s_{kj} = 1$.

Assume a society with K public goods. Take two generic groups, $j = 1, 2$, and assume that their preferred public goods are represented by two generic vectors $\mathbf{s}_1 = (s_{11}, s_{21}, s_{31}, 0, \dots, 0)$ and $\mathbf{s}_2 = (s_{12}, s_{22}, s_{32}, 0, \dots, 0)$, where the constraint that $s_{4j}, \dots, s_{Kj} = 0$ will be shown to be irrelevant. Under this constraint, the policy space can be represented as in figure 2, with only three political dimensions. Remember that $\sum_{k=1}^K s_{kj} = 1$, so that $s_{3j} = 1 - s_{1j} - s_{2j}$. The distance can be written

as:

$$\begin{aligned}
d_{1,2} &= \sqrt{(s_{11} - s_{12})^2 + (s_{21} - s_{22})^2 + (s_{31} - s_{32})^2 + 0 + \dots + 0} = \\
&= \sqrt{(s_{11} - s_{12})^2 + (s_{21} - s_{22})^2 + (1 - s_{11} - s_{21} - 1 + s_{12} + s_{22})^2}.
\end{aligned} \tag{20}$$

Applying a monotonic transformation to this function, we can state that $\min_{\mathbf{s}_1, \mathbf{s}_2} d_{1,2} = \min_{\mathbf{s}_1, \mathbf{s}_2} d_{1,2}^2$, where

$$d_{1,2}^2 = (s_{11} - s_{12})^2 + (s_{21} - s_{22})^2 + ((s_{11} - s_{12}) + (s_{21} - s_{22}))^2. \tag{21}$$

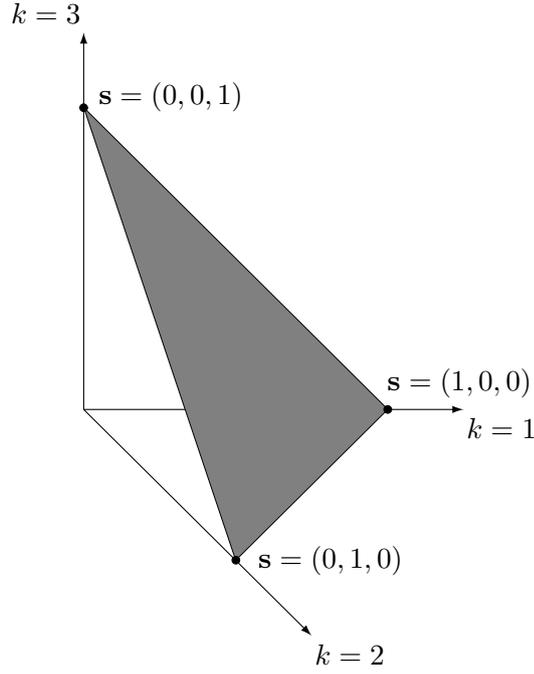


Figure 2: Illustration of possible preferences in case of three public goods

Since the function in eq.(21) is convex (being the sum of convex functions), the minimum can be found through first order conditions:

$$\begin{cases} \frac{\partial d_{1,2}^2}{\partial s_{11}} = 2(s_{11} - s_{12}) + 2((s_{11} - s_{12}) - (s_{22} - s_{21})) = 0 \\ \frac{\partial d_{1,2}^2}{\partial s_{12}} = 2(s_{22} - s_{21}) + 2((s_{11} - s_{12}) - (s_{22} - s_{21})) = 0 \end{cases} \tag{22}$$

that are solved for:

$$\begin{cases} s_{11} = s_{12} \\ s_{21} = s_{22} \end{cases} \Rightarrow \mathbf{s}_1 = \mathbf{s}_2 \tag{23}$$

meaning that if the two vectors are equal the loss function is at the minimum, and the minimum is 0.

Since the function is convex, and the support is finite ($s_{kj} \in [0, 1]$), it reaches the maximum at the

boundary of the support. Moreover, the function in eq.(21) is clearly increasing in both $|s_{11} - s_{12}|$ and $|s_{21} - s_{22}|$. Indeed, if you call the two terms x_1 and x_2 respectively, the first derivatives are always positive:

$$\begin{cases} \frac{\partial d_{1,2}^2}{\partial x_1} = 2x_1 + 2(x_1 + x_2) > 0 \\ \frac{\partial d_{1,2}^2}{\partial x_2} = 2x_2 + 2(x_1 + x_2) > 0. \end{cases} \quad (24)$$

Since $s_{kj} \in [0, 1]$, the maximum value of both $|s_{11} - s_{12}|$ and $|s_{21} - s_{22}|$ is 1. Therefore, the distance is maximum whenever:

$$\begin{cases} |s_{11} - s_{12}| = 1 \\ |s_{21} - s_{22}| = 1. \end{cases} \quad (25)$$

Combining eq.(25) with $\sum_{k=1}^K s_{kj} = 1$, then either

$$s_{11} = 1 \Rightarrow \begin{cases} s_{2,1} = 0 \\ s_{1,2} = 0 \Rightarrow s_{2,2} = 1 \end{cases} \quad (26)$$

or

$$s_{11} = 0 \Rightarrow \begin{cases} s_{12} = 1 \\ s_{21} = 1 \Rightarrow s_{22} = 0. \end{cases} \quad (27)$$

Therefore:

$$\arg \max d_{1,2} = \begin{cases} \mathbf{s}_1 = (0, 1, 0, 0, \dots, 0) \\ \mathbf{s}_2 = (1, 0, 0, 0, \dots, 0) \end{cases} \quad \text{or} \quad \begin{cases} \mathbf{s}_1 = (1, 0, 0, 0, \dots, 0) \\ \mathbf{s}_2 = (0, 1, 0, 0, \dots, 0) \end{cases} \quad (28)$$

and

$$\max d_{1,2} = \sqrt{2}. \quad (29)$$

In any case, $s_{31} = s_{32} = 0$ and therefore the same is true for all s_{4j}, \dots, s_{Kj} . The procedure can be generalized to any triple of public goods k .

