A Contribution to the Positive Theory of Indirect Taxation

Introduction

Taxation has been a much-discussed topic in the economic literature. From previous contributions we know that maximum efficiency is achieved via lump-sum taxes, because they nullify the excess burden of taxation. Nevertheless, such taxation is not desirable because it is considered unjust. Thus, in order to achieve equity goals, benevolent governments must accept that taxpayers distort their economic choices in order to escape the burden of taxation. As a consequence market failures arise, such as a reduction in the labour supply.

In democratic societies, allocation choices for the public sector are made through voting, and through the actions of elected representatives. Economic outcomes must be evaluated in a broader context, one that allows for the possibility of setting tax rates at a candidate's discretion, together with the collective nature of existing political institutions that must be relied on to take decisions on fiscal issues.

Depending on whether the political decision-making mechanism is considered by the analysis or not, the literature on taxation is divided in two main streams of research: the normative and the positive approach.

The *normative approach* seeks efficiency-oriented solutions considering the existence of a benevolent social-planner who avoids any concern regarding collective action. A tenet achieved by this analysis states that a tax system is efficient if it minimizes the total excess burden of raising a given amount of revenue. A typical application of this approach is the inverse elasticity rule associated with Ramsey (1927) who analysed an economy with sales taxes imposed on different commodities. This work concludes by affirming that, in order to minimise the excess burden, higher tax rates should be levied on commodities which have a relatively inelastic demand in the range of the demand function with respect to commodities whose demands are more elastic, so as to raise a given total revenue while avoiding, as far as possible, the excess burdens associated with the substitution away from commodities whose after-tax price has risen. Furthermore, a version of Ramsey's rule modified by Diamond (1975) envisions a planner who takes distributional goals into account, derived from a welfare function where weights are attached to the welfare of different individuals. In order to maximise social welfare the planner equalizes distributionally weighted marginal excess burdens per dollar raised across available tax bases.

Otherwise, the *positive approach* studies collective choice processes and their influence on political and economic outcomes. Works belonging to this second strand not only focus on market failures but also on political failures. Two recent works (Polo, 1998; Svensson, 1997) focus on the role of political competition, where candidates propose policy platforms in order to maximise the probability of winning elections (or the number of votes), under conditions of uncertainty about voters' political preferences. Since individuals aim to maximise their utility influenced by public policies, they react positively to an increase in the amount of public goods and negatively to the payment of taxes and to welfare losses caused by taxation. The maximisation of the probability of winning is achievable if politicians design an equilibrium tax structure which equalises the change in opposition per marginal tax dollar raised across groups.

It is essential to understand equilibrium outcomes produced by well-functioning political processes, and to examine how such outcomes change when imperfections become part of the collective action. This implies that we need a model of *collective choice* as our starting point that allows us to study and demonstrate the existence and stability of political equilibria and to examine the nature of specific equilibrium policies or outcomes. *Probabilistic voting Theory* is able to accomplish this goal, since the resulting Nash equilibria amongst parties are Pareto-optimal. (Hettich and Winer 1999, Chapter 4.) However, the need to take this basic analytical step is not tied to the use of a particular framework; rather, it arises from the fundamental nature of normative analysis itself. Imperfections in private markets have their counterparts in failures of the political process. As a consequence, we must focus on the operation of the collective decision mechanism in order to identify those features that cause it to operate imperfectly. Not only must we begin by modelling a political process that leads to an optimal allocation of resources, but it is also necessary to determine tailored tax policies that will be part of the political equilibrium. Once this has been accomplished, we can then extend the examination to specific imperfections of collective decision-making and trace out their implications for structure of the tax policies. Few authors writing on taxation have concerned themselves with this research programme but unless it is carried out, economists cannot accomplish an analysis of tax policy failures that has the same force as does the well-known work on private market imperfections.

Finally, when we also introduce equity goals into the analysis we must deal with welfare state programmes which transfer resources amongst groups. A question naturally arises: to what extent do voters' preferences influence these programmes? A standard model of redistributive taxation may be found in Romer (1975), Roberts (1977) and Meltzer and Richard (1981); if we suppose that both individual productivity and the availability of leisure differ, it can be demonstrated that political candidates commit to the policy preferred by the median voter. In equilibrium, taxes are higher the greater the distance between median and mean income, a specific measure of income inequality. Nevertheless, in these models the single-peakedness condition, which is necessary for an equilibrium to exist, is very likely to fail, as the authors demonstrated.

In this paper I will analyse how self-interested governments set their taxation policies in a probabilistic voting model. Candidates are pure voter-seekers and aim to maximise the probability of winning elections. Society is divided into groups who assign different weights to consumption of goods, based on their preferences; that is, they have different levels of single-mindedness. Results show how in equilibrium candidates must satisfy the most powerful groups, which do not necessarily represent the median voter, or the middle class, but may be located at extreme positions on the income scale. The introduction of a probabilistic voting model characterized by single-minded groups changes the classic result of median voter models because it is no longer the position on the income scale which drives the choice of candidates but rather the ability of groups to focus on issues they prefer. This ability enables them to achieve a strong political power which candidates cannot help going along with, because they would lose the elections otherwise. Escaping the more single-minded groups is impossible for politicians, as long as they are prisoners of their own self-interest. In this vicious circle, the function of taxation is reduced to one of merely protecting private interests. Results of this model represent the antithesis of classic normative models. Taxation loses its pro-active role as a mechanism to redistribute resources from the rich to the poor or to supply public goods and becomes only a way to win elections, no matter if this means protecting even the richest components of society.

