

Blomquist, Sören; Christiansen, Vidar

Working Paper

Why Politicians Favor Redheads - A Theory of Tactical Horizontal Redistribution.

Working Paper, No. 2000:10

Provided in Cooperation with:

Department of Economics, Uppsala University

Suggested Citation: Blomquist, Sören; Christiansen, Vidar (2000) : Why Politicians Favor Redheads - A Theory of Tactical Horizontal Redistribution., Working Paper, No. 2000:10, Uppsala University, Department of Economics, Uppsala

This Version is available at:

<https://hdl.handle.net/10419/82912>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Why Politicians Favor Redheads

A Theory of Tactical Horizontal Redistribution

by

Sören Blomquist*

and

Vidar Christiansen**

Abstract

This paper studies a very pure form of “vote purchasing”. We consider whether it may be in the interest of a party to discriminate between groups that, possibly except for size, are identical in all welfare relevant aspects, i.e. the groups are assumed to have the same income, needs, etc. To emphasise this aspect we label the groups brown-heads and redheads. The interpretation is that they differ only in some characteristic that is entirely irrelevant from a welfare perspective. There are no systematic differences between people with the same income. Taking two samples of people from an income class their political support will be identically distributed. We will show that even with these uniformity assumptions there can be strong incentives for political parties to undertake vote purchasing by favouring one of the identical groups at the expense of the other.

Keywords: vote purchasing, tactical redistribution, political economy

JEL classification: D72

* Department of Economics, Uppsala University, Box 513, SE-751 20 Uppsala, Sweden.
E-mail: soren.blomquist@nek.uu.se

** Department of Economics, University of Oslo, P.B. 1095 Blindern, N-0317 Oslo, Norway.
E-mail: vidar.christiansen@econ.uio.no

1. Introduction

In all developed countries we observe a multitude of taxes, transfers and public expenditures. Often these policies can be rationalised on efficiency grounds or as measures to achieve distributional gains according to prevalent social preferences. However, we also observe policies that can be interpreted as "vote purchasing". For example a public project (hospital, airport, military base, research institute) may be located to a certain region in order to gain votes in that region. Also contractionary measures may be selective. When there is to be a cut-back of the armed forces in a country, the decision to retain a certain base may be motivated by the wish to buy votes in the area where the camp is located. The vote purchasing can take the form of improving the economic conditions for certain groups like farmers, the elderly, families with children or other groups.¹ Owners of certain capital assets, e.g. owner occupied housing, may be taxed more leniently than those who invest in other forms of capital. The politicians are often seen to remain advocating policies that are dismissed by economists on conventional efficiency and distributional grounds. Actually taking the advice may prove costly in terms of political support.

The purpose of this paper is to obtain a better understanding of the political parties' tactical redistribution between groups in order to enhance the support from the electorate. We shall consider whether it may be in the interest of a party to discriminate between groups that, possibly except for size, are identical in all welfare relevant aspects, i.e. the groups are assumed to have the same income, needs, etc. To emphasise this aspect we label the groups *brown-heads* and *redheads*. The interpretation is that they differ only in some characteristic that is entirely irrelevant from a welfare perspective. It is hard to think of any reason why a party or voter would attach any welfare weight to the colour of a person's hair. The basic assumption is that there are no systematic differences between people with the same income. Taking two samples of people from an income class their political support will be identically distributed. We will show that even with these uniformity assumptions there can be strong incentives for political parties to undertake vote purchasing by favouring one of the identical groups at the expense of the other.

¹ For example, in the 1998 Swedish election the social democrats promised substantially reduced day care fees for families with children. One interpretation of this proposal is that it mainly was designed to buy votes from parents with children in day care ages. In Norway old aged pensioners pay a lower tax than the work active with the same income.

To illustrate the basic mechanism we can think of a simple example. There are two political parties, leftist and rightist, and two classes of people, high income and low income people. Both parties want to redistribute from high income to low income people but the leftist party wants to redistribute more than the rightist party. We assume there are 2000 low income people and 1000 high income people. We also assume that the low income class can be subdivided into two equally sized groups according to a characteristic that is of no distributional or efficiency concern. The groups are labelled redheads and brown-heads. The rightist party would like to pursue redistribution in such a way that each high income person loses \$ 2000 and each low income person gains \$ 1000. The leftist party wants to redistribute so that each high income person loses \$ 3000 and each low income person gains \$ 1500. If the parties were to suggest these policies, and assuming selfish voting behaviour, all low-income people would vote for the leftist party and all high-income people for the rightist party and the leftist party would win the election. The rightist party might then ponder a "vote purchasing" policy. It might propose a policy such that the high-income people lose \$ 2000, low income redheads gain \$ 2000 while low income brown-heads gain nothing. If the leftist party does not react to this policy the rightist party might gain all the votes from redheads and win the election. This policy is considered inferior by both parties to the policy where high income people lose \$ 2000 and all low income people gain \$ 1000. Hence, a policy both parties consider inferior would win. Of course, the leftist party might respond to the "vote purchasing" policy of the rightist party. We do not pursue the example further here but do a formal analysis below.

Redistributional policy with the intention of winning votes has often been coined tactical redistribution. The typical way that tactical redistribution has been modelled is by assuming that society consists of groups that differ in voting behaviour [Lindbeck and Weibull (1987, 1993) and Dixit and Londregan (1995, 1998a, 1998b)]. Our analysis shares a number of features with previous discussions in the literature. Parties care about ideology, but also about the vote they get. In our model, and some others in the literature, the primary concern is with ideology. No weight is given to the vote as such. The vote matters only because a party has preferences over the actual political outcome. Since it prefers its own policy to be implemented, a large vote is a means to increase the likelihood of its own policy winning. Another feature shared with much of the literature is that voters to some extent are selfish and care about their own consumption, while they also care about ideology and the degree of social redistribution achieved by the policy they support. It is common to assume that the parties compete for the votes of groups that are different. The groups are normally distinguished by

three properties: income, size and ideological attachment (party loyalty, depth of political conviction). Some groups can more easily be induced to shift their support from one party to another. They are more responsive to economic favours to themselves and their support is more easily up for “sale” to parties that offer group specific benefits. These people are to a larger extent “swing voters” - voters that swing from one party to another depending on the favours they are offered. A distinction of groups along these lines is a key element in Dixit and Londregan (1995, 1998).²

Although our model in several respects is similar to Dixit and Londregan (1998a) and (1998b), it differs from the Dixit and Londregan (DL) model in four important aspects: *i.* In the DL model two groups with identical properties will be treated identically. In our model two identical groups might be treated unequally. *ii.* In the DL model the size of the groups is of no importance for how they are treated. In our model the size is sometimes of critical importance. *iii.* In the DL model individuals’ preferences for an equal income distribution vary in a continuous way whereas a parameter indicating selfishness is fixed. In our model we also consider variation in the selfishness parameter. *iv.* In the DL model all relevant functions are sufficiently smooth and have concavity properties such that a local analysis is sufficient to characterize an optimum. In our model we account for non-convexities, which means that we have to perform a global analysis in order to give a characterization of the conditions under which discrimination is worthwhile for the political parties.

The rest of the paper is organized as follows. In section 2 we set up our basic model and in section 3 we study its properties under the assumption that it is not possible for the parties to pursue discriminating policies. We study the Nash equilibrium. Section 4 surveys the arguments for and against discriminating. In sections 5 and 6 we introduce the possibility for the parties to discriminate according to some observable characteristic. Section 7 concludes.

2. The Model

The model we use is in many respects similar to the one in Blomquist and Christiansen (1999) and is inspired by Hansson and Stuart (1984), Lindbeck and Weibull (1993) and Dixit and Londregan (1998a). There are two parties denoted by L (leftist party) and R (rightist party). There are two classes of individuals: n_1 low-income individuals and n_2 high-income individuals with exogenous incomes Y_1 and Y_2 , respectively. The low-income class can be

² In Johansson (1999) voters differ in income, ideological preferences and location.

divided into two types: n_r individuals of type r (redheads) and n_b individuals of type b (brown-heads).