A.2 Proof of lemma1

The expected loss of a generic group j is:

$$\mathbb{E} \left[L \left(\frac{d_j(\mathbf{s}_j, \mathbf{s}^*)}{\max d} \right) \right]. \quad (30)$$

Since $E \left[L \left(\frac{d_j(\mathbf{s}_j, \mathbf{s}^*)}{\max d} \right) \right] = E \left[\left(\frac{\sqrt{\sum_{k=1}^K (s_{jk} - s_k^*)^2}}{\sqrt{2}} \right)^2 \right]$, for simplicity and without loss of generality we multiply the distance by $\max d = \sqrt{2}$.

By definition of quadrance (that is the square of Euclidean distance) we can write:

$$E [L (d_j (\mathbf{s}_j, \mathbf{s}^*))] = E \left[\left(\sqrt{\sum_{k=1}^K (s_{jk} - s_k^*)^2} \right)^2 \right] = E \left[\sum_{k=1}^K (s_{jk} - s_k^*)^2 \right] \quad (31)$$

and, developing the quadratic form:

$$E \left[\sum_{k=1}^K (s_{jk} - s_k^*)^2 \right] = E \left[\sum_{k=1}^K s_{jk}^2 - \sum_{k=1}^K 2s_{jk}s_k^* + \sum_{k=1}^K s_k^{*2} \right]. \quad (32)$$

Since the last term can be written as $\sum_{k=1}^K s_k^{*2} = K \frac{\sum_{k=1}^K s_k^{*2}}{K} = K \mu_k [s_k^{*2}]$:

$$E \left[\sum_{k=1}^K s_{jk}^2 - \sum_{k=1}^K 2s_{jk}s_k^* + \sum_{k=1}^K s_k^{*2} \right] = E \left[\sum_{k=1}^K s_{jk}^2 - \sum_{k=1}^K 2s_{jk}s_k^* + K \mu_k [s_k^{*2}] \right]. \quad (33)$$

Adding and subtracting $K \left(\frac{\sum_{k=1}^K s_k^*}{K} \right)^2$ to the previous, and recalling that since $\sum_{k=1}^K s_k^* = 1$, then:

$$K \left(\frac{\sum_{k=1}^K s_k^*}{K} \right)^2 = K (\mu_k [s_k^*])^2 = K \frac{1}{K^2} = \frac{1}{K} \Rightarrow \frac{1}{K} - K (\mu_k [s_k^*])^2 = 0 \quad (34)$$

leads to:

$$E \left[\sum_{k=1}^K s_{jk}^2 - \sum_{k=1}^K 2s_{jk}s_k^* + K \mu_k [s_k^{*2}] \right] = E \left[\sum_{k=1}^K s_{jk}^2 - \sum_{k=1}^K 2s_{jk}s_k^* + K \mu_k [s_k^{*2}] - K (\mu_k [s_k^*])^2 + \frac{1}{K} \right]. \quad (35)$$

Since $K \mu_k [s_k^{*2}] - K (\mu_k [s_k^*])^2 = K \sigma_k^2 [s_k^*]$:

$$E \left[\sum_{k=1}^K s_{jk}^2 - \sum_{k=1}^K 2s_{jk}s_k^* + K \mu_k [s_k^{*2}] - K (\mu_k [s_k^*])^2 + \frac{1}{K} \right] = E \left[\sum_{k=1}^K s_{jk}^2 - \sum_{k=1}^K 2s_{jk}s_k^* + K \sigma_k^2 [s_k^*] + \frac{1}{K} \right]. \quad (36)$$

Since s_{jk} is known and certain, being the vector of preferences of group j itself, for every group j the

only random variable is s_k^* , thus expectation becomes:

$$\mathbb{E} \left[\sum_{k=1}^K s_{jk}^2 - \sum_{k=1}^K 2s_{jk}s_k^* + K\sigma_k^2 [s_k^*] + \frac{1}{K} \right] = \sum_{k=1}^K s_{jk}^2 - \mathbb{E} \left[\sum_{k=1}^K 2s_{jk}s_k^* \right] + K \mathbb{E} [\sigma_k^2 [s_k^*]] + \frac{1}{K}. \quad (37)$$

Similarly as before, adding and subtracting $K \left(\frac{\sum_{k=1}^K s_{jk}}{K} \right)^2 = K (\mu_k [s_{jk}])^2 = K \frac{1}{K^2} = \frac{1}{K}$ (from which $\frac{1}{K} - K (\mu_k [s_{jk}])^2 = 0$) since $\sum_{k=1}^K s_{jk} = 1$ leads to:

$$\begin{aligned} & \sum_{k=1}^K s_{jk}^2 - \mathbb{E} \left[\sum_{k=1}^K 2s_{jk}s_k^* \right] + K \mathbb{E} [\sigma_k^2 [s_k^*]] + \frac{1}{K} = \\ & \sum_{k=1}^K s_{jk}^2 - K (\mu_k [s_{jk}])^2 + \frac{1}{K} - \mathbb{E} \left[\sum_{k=1}^K 2s_{jk}s_k^* \right] + K \mathbb{E} [\sigma_k^2 [s_k^*]] + \frac{1}{K}. \end{aligned} \quad (38)$$

