

The politics of progressive income taxation with incentive effects

Philippe De Donder

IDEI, University of Toulouse 1

and

Jean Hindriks*

Queen Mary and Westfield College, University of London

July 2000

Abstract: This paper studies majority voting over non-linear income taxes when individuals respond to taxation by substituting untaxable leisure to taxable labor. We first show that voting cycle over progressive and regressive taxes is inevitable. This is because the middle-class can always lower its tax burden at the expense of the rich by imposing progressive taxes (convex tax function) while the rich and the poor can reduce their tax burden at the expense of the middle-class by imposing regressive taxes (concave tax function). We then investigate three solutions to this cycling problem: (i) reducing the policy space to the policies that are ideal for some voter; (ii) weakening the voting equilibrium concept; (iii) assuming parties also care about the size of their majority. The main result is that progressivity emerges as a voting equilibrium if there is a lack of polarization at the extremes of the income distribution. Interestingly the poor would prefer regressive taxes.

JEL classification: D72

Keywords: Majority Voting; Income taxation; Tax Progressivity.

* Economics department, QMW, University of London, Mile End Road, London E1 4NS, England. Tel. (44)-020-78827807, Email: j.hindriks@qmw.ac.uk.

1 Introduction

Why do most democratic countries tend to propose marginal progressive income tax schedules? The economic literature has tried to answer this question by resorting either to a normative or a positive approach. The normative approach assumes that tax policies are chosen by a benevolent planner, whose objective is to maximize social welfare under informational and incentive constraints. Unfortunately, this approach has proven very inconclusive on the shape of the optimal tax function (for a recent account, see Myles, 2000). A notable exception is when the planner's objective is to choose a "fair" tax schedule. Indeed, Young (1990) has shown that the equal sacrifice requirement implies progressive taxation.

The positive approach, to whom this paper clearly belongs, stresses the fact that tax schedules are, directly or indirectly, chosen democratically. It is also often assumed that voters are self-interested. The main difficulty in applying voting theory to (non-linear) income taxation resides in the multi-dimensionality of the policy space. It is well known that the aggregation, by means of majority voting, of transitive individual preferences on a multi-dimensional set of options may result in a non-transitive social preference. In other words, majority may give rise to voting cycles, where policy a is majority preferred to policy b which in turn is majority preferred to policy c , which eventually is majority preferred to policy a . No natural majority winner emerges in such situations.

The apparent stability of the democratic demand for progressivity contrasts sharply with this vote cycling. Economists have proposed different explanations to this puzzle. The most common one is to consider that the true policy space is not so large and that it should be possible to apply some kind of median voter theorem on this restricted policy space. The key question is of course what is the natural reduction of the policy space. Romer (1975) reduces the policy space to linear tax schedules and obtains a

Condorcet winner involving average progressivity (see also Roberts (1977)). Berliant and Gouveia (1994) introduce uncertainty over the income distribution and then use the ex-ante budget balance requirement to reduce the policy space so that a Condorcet winner exists. Snyder and Kramer(1988) assume that candidates cannot credibly commit to implement something different from their most-preferred policy and thus if everybody can stand as a candidate (like in Citizen Candidates models of Osborne-Slivinsky (1996) and Besley-Coate (1997)) then the true policy space reduces to the policies that are ideal for some voter. In doing so they obtain a Condorcet winner involving progressive taxes. Roemer (1999) considers that the policy space is reduced not by lack of commitment but rather by the need to reach an intra-party consensus among 'opportunists' whose only objective is to win the elections, and 'militants' who care only about the policy chosen by their party. He shows that there exists a Condorcet winner in this reduced policy space and that it involves progressive taxes.

Recently, Marhuenda and Ortuno-Ortin (1995) have followed another approach. Instead of isolating a majority winning tax schedules, they seek to provide some insights on the democratic demand for progressivity. Indeed they obtain the remarkable result that any marginal progressive tax wins over any regressive one provided that the median is less than the mean income. This *popular support for progressivity* result is obtained with self-interested voters who vote for the tax policy that taxes them less.¹

The problem with Marhuenda and Ortuno-Ortin (1995)'s analysis is that it gives only a partial account of the majority voting process. Indeed Hindriks (2000) has provided a reverse *popular support for regressivity* result according to which any marginal progressive tax can in turn be defeated by a marginal regressive tax (supported by a majority coalition of the extremes of the income distribution). This result holds true for any distribution of

¹The result has been generalized by Mitra et al (1998) to more sophisticated voters who also care about their relative position in the income distribution.

income provided that the median is below the mean income. Of course combining this regressivity result with the progressivity result establishes the inevitable voting cycle. We shall discuss that in Section 3.

The model we develop in this paper involves quadratic tax functions with variable labour supply.² Using a similar model, Cukierman and Meltzer (1991) have shown that the existence of a Condorcet winner can only be obtained under rather strong and mainly unjustifiable conditions. So we investigate three ways of circumventing this non-existence problem. First, we consider a natural reduction of the policy space to tax schedules that are ideal for some voter and we show that a Condorcet winner exists on this policy subset. This is the approach adopted in Section 4. Second we keep the entire policy space but adopt less demanding political equilibrium concepts that are Condorcet-consistent (i.e., they uniquely select the Condorcet winner if any) such as the uncovered set and the bipartisan set.³ This is the approach we follow in Section 5. Third we consider a dynamic voting game due to Kramer (1977) in which two political parties alternate in power (due to the non-existence of a Condorcet winner) and propose policies that maximise their vote shares (i.e., they care about the size of their majority).

The rest of the paper is organized as follows. Section 2 describes the model. Section 3 demonstrates that vote cycling is inevitable. Section 4 studies the reduction of the policy space to voters' blisspoints. Section 5 studies weaker political equilibrium concepts. Section 6 explores the Kramer dynamic voting game. Section 7 concludes.

²This class of non-linear tax schedules accomodates in a simple way both (marginal) regressive and progressive taxes. It has been adopted in Cukierman and Meltzer (1991) as well as in Roemer (1999).

³See De Donder et al.(2000) for a presentation and a set-theoretical comparison of these different solution concepts.

2 The Model

We consider a two-good economy (consumption and labor) populated by a large number of individuals who differ only in their ability. Each individual is characterized by her ability to earn income, $w \in [0, 1]$. The distribution of ability in the population is described by a strictly increasing distribution function F on $[0, 1]$, so that $F(w)$ is the fraction of the population with ability less or equal to w . The average ability level is

$$\bar{w} = \int_0^1 w dF(w) \quad (1)$$

and the median ability level is

$$w_m = F^{-1}\left(\frac{1}{2}\right). \quad (2)$$

We assume throughout that $w_m \leq \bar{w}$. Individuals choose the amount of labor they sell on a competitive market and receive a wage rate equal to their ability. The production sector exhibits constant-returns-to-scale so that the wage rate is constant. Then an individual with ability $w \neq 0$ who supplies y/w units of labor earns pre-tax income y . Her after-tax income is

$$x(y, t) = y - t(y) \quad (3)$$

where $t(y)$ is a continuous tax function $t : R_+ \rightarrow R$. Note that we allow for negative taxes.

A **feasible** tax function satisfies the following conditions.

$$t(y) \leq y \quad \text{for all } y \in R_+ \quad (4)$$

$$t'(y) \leq 1 \quad \text{for all } y \in R_+ \quad (5)$$

$$\int_0^1 t(y(w)) dF(w) = 0 \quad (6)$$

Condition (4) says that tax liabilities cannot exceed taxable income. Condition (5) implies that after-tax income is non-decreasing in pre-tax income.

The budget balance condition (6) means that income taxation is purely redistributive (i.e., zero revenue requirement).

Note that these conditions place few restrictions on the set of admissible tax schedules which is infinitely dimensional. Given that we are mainly interested to explain the prevalence of progressive income taxation, we shall thereafter restrict attention to tax schedules that are either concave, linear or convex everywhere. This class of tax schedules is nicely represented by the quadratic tax function.⁴

$$t(y) = -c + by + ay^2 \quad (7)$$

where $c \geq 0$ is the uniform lump-sum transfer, b is the linear tax parameter (with $0 \leq b \leq 1$) and a is the progressivity tax parameter, with $a > 0$ indicating a progressive tax scheme (i.e. marginal tax rates increasing with pre-tax income) and $a < 0$ representing a regressive one. As will become clear shortly, the feasibility conditions (4) and (5) impose lower and upper bounds on the progressivity parameter ($-\frac{1}{2} \leq a \leq +\frac{1}{2}$). Intuitively, too much progressivity would make the marginal tax rate greater than one at the top of the income distribution while too much regressivity would bankrupt the poor individuals ($t(y) > y$ for low y). Combining (6) and (7) yields

$$c = b\bar{y} + a\bar{y}_2 \quad (8)$$

where $\bar{y} = \int y(w) dF(w)$ and $\bar{y}_2 = \int y^2(w) dF(w)$. Hence, tax policies are two-dimensional, $(a, b) \in [-\frac{1}{2}, \frac{1}{2}] \times [0, 1]$. Plugging (7) and (8) into (3) we get

$$x = \bar{y} + (1 - b)(y - \bar{y}) - a(y^2 - \bar{y}_2) \quad (9)$$

Given the tax parameters (a, b) , individual with ability w chooses pretax income $y(a, b; w)$ that maximises $u(x, \frac{y}{w})$ subject to (9). Throughout we assume quasi-linear preferences over consumption and labour supply.

⁴The quadratic tax function has already been used in a voting context by Cukierman and Meltzer (1991) with variable income and by Roemer (1999) with fixed income.

$$u(x, \frac{y}{w}) = x - \frac{1}{2} \left(\frac{y}{w} \right)^2 \quad (10)$$

With this formulation of preferences there is no income effect on labour supply decision and everyone chooses to work.⁵ While this is a highly restrictive specification of preferences, it captures the incentive effects of taxes (consumption-leisure trade-off), thus making it more general than the models with fixed income (like Roemer, 1999), and it is also considerably more tractable than the general specification of Cukierman-Meltzer (1991), allowing us to obtain clear, intuitive results.

For this specification of preferences and given the tax parameters (a, b) the optimal pre-tax income of individual with ability w is

$$y(a, b; w) = (1 - b) \frac{w^2}{1 + 2aw^2} \quad (11)$$

Therefore marginal tax progressivity ($a > 0$) reduces the pre-tax income of everyone. But because progressivity decreases more the pre-tax income of the rich, the dispersion of pre-tax income decreases. The figure below illustrates the effect of progressivity and regressivity on pre-tax income levels.

⁵See Laslier et al.(2000) for a model where the less productive individuals choose not to work.