Results of this model would also provide a possible explanation for the existence of indirect taxation. This is an old issue addressed by Atkinson and Stiglitz (1976) who demonstrated that the optimal direct-cum-indirect tax problem puts all commodity taxes to zero and raises everything through income tax. More recent works by Laroque (2005) and Kaplow (2006) demonstrated that Atkinson and Stiglitz's result is even stronger because there appears to be no role for taxes on commodities even in the presence of a non-optimally designed tax structure. Then, why are Governments so reluctant to abolish indirect taxation? If we consider that direct taxation is progressive in practice whilst indirect taxation is mildly regressive[10] it might be perfectly possible to see the interest of powerful interest groups in preventing a substantial shift from indirect to direct taxation. If more single-minded groups are found amongst the richest component of society and are not favourable to increase the weight of direct taxation with respect to indirect taxation, Governments could not undertake this reform. As a consequence income distribution is less egalitarian.

A model of indirect taxation

I consider a society divided into H groups, indexed by $h = 1, ..., H$. Groups have size f^h and their members are exactly alike. Two political candidates, $j = D, R$, run for an election. Both candidates have an ideological label (for example, Democrats and Republicans) which is exogenously given. Voters' welfare depends on two components; the first is deterministic and it is represented by the consumption of goods, whilst the second is stochastic and derives from the personal attributes of candidates.

Each individual in group h derives his consumption from n goods x_i^h, indexed by $i = 1, ..., n$. Consumption is a function of the tax policy chosen by candidates and it is perfectly observable. The deterministic component of welfare may be written in a logarithmic fashion, $\sum_{i=1}^{n} \psi_i^h \log x_i^h$, where ψ_i^h represents the weight (importance) that group h attaches to good i.

The stochastic component is denoted by $D^R \cdot \left(\xi^h + \varsigma\right)$, where the indicator function

$$D^R = \begin{cases} 1 & if\ R\ wins \\ 0 & if\ D\ wins \end{cases}$$

The random variable $\varsigma \lesseqgtr 0$ reflects candidate R's popularity amongst voters and it is realized between the announcement of policies and elections. It is not idiosyncratic and it is uniformly distributed as

$$\varsigma \sim U\left[-\frac{1}{2},\frac{1}{2}\right]$$

with mean zero. Otherwise, $\xi^h \lesseqgtr 0$ represents an idiosyncratic random variable which measures voters' preferences for D. It is not perfectly observable by candidates and it is uniformly distributed as

$$\xi^h \sim U\left[-\frac{1}{2s^h},\frac{1}{2s^h}\right]$$

again with mean zero and density s^h.

Therefore, a representative individual in group h maximizes the following utility function:

$$U^h = \sum_{i=1}^{n} \psi_i^h \log x_i^h + D^R \cdot \left(\xi^h + \varsigma\right) \tag{27}$$

under the following budget constraint

$$\sum_{i=1}^{n} q_i^j x_i^h = M^h \tag{28}$$

where M^h is the income of any individual in group h. I denote by $q_i^j = p_i + t_i^j$ the consumption price of good i, by p_i the fixed production price[11] and by t_i^j the unit excise tax levied by candidate j on good i. Hence, $\overrightarrow{x} = [x_1, ..., x_n] \in X \subset \mathbb{R}^n$ denotes the vector of consumption, $\overrightarrow{q^j} = \left[q_1^j, ..., q_n^j\right] \in Q^j \subset \mathbb{R}^n$ the vector of consumption prices, $\overrightarrow{p} = [p_1, ..., p_n] \in P \subset \mathbb{R}^n$ the vector of production prices and $\overrightarrow{t^j} = \left[t_1^j, ..., t_n^j\right] \in T^j \subset \mathbb{R}^n$ the vector of tax rates.

I introduce two important definitions:

Definition 17 *group A is said to be more single-minded than group B with respect to good i if the weight assigned by A to i is greater than the weight assigned by B. That is, if $\psi_i^A > \psi_i^B$.*

This definition states that groups, in attaching weights to goods, are less or more willing to substitute a good with another[12], depending on the preferences they have for every good. As a consequence, there exist some goods whose consumption is more claimed by groups, because its reduction would affect their welfare in a more tangible way.

Definition 18 *group A is said to be more politically powerful than group B if its density is higher than B's. That is if $s^A > s^B$.*

In this case the political power of a group must be intended as the ability of influencing candidates' choices, when they have to take decisions over a policy. In traditional probabilistic voting models this power is expressed by a density function which captures the distribution of the constituency.