Since the first term can be written as $\sum_{k=1}^K s_{jk}^2 = K \frac{\sum_{k=1}^K s_{jk}^2}{K} = K \mu_k [s_{jk}^2]$ and $\mu_k [s_{jk}^2] - (\mu_k [s_{jk}])^2 = \sigma_k^2 [s_{jk}^2]$, the first two terms are K times the variance of s_{jk} , that is:

$$K\sigma_k^2 [s_{jk}] + \frac{1}{K} - \mathbb{E} \left[\sum_{k=1}^K 2s_{jk}s_k^* \right] + K \mathbb{E} [\sigma_k^2 [s_k^*]] + \frac{1}{K}. \quad (39)$$

Recall that s_k^* is random, while s_{jk} is known, meaning that: $\mathbb{E} \left[\sum_{k=1}^K 2s_{jk}s_k^* \right] = \sum_{k=1}^K (2s_{jk} \mathbb{E} [s_k^*])$, from which:

$$K\sigma_k^2 [s_{jk}] + \frac{1}{K} - \sum_{k=1}^K (2s_{jk} \mathbb{E} [s_k^*]) + K \mathbb{E} [\sigma_k^2 [s_k^*]] + \frac{1}{K}. \quad (40)$$

s_k^* is the realization of a randomly chosen set of policies among all the $j = 1, \dots, J+1$. The expected value of s_k^* is therefore $\left(\sum_{j=1}^J p_j s_{jk} + p_{j+1} s_{k,j+1} \right) = \frac{1}{J+1} \left(\sum_{j=1}^J s_{jk} + s_{k,j+1} \right)$ since $p_j = \frac{1}{J+1}, \forall j$. However, since $s_{k,j+1}$ is the mean of the other J elements, we can state that:

$$\mathbb{E} [s_k^*] = \frac{\sum_{j=1}^J s_{jk} + s_{k,j+1}}{J+1} = \frac{\sum_{j=1}^J s_{jk}}{J}. \quad (41)$$

Indeed:

$$\frac{\sum_{j=1}^J s_{jk} + s_{k,j+1}}{J+1} = \frac{\sum_{j=1}^J s_{jk} + \frac{1}{J} \sum_{j=1}^J s_{k,j}}{J+1} = \frac{J \sum_{j=1}^J s_{jk} + \sum_{j=1}^J s_{jk}}{J(J+1)} = \frac{(J+1) \sum_{j=1}^J s_{jk}}{J(J+1)} = \frac{\sum_{j=1}^J s_{jk}}{J}. \quad (42)$$

Since $E[s_k^*] = \frac{\sum_{j=1}^J s_{jk}}{J}$, and $E[\sigma_k^2[s_k^*]]$ is the expected value of all variances, constant across groups:

$$\begin{aligned} K\sigma_k^2[s_{jk}] + \frac{1}{K} - 2 \sum_{k=1}^K \left(s_{jk} \left(\frac{\sum_{j=1}^J s_{jk}}{J} \right) \right) + K \sum_{j=1}^J \frac{\sigma_k^2[s_{jk}]}{J} + \frac{1}{K} = \\ K\sigma_k^2[s_{jk}] + \frac{1}{K} - 2 \sum_{k=1}^K (s_{jk}\mu_j[s_{jk}]) + K\mu_j[\sigma_k^2[s_{jk}]] + \frac{1}{K}. \end{aligned} \quad (43)$$

Since $\sum_{k=1}^K (s_{jk}\mu_j[s_{jk}]) = K \frac{\sum_{k=1}^K (s_{jk}\mu_j[s_{jk}])}{K} = K\mu_k[s_{jk}\mu_j[s_{jk}]]$:

$$K\sigma_k^2[s_{jk}] - 2K\mu_k[s_{jk}\mu_j[s_{jk}]] + K\mu_j[\sigma_k^2[s_{jk}]] + \frac{2}{K}. \quad (44)$$

We know that $\sigma[X, Y] = \mu[XY] - \mu[X]\mu[Y]$ and therefore $\mu[XY] = \sigma[XY] + \mu[X]\mu[Y]$. If we replace X with s_{jk} and Y with $\mu_j[s_{jk}]$, we can write:

$$\begin{aligned} K\sigma_k^2[s_{jk}] - 2K\mu_k[s_{jk}\mu_j[s_{jk}]] + K\mu_j[\sigma_k^2[s_{jk}]] + \frac{2}{K} = \\ K\sigma_k^2[s_{jk}] - 2K(\sigma[s_{jk}, \mu_j[s_{jk}]] + \mu_k[s_{jk}]\mu_k[\mu_j[s_{jk}]] + K\mu_j[\sigma_k^2[s_{jk}]] + \frac{2}{K}. \end{aligned} \quad (45)$$

Since $\mu_k[s_{jk}] = \frac{1}{K}$ and $\mu_k[\mu_j[s_{jk}]] = \frac{\sum_{k=1}^K \sum_{j=1}^J (s_{jk})}{KJ} = \frac{J}{KJ} = \frac{1}{K}$:

$$K\sigma_k^2[s_{jk}] - 2K \left(\sigma[s_{jk}, \mu_j[s_{jk}]] + \frac{1}{K^2} \right) + K\mu_j[\sigma_k^2[s_{jk}]] + \frac{2}{K}. \quad (46)$$

After some very simple algebra, we get:

$$K(\sigma_k^2[s_{jk}] - 2\sigma[s_{jk}, \mu_j[s_{jk}]] + \mu_j[\sigma_k^2[s_{jk}]]) . \quad (47)$$

Recall that we initially multiplied the distance by $\sqrt{2}$; therefore, we conclude that:

$$\begin{aligned} E \left[L \left(\frac{d_j(\mathbf{s}_j, \mathbf{s}^*)}{\max d} \right) \right] \\ = \frac{K}{\sqrt{2}} (\sigma_k^2[s_{jk}] - 2\sigma[s_{jk}, \mu_j[s_{jk}]] + \mu_j[\sigma_k^2[s_{jk}]]) . \end{aligned} \quad (48)$$