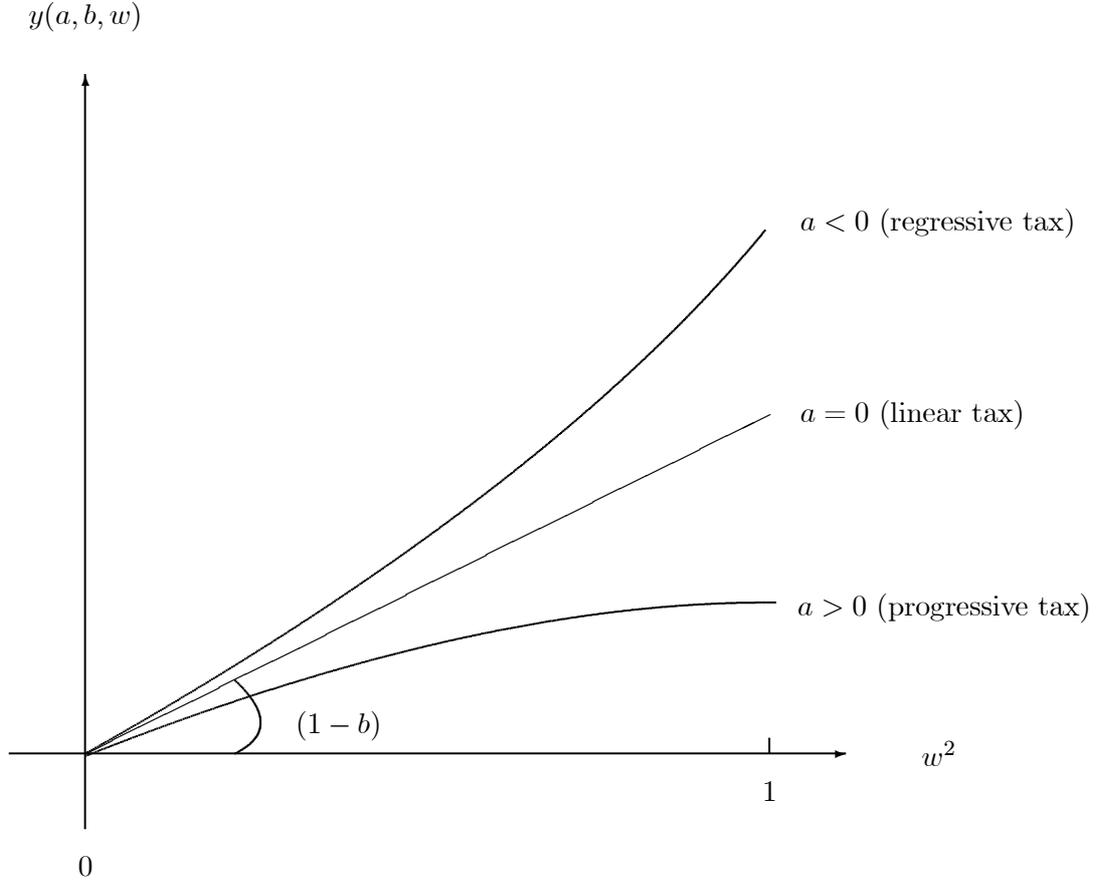


Figure 1. Effect of progressive and regressive taxes on pre-tax income.

Let $\omega(a) = \frac{w^2}{1+2aw^2}$, then slightly abusing notation we obtain

$$y = (1 - b)\omega(a) \quad (12)$$

$$\bar{y} = (1 - b)\bar{\omega}(a) \quad (13)$$

$$\bar{y}_2 = (1 - b)^2\bar{\omega}_2(a) \quad (14)$$

where $\bar{\omega}(a) = \int \omega(a) dF(w)$ and $\bar{\omega}_2(a) = \int \omega(a)^2 dF(w)$.

For later use, we can also compute the (average) incentive effects of the tax parameters as

$$\frac{\partial \bar{y}}{\partial b} = -\bar{\omega}(a) < 0 \quad (15)$$

$$\frac{\partial \bar{y}}{\partial a} = \frac{\partial \bar{y}_2}{\partial b} = -2(1-b)\bar{\omega}_2(a) < 0 \quad (16)$$

$$\frac{\partial \bar{y}_2}{\partial a} = -4(1-b)^2\bar{\omega}_3(a) < 0 \quad (17)$$

where $\bar{\omega}_3(a) = \int \omega^3(a) dF(w) > 0$.

We can now derive the precise restrictions on (a, b) needed to satisfy the feasibility conditions. Condition(5) requires that $t'(y) = b + 2ay \leq 1$ for all $y \in R_+$. Clearly this condition is satisfied for all $a \in [-\frac{1}{2}, 0]$ since $b \in [0, 1]$. It is also satisfied for all $a \in (0, \frac{1}{2}]$ because from (11), $t'(y) = b + 2ay = b + (\frac{2aw^2}{1+2aw^2})(1-b) \leq 1$ for all $w \in [0, 1]$ and for all $b \in [0, 1]$. Since the marginal rate of taxation is less than one everywhere, condition (4) simply reduces to $t(0) \leq 0$, that is

$$\begin{aligned} c &= b\bar{y} + a\bar{y}_2 \geq 0 \\ &= b(1-b)\bar{\omega}(a) + a(1-b)^2\bar{\omega}_2(a) \geq 0 \end{aligned} \quad (18)$$

where the second equation follows from (13) and (14). Hence, regressivity ($a < 0$) imposes the following lower bound on b ,

$$b \geq \underline{b}(a) \equiv \frac{a\bar{\omega}_2(a)}{a\bar{\omega}_2(a) - \bar{\omega}(a)} \quad (19)$$

It can be shown that this lower bound is decreasing with a and

$$\begin{aligned} \underline{b}(a) &> 0 && \text{for } a < 0 \\ &= 0 && \text{for } a = 0 \\ &< 0 && \text{for } a > 0. \end{aligned}$$

The feasible policy set is then $X = \{a, b : -\frac{1}{2} \leq a \leq \frac{1}{2}, 0 \leq b \leq 1, \underline{b}(a) \leq b\}$.

Remark 1: For any distribution $F(w)$, if $b \in [\frac{1}{3}, 1]$ then $\partial c / \partial a \leq 0$ for all $a \in [0, \frac{1}{2}]$.

This is easily seen. Indeed,

$$\frac{\partial c}{\partial a} = b \frac{\partial \bar{y}}{\partial a} + \bar{y}_2 + a \frac{\partial \bar{y}_2}{\partial a}$$

Which using (14),(16) and (17) gives

$$\frac{\partial c}{\partial a} = (1-b)(1-3b)\bar{\omega}_2(a) - 4a(1-b)^2\bar{\omega}_3(a). \quad (20)$$

Since for any $b \in [\frac{1}{3}, 1]$, we have $(1-b)(1-3b) \leq 0$, and for any $a \geq 0$ and any distribution $F(w)$, we have $\bar{\omega}_2(a), \bar{\omega}_3(a) > 0$, it follows that $\partial c / \partial a \leq 0$.

Remark 1 leads to the following:

Remark 2: Any policy involving $a \in [0, \frac{1}{2}]$ and $b \in [\frac{1}{3}, 1]$ is Pareto dominated.

This is a straightforward implication of Remark 1 since decreasing a from $a \in [0, \frac{1}{2}]$ given $b \in [\frac{1}{3}, 1]$ cannot not decrease c but will lower the tax burden of everyone, making everyone better off.

We can also wonder whether the poor would prefer regressive taxes or progressive taxes. Clearly the poorest individual with income $y = 0$ will choose the value of a that maximizes the transfer c . Fixing $b \in [0, 1]$ and setting (20) equal to 0 we get,

$$2a \frac{\bar{\omega}_3(a)}{\bar{\omega}_2(a)} = \left(\frac{1}{2} - \frac{b}{1-b} \right).$$

It follows that,

Remark 3: For any distribution $F(w)$ maximizing the utility of the poorest individual (or maximizing the transfer c) requires progressivity when

$b < 1/3$, *linearity* when $b = 1/3$ and *regressivity* when $b > 1/3$.

Note that this result does not depend on the distribution of ability. Resorting to numerical simulations for various distributions of ability we further get that (i) the transfer maximizing value of a is monotonically decreasing with b and that (ii) transfer is globally maximized with regressive taxation (and thus $b > 1/3$). So the poor would prefer regressive taxes.

We can now derive the indirect utility function. Substituting (3), (7) and (12) into the utility function (10) one obtains,

$$v(a, b; w) = c + (1 - b)^2 \frac{\omega(a)}{2}. \quad (21)$$

Hence for any individual with ability w (which maps to $w(a) = \frac{w^2}{1+2aw^2}$) the marginal rate of substitution between a and b is

$$\begin{aligned} \left[\frac{\partial b}{\partial a} \right]_{dv=0} &= \frac{-[\partial c/\partial a - (1-b)^2 \omega^2(a)]}{\partial c/\partial b - (1-b)\omega(a)} \\ &= \frac{-[\partial c/\partial a - y^2]}{\partial c/\partial b - y} \end{aligned}$$

The marginal rate of substitution between the two tax parameters is the result of their marginal effect on the transfer net of their marginal effect on tax liability. Pareto efficiency requires $\partial c/\partial a > 0$ and $\partial c/\partial b > 0$. Moreover, due to incentive effects (15)-(17), we have that for any $a > 0$, $\partial c/\partial a < \bar{y}_2$ and $\partial c/\partial b < \bar{y}$. Therefore all those with either low income such that both $y < y_1 = \partial c/\partial b$ and $y^2 < y_2 = \partial c/\partial a$ or high income such that both $y > y_1$ and $y^2 > y_2$ will display a negative marginal rate of substitution. Those in the middle of the income distribution with income y such that $(y - y_1)(y^2 - y_2) < 0$ will display a positive marginal rate of substitution.

We now turn to the voting problem over the admissible policies $(a, b) \in X$. It is well known since Plott (1967) that the conditions required to obtain a Condorcet winner (i.e. a policy preferred by a majority of individuals

to any other feasible policy) in multi-dimensional settings are extremely demanding. In the absence of Condorcet winner, the social preference generated by majority voting exhibits cycles. In the next section, we establish the inevitable voting cycle over regressive and progressive taxes.

3 Inevitable Voting Cycle

The inevitable character of vote cycling over quadratic tax functions is best seen by considering fixed pre-tax income. In doing so we can indeed prove the inevitable aspect of vote cycling by means of two propositions. The first proposition is a simple adaptation to quadratic tax functions of a result due to Marhuenda and Ortuno-Ortin (1995).

Proposition 1: (Marhuenda and Ortuno-Ortin, 1995) *Suppose income is fixed and distributed according to $F(y)$ over $[0, Y]$, with median below mean income, $y_m \leq \bar{y}$. Then for any tax scheme $t_2 = -c_2 + b_2y + a_2y^2$ there exists a more progressive tax scheme $t_1 = -c_1 + b_1y + a_1y^2$ with $a_1 > a_2$, $b_1 < (>)b_2$ and $c_1 > c_2$ such that a majority of voters prefer t_1 over t_2 .*

Proof: Consider the two tax schedules $t_1 = -c_1 + b_1y + a_1y^2$ and $t_2 = -c_2 + b_2y + a_2y^2$ and let

$$\begin{aligned} T = t_1 - t_2 &= -(c_1 - c_2) + (b_1 - b_2)y + (a_1 - a_2)y^2 \\ &= -c + by + ay^2 \end{aligned} \tag{22}$$

with $a = a_1 - a_2 > 0$, $b = b_1 - b_2 < (>)0$, and $c = c_1 - c_2 > 0$. From the balanced budget constraint, we have $c = b\bar{y} + a\bar{y}^2$. Clearly, T is strictly convex with a negative intercept. Since T is strictly convex, using Jensen's inequality we get,

$$T(\bar{y}) = T\left(\int_0^Y y dF(y)\right) < \int_0^Y T(y) dF(y) = 0,$$

So, $T(\bar{y}) < 0$ and since T is strictly convex, T must be strictly increasing and therefore $T(y) < 0$ for all $y \in [0, \bar{y}]$, that is all those with income below the mean income pay less taxes under t_1 than under t_2 . Since $y_m \leq \bar{y}$ more than half the voters would prefer t_1 to t_2 and the result follows. QED

This proposition reflects the success of the middle class in minimizing its own tax burden. Greater progressivity in income taxation actually reduces the tax burden on middle-income taxpayers at the expense of high-income taxpayers. This is of course reminiscent of Director's Law of Income Redistribution (Stigler, 1970). The voting cycle follows from the next proposition according to which regressivity can always defeat progressivity with a majority coalition of the extremes (rich and poor).