The demand for goods

Individuals maximize 27 subject to 28. The Lagrangian function for a representative individual in group h is

$$\mathcal{L}^h = \sum_{i=1}^{n} \psi_i^h \log x_i^h + D^R \cdot \left(\xi^h + \varsigma\right) + \lambda^h \left(M^h - \sum_{i=1}^{n} q_i^j x_i^h\right)$$

The set of first order conditions is

$$\begin{pmatrix} \frac{\partial \mathcal{L}^1}{\partial x_1^1} & \cdots & \frac{\partial \mathcal{L}^H}{\partial x_1^H} \\ \vdots & \ddots & \\ \frac{\partial \mathcal{L}^1}{\partial x_n^1} & & \frac{\partial \mathcal{L}^H}{\partial x_n^H} \\ \frac{\partial \mathcal{L}^1}{\partial \lambda^1} & & \frac{\partial \mathcal{L}^H}{\partial \lambda^H} \end{pmatrix} = \begin{pmatrix} \frac{\psi_1^1}{x_1^1} = \lambda^1 q_1^j & \cdots & \frac{\psi_1^H}{x_1^H} = \lambda^H q_1^j \\ \vdots & \ddots & \vdots \\ \frac{\psi_n^1}{x_n^1} = \lambda^1 q_n^j & & \frac{\psi_n^H}{x_n^H} = \lambda^H q_n^j \\ \sum_{i=1}^{n} q_i^j x_i^1 = M^1 & & \sum_{i=1}^{n} q_i^j x_i^H = M^H \end{pmatrix}$$

The resolution of first order conditions yields the Marshallian demand functions $x_i^{h*} = \frac{\psi_i^h M^h}{q_i^j}$ and the indirect utility functions

$$V\left(x\left(q_i^j, M^h\right)\right) = \sum_{i=1}^{n} \psi_i^h \log \frac{\psi_i^h M^h}{q_i^j} + D^R \cdot \left(\xi^h + \varsigma\right)$$

Political Competition

I consider now the problem of candidates. What distinguishes this contribution from previous taxation models in Political Economy is `the existence of a new setting where probabilistic voting and single-mindedness theory fuse together`. In the classic literature governments had always been considered as benevolent planners, who aimed to maximise a Social Welfare Function whose characteristics depended on the preferences of society for equity, perfectly mirrored by policy-maker's preferences. Weights attached to the utility of different agents were higher for the poor and lower for the rich.

Instead, in this model politicians are considered as voter-seekers who aim to maximise the probability of winning elections by choosing an optimal policy vector $\overrightarrow{t^j}$. Each voter in group h votes for R if and only if R's policy provides him with a greater utility than that provided by D. That is a generic voter ι votes for D if and only if:

$$V^h\left(\overrightarrow{t^R}\right)+\xi^{\iota,h}+\varsigma \geq V^h\left(\overrightarrow{t^D}\right) \qquad \forall\iota \tag{29}$$

where $V^h\left(\overrightarrow{t^j}\right)$ represents the indirect utility function which group h derives under the vector of policies chosen by candidate j. Within each group there is a fraction of *swing voters*, denoted by $\widehat{\iota}$, represented by those individuals who are indifferent between D or R. For these voters equation 29 holds with equality:

$$\xi^{\widehat{\iota},h}=V^h\left(\overrightarrow{t^D}\right)-V^h\left(\overrightarrow{t^R}\right)-\varsigma \tag{30}$$

Otherwise, voter ι votes for D if $\xi^{\iota,h}<\xi^{\widehat{\iota},h}$ and for R if $\xi^{\iota,h}>\xi^{\widehat{\iota},h}$. Swing voters are pivotal, since even a small change in the policy vector makes them no longer indifferent to candidates and it forces them to vote for one of two.

The probability of winning elections for candidate j is given by

$$p^j\left(\overrightarrow{t^j},\overrightarrow{t^{-j}}\right)=\frac{1}{2}+\frac{d}{s}\sum_{h=1}^{H}f^h s^h\left[V\left(\overrightarrow{t^j}\right)-V\left(\overrightarrow{t^{-j}}\right)\right] \tag{31}$$

where $V\left(\overrightarrow{t^j}\right):=V\left(p_i+t_i^j,M^h\right)$ and $s:=\sum\limits_h s^h f^h$.

Axiom 19 *the density function of a group is twice differentiable and monotonically increasing in the level of consumption of goods. That is* $s^h=s(x_1^h,...,x_n^h)$, *with* $\frac{\partial s^h}{\partial x_i^h}>0$.

This axiom brings something new with respect to the traditional probabilistic voting models, where the density function was always treated as a constant. The idea to make the density function depend on the consumption of goods is new and deserve to be explained. The classic literature on probabilistic voting models (Persson and Tabellini (2000), Lindbeck and Weibull (1987), and Coughlin (1992)) has always assumed that preferences of voters for political candidates have a distribution where the density function is a constant. Instead, in this model, the density function is increasing in the level of consumption which is in turn affected by the vector of policies. Candidates realize that, should they change a policy, the welfare of groups would change and their political power, captured by the density function, accordingly. Hence, a nexus is created amongst governments' policies, voters' consumption and political power of groups which eventually affects elections' outcome.

Furthermore, as suggested by Lindbeck and Weibull (1987), I assume that

Remark 20 *there exists a balanced-budget constraint*

$$\sum_h f^h \sum_i t_i^j x\left(q_i^j, M^h\right) = 0 \tag{32}$$

which coerces the government to redistribute via transfers all the tax revenues collected.

