Redistribution in a Divided Society[1]

David Austen-Smith
Departments of Political Science and Economics
Northwestern University
Evanston, IL 60208

Michael Wallerstein
Department of Political Science
Northwestern University
Evanston, IL 60208

January 2003

[1]We are grateful to the NSF (SES-0212049) and the MacArthur Foundation for support. All responsibility for the content of the paper resides with the authors.

Abstract

The paper develops an integrated political economy model in which individuals are distinguished by earning ability and an ascriptive characteristic such as race, ethnicity or religion. The policy space is a transfer payment to low-income workers financed by a flat tax on wages and an affirmative action constraint on firms' hiring decisions. The distribution of income and the policy are endogenous, with the latter being the outcome of a legislative bargaining game between three parties. The model provides support for the common claim that racial divisions reduce support for welfare expenditures, even when voters have color-blind preferences. We show that relatively advantaged members of both the majority and minority group benefit from the introduction of a second dimension of redistribution, while the less advantaged members of the majority are the principal losers.

1 Introduction

Many scholars have observed that the politics of redistribution in the US is intertwined with the politics of race. Lipset and Bendix, writing in the 1950s, argued that the "social and economic cleavage" created by discrimination against blacks and Hispanics "diminishes the chances for the development of solidarity along class lines" (1959: 106). Myrdal (1960), Quadagno (1994) and, most recently, Gilens (1999) claim that racial animosity in the US is the single most important reason for the limited growth of welfare expenditures in the US relative to the nations of Western Europe. According to Quadagno (1994), political support for Johnson's War on Poverty was undermined by the racial conflicts that erupted over job training and housing programs. Alesina, Baqir and Easterly (1999) find that localities in the US with high levels of racial fragmentation redistribute less and provide fewer public goods than localities that are racially homogeneous.

The US electorate is not unique in containing important social cleavages that cannot be reduced to class or income differences. Politically-salient social divisions that are less than perfectly correlated with income occur in many countries. Belgium and Canada have come close to breaking up along linguistic lines. In Northern Ireland and in Israel, a person's religious identity is of greater political importance than a person's income. While the left-right division remains politically important in most democracies, in no country is the division along class or income lines the only politically salient cleavage. Alesina, Glaeser and Sacerdote (2001) present cross-national evidence that social spending as a share of GDP declines as ethno-linguistic or racial "fractionalization" increases. Bartolini (2000), in an historical study of 13 European countries, concludes that religious cleavages reduced the votes received by socialist or social democratic parties in the early part of the century, while linguistic cleavages have reduced socialist or social democratic votes in recent decades.

The dominant approach in studies of race and redistributive politics in the US is to focus on the manner in which race affects voters' preferences regarding redistributive policies.

Kinder and Sanders (1996) and Alesina and La Ferrara (2000) find that the sharpest contrast in preferences for redistributive policies in the US today is not between rich and poor or between men and women, but between whites and blacks. Moreover, the racial gap in public opinion towards redistributive policies is not eliminated when personal income or personal experience with unemployment are included as control variables (Kinder and Sanders 1996). Gilens (1999) and Luttmer (2001) find evidence that American voters are more willing to support redistributive policies if the perceived beneficiaries are of the same race.

In this paper, we take the complementary approach of studying the effects of social cleavages on the politics of redistribution that are due to the introduction of additional dimensions of potential redistribution. According to the standard model of redistributive politics, as developed formally by Romer (1975), Roberts (1977) and Meltzer and Richard (1981), redistributive policies reflect political conflict between rich and poor, or between voters with above average incomes and voters with below average incomes. In fact, there are a multiplicity of divisions other than the division between rich and poor that can fuel redistributive demands, of which ethnic or racial divisions are among the most salient. Yet, for the most part, racial or ethnic divisions have not been incorporated into the formal theory of redistributive politics. Gelbach and Pritchett (1997) examine the political consequences of using of ethnic identity to target redistributive benefits in an environment where targeting benefits by income is assumed to be infeasible. In addition, there are a variety of studies of redistribution among electoral districts or among industries (Grossman and Helpman 1999, Dixit and Londregan 1998). The dynamics of political conflict among many narrow interests, however, is likely to be quite different from political conflict among a few, large social groups defined by non-economic criteria.