Both parties want to redistribute from the high to the low income class according to preferences that will be described below. For simplicity we assume that redistribution can be done without distortions of individuals' behavior. Hence, all allocations considered are Pareto efficient in the traditional sense of the concept. We instead focus on "political distortions". Let C_i^p denote the consumption expenditure (disposable income) of an individual in class i that party p proposes. In a non-discriminating equilibrium a policy consists of a pair $\{C_1^p, C_2^p\}$; $p = L, R$. A discriminating policy consists of a triple $\{C_r^p, C_b^p, C_2^p\}$; $p = L, R$. A feasible policy must satisfy the budget constraint $n_r C_r^p + n_b C_b^p + n_2 C_2^p = n_1 Y_1 + n_2 Y_2$. As a shorthand we denote the policies as z^p ; $p=L,R$.

The parties

Each party proposes a policy in order to maximise its expected welfare as defined below. We assume that the policy that is actually implemented is the policy of the party that wins the election. The underlying presumption is that each party is able to credibly commit to a certain policy. Thus, we rule out the possibility that a party does not keep its promises from the election campaign³. It is common in the literature to make some kind of assumption to the effect that the parties implement their promised policies if winning. Sometimes it is just imposed as an assumption (e.g. Dixit and Londregan (1998)), sometimes the assumption is implicit (e.g. Lindbeck and Weibull (1987) or Hansson and Stuart (1984)), and sometimes there is an appeal to underlying assumptions. Lindbeck and Weibull (1993) assume that at least all voters *believe* that the winning party will implement its announced policy. They also appeal to the existence of future elections (otherwise not appearing in the model) to suggest that cheating may be prohibitively costly for the parties in terms of future loss of credibility. Besley and Coate (1997) point out that it is natural to assume that the citizen who wins the election implements his

³ A complication in practice may be that at the time the policy is designed the state of the world that will materialise in the period of office is not known with certainty. Since there is limited scope for state contingent election manifestos, it may be a matter of interpretation whether the actual policy is according to the pre-election platform when allowing for the need to adjust to the circumstances that materialise. We abstract from uncertainty of this kind.

preferred policy (i.e. the policy he would choose as a dictator), while commitment to other policies is more problematic due to lack of credibility.⁴

At the pre-election stage the outcome of the election is perceived as random. We explain this in detail below. Once the outcome of the election is known, each party evaluates the winning policy according to its social welfare function. We assume there is no benefit from winning the election as such (no "ego-rent"). It follows that prior to the election the objective function of a party is the expected welfare according to the preferences of the party.

We assume that the social welfare function of each party is a weighted sum of the utilities obtained by each class. The utility that an individual derives from a consumption level C is denoted by $V(C)$. Letting \mathbf{a}_p denote the weight given to class one, the social welfare for party p of a policy z^q is given by

$$W^p(z^q) = \mathbf{a}_p n_r V(C_r^q) + \mathbf{a}_p n_b V(C_b^q) + (1 - \mathbf{a}_p) V(C_2^q), \quad p = L, R; \quad q = L, R. \quad (1)$$

The party's expected welfare is then:

$$\mathbf{p}(z^L, z^R) W^p(z^R) + (1 - \mathbf{p}(z^L, z^R)) W^p(z^L) \quad p = L, R \quad (2)$$

where $\mathbf{p}(z^L, z^R)$ is the probability that party R will win the election. This probability depends on the policy choices of the two parties. Below we describe how this probability is determined. We assume the leftist party assigns a higher weight to the welfare of the low-income class and that both parties assign a smaller weight to class 1 than to class 2, i.e. $0.5 > \mathbf{a}_L > \mathbf{a}_R$.⁵ We shall address below the behaviour of the parties on the basis of these preferences.

Individuals' voting behavior

We assume that a voter choosing ballot takes into account both his own utility and his notion of what is good for society as a whole. More precisely, the *policy preference function* of an individual i in class 1 is assumed to be a weighted average of a selfish part $V(C_i^p)$ and an "unselfish" part (a social welfare function): $W^p = \mathbf{a}_i (n_r V(C_r^p) + n_b V(C_b^p)) + (1 - \mathbf{a}_i) n_2 V(C_2^p)$, where \mathbf{a}_i is the weight given to the aggregate utility of class one. We write the policy

⁴ See Alesina (1988).

⁵ In our model there is no explicit modelling of disincentives. But we have to acknowledge that a tax policy that gives high-skill individuals a higher before tax income but lower after tax income than low-skill people would be unfeasible. A crude way to avoid such a solution is to assume that both parties assign a sufficiently higher welfare weight to high-skill persons.

preference function as $\mathbf{g}_i V(C_i^P) + (1 - \mathbf{g}_i)W^P$ where \mathbf{g}_i is the weight given to the selfish part. The function $P_i^R = \mathbf{g}_i V(C_i^R) + (1 - \mathbf{g}_i)W^R - \mathbf{g}_i V(C_i^L) - (1 - \mathbf{g}_i)W^L$ then shows the advantage to individual i of the R policy rather than the L policy being implemented. Individual i obviously prefers the policy of R if $P_i^R > 0$. In general we will refer to weights \mathbf{a} and \mathbf{g} without a subscript. It is implicit that the values vary across individuals.

We assume that a voter in class 2 is entirely selfish and votes for the party that offers him the higher income C_2 . Hence, for individuals in class two the policy preference function takes the form $W_2 = V(C_2^P)$, which is only the selfish part.⁶ It is purely for analytical convenience that we assume there is no “unselfish” part in the policy preference function of class 2 individuals. Voter aversion towards discrimination is already in the model via the unselfish part of the policy preference function of class 1 individuals. Adding unselfishness to the policy preference function of class two people would maybe increase the realism of our model, but would not change its qualitative character.

We can argue that it is optimal for the individuals to vote sincerely⁷.

Let m_1 and m_2 be the number of type 1 and type 2 individuals, respectively, that prefer the policy of party R and let $m(z^L, z^R) = m_1 + m_2$.⁸ If the voting behaviour is entirely deterministic each party can calculate for any configuration of policies whether it is going to win the election. The policies will then converge as from the perspective of any party a winning policy is always better than a losing policy as long as the winning policy is considered at least slightly preferable to the policy of the opponent. This scenario is not very realistic. In real life we never observe complete policy convergence. Also, in practice there are numerous elements which are beyond the control of the parties and which they will treat as random. In the literature several approaches have been adopted to model random voting behaviour.⁹ But it seems that the

⁶ Formulating the policy preference function for class 2 people in this way implies that all individuals in group 2 are identical and that they all will vote in the same way. All votes of this group will go to one party. The formulation also implies that class two people do not care about class one people and do not mind discriminatory policies towards groups within class 1.

⁷ The same case can be established as in Blomquist and Christiansen (1999). The main argument is that a voter cannot credibly commit to vote for the party he likes less.

⁸ In sections 3 we show how $m(z^L, z^R)$ is determined when \mathbf{g} varies across the population and in section 4 we show how $m(z^L, z^R)$ is determined when \mathbf{a} varies.

⁹ Sometimes it is just postulated that probabilities depend on policy choices. Sometimes more specific assumptions are made. For instance Dixit and Londregan (1998) assume that there are groups of extremist voters supporting their respective parties irrespective of election manifestos, but the size of each group is unknown to and is treated as random by the parties.

exact way random voting behaviour is modelled is of less importance. What is central is that the parties can only affect the probability of winning and the *expected* outcome in terms of policy and welfare. Accepting this premise we make the assumption about the probability of victory for party R that $\mathbf{p}(z^L, z^R) = \mathbf{p}(m(z^L, z^R))$ with $\mathbb{1}\mathbf{p} / \mathbb{1}m > 0$ ¹⁰ In appendix A we show one simple way to derive such a relationship between m and \mathbf{p} .