Proposition 2: (Hindriks, 2000) *Suppose income is fixed and distributed according to $F(y)$ in $[0, Y]$, with $y_m < \bar{y}$ and Y large enough. Then for any tax scheme $t_2 = -c_2 + b_2y + a_2y^2$ there exists a less progressive tax scheme $t_1 = -c_1 + b_1y + a_1y^2$ with $a_1 < a_2$, $b_1 > b_2$ and $c_1 > c_2$ such that a majority of voters prefers t_1 over t_2 .*

Proof: see Hindriks (2000).

Figure 2 illustrates this proposition. t_1 is preferred by a majority to t_2 if the probability mass on the interval $[y_1, y_2]$ is less than $1/2$. Notice that the result is true irrespective of the form of the income distribution, provided that the median is less than the mean income.⁶ This result is not trivial. Indeed the average income \bar{y} is always lying within $[y_1, y_2]$ and the lower the probability mass around \bar{y} , the greater the length of $[y_1, y_2]$.

⁶It can be shown that the result does not hold for symmetric distributions.

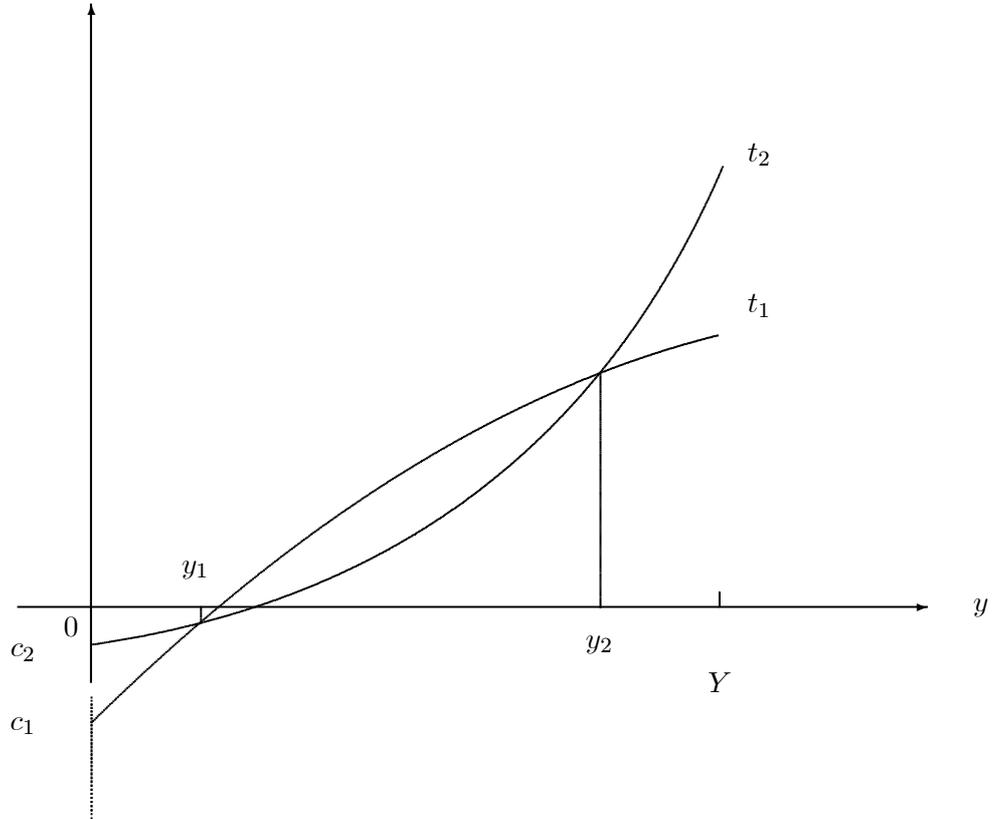


Figure 2. Regressive tax t_1 preferred by a majority to progressive tax t_2 .

Introducing variable income (as formalized in the previous section) does not break the voting cycle down as Figures 3 a, b and c show respectively for the uniform, the triangular and the linear decreasing distributions of ability.⁷ Notice that for each distribution, the median is less than the mean *income* even though the distribution of *ability* is symmetric. The cycle ABC (with A preferred to B, B preferred to C and C preferred to A) arises from alternating majority coalitions consisting of the poor and the middle class supporting more progressivity (as in Proposition 1) and then the poor and the rich supporting more regressivity (as in Proposition 2).

⁷For each figure, a and b vary by increments of 0.05 respectively in $[-0.5, 0.5]$ and $[0, 1]$.

Insert Figures 3 a, b, c

In the following sections we shall see how we can escape this vote cycling difficulty and gain insight on which tax functions have a chance to emerge as political equilibria: first by reducing naturally the voting space (Section 4), then by applying weaker political equilibrium concepts (Section 5); and lastly by assuming that political parties also care about the size of their majority (Section 6).

4 Natural reduction of the policy space

In this section, we restrict the policy space to the set of feasible tax schedules that are most-preferred by some voter. This assumption seems a natural way to reduce the policy space, and has been successfully adopted by Snyder and Kramer (1988) enabling them to obtain a Condorcet winner with progressive taxation.⁸ This reduction of the policy space can also be motivated in the line of the Citizen Candidates models as a lack of commitment from candidates selected from the set of voters and who can only credibly propose their most-preferred policy (see Osborne-Slivinsky, 1996, and Besley-Coate, 1997).

In Figures 4 a,b,c we have represented the bliss points of the electorate for the three distributions of ability we are considering.⁹ Surprisingly the distribution of bliss points has a similar form for each ability distribution. The lowest-ability individual prefers the policy that maximizes c (i.e., the peak of the Laffer surface). This is the regressive tax schedule, with the highest b and the lowest $a < 0$, namely the top of the upper curve in the distribution of bliss points. The second lowest-ability individual prefers a

⁸However, their model differs from ours in the sense that individuals do not respond to taxation by substituting untaxable leisure to taxable labor, but rather by working in an untaxed sector with lower wage rate.

⁹For each figure, a and b vary by increment of 0.005 respectively in $[-0.5, 0.5]$ and $[0, 1]$, that is they take on 200 different values each.

slightly less regressive tax (lower b for a slightly higher a), and so on moving downwards along the upper curve as ability increases, until we reach a voter with an ability close to the median who prefers the progressive tax policy with $b = 0$ and the highest $a > 0$. From that point on we move to the left along the horizontal axis as ability increases, involving less and less progressivity as a decreases. Then we reach the median ability individual who prefers $(a_m, b_m) = (0.285, 0)$, $(0.25, 0)$ and $(-0.15, 0.4)$ for the uniform, the triangular and the linear decreasing distributions, respectively. The move to the left (lower a) continues as ability further increases, moving progressively upwards along the lower curve. This lower curve is in fact the lower bound of the feasibility condition (19) which means that all these bliss points are in fact constrained by the feasibility condition $b > \underline{b}(a)$.

Figures 4 a,b,c

We can now put the bliss points by increasing order of ability and derive the (indirect) utility function of each voter over this distribution of bliss points. Interestingly enough we obtain that for the three distributions of ability, each voter displays single-peaked preference over this ordered set of bliss points. Therefore it follows from the median voter theorem that the bliss point of the median voter is a Condorcet winner.¹⁰ Since the median voter is also the more favourable to progressivity we get a first possible explanation for the prevalence of progressive taxation.¹¹

In the following section, instead of reducing the policy space we study weaker solution concepts in the context of a standard Downsian political competition game.

¹⁰Roell (1997) obtains a similar result with quasi-linear preferences. We thank Robin Boadway for pointing out this paper to our attention.

¹¹In the case of the linearly decreasing distribution of ability, the median ability is so low that the median voter prefers regressivity like the low-ability voters.

5 The simultaneous two-party competition game

We consider a Downsian voting game with two political parties competing to win the election. Both parties simultaneously announce their tax schedule (a pair (a, b)) which they commit to impose if elected. Each individual then votes for the party whose platform is better according to his/her preferences. The party receiving the most votes wins the election and imposes its platform as the choice of the polity. In the event of ties, a fair coin decides which party wins the election. Note that in contrast with Section 4, it is assumed here that candidates can commit to any policy.

Formally, we denote the (finite) set of feasible policies by X . An element of X is thus a pair (a, b) . Let P denote the majority preference relation on this set X . The *majority preference* P over any policy pair $(x, y) \in X^2$ is given by,

$$\begin{aligned} xPy & : n(x, y) > n(y, x) \\ yPx & : n(x, y) < n(y, x), \end{aligned}$$

where $n(x, y) = \#\{w \in [0, 1] : v(x; w) > v(y; w)\}$ is the number of voters who (weakly) prefer x to y , and $n(y, x) = \#\{w \in [0, 1] : v(x; w) < v(y; w)\}$ is the number of voters who prefer y to x .¹²

Assuming an odd number of voters, the majority preference relation is a binary relation satisfying the asymmetry and completeness properties of a tournament.

The objectives assigned to the parties are crucial. We suppose that parties are *only* interested in winning the election and that they derive no intrinsic utility from the platform chosen (i.e., no ideology). Moreover parties are indifferent about the size of their majority: having a bare majority they attach no utility to any extra vote. Given that parties can choose among

¹²We rule out indifference since it is a very unlikely event.

the same set of feasible policies, we can represent this electoral competition by a symmetric two-player zero-sum game $G = (X, X, U)$ where:

$$U(x, y) = \begin{cases} 1 & : & xPy \\ -1 & : & yPx \\ 0 & : & x = y \end{cases}$$

This game, called the majority game, has a unique Nash equilibrium in pure strategies if and only if there exists a Condorcet winner. Formally, $x \in X$ is a *Condorcet winner* if and only if xPy for all $y \in X \setminus \{x\}$. In this case, both parties choose the Condorcet winner as a strategy.

But we know from Section 3 that there is no Condorcet winner in the (unrestricted) two-dimensional policy space. It follows that the game does not have any equilibrium in pure strategies: each party could win the election if it knew which policy is chosen by the other party.¹³

However, the absence of a Condorcet winner does not imply that any policy is equally likely to be selected by any party. First, we do not expect any party to select a Pareto dominated policy x (like $a > 0$ and $b > 1/3$).