The work that is most similar to this paper, as far as we know, is Roemer's (1998) study of the political demand for redistributive taxation when voters' preferences differ along two dimensions. In Roemer's study, support for redistribution is influenced by political disagreement in a second, non-economic dimension. In contrast, we analyze redistributive policy in a democracy where voters are identical in their (self-interested) preferences, but differ by both income and by some ascriptive characteristic such as race. To be clear about our research strategy, we do not deny the importance of political disagreements over non-economic issues. Our purpose in assuming that voters vote in a color-blind fashion to maximize their post-tax and transfer income is to highlight the "pure" effect of introducing the possibility of redistribution by race as well as by income on the type and extent of redistribution that occurs in equilibrium. In particular, we study the case in which the government may redistribute according to income with taxes and transfers or redistribute according to ascriptive characteristic by affirmative action policies or both.

The importance of developing a model such as the one we outline below is to gain the ability to address theoretically a variety of questions concerning redistributive politics in a divided society that cannot be addressed with existing models. Does the existence of a racial or ethnic cleavage reduce support for redistribution according to income? Who benefits and who loses when poorer minority groups achieve independent political representation? How do the policies selected in equilibrium change as the relative size of the two groups changes or as the distribution of earnings within the minority group changes? We return to these questions after describing our model of the economy and our characterization of the political equilibrium.

2 A labor market with redistribution

To study political conflict over affirmative action policies and taxes and transfers, we need to develop a model of the labor market with heterogenous workers and heterogeneous jobs. In the standard competitive model of the labor market, each worker receives a wage that is just equal to the worker's best alternative and there is nothing to be gained from affirmative action. For affirmative action to make sense, there must be jobs that at least some individuals consider more desirable than others. We therefore need to start with a model of the labor market in which workers with good jobs have something to lose while workers in bad jobs have something to gain. In this and the subsequent section respectively, we describe our model of equilibrium in the labor market and voters' induced preferences over redistributive policies.

2.1 Demographics, human capital and jobs

Assume that society is divided into two groups on the basis of some ascriptive characteristic such as race, language or religion. The split could be between Whites and Blacks, Protestants and Catholics, or French-speakers and English-speakers; any case, in short, where some ascriptive characteristic is correlated with economic opportunity such that the minority group is disadvantaged. Let $p < 1/2$ denote the share of the population who belong to the minority group. With the US in mind, we refer to the majority group as white and the minority group as black, but readers should remember that the terms "white" and "black" can refer to any salient, ascriptive social division.

Assume that workers have one of two levels of human capital, $H = \{0, h\}$ where $h > 0$. The two levels might be interpreted as having a college degree or only a high school degree, for example. Thus, h refers to the additional human capital that educated workers have above the basic education that all workers share. Let the subscript W (for White) denote the majority group and the subscript B (for Black) denote the minority group. Let θ_i be the share of group $i = W, B$ with the high level of human capital and let θ be the share of the population with the high level of human capital, or

$$\theta = p\theta_B + (1 - p)\theta_W$$

Finally, we assume that blacks are disadvantaged in the labor market because they have less human capital on average than whites, or that $\theta_B < \theta_W$. An alternative assumption would be that blacks are disadvantaged because they face discrimination in the job market. In recent studies of the racial gap in earnings in the US, the weight of the evidence indicates that the unequal pay received by white and black workers today reflects unequal education and training received prior to entering the labor market rather than unequal pay for equal work (Altonji and Blank 1999). Therefore, we assume the racial gap in average earnings reflects a racial difference in average human capital.

There are two types of jobs in the economy, $j = \{good, bad\}$. All workers are equally productive in bad jobs, regardless of their level of human capital. Workers' productivity in good jobs, however, is assumed to depend on the worker's level of human capital and on a random-variable that reflects the productivity of the match between the particular worker and the particular job. Let $y(H, j)$ be the marginal product of a worker with human capital H in job j. Assume

$$y(H,j) = \begin{cases} 0 & \text{if } j = bad \\ H + x & \text{if } j = good \end{cases}$$

where x is a match-specific component of productivity with a CDF of $F(x)$ and PDF of $F'(x) \equiv f(x)$. Several of the results to follow depend in part on value of the density, f, at various points. The formal analysis is made easier and the substantive qualitative properties of the model are made more transparent by assuming that the distribution F is approximately uniform over a suitable interval, although such an assumption is not necessary for the results to hold. For the sake of transparency, therefore, we assume F is approximately uniform.