A.3 Proof of proposition 1

Since the only random term of the utility function in eq.(8) is the vector \mathbf{s}^* , we can restate the expected utility function in eq.(12) as:

$$E[u_{ij}] = (1-t)y_i + (t\bar{y})^{1-\gamma} E[(1 - [L(d_j(\mathbf{s}_j, \mathbf{s}^*))])^\gamma]. \quad (49)$$

The maximization of expected utility in eq.(49) leads to the following first order condition:

$$t_{ij}^* = \frac{(1-\gamma)^{\frac{1}{\gamma}}}{\bar{y}} \left(\frac{\bar{y}}{y_i} \right)^{\frac{1}{\gamma}} (E[(1-L(d_j(\mathbf{s}_j, \mathbf{s}^*)))^\gamma])^{\frac{1}{\gamma}}. \quad (50)$$

Let us focus on $(E[(1-L(d_j(\mathbf{s}_j, \mathbf{s}^*)))^\gamma])^{\frac{1}{\gamma}}$. If we define $X = (1-L(d_j(\mathbf{s}_j, \mathbf{s}^*)))^\gamma$, we can re-write the random component of the optimal t_{ij}^* as $E[X]^{\frac{1}{\gamma}} = g(E[X])$, where, given $\gamma < 1$, g is a convex function. From Jensen's inequality we know that a convex transformation of a mean is necessarily less than or equal to the mean after convex transformation, i.e. $g(E[X]) \leq E[g(X)]$. Therefore, we can write $g(E[X]) = E[g(X)] - \delta$, where δ is an unknown non-negative scalar. In our case, since $X \in (0, 1)$, then $\delta \in (0, 1)$. Going back to our specific function we can thus write:

$$E[(1-L(d_j(\mathbf{s}_j, \mathbf{s}^*)))^\gamma]^{\frac{1}{\gamma}} = E[(1-L(d_j(\mathbf{s}_j, \mathbf{s}^*)))] - \delta = (1-\delta) - E[L(d_j(\mathbf{s}_j, \mathbf{s}^*))]. \quad (51)$$

Following lemma 1, we study the sign of the first derivatives of eq.(14). We know that mean income \bar{y} and the parameter γ are positive and that $(1-E[L(d_j(\mathbf{s}_j, \mathbf{s}^*))])^{\frac{1}{\gamma}}$ is always non-negative, as

$$0 \leq L(d_m(\mathbf{s}_m, \mathbf{s}^*)) = \left(\frac{\sqrt{\sum_{k=1}^K (s_{jk} - s^*)^2}}{\sqrt{2}} \right)^2 \leq 1, \text{ from the proof in A.1. Therefore, given that the loss function is bounded between 0 and 1, so it is its expected value.}$$

First,

$$\frac{\partial t_{ij}^*}{\partial y_i} = -\frac{1}{\gamma \bar{y} y_i} \gamma^{\frac{1}{\gamma}} \left(\frac{\bar{y}}{y_i} \right)^{\frac{1}{\gamma}} \left((1-\delta) - \frac{K}{\sqrt{2}} (\sigma_k^2[s_{jk}] - 2\sigma[s_{jk}, \mu_j[s_{jk}]] + \mu_j[\sigma_k^2[s_{jk}]]) \right) \leq 0. \quad (52)$$

Second,

$$\frac{\partial t_{ij}^*}{\partial \sigma_k^2[s_{jk}]} = -\frac{\gamma^{\frac{1}{\gamma}}}{\bar{y}} \left(\frac{\bar{y}}{y_i} \right)^{\frac{1}{\gamma}} \frac{K}{\sqrt{2}} < 0. \quad (53)$$

Third,

$$\frac{\partial t_{ij}^*}{\partial \mu_j[\sigma_k^2[s_{jk}]]} = -\frac{\gamma^{\frac{1}{\gamma}}}{\bar{y}} \left(\frac{\bar{y}}{y_i} \right)^{\frac{1}{\gamma}} \frac{K}{\sqrt{2}} < 0. \quad (54)$$

Fourth,

$$\frac{\partial t_{ij}^*}{\partial \sigma[s_{jk}, \mu_j[s_{jk}]]} = \frac{\gamma^{\frac{1}{\gamma}}}{\bar{y}} \left(\frac{\bar{y}}{y_i} \right)^{\frac{1}{\gamma}} \frac{2K}{\sqrt{2}} > 0. \quad (55)$$

A.4 Proof of corollary 1

We can write the solution in eq.(14) as follows:

$$t_{ij}^* = A (B(1-\gamma))^{\frac{1}{\gamma}} \quad (56)$$

where $A = \frac{((1-\delta) - \frac{K}{\sqrt{2}} (\sigma_k^2[s_{jk}] - 2\sigma[s_{jk}, \mu_j[s_{jk}]] + \mu_j[\sigma_k^2[s_{jk}]])}{\bar{y}} \geq 0$, $B = \frac{\bar{y}}{y_i} > 0$ and $\gamma \in (0, 1)$.

We need to compute the derivatives in order to study the shape of the function. The first derivative is

$$\frac{\partial t_{ij}^*}{\partial \gamma} = \frac{A(B(1-\gamma))^{\frac{1}{\gamma}} \left(-\frac{\gamma}{1-\gamma} - \ln(1-\gamma) - \ln(B) \right)}{\gamma^2}. \quad (57)$$

Since A , B , and γ are all positive, the sign of the derivative depends on the sign of the term in parenthesis: $\text{sgn} \left(\frac{\partial t_{ij}^*}{\partial \gamma} \right) = \text{sgn} \left(-\frac{\gamma}{1-\gamma} - \ln(1-\gamma) - \ln(B) \right)$. We define the second term in parentheses as D and we focus on this.