Second, we feel confident that any party is unlikely to propose weakly dominated policies. A policy x is *weakly dominated* by policy y if $U(x, z) \geq U(y, z)$ for all $z \in X$, with a strict inequality for at least one z . Reformulated in our voting context, y weakly dominating x means that y beats x as well as any policy z that x can beat.

In the social choice literature, this dominance relation is also called the covering relation (see Miller, 1980). Formally, given a tournament (X, P) , a policy $x \in X$ is *covered* whenever there exists some other policy $y \in X$ such that yPx and $\{z : xPz\} \subseteq \{z : yPz\}$. The set of non-covered options for the majority preference relation P is called the uncovered set, denoted $UC(X, P)$. Since the covering relation is equivalent to the weak dominance

¹³This is in contrast to Cukierman and Meltzer (1991) who make very restrictive assumptions on the ability distribution and voters' preferences to get the existence of a Condorcet winner over quadratic income tax schemes. In our model even with very simple preferences and ability distributions the Condorcet winner fails to exist.

relation in our setting, the uncovered set is precisely the set of weakly undominated pure strategies of the two-party zero-sum game $G = \{X, X, U\}$. Given that any Pareto-dominated policy is covered, the uncovered set is a subset of the Pareto set.¹⁴

We can further restrict the set of interesting strategies by looking at the Nash equilibrium in pure strategies of this game.¹⁵ Laffond et al. (1993) have shown that the finite and symmetric majority game $G = (X, X, U)$ has a unique equilibrium in mixed strategies. They call the support of this unique equilibrium the *bipartisan set* denoted $BP(X, P)$.¹⁶ It can be shown that the bipartisan set is a subset of the uncovered set and thus a more discriminating solution concept (see Banks et al., 1996). Furthermore, both the uncovered set and the bipartisan set reduce to the Condorcet winner when the latter exists.

Given these definitions, we can now compute the Pareto set, the uncovered set and the bipartisan set for the three distributions of ability. The results are reported below.¹⁷

Insert Figures 5 a,b,c

As Figures 5 a,b,c indicate the distribution of abilities matters a lot in determining whether progressive taxation is likely to emerge in equilibrium. With a triangular distribution (Figure 5b), regressive taxes are weakly dominated (and a fortiori are never played in the mixed strategies Nash equilibrium). The reason is that this symmetric distribution implies a large

¹⁴If x is Pareto-dominated by y , y has a better position than x in every individual preference ordering. Hence, using majority voting, y beats all the options that x beats, which means that y covers x .

¹⁵See Laslier(2000) for an interpretation of electoral mixed strategies

¹⁶The bipartisan set is thus the set of options played with a strictly positive probability at the equilibrium.

¹⁷For each figure, a and b vary by increment of 0.05 respectively in $[-0.5, 0.5]$ and $[0, 1]$.

middle income group who can impose progressivity in order to reduce its own tax burden at the expense of the high-income group. The reverse holds when we shift the probability mass to the low ability levels as with the linear decreasing distribution (Figure 5c). In this case, the low-income group becomes sufficiently large to impose regressive taxes in order to maximize tax revenue (and thus redistribution). When ability levels are uniformly distributed (Figure 5a), then both regressive and progressive taxation can emerge in equilibrium. Therefore, these results suggest that the prevalence of progressive taxation can only be explained by the predominance of the middle class in the income distribution or equivalently by a lack of polarization at the extremes of the income distribution. Figures 5a,b,c also reveal that the Uncovered set and the Bipartisan set are very discriminating solution concepts. This is in contrast to Epstein (1997) who shows that for games of purely distributive politics the Uncovered set coincides approximately with the Pareto set. Our results demonstrate that for the game studied here, the Uncovered set (and a fortiori the Bipartisan set) can give rather sharp predictions on equilibrium outcomes.

In the next section we analyse how it is possible to get better predictions on the equilibrium outcomes by changing the rules which dictate the play of the game instead of changing the equilibrium concept.

6 The sequential two-party competition game

We use the dynamic version of the two-party electoral competition game due to Kramer (1977). In this voting game, two political parties repeatedly compete for the votes with the peculiarity that when a party is elected, it is committed to keep the same political platform for the next election (reputation inertia). In absence of Condorcet winner, this assumption implies that both parties alternate in office because each party can win the election once it knows the policy chosen by the other party. It is further assumed

that parties are interested in maximizing the size of their majority (or net plurality defined as $n(x, y) - n(y, x)$), and thus the opposition party will always select a policy which maximizes its voting share given the incumbent's policy. This dynamic voting process can be represented by the sequence of winning policies, which Kramer calls a *vote-maximizing trajectory* (i.e., a sequence of policies such that each policy along the sequence beats with a maximal number of votes its predecessor). Formally, for any two adjacent policies (x_t, x_{t+1}) along the sequence, $x_{t+1} \in \operatorname{argmax}_{y \in X} n(y, x_t)$.

Using our model, we have simulated the Kramer trajectories for the three distributions of ability. The results are reported in Figures 6a,b,c.¹⁸ For each distribution, we get that the Kramer trajectory converges to a cycle which is independent of the initial point. For each distribution we move clockwise along the cycle with a coalition of the extremes alternating with a coalition of the middle and low-income groups.

Figures 6 a,b,c

We have also represented the *minmax set* which is defined as the set of policies whose maximal opposition is minimal. Formally, $\operatorname{minmax}(X) = \operatorname{arg min}_x \max_{y \neq x} n(y, x)$. Kramer has demonstrated that for Euclidean preferences the minmax set behaves like a basin of attraction in the sense that any vote-maximizing trajectory converges to the minmax set. As shown below, for the model used here (with non-Euclidian preferences), we obtain that Kramer cycles always pass through the minmax set (which may reduce to a singleton).

According to these simulation results, the prevalence of progressive taxation can only be explained, as for the Bipartisan and the Uncovered sets, by the predominance of the middle class in the income distribution. Indeed only

¹⁸For each figure, a and b vary by increment of 0.025 respectively in $[-0.5, 0.5]$ and $[0, 1]$.

the triangular distribution (Figure 6b) with a lack of polarization at the extremes produces a Kramer cycle that only contains progressive taxes. Shifting the probability mass of individual abilities evenly towards the extremes (Figure 6a) makes the vote maximizing trajectory cycles evenly between regressive and progressive taxes; while a high probability mass at low-ability levels (Figure 6c) induces a Kramer cycle that only contains regressive taxes.

7 Conclusion

This paper is an attempt to explain the prevalence of income tax progressivity in a positive rather than normative perspective. We have used a highly stylized model that nevertheless includes the salient aspects of voting over non-linear income tax schedules. We have first shown that voting cycles over the set of progressive and regressive taxes is inevitable when taxation is not distortionary: progressive taxes (or convex tax function) enables the middle class to reduce its tax burden at the expense of the high-income group; while regressive taxes (concave tax function) reduces the tax burden at the extremes at the expense of the middle-class.

Introducing incentive effects allows us to take account of the fact that progressivity discourages effort which to some extent ticks the model in favor of regressive taxes. However we show that vote cycling over progressive and regressive taxes still prevails in the presence of incentive effects. Then we consider three different ways to give better predictions on the voting outcome and more importantly to explain the prevalence of progressive taxes. The first approach reduces the policy space to the tax schedules that are ideal for some voter. In that case we obtain a Condorcet winner corresponding to the policy preferred by the median voter. The second approach considers the entire policy space, but adopts weaker solution concepts in the context of a simultaneous two-party competition game. The third approach

considers a sequential two-party competition game in which parties not only care about winning, but also about the size of their majority.

The main result is that whatever the approach adopted, progressive taxes emerge as the only possible voting outcome when there is a lack of polarization at the extremes of the income distribution. In this exercise, we do not pretend we have succeeded in capturing the reason of the revealed preference for progressive income taxation, but perhaps we have captured some of its ground. Further research is of course needed to solve this demand for progressivity puzzle.

Acknowledgements. This paper was prepared for presentation at the ISPE conference on Public Finance and Redistribution at CORE, Belgium, June 2000. We wish to thank our discussants Gerard Roland and David Wildasin for helpful comments. We also thank participants at the Public Economics conference of the AFSE in Marseille, May 2000. This paper was initiated while Jean Hindriks was visiting the Institut d'Economie Industrielle (IDEI-CERAS-GREMAQ) in Toulouse, whose financial support and hospitality are gratefully acknowledged. The usual disclaimer applies.

References

- [1] Banks, J., J. Duggan and M. Le Breton, 1998, Bounds for mixed strategies equilibria and the spatial theory of elections, mimeo, University of Rochester.
- [2] Besley, T. and S. Coate, 1997, An Economic Model of Representative Democracy, *Quarterly Journal of Economics*, 112, 85-114
- [3] Cukierman A., A. Meltzer, 1991, A political theory of progressive income taxation, in *Political Economy*, by A. Meltzer, A. Cukierman and S.F. Richard.(Eds.). Oxford University Press. New York.
- [4] De Donder, P., J. Hindriks (1998). The political economy of targeting, *Public Choice*, 95, 177-200.
- [5] De Donder, P., M. Le Breton and M. Truchon, 2000, Choosing from a weighted tournament, *Mathematical Social Science*, forthcoming.
- [6] Epstein, D., 1997, Uncovering some subtleties of the uncovered set: social choice theory and distributive politics, *Social Choice and Welfare*, 15, 81-93.
- [7] Hindriks, J., 2000, Is there a demand for progressive taxation?, mimeo QMW.
- [8] Kramer, G., 1972, Sophisticated voting over multidimensional choice spaces, *Journal of Mathematical Sociology*, 2, 165-80.
- [9] Kramer, G., 1977, A dynamic model of political equilibrium, *Journal of Economic Theory*, 16, 310-34.
- [10] Kramer, G., 1983, Is there a demand for progressivity?, *Public Choice*, 41, 223-28.
- [11] Laffond, G., J.F. Laslier and M. Le Breton, 1993, The bipartisan set of a tournament game, *Games and Economic Behavior*, 5, 182-201.

- [12] Laslier, J.F., 2000, Interpretation of Electoral Mixed Strategies, *Social Choice and Welfare*, 17-2, 283-292.
- [13] Laslier, J.-F., A. Trannoy, K. Van der Straten, 2000, Voting under ignorance of unemployed job skills: the overtaxation bias, mimeo THEMA.
- [14] Marhuenda, F., I. Ortuno-Ortin, 1995, Popular support for progressive taxation, *Economics Letters*, 48, 319-24.
- [15] Miller, N., 1980, A new solution set for tournaments and majority voting, *American Journal of Political Science*, 24, 68-96.
- [16] Mitra, T., E. Ok, L. Kockesen, 1998, Popular support for progressive taxation and the relative income hypothesis, *Economics Letters*, 58, 69-76.
- [17] Myles, G., 2000, On the optimal marginal rate of income taxation, *Economics Letters*, 113-19.
- [18] Osborne, M.J., A. Slivinski, 1996, A Model of Political Competition with Citizen candidates, *Quarterly Journal of Economics*, 111, 65-96.
- [19] Plott, C.R., 1967, A Notion of Equilibrium and its Possibility under Majority Rule, *American Economic Review*, 57, 787-806
- [20] Roberts, K., 1977, Voting over income tax schedules, *Journal of Public Economics*, 8, 329-40.
- [21] Roell, A., 1997, Voting over non-linear income tax schedules, mimeo Tilburg University.
- [22] Roemer, J., 1999, The democratic political economy of progressive income taxation, *Econometrica*, 67, 1-19.
- [23] Romer, T., 1975, Individual Welfare, Majority Voting, and the Properties of a Linear Income Tax, *Journal of Public Economics*, 4, 163-185.