This assumption allows us to treat the model as `purely redistributive`, which has the advantage of clearly showing the redistributive effects, neglecting any concern about the existence of public expenditure. In turn, equation 32 says that all the revenues collected via taxation are used to redistribute wealth amongst groups. As a consequence, if some groups are better off by the achievement of a net transfer, some others must necessarily be worse off because they have to bear the entire payment of these transfers.

Finally, notice how this political game is a two-person, constant-sum and symmetric game where a pair of policies is an `equilibrium pair` if and only if it is a saddle point for

$$\Gamma = \left(T^D, T^R; p^D, 1 - p^D\right)$$

The equilibrium

To solve the problem I write the Lagrangian function for candidate D (the same holds for candidate R):

$$\mathcal{L}^D = \frac{1}{2} + \frac{d}{s}\sum_h f^h s^h \left[V\left(\overrightarrow{t^D}\right) - V\left(\overrightarrow{t^R}\right)\right] + \mu^D \left(\sum_h f^h \sum_i t_i^D x\left(q_i^D, M^h\right)\right) \tag{33}$$

The set of first order conditions is:

$$\begin{cases} \frac{\partial \mathcal{L}^D}{\partial q_1^D} = \frac{\partial}{\partial q_1^D}\left(\frac{1}{s}\right) d\sum_h f^h s^h \left[V\left(\overrightarrow{t^D}\right) - V\left(\overrightarrow{t^R}\right)\right] + \frac{d}{s}\sum_h f^h \frac{\partial s^h}{\partial q_1^D}\left[V\left(\overrightarrow{t^D}\right) - V\left(\overrightarrow{t^R}\right)\right] + \\ + \frac{d}{s}\sum_h \frac{\partial V^h}{\partial q_1^D} f^h s^h + \mu^D \left(\sum_o t_o^D \sum_h f^h \frac{\partial x_o^h}{\partial q_o^D} + x_o^h\right) = 0 \qquad o \neq i \\ \vdots \\ \frac{\partial \mathcal{L}^D}{\partial q_n^D} = \frac{\partial}{\partial q_n^D}\left(\frac{1}{s}\right) d\sum_h f^h s^h \left[V\left(\overrightarrow{t^D}\right) - V\left(\overrightarrow{t^R}\right)\right] + \frac{d}{s}\sum_h f^h \frac{\partial s^h}{\partial q_n^D}\left[V\left(\overrightarrow{t^D}\right) - V\left(\overrightarrow{t^R}\right)\right] + \\ + \frac{d}{s}\sum_h \frac{\partial V^h}{\partial q_n^D} f^h s^h + \mu^D \left(\sum_o t_o^D \sum_h f^h \frac{\partial x_o^h}{\partial q_o^D} + x_o^h\right) = 0 \\ \frac{\partial \mathcal{L}^D}{\partial \mu^D} = \sum_h f^h \sum_i t_i^D x\left(q_i^D, M^h\right) = 0 \end{cases}$$

In this game, the existence of an equilibrium is guaranteed by the concavity of the utility functions. This proof was used for voting models by Hinich et al. (1973). An easy proof is also provided for a special case of redistributive models by Coughlin (1985).

Proposition 21 *In a constant-sum game an equilibrium is achieved via a convergence of policy; that is:* $\overrightarrow{t^{D*}} = \overrightarrow{t^{R*}}$.

Proof. First of all, we have defined Γ as a constant-sum game, since $p^R\left(\overrightarrow{t^D}, \overrightarrow{t^R}\right) = 1 - p^D\left(\overrightarrow{t^D}, \overrightarrow{t^R}\right)$. Suppose now that the pair $\left(\overrightarrow{t^{D\diamond}}, \overrightarrow{t^{R\diamond}}\right) \in T \times T$ is an equilibrium of the game. Suppose also that $\overrightarrow{t^{D\diamond}} \neq \overrightarrow{t^{R\diamond}}$. We know from Proposition 6 that $p^D\left(\overrightarrow{t^R}, \overrightarrow{t^R}\right) = \frac{1}{2}$. Therefore, by the definition of a Nash Equilibrium it must be

$$p^D\left(\overrightarrow{t^{D\diamond}}, \overrightarrow{t^{R\diamond}}\right) > p^D\left(\overrightarrow{t^R}, \overrightarrow{t^R}\right) = \frac{1}{2} \tag{34}$$

By the definition of a constant-sum game we also know that $p^R\left(\overrightarrow{t^D}, \overrightarrow{t^D}\right) = 1 - p^D\left(\overrightarrow{t^D}, \overrightarrow{t^D}\right) = \frac{1}{2}$ and again by the definition of a Nash Equilibrium, it must be

$$p^R\left(\overrightarrow{t^{R\diamond}}, \overrightarrow{t^{D\diamond}}\right) > p^R\left(\overrightarrow{t^D}, \overrightarrow{t^D}\right) = \frac{1}{2} \tag{35}$$

Since $p^R\left(\overrightarrow{t^{R\diamond}}, \overrightarrow{t^{D\diamond}}\right) = 1 - p^D\left(\overrightarrow{t^{R\diamond}}, \overrightarrow{t^{D\diamond}}\right)$, this implies that $p^D\left(\overrightarrow{t^{R\diamond}}, \overrightarrow{t^{D\diamond}}\right) < \frac{1}{2}$. By 34, this implies that $p^D\left(\overrightarrow{t^{R\diamond}}, \overrightarrow{t^{D\diamond}}\right) > \frac{1}{2}$, a contradiction. Therefore, $\overrightarrow{t^{D\diamond}} = \overrightarrow{t^{R\diamond}}$. ■

Corollary 22 *In equilibrium,* $V\left(\overrightarrow{t^{D\diamond}}\right) = V\left(\overrightarrow{t^{R\diamond}}\right)$.