First of all, we show that $-\frac{\gamma}{1-\gamma} - \ln(1-\gamma) < 0$. The limit of the function at the lower bound is

$$\lim_{\gamma \rightarrow 0^+} -\frac{\gamma}{1-\gamma} - \ln(1-\gamma) = 0 \quad (58)$$

and its first derivative is always negative

$$\frac{\partial \left[-\frac{\gamma}{1-\gamma} - \ln(1-\gamma) \right]}{\partial \gamma} = \frac{\gamma - 2}{(1-\gamma)^2} < 0 \quad (59)$$

since $\gamma \in (0, 1)$. Moreover, $-\ln(B) \leq 0$, $\forall B \geq 1$. Therefore, $D < 0$ for any $B \geq 1$, that is, for all individuals poorer than (or equal to) the mean income.

For individuals richer than the mean ($B < 1$), the $\text{sgn} \left(\frac{\partial t_{ij}^*}{\partial \gamma} \right) < 0$ only if their income is such that $B < e^{-\frac{\gamma}{1-\gamma} - \ln(1-\gamma)}$, and positive otherwise.

Summarizing, as we can see in figure 3, we found that:

- for $B \geq 1 \Rightarrow \frac{\partial t_{ij}^*}{\partial \gamma} < 0 \forall \gamma \in (0, 1) \Rightarrow t_{ij}^*$ is decreasing with γ ;
- for $B < 1 \Rightarrow \frac{\partial t_{ij}^*}{\partial \gamma}$ is increasing up to $B = e^{-\frac{\gamma}{1-\gamma} - \ln(1-\gamma)}$ and decreasing afterwards.

A.5 Proof of lemma 2

Within any group j , we can rewrite individuals' preferences as follows:

$$u_i(t; y_i) = \Phi(t) + \Psi(t)F(y_i) \quad (60)$$

where $F(y_i) = y_i$ is monotonic in the income type and $\Phi(t) = (t\bar{y})^{(1-\gamma)}(1-L)^\gamma$ and $\Psi(t) = (1-t)$ are common to all voters.

We can show that for any two individuals a, b such that $y_a < y_b$, and for any two policy alternatives t and t' such that $t < t'$, if $u_a(t) > u_a(t')$ then $u_b(t) > u_b(t')$, but if $u_b(t) < u_b(t')$ then $u_a(t) < u_a(t')$.

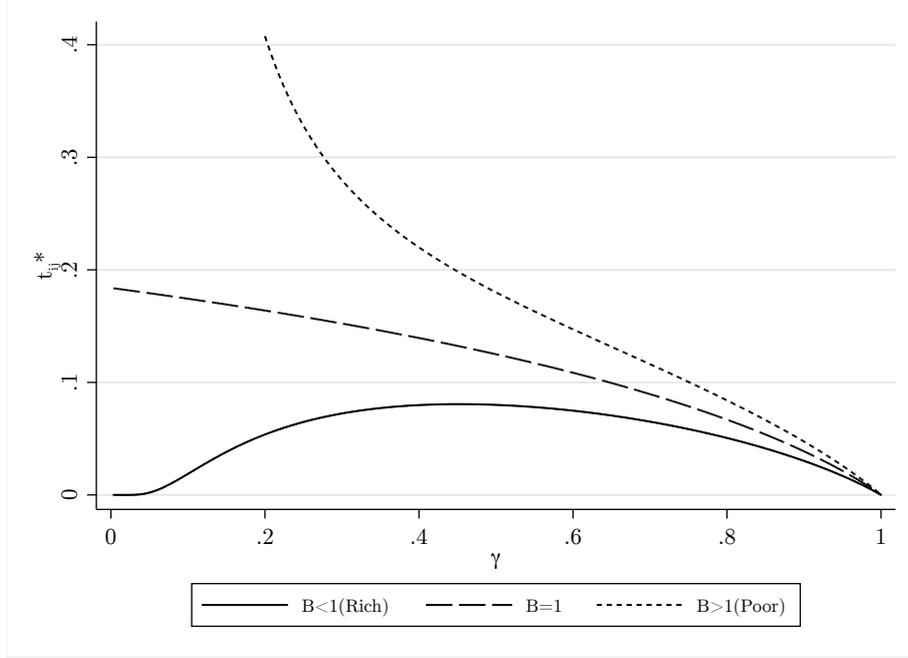


Figure 3: Optimal tax rate for different levels of B

Indeed, since $u_a(t) > u_a(t')$, then:

$$\Phi(t') + \Psi(t')F(y_a) < \Phi(t) + \Psi(t)F(y_a) \quad (61)$$

$$F(y_a) [\Psi(t') - \Psi(t)] < \Phi(t) - \Phi(t') \quad (62)$$

thus

$$F(y_a) > \frac{\Phi(t) - \Phi(t')}{[\Psi(t') - \Psi(t)]} \text{ as } \Psi(t) > \Psi(t'). \quad (63)$$

Let assume that the condition is not satisfied. Therefore, for voter b :

$$\Phi(t') + \Psi(t')F(y_b) > \Phi(t) + \Psi(t)F(y_b) \quad (64)$$

thus

$$F(y_b) < \frac{\Phi(t) - \Phi(t')}{[\Psi(t') - \Psi(t)]} \text{ as } \Psi(t) > \Psi(t'). \quad (65)$$

Combining eq.(63) and eq.(65), we get:

$$F(y_b) < \frac{\Phi(t) - \Phi(t')}{[\Psi(t') - \Psi(t)]} < F(y_a) \quad (66)$$

Given the monotonicity of $F(y_i)$, conditions (63) and (65) are incompatible, thus $u_a(t) > u_a(t')$ implies $u_b(t) > u_b(t')$ and by the same means can be easily shown that $u_b(t) < u_b(t')$ then $u_a(t) < u_a(t')$.