- [24] Snyder, J. and G. Kramer, 1988, Fairness, self-interest, and the politics of the progressive income tax, *Journal of Public economics*, 36, 197-230.
- [25] Stigler, G.J., 1970, Directors' law of income distribution, *Journal of Law and Economics*, 13, 1-10.
- [26] Young, P., 1990, Progressive taxation and equal sacrifice, *American Economic Review*, 80, 253-66.

Figure 3a: Majority voting cycle with uniform distribution

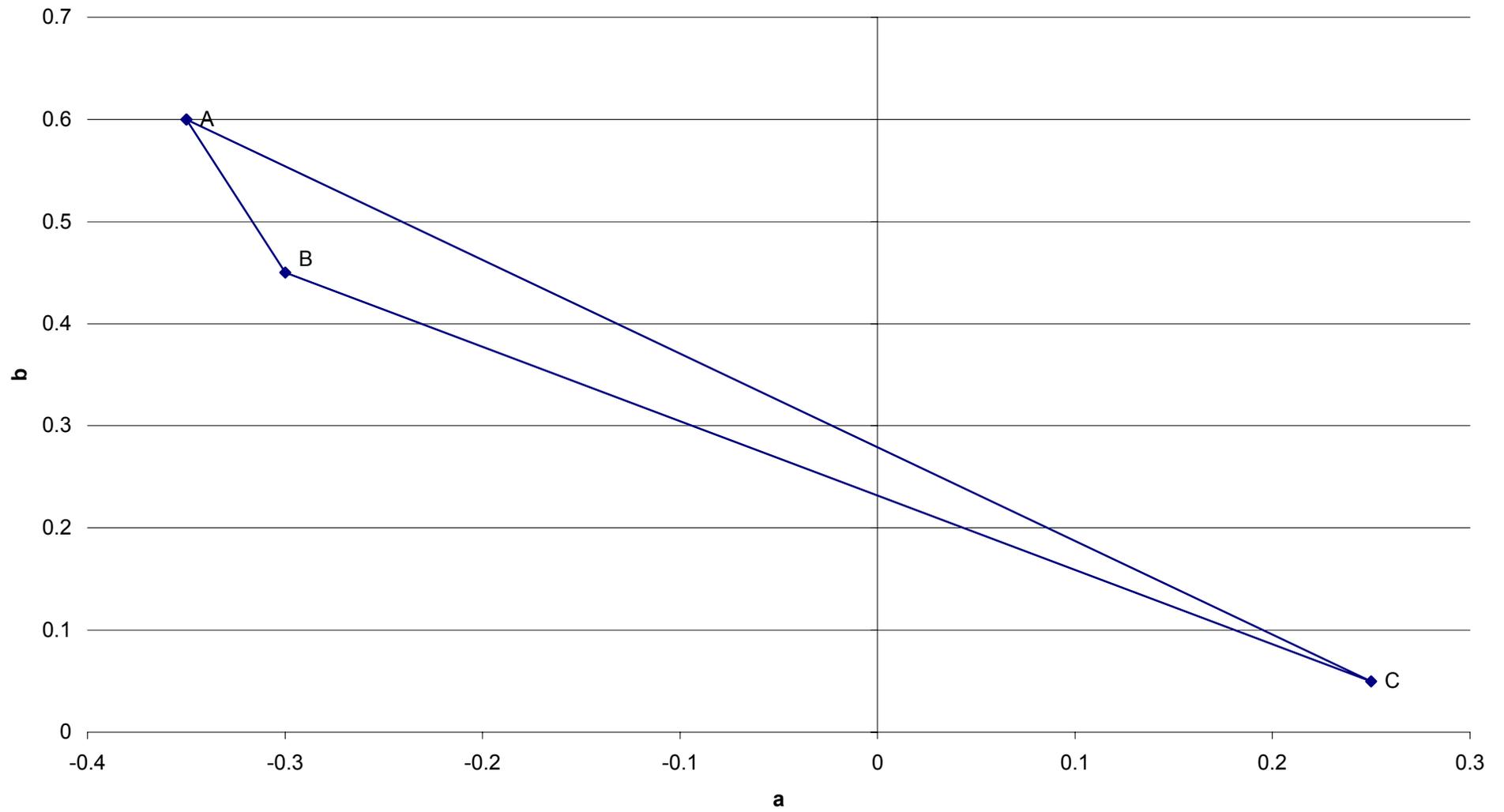