Proof. By the meaning of Proposition 6, $\overrightarrow{t^{D\diamond}} = \overrightarrow{t^{R\diamond}}$. Therefore, $V\left(\overrightarrow{t^{D}}\right) = V\left(\overrightarrow{t^{R}}\right)$. ■

Exploiting Corollary 7, we may re-write the first order conditions in the following manner:

$$\begin{cases} \frac{\partial \mathcal{L}^D}{\partial q_1^D} = \frac{d}{s}\sum\limits_h \frac{\partial V^h}{\partial q_1^D} f^h s^h + \mu^D \left(\sum\limits_o t_o^D \sum\limits_h f^h \frac{\partial x_o^h}{\partial q_o^D} + x_o^h\right) = 0 & o \neq i \\ \vdots \\ \frac{\partial \mathcal{L}^D}{\partial q_n^D} = \frac{d}{s}\sum\limits_h \frac{\partial V^h}{\partial q_n^D} f^h s^h + \mu^D \left(\sum\limits_o t_o^D \sum\limits_h f^h \frac{\partial x_o^h}{\partial q_o^D} + x_o^h\right) = 0 \\ \frac{\partial \mathcal{L}^D}{\partial \mu^D} = \sum\limits_h f^h \sum\limits_i t_i^D x\left(q_i^D, M^h\right) = 0 \end{cases}$$

From `Roy's Identity` we know that $\frac{\partial V^h}{\partial q_i^D} = -\lambda^h x_i^h$ where λ^h is the marginal utility of income. Applying `Slutzky decomposition` we obtain the `Slutzky matrix`

$$D_{q^j} x\left(q^j, M^h\right) = D_{q^j} h\left(q^j, U^h\right) - D_{M^h} x\left(q^j, M^h\right) x\left(q^j, M^h\right)^\top$$

An element of the matrix is $\frac{\partial x_i^h}{\partial q_i^D} = \frac{\partial \left(x_i^h\right)^c}{\partial q_i^D} - \frac{\partial x_i^h}{\partial M^h} x_i^h$, where $\frac{\partial \left(x_i^h\right)^c}{\partial q_i^D}$ is the change in the Hicksian demand with a change in price, representing the *substitution effect*, and $\frac{\partial x_i^h}{\partial M^h} x_i^h$ is the *income effect*. Under the hypothesis of normal goods, $\frac{\partial x_i^h}{\partial q_i^D} < 0$, for every i. Substituting these two expressions in the set of first order conditions we obtain:

$$\frac{\partial \mathcal{L}^D}{\partial q_i^D} = -\sum_h \left(\lambda^h f^h s^h \frac{d}{s} + \mu^D f^h \sum_o t_o^D \frac{\partial x_o^h}{\partial M^h}\right) x_i^h + \\ + \mu^D \left(\sum_o t_o^D \sum_h f^h \xi_{oi}^h + x_o^h\right) = 0 \tag{36}$$

Expression

$$\alpha^{h,D} := \lambda^h f^h s^h \frac{d}{s} + \mu^D f^h \sum_o t_o^D \frac{\partial x_o^h}{\partial M^h}$$

denotes the *marginal probability of winning of* D for group h. It measures the weight that D attaches to a group as a function of its political power, represented by two parameters: density and size. A suitable interpretation for this expression is the following: Government's transfers are a function of the weight that candidates attach to groups, which depends on the effect that a change in the utility of the group, due to a change in the policy vector, has on the probability of winning elections at the margin. Hence, groups are assigned with a weight which is higher the more powerful the group. Furthermore, $\phi_{oi}^h := \frac{\partial \left(x_i^h\right)^c}{\partial q_o^D}$ represents the effect of a variation in price of good o on the compensated demand of good i for group h. Equation 36 may be re-written as follows:

$$\frac{\partial \mathcal{L}^D}{\partial q_i^D} = -\sum_h \alpha^h x_i^h + \mu^D \left(\sum_o t_o^D \sum_h f^h \phi_{oi}^h + x_o^h \right) = 0 \tag{37}$$

Dividing both sides by μ^D and x_i^h and re-arrange terms we finally obtain:

$$-\frac{\sum_o t_o^D \sum_h f^h \phi_{oi}^h}{x_i^h} = -\frac{\Delta x_i^{ch}}{x_i^h} = \frac{\mu^D - \chi_i^h}{\mu^D} \tag{38}$$

$\forall i$

$\chi_i^{h,D} := \frac{\sum_h \alpha^h x_i^h}{x_i^h}$ represents what in literature is known as the *ditributive characteristic* of good i for group h and for candidate D. $-\frac{\Delta x_i^{ch}}{x_i^h}$ measures the approximate proportional variation in the compensated aggregate demand of good i.