A.6 Proof of lemma 3

The existence of a Condorcet winner in a similar context has been proven by De Donder (2013). Following the same line of reasoning and applying to our setting, let us denote the set of most preferred tax rate t of individuals of income y belonging to group J by $M_j(t, y) = \arg \max_t u_{ij}(t, y)$. $M_j(t, y)$ is deemed to be continuous in $y \forall j = 1, \dots, J$ and $\forall t$ in the feasible set $(0, 1)$.

Let us define $\bar{\tau}_j$ and $\underline{\tau}_j$ as the highest and the lowest most preferred tax rate of group j , respectively, and thus, $\bar{\tau}$ and $\underline{\tau}$ as the minimax and the maximin of the most preferred tax rate across groups: $\underline{\tau} = \max \underline{\tau}_j < \bar{\tau} = \min \bar{\tau}_j$.

$[\underline{\tau}, \bar{\tau}]$ is then the largest interval where there are some voters' blisspoints in all groups. Therefore, given that $M_j(t, y)$ is continuous and decreasing in y in the range $[\underline{\tau}, \bar{\tau}]$ we know from the intermediate value theorem that there exist at least one $t_{CW}^* \in [\underline{\tau}, \bar{\tau}]$ such that $\sum \mu_j F_j \left(y_j^{t_{CW}^*} \right) = \frac{1}{2}$.

Since $t_{CW}^* \in [\underline{\tau}, \bar{\tau}]$, we know that there exists at least one individual with income type $y_j^{t_{CW}^*}$ in all groups j such that (s)he most prefers t^* to any other feasible option, i.e. $u_{ij}(t_{CW}^*, y_j^{t_{CW}^*}) \geq u_{ij}(t^*, y_j^{t_{CW}^*}), \forall t^* \in (0, 1)$.

Assume that $t'^* > t_{CW}^*$. Recalling lemma 2, this means that, in each group j , $u_{ij}(t_{CW}^*, y) \geq u_{ij}(t^*, y), \forall y \geq y_j^{t_{CW}^*}$. In turn, this guarantees that at least $1 - F_j \left(y_j^{t_{CW}^*} \right)$ in each group j will support t_{CW}^* when voted against t^* . Given how we defined t_{CW}^* this support aggregates one half of the of the individuals over all groups so that it cannot be defeated by majority voting by t^* . The opposite case can be proven likewise.

In our context the validity of the continuity and monotonicity assumptions can be represented as in figures 4 and 5, where we present the results of a simulation run on a randomly generated sample of 40,000 individuals, randomly assigned to four preferences-type groups.³⁴

A.7 Proof of proposition 2

Let us consider a transfer of magnitude x from individual 1 to individual 2, in group A , with income respectively y_1 and y_2 where $y_1 > y_m > y_2$ and y_m is the median income on the overall population. For the median income to increase, x has to be such that $y_1 - x > y_2 + x > y_m$.³⁵ Define $y'_1 = y_1 - x$

³⁴In this simulation $\gamma = .5$.

³⁵Notice that a Pigou-Dalton transfer rules out re-ranking.