Figure 3b: Majority voting cycle with triangular distribution

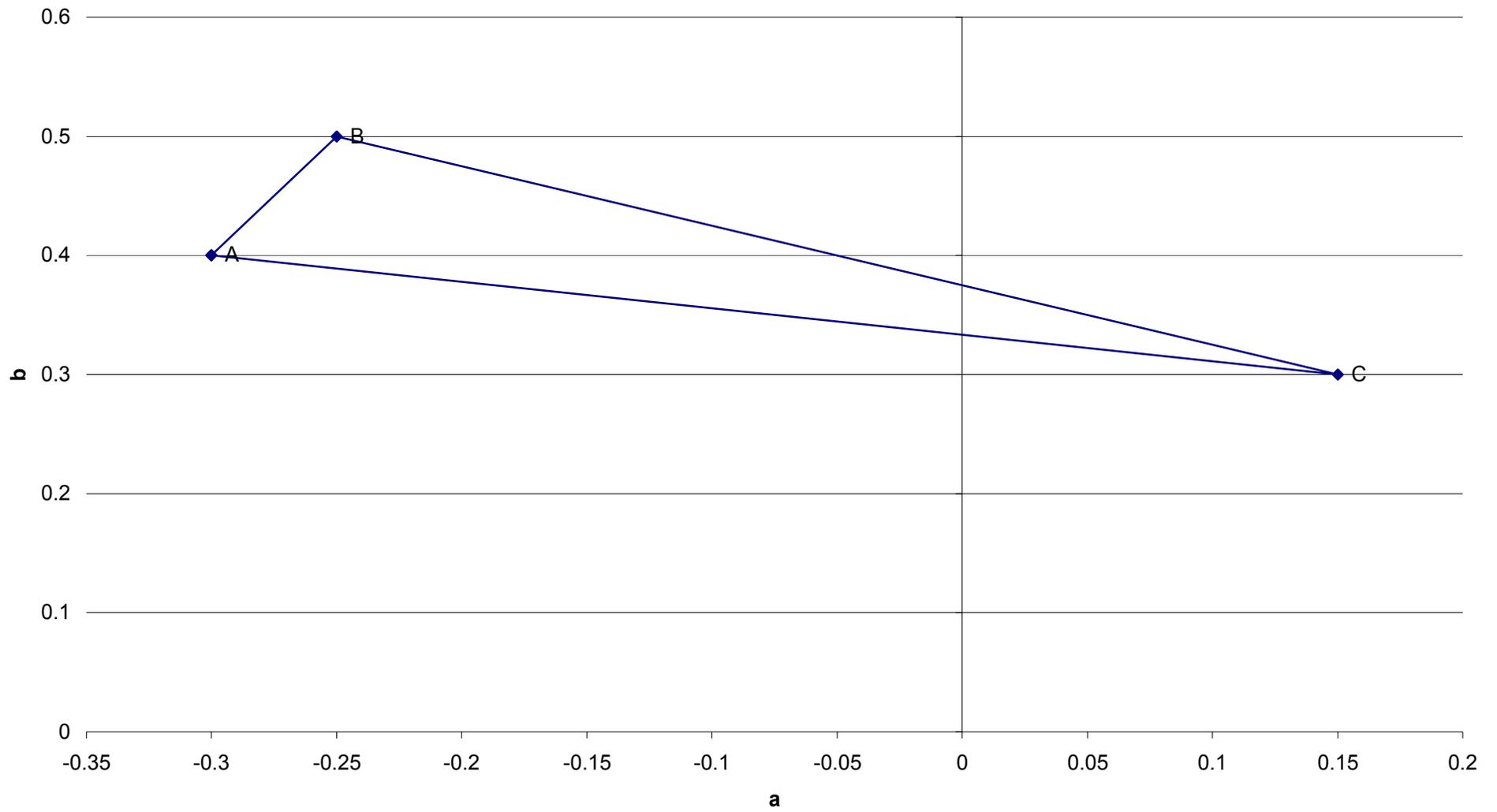


Figure 3c: Majority voting cycle with linearly decreasing distribution

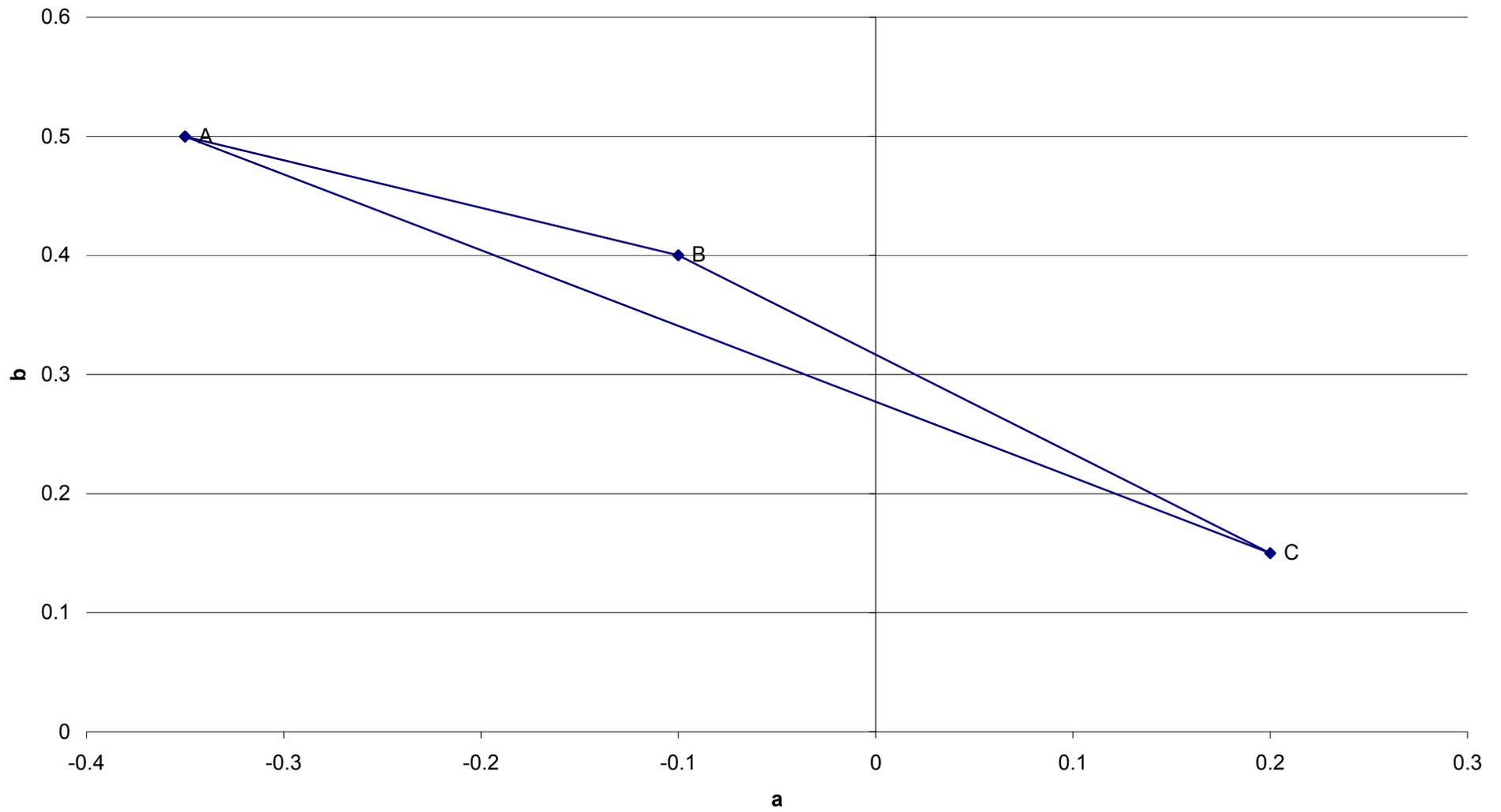


Figure 4a: Voters' blisspoints with uniform distribution of abilities

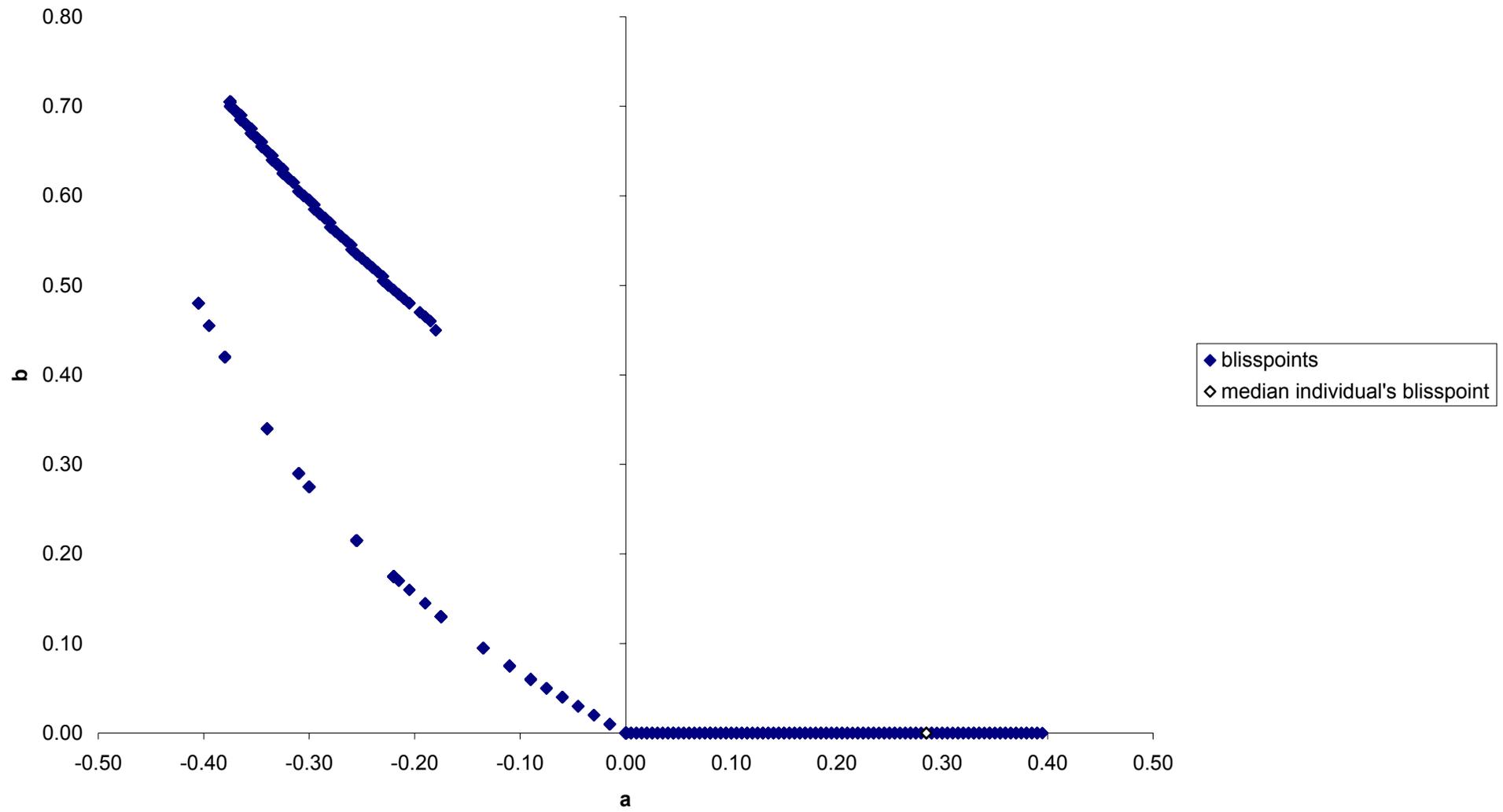


Figure 4b: Voters' blisspoints with triangular distribution of abilities

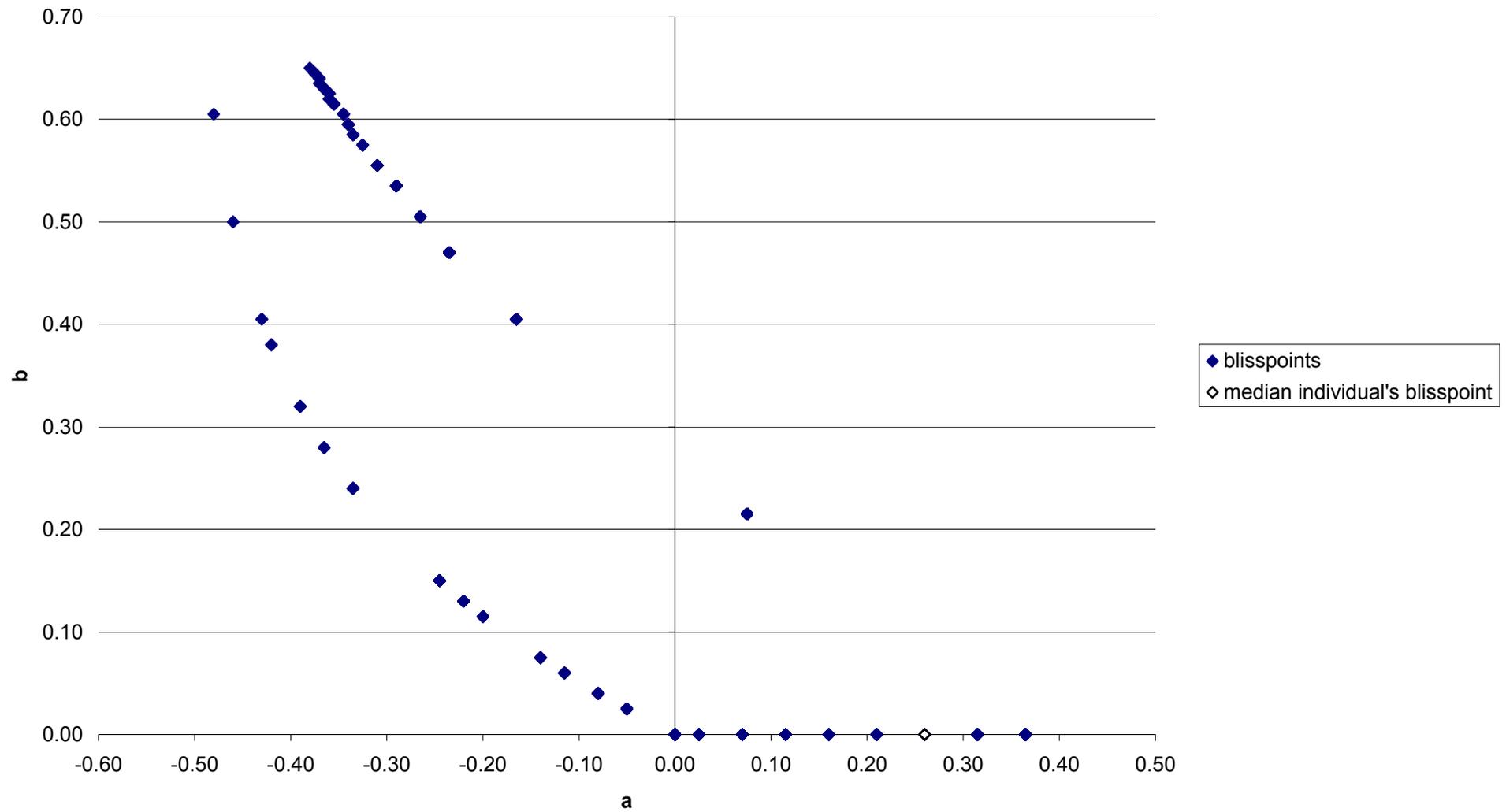


Figure 4c: Voters' blisspoints with linearly decreasing distribution of abilities