Proposition 23 *The distributive characteristic is higher the higher is the amount of good consumed by groups which receive a higher weight by candidates, that is the more single-minded.*

Proof. the distributive characteristic of good i for group h and for candidate D is obtained by multiplying the *marginal probability of winning of candidate* j for group h by the consumption of a good by group h with respect to the total consumption of good i. Notice that $\chi_i^{h,D}$ is increasing in $\alpha^{h,D}$, being $\frac{\partial \chi_i^{h,D}}{\partial \alpha^{h,D}} = 1$. We also know that $\alpha^{h,D}$ increases with respect to an increase in the group's density

$$\frac{\partial \alpha^h}{\partial s^h} = \lambda^h f^h \frac{d}{s} \left(1 - \frac{\overbrace{f^h s^h}^{<1}}{s} \right) > 0$$

By Axiom 19 we know that the density function is monotonically increasing in the consumption of goods. Finally, the first order derivative of the Marshallian functions is increasing in the level of single-mindedness, since $\frac{\partial x_i^{h*}}{\partial \psi_i^h} = \frac{M^h}{q_i^j} > 0$. ■

More single-minded groups provide the candidates with a higher *marginal probability of winning,* which translates the consumption of goods in higher level of the distributive characteristic. Therefore, we have found a precise linkage between single-mindedness and distributive characteristic represented by the following transmission mechanism:

$$single - mindedness\ (\uparrow) \Longrightarrow consumption\ of\ good\ (\uparrow) \Longrightarrow$$

$$density\ function\ (\uparrow) \Longrightarrow distributive\ characteristic\ (\uparrow)$$

Proposition 24 *The optimal tax structure induces a lower reduction in the consumption of those goods which are the most preferred by more single-minded groups.*

Proof. a reduction in the consumption is captured by the left-hand side of 38, which is negative. This expression is lower the lower is the right-hand side, which is lower the smaller the difference between μ^j and $\alpha^{h,j}$. By proposition 23 we know that the distributive characteristic is higher the higher the single-mindedness of a group and hence the right-hand side reduces as well. ■

To what extent do the taxation of goods obtained in this Political Economy framework differ from the classic taxation *à la* Ramsey? To answer, we must compare the many-person Ramsey's rule (Diamond, 1975) with equation 38. In the former optimal tax rates induce a lower reduction in the consumption of those goods which are more consumed by the poor, because they gain a higher weight by society. Instead, in 38, the weight attached by the Government does not only depend on individuals' income but also on groups' political power. That is, the more powerful groups obtain a higher political consideration by candidates. As a consequence, candidates do not take equity goals into account as in the classic Ramsey rule, and this attitude represents the real political failure of the model. The difference between the traditional Ramsey rule $-\frac{\sum_o t_o \sum_h \phi_{oi}^h}{x_i^h} = \frac{\mu - \chi_i^h}{\mu}$ and 38 can be calculated taking the difference of the two expressions. This difference, equal to $\lambda^h \left(\frac{\partial W}{\partial V^h} - f^h s^h \frac{d}{s}\right) + \mu^j \left(1 - f^h\right) \sum_o t_o^j \frac{\partial x_o^h}{\partial M^h}$, is higher the lower f^h and s^h; this means that the less single-minded groups receive a lower weight by candidates, whilst under Ramsey the social weight assigned by the Government depended on the effect which an increase in the utility of group h has on the social welfare at the margin, $\frac{\partial W}{\partial V^h}$. This weight is generally higher for the poorest as long as the Social Welfare Function is strictly concave. Notice that 38 does not say that candidates totally neglect the welfare of the poor because $\alpha^{h,j}$ is higher the higher is the marginal utility of income, λ^h, which is higher for the poorest[13]. Notice also that the classic Ramsey rule and 38 coincide if $\frac{\partial W}{\partial V^h} = f^h s^h \frac{d}{s}$; that is, *if the importance attributed by society to the increase in the welfare of group h is exactly equal to the political importance attributed by candidates to the same group.* In this case, and only in this case, the normative and the positive approaches achieve the same results. Nevertheless, a tenet taken from the theory of optimal taxation still holds: in equilibrium, tax rates chosen by candidates are differentiated, even though the redistribution does not take place between the rich and the poor but between the strongest and the weakest groups of society. The following table compares results obtained under the classic Ramsey rule and 38.

	Classic Ramsey rule	Single-mindedness rule
General formula	$-\frac{\sum_o t_o \sum_h \phi_{oi}^h}{x_i^h} = \frac{\mu - \chi_i^h}{\mu}$	$-\frac{\sum_o t_o^D \sum_h f^h \phi_{oi}^h}{x_i^h} = \frac{\mu^D - \chi_i^h}{\mu^D}$
Distributive characteristic	$\frac{\sum_h \left(\lambda^h + \mu \sum_o t_o \frac{\partial x_o^h}{\partial M^h}\right) x_i^h}{x_i^h}$	$\frac{\sum_h \left(\lambda^h f^h s^h \frac{d}{s} + \mu^D f^h \sum_o t_o^D \frac{\partial x_o^h}{\partial M^h}\right) x_i^h}{x_i^h}$
Distortion on consumption	yes	yes
Political failure	no	yes
Achievement of equity goals	yes	depending on the location of single-minded groups on the income scale
Better off groups	poor	more single-minded
Worse off groups	rich	less single-minded
Highest weight assigned	poor	more single-minded