3. The non-discriminating political equilibrium

Before proceeding to political equilibria it is helpful to establish as a benchmark the parties' *dictator policies* which we define as the policies derived from maximising the respective welfare functions of the parties expressed by (1). In policy space the dictator policies are denoted as $\bar{z}^L = \{\bar{C}_1^L, \bar{C}_2^L\}$ and $\bar{z}^R = \{\bar{C}_1^R, \bar{C}_2^R\}$. We denote the corresponding values of V^2 as \bar{V}_R^2 and \bar{V}_L^2 , respectively. It is clear that even if discrimination were allowed the dictator policies would be non-discriminatory¹¹.

A few properties of the dictator policies are easily established. Assume that a party maximises $W = \mathbf{a}n_1V(C_1) + (1-\mathbf{a})n_2V(C_2)$ subject to the budget constraint $n_1C_1 + n_2C_2 = Y$. The first order conditions imply that

$$\frac{V'(C_2)}{V'(C_1)} = \frac{\mathbf{a}}{(1-\mathbf{a})} \quad (3)$$

We see immediately that if $\mathbf{a} = 0.5$, then $C_1 = C_2$. It also follows that due to the concavity of the utility function and the budget constraint, an increase in \mathbf{a} implies that C_1 will increase and C_2 will decrease. Two further implications are then immediate. If $\mathbf{a} < 0.5$, then $\bar{C}_1 < \bar{C}_2$. As a benevolent dictator the party with the larger value of \mathbf{a} will allocate more income to class one. Since we have assumed that $\mathbf{a}_L > \mathbf{a}_R$ it follows that $\bar{C}_2^L < \bar{C}_2^R$.

When discussing possible incentives for a party to start discriminating between redheads and brown-heads we take as our point of departure a Nash equilibrium where discrimination is ruled out. At a Nash equilibrium party L maximizes the expected welfare w.r.t. C_2^L for a given C_2^R and party R maximizes the expected welfare w.r.t. C_2^R for a given

¹⁰ This is not an entirely trivial assumption as supporting a party is not necessarily the same as actually going to the polling-both to vote for it.

¹¹ For a fixed income to be allocated between the r and b groups welfare is maximised by giving the same income to everybody in those groups.

C_2^L . Note that in each case it is sufficient to determine C_2 as C_1 then follows from the budget constraint. We denote a Nash equilibrium as $\{\hat{C}_2^L, \hat{C}_2^R\}$.

Proposition 1: If discrimination is not allowed, then there exists a Nash equilibrium such that

$$\bar{C}_2^L < \hat{C}_2^L < \hat{C}_2^R < \bar{C}_2^R.$$

See appendix B for proof.

At the nondiscriminating equilibrium the R party is more generous than the L party towards class 2, but the policies converge to some extent compared to the dictator policies.

In general the voting problem might be degenerate in the sense that the parameters of the economy might be such that the dictator policy of one of the parties would win with certainty. If class two were a majority of the population, the party proposing the higher income for class two would win with certainty. The R party could then propose its dictator policy, and the L party would have no incentive to outbid or even match the R party's offer to class two. On the other hand the class sizes and the distribution of parameters might be such that no outcome with a positive probability would induce a majority to vote against the dictator policy of party L . We shall assume, without spelling out the conditions in detail, that the parameters of the economy are such that a degenerate case does not materialise.

4. Discrimination – pros and cons.

Taking the non-discrimination equilibrium as our point of departure we shall consider whether and under what conditions party R will have incentives to start discriminating by unequal treatment of redheads and brown-heads. By convention we label as redheads the group that is favoured in the case of differential treatment. We confine ourselves to discrimination from the perspective of the R party. The analysis would be similar by considering discrimination from the point of view of the L party.

To study the model in full generality is quite complicated. We shall therefore throughout the discussion of potential discrimination keep the income of class two fixed so the question is whether a fixed income to class one is to be shared equally by redheads and brown-heads or not. Thus the assumption is that party R considers a discriminating policy such that C_1^R is unchanged but $C_r^R > C_b^R$.

Discrimination will have two effects on the expected welfare of party R . First, the party's own welfare, $W^R(z^R)$, decreases. This follows from the concavity of the V function.

Splitting a given amount unevenly between redheads and brown-heads gives lower welfare than splitting it evenly. Second, the probability of winning the election changes. A brown-head dislikes the discriminating policy on two accounts. First, the selfish part $V(C_b^R)$ is less than in the initial situation. Second, the value of the unselfish part has also decreased since a given amount of consumption for the low-income class now is split unevenly. Hence, the political support from brown-heads to the R party will decrease. For a redhead in class one there are two opposing forces. The redhead likes the new R policy better because the selfish part $V(C_r^R)$ is now more favourable than before. However, the redhead likes the policy less because the value of the unselfish part has gone down.

Let z^p denote the policy of party p and let subscripts u and d indicate uniform and discriminating policies versus the groups r and b . Assume that R substitutes a discriminating policy for a uniform one. Then the welfare for party R changes as follows

$$\begin{aligned} D &= \mathbf{p}_d W^R(z_d^R) + (1 - \mathbf{p}_d) W^R(z^L) - [\mathbf{p}_u W^R(z_u^R) + (1 - \mathbf{p}_u) W^R(z^L) - \mathbf{p}_d W^R(z_u^R) + \mathbf{p}_d W^R(z_u^R)] \\ &= (\mathbf{p}_d - \mathbf{p}_u) (W^R(z_u^R) - W^R(z^L)) - \mathbf{p}_d (W^R(z_u^R) - W^R(z_d^R)) \end{aligned} \quad (4)$$

It is perfectly possible that $D \leq 0$, which implies that the uniform equilibrium remains in place even when discrimination is possible. We see that the following conditions are conducive to party R discriminating

1. R considers its uniform policy as significantly better than that of party L
 $[(W^R(z_u^R) - W^R(z^L)) \text{ is large}]$
2. Discrimination significantly increases the probability of winning $[(\mathbf{p}_d - \mathbf{p}_u) \text{ is large}]$
3. Party R finds the cost of discrimination small $[(W^R(z_u^R) - W^R(z_d^R)) \text{ is small.}]$

Whether a redhead prefers the discriminating policy or not depends on the values of \mathbf{a} and \mathbf{g} . In principle we could invoke a bivariate distribution over \mathbf{a} and \mathbf{g} . However, this would make the model very complicated. Below we therefore consider two special cases; one where \mathbf{a} is constant but \mathbf{g} varies across the population according to a pdf $g(\mathbf{g})$. In the second case we let \mathbf{g} be fixed and \mathbf{a} varies across the population according to a pdf $f(\mathbf{a})$.

In the formal analysis below we shall start by considering whether the R -party has incentives to deviate marginally from the non-discriminating equilibrium. We will denote this as *marginal discrimination*. We then move on to consider non-marginal discrimination implying a finite increase in redhead income and a corresponding decrease in brown-head income. This can lead to two kinds of discrimination. With *non-extreme discrimination* we

mean discrimination such that at least one brown-head prefers the R -party. With *extreme discrimination* we mean discrimination such that the R -party forfeits the votes of all brown-heads.

5. Incentives for discrimination when the selfishness parameter, g , varies.

Two parameters characterise the preferences of each voter in class one. a determines the weight assigned to class one in the social welfare consideration, while g reflects the selfishness of a voter. It is convenient to consider variation in one parameter at the time. We start by addressing the case where all individuals in class 1 have the same value of a , while the value of g differs across individuals. People are not equally selfish. The distribution of g is described by the probability density function $g(g)$ with support $0 \leq g \leq 1$. As discussed above an individual in class 1 supports party R if $P^R = gV(C_1^R) + (1-g)W^R - gV(C_1^L) - (1-g)W^L$ is greater than zero and otherwise supports party L . $P^R = 0$ defines indifference between the two parties and also a cut-point value of g separating the supporters of the two parties in case both parties obtain some support.