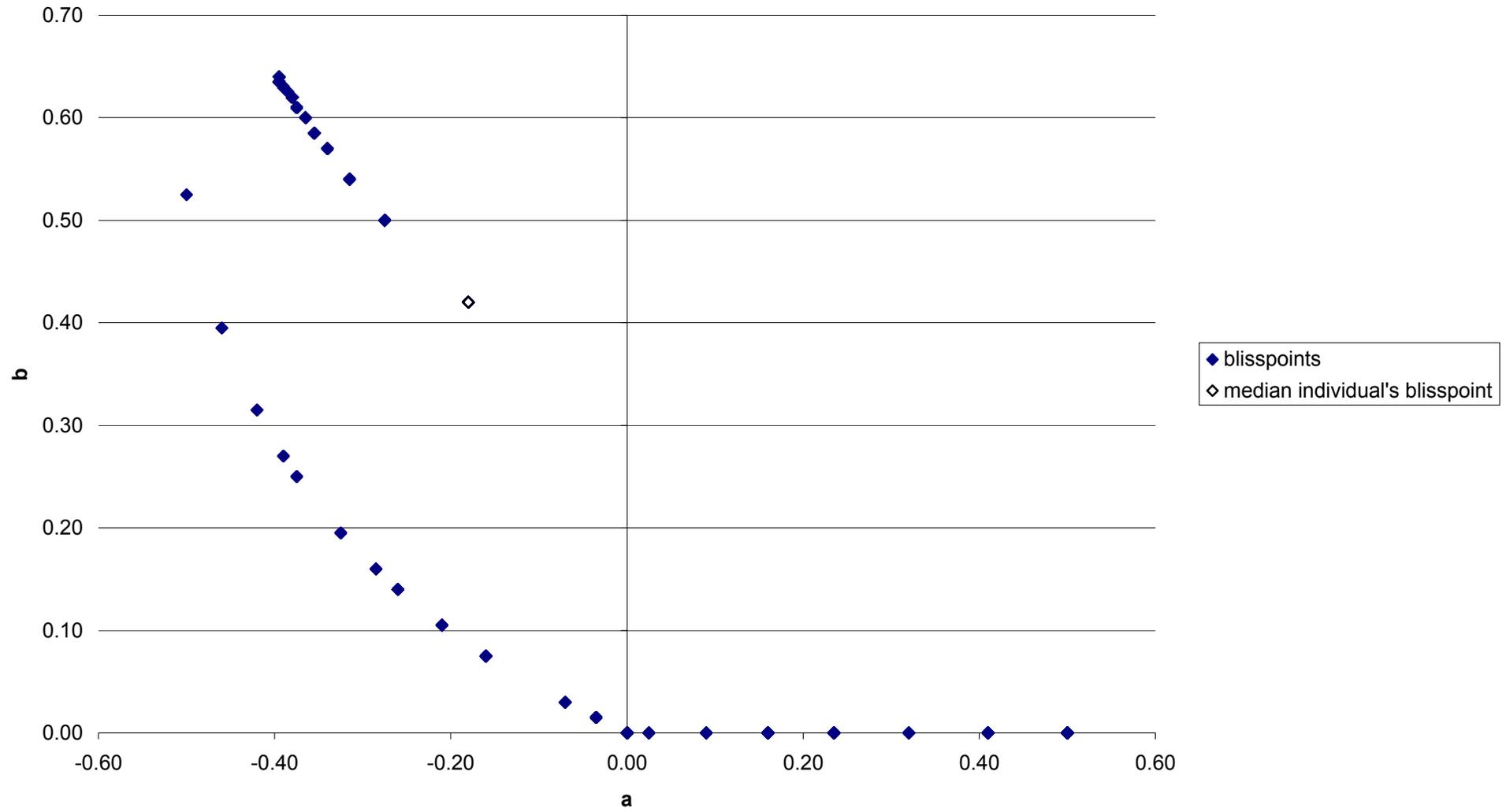


Figure 5a: Pareto set, Uncover set and Bipartisan set for the uniform distribution of abilities

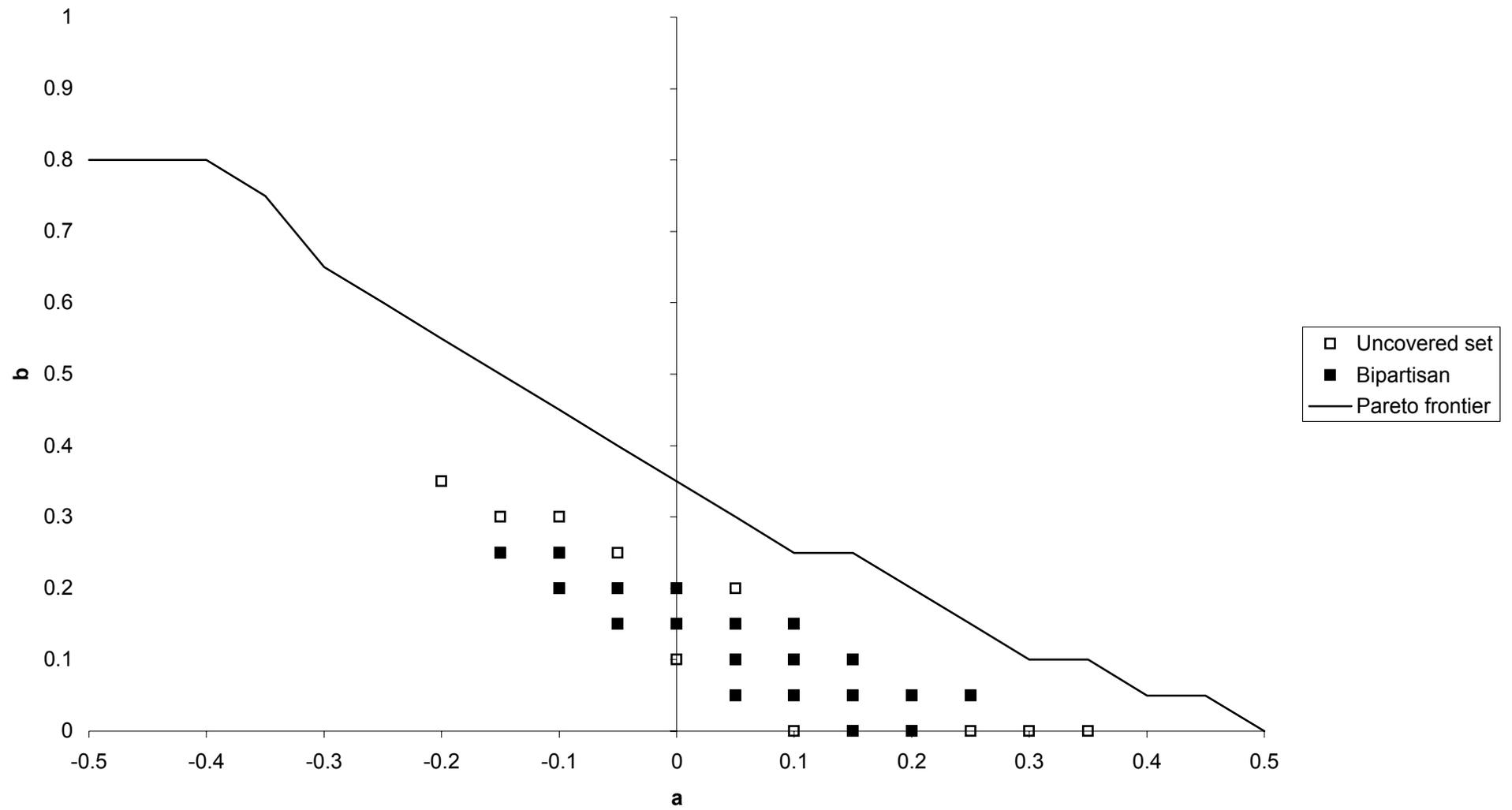


Figure 5b: Pareto set, Uncover set and Bipartisan set for the triangular distribution of abilities

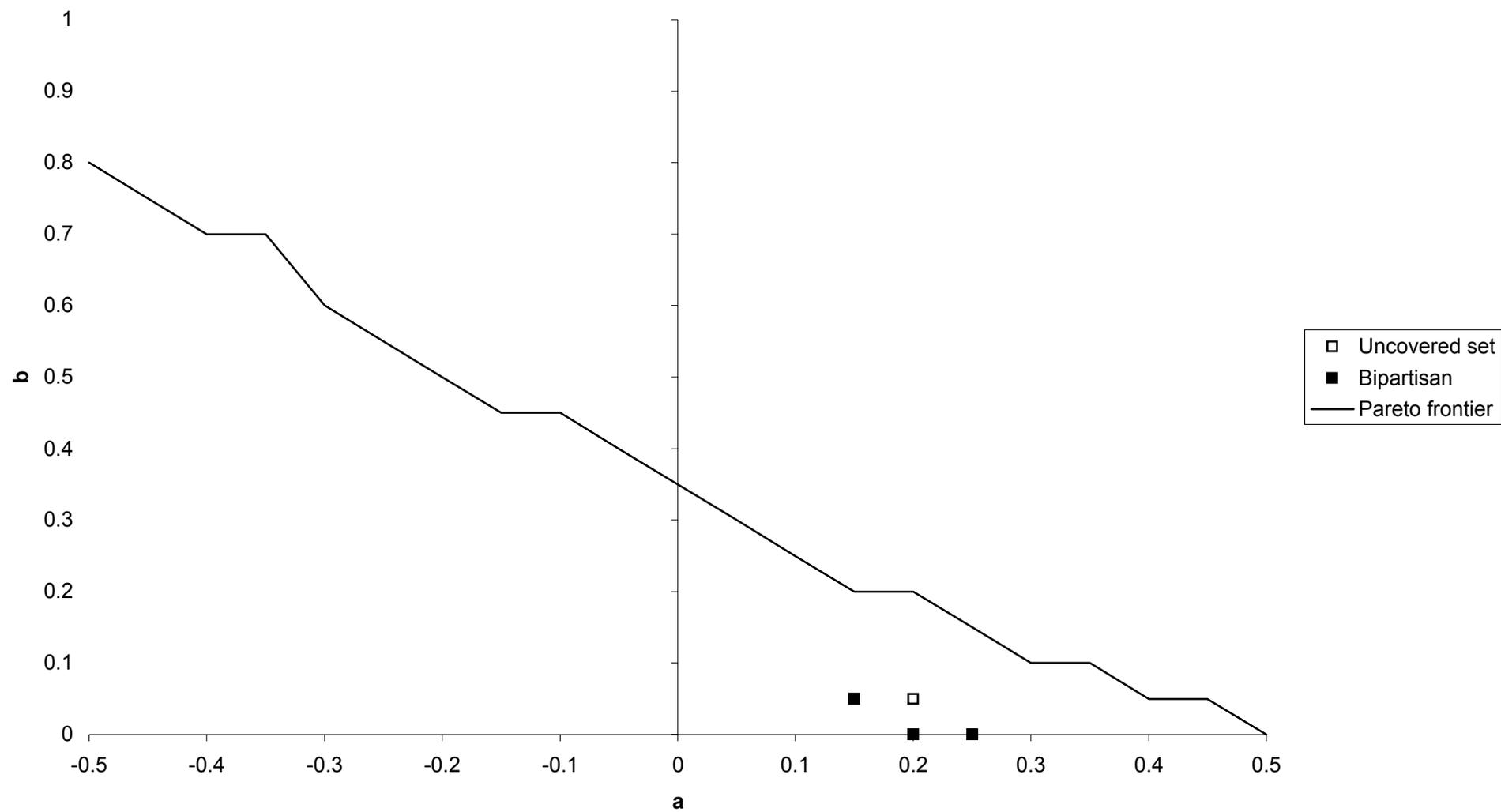


Figure 5c: Pareto set, Uncover set and Bipartisan set for the linearly decreasing distribution of abilities