Legend of symbols	
$h = 1, ..., H$	social groups
f^h	group's size
$j = D, R$	political candidates
$i = 1, ..., n$	goods
x_i^h	consumption of goods
ψ_i^h	preference for goods/level of single-mindedness
ξ^h	idiosyncratic stochastic variable
ς	non-idiosyncratic stochastic variable
$q_i^j = p_i + t_i^j$	consumption price = production price + tax rate
s^h	density function of idiosyncratic variable/political power of a group
d	density function of non-idiosyncratic variable
λ^h	marginal utility of income
α^j	marginal probability of winning of D for group h
ϕ_{oi}^h	effect of a variation in price of good o on the compensated demand of good i for group h
$\chi_i^{h,j}$	distributive characteristic
φ^h	preferences for public good
G^j	public good

Endogenous public expenditure

I analyse now an extension of the previous model considering a Government which must choose both the tax rates and the provision of a public good. The introduction of public goods in probabilistic voting models with single-minded groups raises two fundamental questions:

1. to what extent is the optimal provision of public goods influenced by distortionary taxation?
2. to what extent is the traditional Samuelson rule modified when the Government is not benevolent but aims to maximise the probability of winning elections?

The problem of the individual may be re-written in the following log-linear fashion:

$$\max_{\{x_i^h\}} \sum_{i=1}^{n} \psi_i^h \log(x_i^h) + \varphi^h \log G^j + D^R \cdot \left(\xi^h + \varsigma\right)$$

$$s.t. \sum_{i=1}^{n} q_i^j x_i^h = M^h$$

where G^j denotes the per-capita level of provision of a public good chosen by candidates and φ^h the idiosyncratic preference of group h for the provision of the public good, or in other words, the mindedness of the group for the amount of the public good. The production of this good is entirely financed by taxes levied on the tax-payers. Thus, individuals' choices are influenced by the amount of the public good. On the one hand, G reduces the individuals' disposable income, since the higher G the higher the taxes which individuals must pay to balance the budget. In turn, public expenditure crowds out private consumption. On the other hand, the arising substitution effect depends on the degree of complementarity or substitutability between private and public goods; the effect of a change in the amount of public good on private goods is higher the higher is the degree of complementarity between private and public goods.

Solving the individual maximization problem we obtain the Marshallian functions $x_i^{h*} = \frac{\psi_i^h M^h}{q_i^j}$ and the Indirect Utility Function

$$U\left(x_i^{h*}, G^j\right) = V\left(x\left(q_i^j, M^h\right), G^j\right)$$

The Government's budget constraint is:

$$C\left(G^{j}\right)\sum_{h}f^{h}=\sum_{h}f^{h}\sum_{i}t_{i}^{j}x\left(q_{i}^{j},M^{h}\right)$$

where $C\left(G^{j}\right)$ denotes the per-capita cost function of the public good. I assume that $C\left(G^{j}\right)$ is a twice differentiable function, with $C_{G^{j}}:=\frac{\partial C\left(G^{j}\right)}{\partial G^{j}}>0$ and $C_{G^{j}G^{j}}:=\frac{\partial^{2}C\left(G^{j}\right)}{\partial^{2}G^{j}}>0$; that is, the production of the public good has marginal decreasing costs. Furthermore, $C_{G^{j}}$ measures the Marginal Rate of Transformation (MRT) and in order to emphasise this fact I will define $C_{G^{j}}:=MRT^{j}$.

Secondly, I solve the candidate's problem, which is the same as before, modified only by the presence of the public good. I will denote the new candidate policy vector by $\eta^{j}=\left[t_{1}^{j},...,t_{n}^{j},G^{j}\right]\in\Phi^{j}\subset\mathbb{R}^{n+1}$ and I write the Lagrangian function:

$$\mathcal{L}^{j}=\frac{1}{2}+\frac{d}{s}\sum_{h}f^{h}s^{h}\left[V\left(\eta^{j}\right)-V\left(\eta^{-j}\right)\right]+$$

$$+\mu^{j}\left(\sum_{h}f^{h}\sum_{i}t_{i}^{j}x\left(q_{i}^{j},M^{h}\right)-C\left(G^{j}\right)\sum_{h}f^{h}\right) \tag{39}$$

First, notice that 38 does not change even in the presence of public expenditure which finances public goods.

Proposition 25 *the marginal rate of transformation is equal to the sum of idiosyncratic preferences for the public good of groups weighted by their size and density.*

Proof. The first order conditions for 39 are:

$$\frac{\partial\mathcal{L}^{j}}{\partial q_{1}^{j}}=\frac{d}{s}\sum_{h}\frac{\partial V^{h}}{\partial q_{1}^{j}}f^{h}s^{h}+\mu^{j}\left(\sum_{o}t_{o}^{j}\sum_{h}f^{h}\frac{\partial x_{o}^{h}}{\partial q_{1}^{j}}+x_{o}^{h}\right)=0$$

$$\vdots$$

$$\frac{\partial\mathcal{L}^{j}}{\partial q_{n}^{j}}=\frac{d}{s}\sum_{h}\frac{\partial V^{h}}{\partial q_{n}^{j}}f^{h}s^{h}+\mu^{j}\left(\sum_{o}t_{o}^{j}\sum_{h}f^{h}\frac{\partial x_{o}^{h}}{\partial q_{n}^{j}}+x_{o}^{h}\right)=0 \tag{40}$$

$$\frac{\partial\mathcal{L}^{j}}{\partial G^{j}}=\frac{d}{s}\sum_{h}\frac{\partial V^{h}}{\partial G^{j}}f^{h}s^{h}-\mu^{j}\left(MRT^{j}\sum_{h}f^{h}\right)=0 \tag{41}$$