If differential treatment of redheads and brown-heads is in general possible, there is in each group (redheads and brown-heads) a cut-point value of g , g_r and g_b , respectively, between those supporting the two parties.

$$g_i = \frac{W_1^R - W_1^L}{W_1^R - W_1^L + V_i^L - V_i^R} \quad i=r,b \quad (5)$$

If we consider a situation in which the L party is the more generous towards class one individuals, the latter will have a selfish reason to support L . There will only exist persons who are indifferent between the two parties if the R party is preferred on social welfare grounds. Then the least selfish individuals (having a small g) will support the R party while those with a selfishness parameter above the cutpoint value will support the L party. The number of individuals in class 1 preferring party R 's policy is then given by

$$m_1(z^L, z^R) = n_a \int_0^{g_r} g(t) dt + n_b \int_0^{g_b} g(t) dt \quad (6a)$$

If the distribution function of g is uniform with support $[0,1]$ the function simplifies to:

$$m_1(z^L, z^R) = n_r g_r + n_b g_b \quad (6b)$$

However, if the R party outbids the L party in pursuit of supporters from a group without winning it entirely it will be the more selfish who become R supporters.

Individuals in class 2 will vote for party R if $V(C_2^R) > V(C_2^L)$. The number of individuals in class 2 preferring party R is therefore given by:

$$m_2 = n_2 \quad \text{iff} \quad C_2^R > C_2^L$$

$$m_2 = 0 \quad \text{iff} \quad C_2^R < C_2^L$$

We neglect the possibility that $C_2^L = C_2^R$.

Marginal discrimination

Whether the R party can benefit from a marginal discrimination depends on the initial situation and the distribution of \mathbf{g} . It follows from Proposition 1 that at the initial non-discriminating equilibrium $V_1^L > V_1^R$, and nobody in class one will support the R party for purely selfish reasons. However, the equilibrium may be such that $W_1^L \geq W_1^R$ or the other way round, i.e. class one people may favour either party on welfare grounds. In general various regimes are possible at the initial equilibrium.

Proposition 2: Suppose \mathbf{a} is constant. If everybody in class one prefers the L party on welfare grounds ($W_1^L \geq W_1^R$) the R party can gain nothing by marginal discrimination.

Proof: Given the assumptions in the proposition $P^R < 0$ for all feasible values of \mathbf{g} . Everybody finds the L policy socially superior as well as preferable from a selfish perspective. There is a strict preference in favor of the L party and the R party can gain nothing by marginal discrimination.

If class one people prefer the R party on welfare grounds ($W_1^L < W_1^R$), the effect of marginal discrimination depends on the distribution of \mathbf{g} and is in general unsettled. We provide one example where marginal discrimination is worthwhile and one in which it is not.

Proposition 3: If $W_1^R > W_1^L$ and everybody in class one is sufficiently selfish to be supporting the L party, but some only marginally [$g(\mathbf{g}) = 0$ for $0 \leq \mathbf{g} < \bar{\mathbf{g}}$, while $g(\mathbf{g}) > 0$ for $\mathbf{g} \geq \bar{\mathbf{g}}$], the R party will gain by marginal discrimination.

Proof: Given the assumptions of the proposition there is nobody with a value of γ sufficiently low to support the R party. Everybody is so selfish that he supports the L party, and it gains all the votes from class 1. The marginal supporter of L prefers the R party from a welfare perspective, but just refrains from supporting R for selfish reasons.

In this special case the R party can always gain votes by a marginal discrimination. The reason is simple. By slightly favouring redheads the R party will induce the marginal L -supporters to switch their support to the R party. The selfish reason for supporting L has been eroded. However, the brown-heads who are discriminated against are already L -supporters, so from the perspective of the R party there is no loss of support from this group. At the same time there is no welfare loss from a marginal redistribution as the initial welfare weight given to a unit of income is the same whether the unit is allocated to redheads or brown-heads. The first order effect on welfare is zero, but since the party gains more support it will be better off in expected terms.

Proposition 4: Suppose that \mathbf{a} is fixed while there is a uniform distribution of \mathbf{g} . The R -party can gain no additional support from marginal discrimination.

Proof: The proof of this proposition is given in the section below.

Non-marginal discrimination

Even if no voter is indifferent in the first place the R party may gain votes by redistributing enough income from brown-heads to redheads. The precondition is that the R party allocates sufficient income to class 1 initially so that there is enough income to redistribute to redheads to outbid the L party's offer. In that case the R party will buy the more selfish voters among the redheads as those are the people who are more easily induced shift their support in response to personal favours. However, it is an open question whether the party will indeed want to discriminate as by doing so it incurs a welfare loss if in any case it were to win.

Let us then revert to the case where the parties share the support of class one. Assume \mathbf{g} is uniformly distributed on the interval $[0,1]$. Then the number of R supporters in class 1 is given by (5b): $m_1 = \mathbf{g}_r n_r + \mathbf{g}_b n_b$. We then consider the effects of simultaneously increasing C_r and decreasing C_b .

$$\frac{dm_1}{dC_r^R} = n_r \frac{d\mathbf{g}_r}{dC_r^R} + n_b \frac{d\mathbf{g}_b}{dC_r^R} = \quad (7)$$

$$n_r \frac{\partial \mathbf{g}_r}{\partial W_1^R} \frac{dW_1^R}{dC_r^R} + n_b \frac{\partial \mathbf{g}_b}{\partial W_1^R} \frac{dW_1^R}{dC_r^R} \quad \text{I}$$

$$+ n_r \frac{\mathcal{I} \mathbf{g}_r}{\mathcal{I} V_r^R} \frac{\mathcal{I} V_r^R}{\mathcal{I} C_r^R} + n_b \frac{\mathcal{I} \mathbf{g}_b}{\mathcal{I} V_b^R} \frac{\mathcal{I} V_b^R}{\mathcal{I} C_b^R} \left(-\frac{n_r}{n_b} \right) \quad \text{II}$$

Decomposing the effect into the expressions I and II is helpful for further analysis. The two expressions are both zero at departure from a non-discriminating equilibrium. Marginal utilities of income are the same and there is no first order effect on welfare. When we move to a non-marginal discrimination I becomes negative. (The social welfare function does not approve of horizontal inequity.) To see what happens to II let us consider its second derivative, which consists of four terms:

$$n_r \frac{\mathcal{I} \mathbf{g}_r}{\mathcal{I} V_r^R} V''(C_r^R) \quad \text{(i)}$$

$$+ n_r \frac{\mathcal{I}^2 \mathbf{g}_r}{\mathcal{I} (V_r^R)^2} (V'(C_r^R))^2 \quad \text{(ii)}$$

$$+ n_b \frac{\mathcal{I} \mathbf{g}_b}{\mathcal{I} V_b^R} (V''(C_b^R)) \left(-\frac{n_r}{n_b} \right)^2 \quad \text{(iii)}$$

$$+ n_b \frac{\mathcal{I}^2 \mathbf{g}_b}{\mathcal{I} (V_b^R)^2} (V'(C_b^R))^2 \left(\frac{n_r}{n_b} \right)^2 \quad \text{(iv)}$$

It follows immediately from (4) that

$$\frac{\mathcal{I} \mathbf{g}_i}{\mathcal{I} V_i^R} > 0 \quad i=r,b$$

$$\frac{\mathcal{I}^2 \mathbf{g}_i}{\mathcal{I} (V_i^R)^2} > 0 \quad i=r,b$$

It follows that if $|V''|$ is sufficiently small relative to V' , the second derivative above is positive and II becomes positive when non-marginal discrimination is pursued. Moreover, II will dominate over I, and m_1 increases. If $|V''/V'|$ is sufficiently large, m_1 decreases.

To consider the effect of discrimination on party welfare we differentiate

$$W = \mathbf{a}(n_r V(C_r) + n_b V(C_b)) + (1 - \mathbf{a})V(C_2)$$

$$\frac{dW}{dC_r} = -\mathbf{a}n_r (V'(C_b) - V'(C_r)) \quad (8)$$

which is negative for $C_r > C_b$.