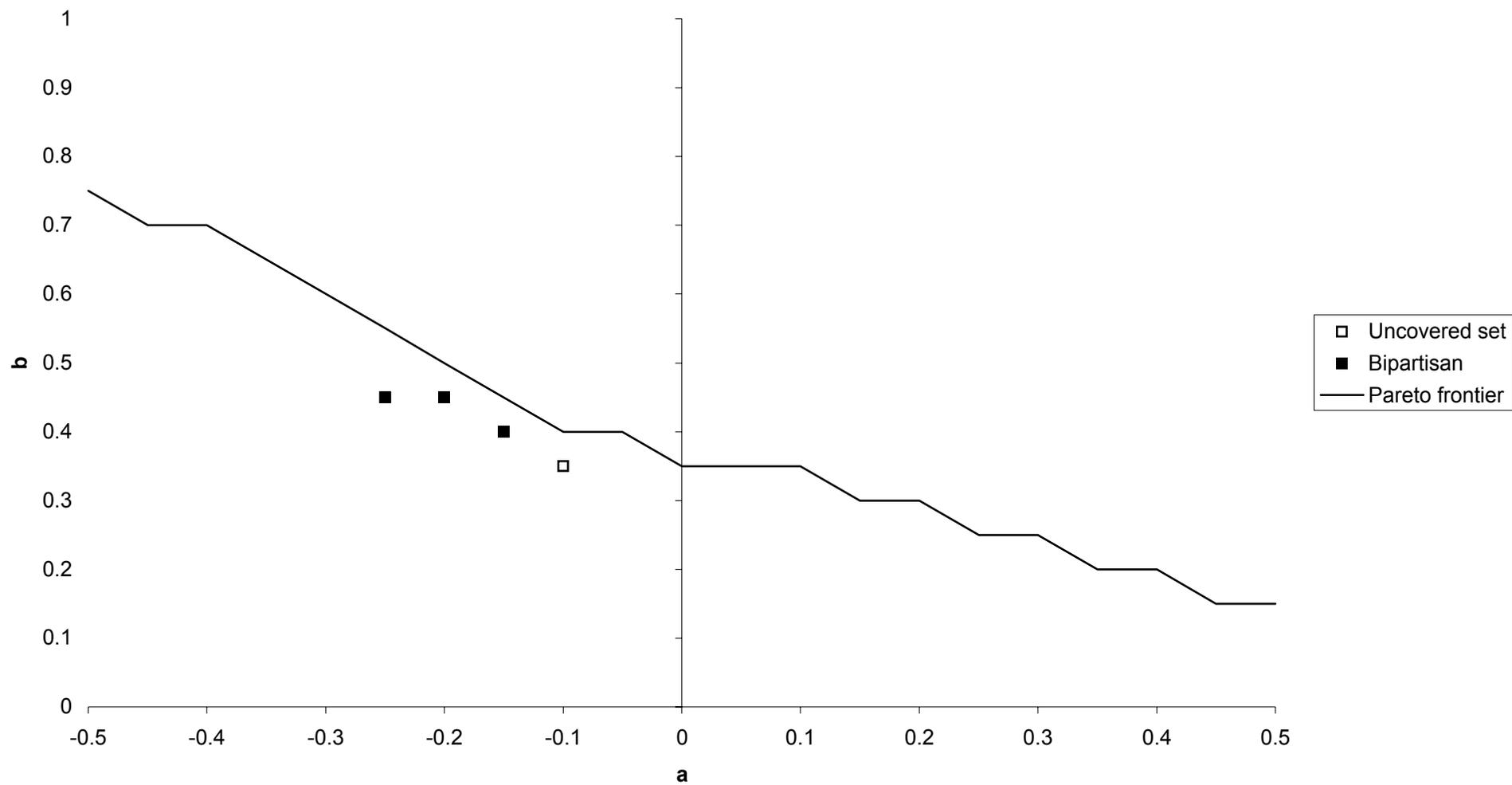


Figure 6a: Kramer cycle and minmax with uniform distribution of abilities

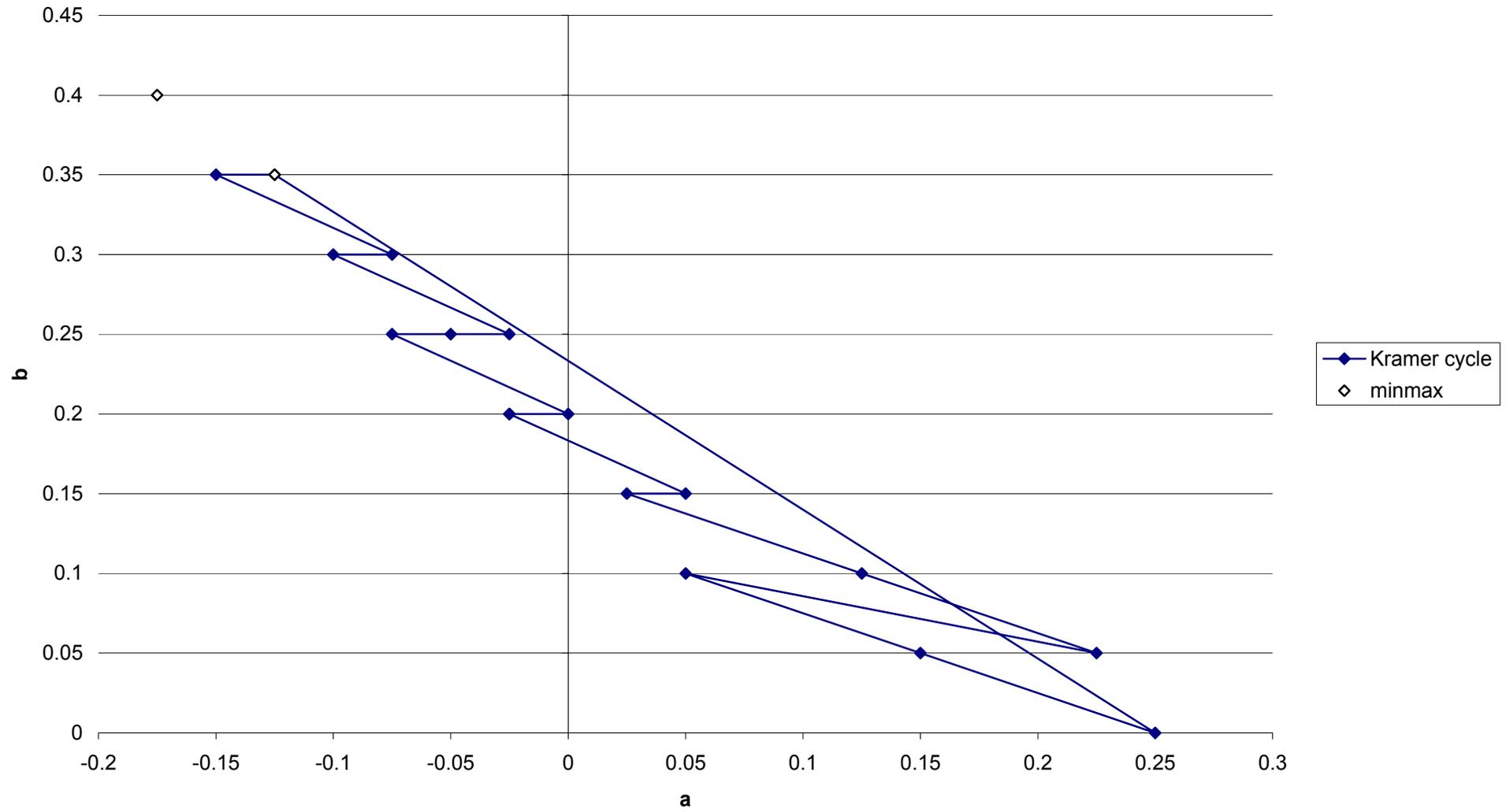


Figure 6b: Kramer cycle and minmax with triangular distribution of abilities

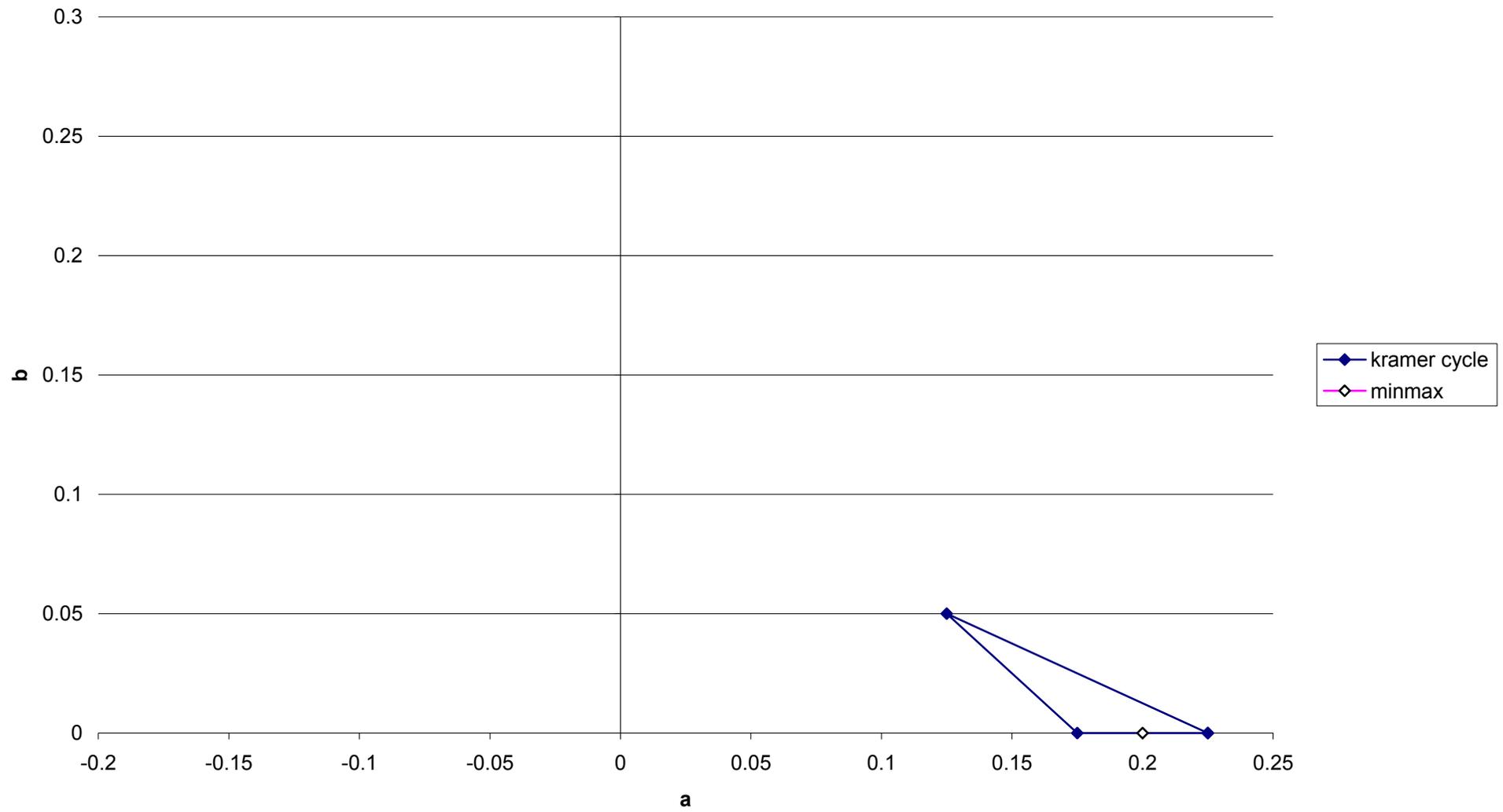


Figure 6c: Kramer cycle and minmax with linearly decreasing distribution of abilities