Since in equation 41 $\frac{\partial V^h}{\partial G^j} = \frac{\varphi^h}{G^j}$ we obtain a final version of the equation which refers to the choice of public good:

$$\frac{d\sum_h \varphi^h f^h s^h}{sG^j \sum_h f^h} = \mu^j \left(MRT^j\right) \tag{42}$$

■

Suppose now, without loss of generality, that $C\left(G^j\right) = \left(G^j\right)^2$, with $MRT^j = 2G^j$. Equation 42 becomes

$$\frac{d\sum_h \varphi^h f^h s^h}{sG^j \sum_h f^h} = 2\mu^j G^j \tag{43}$$

which, solved with respect to G^j yields:

$$G^{j*} = \left(\frac{d\sum_h \varphi^h f^h s^h}{2s\mu^j \sum_h f^h}\right)^{\frac{1}{2}} \tag{44}$$

This expression clearly shows how the provision of public good depends on the mindedness of groups, that is on the idiosyncratic parameter φ^h.

In this expression μ^j represents the `marginal cost of public funds`, defined as the social cost of spending one extra dollar on any given public good and it measures the distortionary effect of taxation.

Proposition 26 *The provision of public good is strictly increasing in the single-mindedness of the group, weighted by its density and size and decreasing in the marginal cost of public fund.*

Proof. Performing some comparative statics we can see that

$$\frac{\partial G^{j*}}{\partial \varphi^h} = \frac{1}{2}\left(\frac{d\sum_h \varphi^h f^h s^h}{2s\mu^j \sum_h f^h}\right)^{-\frac{1}{2}} \frac{df^h s^h}{2s\mu^j \sum_h f^h} > 0$$

$$\frac{\partial G^{j*}}{\partial s^h} = \frac{1}{2}\left(\frac{d\sum_h \varphi^h f^h s^h}{2s\mu^j \sum_h f^h}\right)^{-\frac{1}{2}} \frac{df^h \varphi^h}{2s\mu^j \sum_h f^h} > 0$$

$$\frac{\partial G^{j*}}{\partial f^h} = \frac{1}{2}\left(\frac{d\sum_h \varphi^h f^h s^h}{2s\mu^j \sum_h f^h}\right)^{-\frac{1}{2}} \frac{d}{2s\mu^j} \frac{\varphi^h s^h \sum_h f^h - \sum_h \varphi^h f^h s^h}{\left(\sum_h f^h\right)^2} > 0$$

$$\frac{\partial G^{j*}}{\partial \mu^j} = -\frac{1}{2}\left(\frac{d\sum_h \varphi^h f^h s^h}{2s\mu^j \sum_h f^h}\right)^{-\frac{1}{2}} \frac{df^h s^h}{2s\left(\mu^j\right)^2 \sum_h f^h} < 0$$

■

Otherwise, the Ramsey rule does not change with respect to the previous case and the reason is simple. If the Ramsey rule detects the most efficient way to finance a certain level of expenditure, *for every level of expenditure*, all the more so it must detect the most efficient way to finance the level of expenditure when this is chosen in an optimal way to finance G. Of course tax rates differ, depending on the level of G, since higher level of G entails higher level of tax revenues, but the optimal tax rate *structure* does not change with respect to the previous case.

Concluding, the Single-mindedness Theory states again that, in order to win elections, candidates must content more single-minded groups who are the real winners of the political game. In this case the provision of public good is higher the higher the presence of more single-minded groups which ask for it. With respect to the classic theory and Samuelson rule, a model with single-minded groups tells us that the provision of public goods is not only inefficient because of the presence of distortionary taxation, but also because of the political failure which society falls into, due to the presence of powerful interest groups.

Conclusions

In this paper I analysed how voter-seeking candidates decide indirect taxation policies in a Probabilistic Voting model. Results show that candidates are captured by the most powerful (single-minded) groups, which not necessarily coincide with the median voter, but may represent even the richest components of society. These results are at odds with the classic results achieved by using the median voter theorem, because it is no longer the median position on the income scale which drives the equilibrium policy, but the ability of groups to focus on their preferred issues instead.

Secondly, this model provides a possible explanation to the existence of indirect taxation, since we perfectly know that the optimal direct-*cum*-indirect tax problem puts all commodity taxes to zero and raises everything through income tax (Atkinson and Stiglitz, 1976). Instead, in the model I suggested there is more than one interest by powerful single-minded groups to prevent a substantial shift from indirect to direct taxation. Since indirect taxation is mostly regressive whilst direct taxation mostly progressive, richest single-minded groups would stand up for this shift. The direct-*cum*-indirect tax problem can be perfectly studied using Probabilistic Voting and Single-mindedness theory and I hope this could be done in future contributions.