To get a measure which is independent of units of measurement for utility we may measure the welfare loss due to discrimination in terms of the incremental income that the party would have to give to class two in order to offset the effect of discrimination. We can express this measure as

$$\frac{dC_2}{dC_r} = \frac{-\mathbf{a}n_r (V'(C_b) - V'(C_r))}{(1 - \mathbf{a})V'(C_2)}$$

As further discrimination is pursued this measure will change as follows

$$\frac{d(dC_2/dC_r)}{dC_r} = \frac{\mathbf{a}n_r}{1 - \mathbf{a}} \left[\frac{V''(C_b) V'(C_b) n_r}{V'(C_b) V'(C_2) n_b} + \frac{V''(C_r) V'(C_r)}{V'(C_r) V'(C_2)} \right] < 0$$

If there is a non-marginal discrimination, the loss of welfare is larger, the larger is $|V''/V'|$.

If $|V''/V'|$ becomes sufficiently small the welfare effect tends to vanish.

We note that if the marginal utility of income is diminishing at a sufficiently low rate it is possible for the R party to gain additional support by non-marginal discrimination, and there is little concern with the horizontal inequity it causes. Thus the party will have an incentive to discriminate. If the marginal utility of income diminishes at a sufficiently high rate the concern with equity does not only grow stronger, but the party will also lose support by unequal treatment of redheads and brown-heads. To sum up our findings we can state.

Proposition 5: Suppose that \mathbf{a} is fixed, while there is a uniform distribution of \mathbf{g} . Taking a non-discriminating policy of the L party as given, the R party can gain additional support and possibly increase its expected welfare by non-marginal discrimination if the marginal utility of income is diminishing at a sufficiently low rate.

To get some intuition for the result we may note that when the R party gives a small income to class one individuals the marginal supporter will only give a small weight to his own utility otherwise he would not be supporting a party treating him so unfavourably. But then the only slightly more selfish will also give little weight to the own utility. These are the potential new supporters the R party can hope to win. The implication is that buying additional

supporters is expensive in the following sense. Since they do not give much weight to extra own income a large “bribe” is required to induce them to switch their support to the R party. As the R party increases the income allocated to the redheads of class one the marginal supporters and marginal non-supporters become more selfish. As these are persons giving more weight to additional own income they are cheaper to “buy” than the previous new supporters. In this sense discrimination becomes a more efficient tool for attracting voters as its magnitude increases. This explains why a non-marginal discrimination can achieve what a marginal discrimination fails to achieve.

On the other hand one should note that the voters dislike violation of horizontal equity, as does the party, and even more so the higher is the rate at which the marginal utility of income diminishes. In this sense there is an additional cost which explains why it is only with certain qualifications that a non-marginal discrimination is worthwhile.

If the distribution is non-uniform there will be an additional effect in either direction. If the distribution is skewed in such a way that the density is higher in the neighbourhood of g_r and lower in the neighbourhood of g_b , there will be a tendency to win more supporters as g_r increases and g_b decreases. This would strengthen the case for discrimination even between groups that are identical – redheads and brown-heads.

Extreme and non-extreme discrimination.

Let us assume that the two groups differ in size and that it is possible for party R to “buy” either group. Which group would it prefer to “buy”? Assuming that there are voters of all degrees of selfishness, there is a positive density even for $g = 1$ (complete selfishness). In order to buy all the voters in class one the R party must then offer the members of the group an income that matches that offered by the L party. In order to gain the support of the r group the R party must set $C_r^R = C_1^L$. The cost in terms of additional income given to the r group is then $n_r(C_1^L - C_1^R)$.

The loss of income per person in the b -group is then $\frac{n_r}{n_b}(C_1^L - C_1^R)$. Then

$$C_b^R = C_1^R - \frac{n_r}{n_b}(C_1^L - C_1^R).$$

It can be shown that the loss of welfare from the perspective of the party is increasing in the size of the group that is “bought”. On the other hand buying a larger group is more attractive from the point of view of winning support. (The additional support that is obtained is

$(1-\bar{g})(n_w - n_a)$ where n_w is the size of the group that is won and n_a is the size of the other group.

Suppose the R -party buys the support of the entire r group. Then the loss of welfare from the perspective of R is \mathbf{a}_R times the following expression (as easily seen from the welfare function):

$$K = V(C_1^R) - V\left(C_1^R - \frac{n_r}{n_b}(C_1^L - C_1^R)\right)(n_1 - n_r) - [V(C_1^R - (C_1^L - C_1^R)) - V(C_1^R)]n_r \quad (9)$$

To examine the effect of group-size consider

$$\frac{\partial K}{\partial n_r} = -\left[V(C_1^R + (C_1^L - C_1^R)) - V\left(C_1^R - \frac{n_r}{n_b}(C_1^L - C_1^R)\right)\right] + V\left(C_1^R - \frac{n_r}{n_b}(C_1^L - C_1^R)\right)(C_1^L - C_1^R)\left(1 + \frac{n_r}{n_b}\right) \quad (10)$$

Since V is a strictly concave function

$$\left[V(C_1^R + (C_1^L - C_1^R)) - V\left(C_1^R - \frac{n_r}{n_b}(C_1^L - C_1^R)\right)\right] < V\left(C_1^R - \frac{n_r}{n_b}(C_1^L - C_1^R)\right)\left(1 + \frac{n_r}{n_b}\right)(C_1^L - C_1^R)$$

which implies that $\frac{\partial K}{\partial n_r} > 0$. A larger group is more expensive to “buy” in terms of welfare.

This is true for an arbitrary value of \mathbf{a}_R . So also from the perspective of a voter (characterised by some \mathbf{a}_i) the welfare loss is larger if a larger group is “bought”.

When party R starts discriminating the point of departure is a situation where R gives the same to r - and b -individuals which is less than what L offers. Party R is welfare preferred to party L . $C_r^R = C_b^R = C_1^R < C_1^L$ and $W_1^R > W_1^L$. For this situation to occur it must be the case that $\mathbf{a}_L > \mathbf{a}_1$, otherwise class one individuals would derive a higher welfare from the L -party’s policy.

Now let C_r^R be increased and C_b^R be decreased, which implies a decline in W_1^R . If this process is allowed to continue we can conceive of the following outcomes.

1. A stage is reached where $C_r^R = C_1^L$ while still $W_1^R > W_1^L$. Then R wins the support of the entire r group, while retaining some b supporters (who support R on welfare grounds despite the selfish argument to the contrary). This will happen more easily if r is the smaller group as the welfare loss is then smaller, and W_1^R will less easily fall below W_1^L .

2. Before stage 1 is reached, the party may not find further discrimination worthwhile and it will be content to have some of the redheads and some of the brown-heads as supporters.
3. A stage is reached where $W_1^R = W_1^L$ while still $C_r^R < C_1^L$. Then R has lost the entire r group and obviously also the whole b group. A further redistribution will then have no effect on the party's support until $C_r^R = C_1^L$. As one moves beyond this point and $C_r^R > C_1^L$, the R party will start winning supporters among the most selfish redheads, while having no support from brown-heads. There is extreme discrimination.

It is of interest to note that if there is non-extreme discrimination the R -party will win the support of the least selfish, while if there is extreme discrimination the R -party will win the support of the most selfish. However, what both kinds of discrimination have in common is that one attracts more selfish supporters than one would otherwise get.

In terms of pure support each stage may be a local optimum. We may note that if people exist who are completely unselfish, the R party does not entirely lose the support of a group until $W_1^R = W_1^L$, and we are back to stage 3. If everybody is selfish to some extent ($g(\mathbf{g})$ is zero for sufficiently low values of \mathbf{g}), the R party entirely loses the support of the b group for sufficiently low values of C_b^R and W_1^R even with $W_1^R > W_1^L$.

We can show the following result.