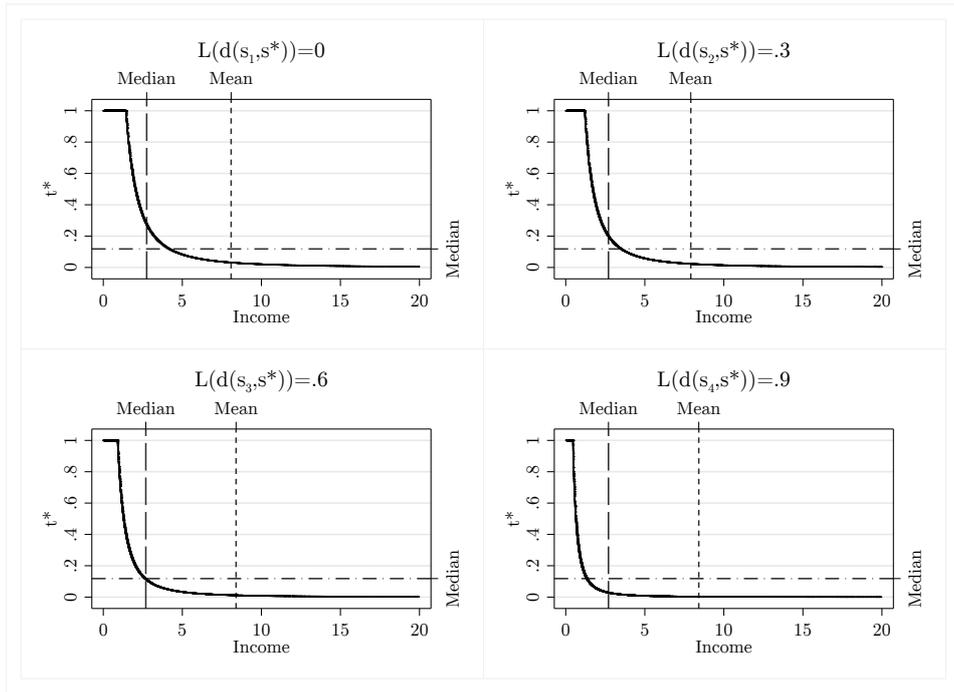


Figure 4: Optimal tax rate for different type

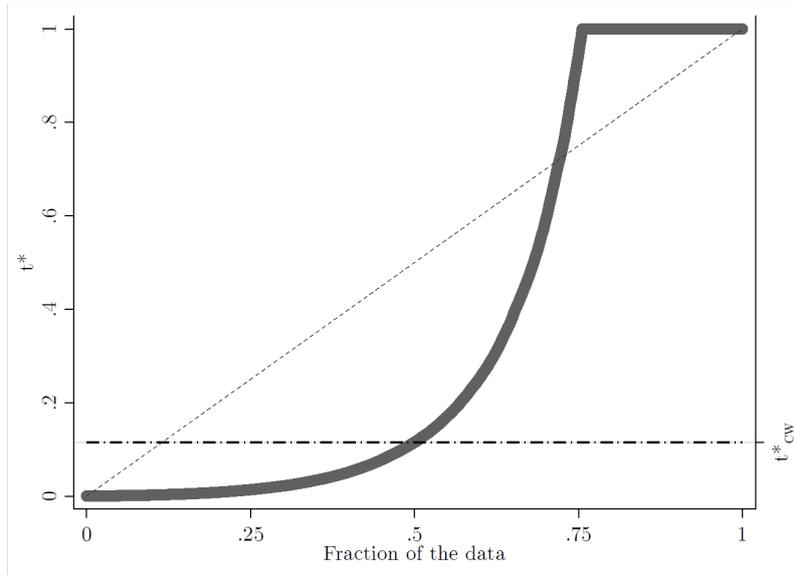


Figure 5: Optimal tax rate

and $y'_2 = y_2 + x$, recalling proposition 1 we know that:

$$t_1^{*'} = t_1^* + \left[\left(\frac{1}{y_1 - x} \right)^{\frac{1}{\gamma}} - \left(\frac{1}{y_1} \right)^{\frac{1}{\gamma}} \right] \Phi \Rightarrow t_1^{*'} > t_1^* \quad (67)$$

$$t_2^{*'} = t_2^* + \left[\left(\frac{1}{y_2 + x} \right)^{\frac{1}{\gamma}} - \left(\frac{1}{y_2} \right)^{\frac{1}{\gamma}} \right] \Phi \Rightarrow t_2^{*'} < t_2^* \quad (68)$$

$$t_2^{*'} = t_1^{*'} + \left[\left(\frac{1}{y_2 + x} \right)^{\frac{1}{\gamma}} - \left(\frac{1}{y_1 - x} \right)^{\frac{1}{\gamma}} \right] \Phi \Rightarrow t_2^{*'} > t_1^{*'} \quad (69)$$

$$t_2^{*'} = t_m^* + \left[\left(\frac{1}{y_2 + x} \right)^{\frac{1}{\gamma}} - \left(\frac{1}{y_m} \right)^{\frac{1}{\gamma}} \right] \Phi \Rightarrow t_2^{*'} < t_m^* \quad (70)$$

with $\Phi = (\bar{y})^{\frac{1}{\gamma}-1} (1-\gamma)^{\frac{1}{\gamma}} \left((1-\delta) - \frac{K}{\sqrt{2}} (\sigma_k^2 [s_{Ak}] - 2\sigma [s_{Ak}, \mu_A [s_{Ak}]] + \mu_A [\sigma_k^2 [s_{Ak}]])) \right) > 0$ if $E \left[L \left(\frac{d_A(\mathbf{s}_A, \mathbf{s}^*)}{\max d} \right) \right] < (1-\delta)$. Therefore, it is possible to order the preferred tax rates as follows:

$$t_2^* > t_m^* > t_2^{*'} > t_1^{*'} > t_1^* \quad (71)$$

Given lemma 2, we know that:

$$F_A(y^{t_2^{*'}}) < F_A(y^{t_m^*}) < F_A(y^{t_2^*}) < F_A(y^{t_1^{*'}}) < F_A(y^{t_1^*}) \quad (72)$$

Denote t_{CW}^{*a} the optimal tax rate before the change. If $t_1^* \notin [t_{CW}^{*a} - \frac{\partial t_1^*}{\partial x}, t_{CW}^{*a}]$, we may be in one of the following cases:

1. If $t_2^* < t_{CW}^{*a}$, or $t_{CW}^{*a} < t_1^{*'}$, or $t_2^{*' > t_{CW}^{*a} > t_1^{*'}$, nothing changes as if before $F_A(y^{t_{CW}^{*a}}) = \phi$ was supporting t_{CW}^{*a} when voted against a lower t^* , now:

$$\sum \nu_j F_j(y_j^{t_{CW}^{*a}}) = \nu_A F_A(y^{t_{CW}^{*a}}) + \sum_{j \neq A} \nu_j F_j(y_j^{t_{CW}^{*a}}) \quad (73)$$

$$= \nu_A \phi + \sum_{j \neq A} \nu_j F_j(y_j^{t_{CW}^{*a}}) = \frac{1}{2} \quad (74)$$

2. If instead $t_2^{*' \leq t_{CW}^{*a} < t_2^*$, after the transfer $\phi' < \phi$ will support t_{CW}^{*a} when voted against a lower t^* , as individual 2 now has moved to the right of individual with income $y^{t_{CW}^{*a}}$, and individual 1 did not change position in the ranking of individuals income. Thus, after the transfer:

$$\sum \nu_j F_j(y_j^{t_{CW}^{*a}}) = \nu_A \phi' + \sum_{j \neq A} \nu_j F_j(y_j^{t_{CW}^{*a}}) < \frac{1}{2} \quad (75)$$

Therefore given lemma 3 and the definition of Condorcet Winner, $t_{CW}^{*p} < t_{CW}^{*a}$.³⁶ As \bar{y} remain

³⁶If $t_{CW}^{*a} - \frac{\partial t_1^*}{\partial x} < t_1^* < t_{CW}^{*a}$ we have that:

unchanged, the total provision of public goods $G_p^* = t_{CW}^{*p} \bar{y}$ is lower.