Proposition 6: If $\mathbf{a}_R > \mathbf{a}_1$ and there are unselfish people, extreme discrimination will not happen.

Proof: Let $D(\mathbf{a})$ be the difference between the welfare derived from the R -policy and that obtained at the L -policy for a preference parameter \mathbf{a} .

$$D(\mathbf{a}) = \mathbf{a}(n_r V_r^R + n_b V_b^R - n_2 V_2^R - n_r V_r^L - n_b V_b^L + n_2 V_2^L) + n_2 V_2^R - n_2 V_2^L$$

special cases are, of course, that only one or neither party discriminates. When the parties propose policies each party must prefer its own policy to that of its opponent. The implication is that $D(\mathbf{a}_R) > 0$, $D(\mathbf{a}_L) < 0$ and $D'(\mathbf{a}) < 0$. It follows that a preference $\mathbf{a}_1 < \mathbf{a}_R$ will prefer R 's policy as $D(\mathbf{a}_1) > D(\mathbf{a}_R)$. Provided there are sufficiently unselfish people the R party will never lose the support of a whole group.

In general it may happen that the party wants to win the support of a whole group (outcome 1) or forfeit the support of an entire group (outcome 3). These outcomes are more likely the more selfishness matters, because then it is "cheaper" to buy votes. If horizontal inequity does not reduce welfare very strongly, it may well happen that an entire group is bought, but it will

be the larger group only if the welfare effect of horizontal inequity is very moderate. It is clear that one will never buy an entire very large group. This is easy to see once we realise that if a group is large enough the transfers that are needed to “buy” this group may not be feasible. Quite plainly there is not enough income to confiscate from the smaller group. Even without such extreme cases it is costly to buy the support of a large group as we discussed above.

6. Incentives for discrimination when \mathbf{a} varies across group 1 individuals

In this section we will study the case where people have the same selfish concern but differ in their social attitudes to the classes. In terms of the parameters of the model \mathbf{g} is fixed but \mathbf{a} varies across the class 1 population. Whether it pays for the R -party to discriminate or not depends both on how the cut-off values of \mathbf{a} are affected by the discrimination and on the distribution of \mathbf{a} . We will try and isolate these effects. First we study the effect of changing cut-off values. We do this under the assumption that \mathbf{a} is uniformly distributed, which means that there is no bias in the distribution either in favor or against discrimination. We thereafter study the effect of the distribution of \mathbf{a} .

As our point of departure we take a non-discriminating equilibrium such that the R -party has all the votes of class 2 and part of the votes in class 1. This implies that $V_2^R > V_2^L, V_1^L > V_1^R$. Also, for the marginal voter preferring the R -party it must be true that $W_u^R > W_u^L$.

Non-extreme discrimination

As expressed in proposition 7 below, when \mathbf{a} is uniformly distributed the incentives for the R -party to favour the redheads and discriminate against the brown-heads are quite different from the case where \mathbf{a} is fixed and there is a uniform distribution over \mathbf{g} .

Proposition 7. Let the non-discriminating equilibrium be such that party R gets all the votes of class 2 and part of the votes in class 1. Let $V(\cdot)$ be concave, \mathbf{g} constant and the distribution of \mathbf{a} uniform. Then it is never in the interest of the R -party to perform non-extreme discrimination.

The proposition covers both marginal and nonmarginal discrimination. It can be shown that marginal discrimination will not affect the number of potential supporters of the R -party nor the expected welfare of the R -party as the first order effects are all zero. It is more complicated to show that nonmarginal discrimination is not worthwhile. A *necessary* condition for discrimination to be worthwhile for the R -party is that the number of individuals in class 1

voting for the R party increases while the number of L -party supporters decreases. We will show that the number of class 1 individuals that vote for party R will not increase as the R -party discriminates against the brown-heads, given the conditions in the proposition.

Let us define P_{ij} as:

$$P_{ij} = (1-\mathbf{g})\mathbf{a}_i n_r V(C_r^R) + (1-\mathbf{g})\mathbf{a}_i n_b V(C_b^R) + (1-\mathbf{g})(1-\mathbf{a}_i)n_2 V(C_2^R) + \mathbf{g}V(C_j^R) - \\ [(1-\mathbf{g})\mathbf{a}_i n_r V(C_r^L) + (1-\mathbf{g})\mathbf{a}_i n_b V(C_b^L) + (1-\mathbf{g})(1-\mathbf{a}_i)n_2 V(C_2^L) + \mathbf{g}V(C_j^L)] \quad j=r,b \quad (11a)$$

P_{ij} expresses the advantage to individual i in group j of party R winning the election. An individual i in group j will vote for party R if $P_{ij} > 0$.

Defining $\tilde{P} = [n_r(V_r^R - V_r^L) + n_b(V_b^R - V_b^L) + n_2(V_2^L - V_2^R)]$, we can write

$$P_{ij} = \mathbf{a}_i(1-\mathbf{g})\tilde{P} + \mathbf{g}(V_j^R - V_j^L) + (1-\mathbf{g})n_2(V_2^R - V_2^L), \quad j=r,b \quad (11b)$$

$$\frac{\partial P_{ij}}{\partial \mathbf{a}_i} = (1-\mathbf{g})\tilde{P} \quad (12)$$

We note that \tilde{P} and $\partial P_{ij} / \partial \mathbf{a}_i$ do not depend on j . At the non-discriminating equilibrium $(V_1^R - V_1^L) < 0$ and $(V_2^L - V_2^R) < 0$. Hence, at this equilibrium \tilde{P} and $\partial P_{ij} / \partial \mathbf{a}_i$ are negative.

$P_{ij} = 0$ defines a critical value, $\bar{\mathbf{a}}$, such that individuals with $\mathbf{a}_i < \bar{\mathbf{a}}$ will vote for party R . Let us define $\mathbf{q} = \mathbf{g}/(1-\mathbf{g})$. The value $\bar{\mathbf{a}}_j$ can then be written as:

$$\bar{\mathbf{a}}_j = \frac{n_2 V_2^R - n_2 V_2^L + \mathbf{q}(V_j^R - V_j^L)}{-\tilde{P}} = \frac{\mathbf{k} + \mathbf{q}V_j^R}{-\tilde{P}} \quad j=r,b \quad (13)$$

where $\mathbf{k} = n_2 V_2^R - n_2 V_2^L - \mathbf{q}V_j^L$. We note that it is only the term $\mathbf{q}V_j^R$ that differs between redheads and brown-heads. We also note that if there is no discrimination the redheads and the brown-heads will have the same P_{ij} curve and the same value of $\bar{\mathbf{a}}_j = \bar{\mathbf{a}}$.

As before m_1 denotes the number of voters from class one preferring the policy of the R -party. In the initial situation $m_1^u = (n_r + n_b)\bar{\mathbf{a}}$. (We let u denote uniform treatment of the r - and b -groups.) As long as $0 < \bar{\mathbf{a}}_j < 1$ for both $j=r$ and $j=b$ and \mathbf{a} is uniformly distributed the change in m_1 is given by $\Delta m_1 = n_r \Delta \bar{\mathbf{a}}_r + n_b \Delta \bar{\mathbf{a}}_b$. To calculate the changes in $\bar{\mathbf{a}}_j$ we can use the relationship $\Delta C_b = -\frac{n_r}{n_b} \Delta C_r$, which follows from the assumption of a fixed income

available for class 1. Let m_1^d denote the number of class one voters who prefer the R -party when there is discrimination. Noting that \mathbf{k} is unaffected by discrimination and using (13) we can write the change in m_1 as party R starts to discriminate as:

$$\begin{aligned} \Delta m_1 &= m_1^d - m_1^u = n_r \left(\frac{\mathbf{k} + \mathbf{q}V_r^{Rd}}{-\tilde{P}^d} \right) + n_b \left(\frac{\mathbf{k} + \mathbf{q}V_b^{Rd}}{-\tilde{P}^d} \right) - (n_r + n_b) \left(\frac{\mathbf{k} + \mathbf{q}V_1^{Ru}}{-\tilde{P}^u} \right) \\ &< n_r \left(\frac{\mathbf{k} + \mathbf{q}V_r^{Rd}}{-\tilde{P}^u} \right) + n_b \left(\frac{\mathbf{k} + \mathbf{q}V_b^{Rd}}{-\tilde{P}^u} \right) - (n_r + n_b) \left(\frac{\mathbf{k} + \mathbf{q}V_1^{Ru}}{-\tilde{P}^u} \right) = \\ &- \mathbf{q} [n_r V_r^{Rd} + n_b V_b^{Rd} - (n_r + n_b) V_1^{Ru}] / \tilde{P}^u < 0 \end{aligned} \quad (14)$$

The first inequality follows from the fact that, due to the concavity of $V(\cdot)$, $-\tilde{P}^d > -\tilde{P}^u$. The second inequality follows from the fact that, due to the concavity of $V(\cdot)$, $(n_r + n_b) V_1^{Ru} > n_r V_r^{Rd} + n_b V_b^{Rd}$.

To get some intuition for the result it is helpful to take as our point of departure the insight that a marginal discrimination makes no difference. It has an impact neither on the number of supporters nor the expected welfare of the R party. The number of brown-head supporters foregone equals the number of redhead supporters gained. As further discrimination is pursued it is redheads attaching an ever lower value to their welfare concerns who are the target of the R party's attempt to win supporters. However, those assigning a higher weight to social welfare and hence a lower *relative* weight to their selfish concern are more expensive to "bribe" into becoming R supporters than the previous converts. As they give relatively less weight to their own income they need to be given more additional income to be induced to switch their support. But since no gain was achieved initially an attempt to "buy" voters demanding an even higher "bribe" cannot be worthwhile. On top of that the party of course incurs a welfare loss from violating its principle of horizontal equity.

Non-uniform distribution of \mathbf{a}

Above we made the simplifying assumption that \mathbf{a} is uniformly distributed. If \mathbf{a} is drawn from a general distribution the analysis will be more complicated. To study the effect of the distribution of \mathbf{a} we will make the simplifying (but false) assumption that an increase in C_r^R will increase $\bar{\mathbf{a}}_r$ by the same amount as $\bar{\mathbf{a}}_b$ will decrease. How the number of voters change will then depend entirely on the distribution of \mathbf{a} . Since we have assumed that red and brown heads are identical except for the hair color, the distribution of \mathbf{a} is the same in the two groups. Let $F(\mathbf{a})$ denote the cumulative distribution of \mathbf{a} and $f(\mathbf{a})$ the pdf. At the non-

discriminating equilibrium $F(\bar{a}_r) = F(\bar{a}_b) = F(\bar{a})$. A marginal discrimination will then increase the number of votes for the R party if $F''(\bar{a}) > 0$, i.e. if $f'(\bar{a}) > 0$. For symmetric bell shaped distributions this holds true for \bar{a} such that $F(\bar{a}) < 0.5$. This implies that it is more likely that discrimination pays off for the R party if it originally holds a small proportion of the class 1 votes.¹² We also want to point out an important asymmetry. If we consider non-marginal discrimination the number of brown-heads voting for the R party will eventually be zero. Further discrimination will then not lead to any decrease in the number of votes from brown heads while the number of votes from red heads can continue to increase. We study this case below.

Extreme discrimination

We have seen that non-extreme discrimination is not worthwhile for the R-party if the distribution of \mathbf{a} is uniform. Could it be worthwhile with extreme discrimination?

We can think of two types of extreme discrimination. One would be where all redheads has been won and there are still some brown-heads who support the R-party. Obviously, it is not beneficial in such a situation to pursue further discrimination of the brown-heads.¹³ A second type of extreme discrimination is where the R-party loses all the brown-heads. Assuming a point is reached at which the R-party loses its last brown-head supporter without appropriating all the resources of brown-heads and without winning all redheads, further transfers from brown-heads to redheads will unambiguously win more supporters for the R-party. However, in reaching that point the R-party has forfeited votes as well as welfare. Further discrimination will add to the loss of welfare. Clearly, a condition for extreme discrimination to be worthwhile in terms of votes is that the loss of brown-head votes $n_b \bar{a}$ is less than the potential number of redheads to win, $n_r(1 - \bar{a})$. This implies that the non-discriminating equilibrium must be such that $\bar{a} < n_r / (n_b + n_r)$. Provided this necessary condition is fulfilled the question is whether the increase in probability of winning the election

¹² This result was obtained under the assumption that \bar{a}_b decreases by the same amount as \bar{a}_r increases. It is in fact true that, if the groups are of equal size, \bar{a}_b will decrease more than \bar{a}_r will increase. This will not affect the conclusion that it is more likely that discrimination pays off for the R-party if it originally holds a small proportion of the class 1 voters. However, it has a bearing on whether discrimination is worthwhile or not.

¹³ Of course, it follows from proposition 5 that if \mathbf{a} follows a uniform distribution it is not advantageous for the R-party to perform discrimination such that it obtains all redhead votes while still having some brownhead supporters. According to proposition 5 such discrimination leads to more lost brownhead supporters than redheads gained.

outweighs the price one has to pay in terms of welfare foregone. In principle either outcome is possible.

7. Summary and conclusions

There exist several earlier models studying the incentives for political parties to purchase votes. A common property of these models is that there are identifiable groups that can be treated differently. Often the groups differ in terms of income and the distribution of political preferences. An important result of influential studies by Dixit and Londregan is that groups with many "swing-voters" will get a preferential treatment. Our model is in some respects similar to the Dixit and Londregan model. However, there are also important differences. One important difference between our study and earlier studies is that in our model groups do not differ, except possibly with respect to size. Another important difference is that we let a selfishness parameter vary continuously across voters. Earlier studies has usually treated selfishness as a fixed parameter and allowed a parameter showing the relative weight on low versus high income individuals to vary.

First, we find that a political party can have strong incentives to discriminate against one group and favor another even if the groups are identical in terms of the distribution of preferences, income and other welfare relevant characteristics. This is in stark contrast to results of earlier studies where identical groups get identical treatment. Second, if groups differ in size, this is of importance for which group is favored and which one that is discriminated. This is also in contrast to earlier studies where group size, somewhat surprisingly, does not matter for how a group is treated. Third, we find that it is of large importance whether the parameter showing selfishness is fixed or follows a continuous distribution. If the selfishness parameter follows a uniform distribution, while a parameter indicating the relative welfare weight on low income individuals versus high income individuals is fixed, the incentives for discrimination is strong. If on the other hand the welfare weight follows a uniform distribution and the selfishness parameter is fixed, then there are no incentives for non-extreme discrimination. Fourth, the concavity of individuals' utility functions are also of large importance. For sufficiently strong concavity, individuals aversion against discrimination will be so strong that the party that starts to discriminate will lose votes. Fifth, the form of the non-discriminating equilibrium is of importance whether discrimination is worthwhile or not.

References

- Alesina, A. (1988). "Credibility and Policy Convergence in a Two-Party System with Rational Voters." *American Economic Review* 28, 796-806.
- Besley, T. and S. Coate. (1997). "An Economic Model of Representative Democracy." *Quarterly Journal of Economics* 112, 85-114.
- Blomquist, S. and V. Christiansen. (1999). "The Political Economy of Publicly Provided Private Goods." *Journal of Public Economics* 73, 31-54.
- Dixit, A. and J. Londregan (1995). "Redistributive politics and economic efficiency." *American Economic Science Review* 89(?) 856-866.
- Dixit, A. and J. Londregan. (1998a). "Ideology, Tactics, and Efficiency in Redistributive Politics." *Quarterly Journal of Economics* 113, 497-529.
- Dixit, A. and J. Londregan. (1998b). "Fiscal Federalism and Redistributive Politics." *Journal of Public Economics* 68, 153-180.
- Fiorini, M.P. (1981). Retrospective Voting in American Elections (Yale University Press).
- Hansson, I. and C. Stuart. (1984). "Voting Competitions with Interested Politicians: Platforms Do not Converge to the Preferences of the Median Voter." *Public Choice* 44, 431-441.
- Johansson, E. (1999), "Tactical Redistribution Between Regions when Parties and Voters Care about Ideology." Chapter 4 in Essays on Local Public Finance and Intergovernmental Grants. Ph.D. dissertation, Department of Economics, Uppsala University.
- Lindbeck, A. and J. W. Weibull. (1987). "Balanced-budget Redistribution as the Outcome of Political Competition." *Public Choice* 52, 273-297.
- Lindbeck, A. and J. W. Weibull. (1993). "A Model of Political Equilibrium in a Representative Democracy." *Journal of Public Economics* 51, 195-209.