A.8 Proof of proposition 3

Suppose that $\bar{\omega}$ decreases, given a change in the preferences of any two groups that leaves the average preferences for the type of public goods unchanged.³⁷ Define ω_j^a as the covariance before the change and ω_j^p as the covariance after the change (where a and p stand for *ante* and *post*, respectively). Define also $I_1, I_2 \subset \{1 \dots J\}$ with $I_1 \cup I_2 = \{1 \dots J\}$, s.t. $\omega_j^p < \omega_j^a$ if $j \in I_1$ and $\omega_j^p = \omega_j^a$ if $j \in I_2$, then:

$$\begin{aligned} \sum_{j \in I_1} (\omega_j^a - \omega_j^p) \nu_j &> \sum_{j \in I_2} (\omega_j^p - \omega_j^a) \nu_j \\ \sum_{j \in I_1} (\omega_j^a - \omega_j^p) \nu_j &> 0. \end{aligned} \quad (76)$$

Thus, recalling proposition 1, we know that:

$$t_{ij}^{*p} = t_{ij}^{*a} - \frac{(1-\gamma)^{\frac{1}{\gamma}}}{\bar{y}} \left(\frac{\bar{y}}{y_i} \right)^{\frac{1}{\gamma}} \frac{2K}{\sqrt{2}} (\omega_j^a - \omega_j^p), \forall i. \quad (77)$$

For each individual of income y_i in group $j \in I_1$, $t_{ij}^{*p} < t_{ij}^{*a}$. Thus, for any t :

$$(1 - F_j(y_j^t))_{post} > (1 - F_j(y_j^t))_{ante} \text{ if } j \in I_1 \quad (78)$$

$$(1 - F_j(y_j^t))_{post} = (1 - F_j(y_j^t))_{ante} \text{ if } j \in I_2. \quad (79)$$

Following eq.(76) and eq.(77):

$$\sum_{j \in I_1} \nu_j [(1 - F_j(y_j^t))_{post} - (1 - F_j(y_j^t))_{ante}] > \sum_{j \in I_2} \nu_j [(1 - F_j(y_j^t))_{ante} - (1 - F_j(y_j^t))_{post}] \quad (80)$$

Denote t_{CW}^{*a} the optimal tax rate before the change. Recalling lemma 2 we know that at least $1 - F_j(y_j^{t_{CW}^{*a}})$ in each group j were supporting t_{CW}^{*a} when voted against a higher t^* . We also know by definition of Condorcet Winner that $\sum \nu_j F_j(y_j^{t_{CW}^{*a}}) = \frac{1}{2}$.

$$\begin{aligned} \frac{\partial t_1^*}{\partial x} &= \left[\left(\frac{1}{y_1 - x} \right)^{\frac{1}{\gamma}} - \left(\frac{1}{y_1} \right)^{\frac{1}{\gamma}} \right] \Phi \\ t_1^* + \frac{\partial t_1^*}{\partial x} &= t_1^{*'} \Rightarrow t_1^* = t_1^{*'} - \frac{\partial t_1^*}{\partial x} \Rightarrow t_1^{*'} > t_1^* > t_{CW}^{*a} \\ &\Rightarrow t_1^{*'} > t_{CW}^{*a} > t_1^* \end{aligned}$$

so that, after the transfer, $\phi'' > \phi$ will support t_{CW}^{*a} when voted against a lower t^* , as individual 1 now has moved to the left of individual with income $y^{t_{CW}^{*a}}$. Thus, after the transfer:

$$\sum \nu_j F_j(y_j^{t_{CW}^{*a}}) = \nu_A \phi'' + \sum_{j \neq A} \nu_j F_j(y_j^{t_{CW}^{*a}}) > \frac{1}{2} \Rightarrow t_{CW}^{*p} > t_{CW}^{*a}.$$

³⁷This implies that the preferences move towards the extremes.

Then:

$$\begin{aligned} \sum_{j \in I_1} \nu_j \left(1 - F_j \left(y_j^{t_{CW}^{*a}}\right)\right)_{post} + \sum_{j \in I_1} \nu_j \left(1 - F_j \left(y_j^{t_{CW}^{*a}}\right)\right)_{ante} + \\ + \sum_{j \in I_2} \nu_j \left(1 - F_j \left(y_j^{t_{CW}^{*a}}\right)\right)_{post} > \sum_{j \in I_2} \nu_j \left(1 - F_j \left(y_j^{t_{CW}^{*a}}\right)\right)_{ante} \end{aligned} \quad (81)$$

$$\sum_j \nu_j \left(1 - F_j \left(y_j^{t_{CW}^{*a}}\right)\right)_{post} > \sum_j \nu_j \left(1 - F_j \left(y_j^{t_{CW}^{*a}}\right)\right)_{ante} \quad (82)$$

$$\sum_j \nu_j F_j \left(y_j^{t_{CW}^{*a}}\right)_{post} < \sum_j \nu_j F_j \left(y_j^{t_{CW}^{*a}}\right)_{ante} \quad (83)$$

from which we get: $\sum_j \nu_j \left(F_j \left(y_j^{t_{CW}^{*a}}\right)\right)_{post} < 1/2$ and, given lemma 3 and the definition of Condorcet Winner, $t_{CW}^{*p} < t_{CW}^{*a}$.

As \bar{y} remain unchanged, the total provision of public goods $G_p^* = t_{CW}^{*p} \bar{y}$ is lower.