Appendix A. Introducing a random element into the voting decision.

We know that in most elections not everyone will vote and that an important aspect of winning an election is to motivate the individuals supporting the party to actually vote. We do not intend to go into this in any depth, but use a quite stylized model to capture the fact that many people do not vote. Without lack of generality we assume that individuals supporting party R will always vote. However, only a fraction of the individuals supporting party L will actually vote¹⁴. This fraction is a random variable q . As a simplification q is assumed to follow a

¹⁴ The inclination to go and vote may depend on weather conditions, whether there is an influenza epidemic, the encouragement offered by the mass media, etc. One may argue that it is a bit special to link voting propensity to party adherence as such. The important assumption is that the probability of voting differs

uniform distribution with support $0 < a < q < b < 1$. Let $n = n_1 + n_2$ and $m = m/(n - m)$. If $m \leq a$ the probability that party R will win is 0. If $m \geq b$ the probability that party R will win is 1. For $a < m < b$ the probability that party R will win is given by:

$$Pr[\text{party } R \text{ wins}] = Pr[m(z^L, z^R) > q(n - m(z^L, z^R))] = Pr[q < (m(z^L, z^R)/(n - m(z^L, z^R)))] = \int_a^m \frac{1}{b-a} dq = \frac{m-a}{b-a}.$$

Rewriting this probability we obtain: $Pr(\text{party } R \text{ wins}) =$

$$p(z^L, z^R) = \frac{m - a(n - m)}{(n - m)(b - a)}.$$

The important feature of this expression is that $\partial p / \partial m > 0$.

Appendix B. Proof of Proposition 1

If discrimination is not allowed, then there exists a Nash equilibrium such that

$$\bar{C}_2^L < \hat{C}_2^L < \hat{C}_2^R < \bar{C}_2^R.$$

Proof:

a) Existence

In order for a Nash equilibrium to exist there must be at least one point of intersection between the response curves of the two parties. Let us as above consider the choice of policy as a choice of C_2 . Let us assume that party L gives all the income to class one and nothing to class two. Then the best response of party R is to give at least some income to class two. First this is a better policy from the perspective of party R and the party can have nothing to lose by proposing a policy it considers as superior. It is even likely to gain as the party will then win the support of the entire class two and may hardly lose support in class one at least if the amount of income given to class two is sufficiently small. The reason is that class one are not entirely selfish, but also care about the welfare of class one. Let us then consider the opposite extreme that class L proposes giving all the income to class two. As the R party is assumed to be a benevolent party that also cares about the welfare of class one it will propose allocating at least some income to class one. If that policy wins the R party will feel better off. If that policy loses, the R party is no worse off than it would have been if it had adopted the winning L

between groups otherwise the stochastic element would just be similar to having an electorate of random size but with deterministic shares of party support. We have opted for the simplest way to model a systematic difference between classes. Our approach may even have a claim for realism as it is known from several countries that conditions inducing a high turn-out tend to favour certain parties.

policy. In an analogous way we can show that if the R Party proposes giving nothing to one of the classes the L party will respond by proposing to give at least some income to that class.

It is easy to see that these properties are sufficient to guarantee that the response curves, being continuous, will in fact intersect. In detail one can show this by depicting the response curves in box measuring C_2^L along the horizontal side and C_2^R along the vertical side and with the length of each side reflecting the maximum value of C_2 .

$$\text{b) } \bar{C}_2^L < \hat{C}_2^L < \hat{C}_2^R < \bar{C}_2^R$$

Let us first establish that

$$\hat{C}_2^L < \bar{C}_2^R, \bar{C}_2^L < \hat{C}_2^R$$

L will never allocate more income to class two than what is implied by R's dictator policy. If it did the party could rather choose R's dictator policy and make sure that a superior policy according to its own preferences would win. Moreover, the L party would never adopt R's dictator policy because by rather adopting a policy it would find at least slightly preferable, there would be at least some probability that a better policy from L's perspective would win. (The precondition is that $\mathbf{a}_L > \mathbf{a}_R$). By analogy the R party will always propose a lower income for class one (higher income for class two) than that implied by the dictator policy of party L.

We then establish that

$$\bar{C}_2^L < \hat{C}_2^L, \hat{C}_2^R < \bar{C}_2^R$$

The voters are assumed to have preferences that are quasi-concave in C_1, C_2 -space. The value of a voter's preference function is monotonically decreasing along the budget line in C_1, C_2 -space as one moves further away from the preferred point. Let \hat{z}^R denote the policy chosen by the R party. If the L party were to choose a policy z^L to the right of $\bar{z}^L = \{\bar{C}_1^L, \bar{C}_2^L\}$ along the budget line (with C_1 along the horizontal axis) new supporters would have to be attracted. There could be no other reason for deviating from the dictator policy than the prospect of winning votes. We use the notation \succ for "preferred to". Then there would have to be voters who simultaneously have preferences $(*) z^L \succ \hat{z}^R \succ \bar{z}^L$. First, unless z^L is preferred to \hat{z}^R , the policy would not win support for the L party. Second, unless the voters prefer \hat{z}^R to \bar{z}^L the supporters would not be *new* supporters attracted by deviating from the dictator policy.

Now consider a policy z^L to the right of $\bar{z}^L = \{\bar{C}_1^L, \bar{C}_2^L\}$. We can infer the following conclusions. If the preferred point of the voter is to the left of \hat{z}^R it follows that $\hat{z}^R \succ \bar{z}^L \succ z^L$ which is incompatible with the required preferences (*). If the preferred point of the voter is between \hat{z}^R and \bar{z}^L , $\bar{z}^L \succ z^L$ which is also inconsistent with the required preferences (*). If the preferred point of the voter is between \bar{z}^L and z^L , $\bar{z}^L \succ \hat{z}^R$ which violates the condition (*). Finally, if the preferred point is to the right of z^L , $z^L \succ \bar{z}^L \succ \hat{z}^R$ which is also at odds with (*). Hence we can rule out that a policy z^L to the right of $\bar{z}^L = \{\bar{C}_1^L, \bar{C}_2^L\}$ will be chosen. There are no potential supporters for the L party to be gained by deviating to the right of the dictator policy (in C_1, C_2 -space). Similarly there is nothing to gain for the R party by deviating to the left from its dictator policy.

Finally we establish

$$\hat{C}_2^L < \hat{C}_2^R$$

At a political equilibrium the L party will not allocate more income to class 2 than does the R party.

Assume that L wants to allocate a larger income to class 2 than does the R party. Given the budget constraint $n_1 C_1 + n_2 C_2 = \text{constant}$, and the properties of the utility function, the welfare function is decreasing as C_1 (and C_2) moves further away from the level implied by the dictator policy. Since the policies of both parties will lie between the dictator policies, the L party will then prefer R's policy and *vice versa* according to the preference function (1). Each party would then rather adopt the policy of the other party and make sure that an ideologically preferred policy would win. The assumed situation was no equilibrium.

We may note that if everybody in class one strictly prefers the L party the L party must have proposed its dictator policy. If not, it could have moved closer to its dictator policy without losing support.