On average, workers with $H = h$ will be more productive in good jobs than workers with $H = 0$, but the most productive workers with $H = 0$ may be more productive in good jobs than the least productive workers with $H = h$. Note that having a bad job may be interpreted as being unemployed, although such an interpretation is not necessary. Similarly, $y(H, j)$ may measure the difference between productivity in job j filled by a person with human capital H and some benchmark level of productivity that all workers can achieve. Firms are assumed able to create good jobs at a cost of $q > 0$ while bad jobs are created at zero cost.

Finally, assume two types of policy: social insurance and affirmative action. The social insurance policy is assumed to provide workers in bad jobs with a benefit of b, financed by a flat tax on wages, t. The affirmative action policy sets a lower bound, α, on the share of good jobs filled by minority workers.

The decision sequence in the polity is as follows :

1. The legislature chooses α and b simultaneously.

2. Workers and employers are randomly matched and the match-specific component of productivity, x, is revealed to both.

3. Employers decide whether to create a good job at a cost of q or a bad job at zero cost.

4. Workers and their employers bargain (individually) as necessary over the wage.

2.2 Wage bargaining and job creation

As usual, we work backwards. The labor market for unskilled work is presumed competitive in that workers receive a wage equal to their marginal product of their labor and firms receive zero profits from allocating a bad job. Thus, workers in bad jobs receive a wage (or wage premium) of zero and a social insurance benefit of $b \geq 0$. Workers in good jobs receive a wage (or wage premium) of $w(y)$ which depends on the productivity of the match. For expository convenience, we assume that both wages and welfare benefits are taxed at a flat tax rate of t; adopting the alternative assumption that welfare benefits are not taxed makes no difference in the analysis. Thus, workers' consumption (assumed to define utility here) is given by

$$c_j(y) = \begin{cases} (1-t)b & \text{if } j = bad \\ (1-t)w(y) & \text{if } j = good \end{cases} .$$

Assume the consequence of failing to agree in wage negotiations is that the worker obtains a bad job and the good job remains vacant. Thus, the worker's gain from an agreement is $(1-t)\left[w(y) - b\right]$.

The firm's profit is given by

$$\pi_j(y) = \begin{cases} 0 & \text{if } j = bad \\ -q & \text{if } j = good \text{ and the job remains vacant} \\ y - w(y) - q & \text{if } j = good \text{ and the worker is hired} \end{cases}$$

Note that the cost q of creating a good job is a sunk cost when wage bargaining occurs. This assumption is a simple way of representing the common situation where wages are set for shorter periods of time than the life-span of the plant and equipment that must be purchased by the firm to usefully employ a worker in a good job. Further, while profits are assumed not to be taxed, profits are affected by taxes and transfers since taxes and transfers affect the wage that firms must pay.[1] Once a good job is created, the firm's gain from an agreement is $y - w(y)$.

Using the generalized Nash bargaining solution to represent the outcome of wage bargaining, we have $w(y) = \arg\max(y - w)^{1-\beta}(w - b)^{\beta}$ or

$$w(y) = \beta y + (1 - \beta)b \tag{1}$$

as the wage offered in good jobs and

$$\pi(y) = (1 - \beta)(y - b) - q \tag{2}$$

as the profit earned from the creation of a good job, where $\beta \in (0, 1)$ represents the worker's share of the joint gains.

In the absence of affirmative action policies, firms create good jobs only if it is profitable to do so; that is, if and only if $\pi(y) \geq 0$ or, equivalently, if and only if

$$y \geq y_0 \equiv \frac{q}{1 - \beta} + b. \tag{3}$$

[1]Taxing profits at the same rate as wages has virtually no consequences for the qualitative results to follow. This is not true under the plausible assumption that profits are taxed at a different rate; but this greatly complicates the political decision-making problem. So, for now at least, we simply assume these complications away by having only wages be taxed.

The fraction of group $i = W, B$ with good jobs, denoted σ_i, is given by

$$\sigma_i = [1 - G_i(y_0)]$$

where $G_i(y) \equiv \theta_i F(y-h) + (1-\theta_i)F(y)$ is the fraction of group i with productivity less than y. Without affirmative action, the share of good jobs held by minority workers, denoted α_0, is given by

$$\alpha_0 = \frac{p\sigma_B}{p\sigma_B + (1-p)\sigma_W}.$$

By assumption, $\theta_B < \theta_W$ so $G_B(y_0) > G_W(y_0)$ and, consequently, $\sigma_B < \sigma_W$. Therefore, lower average human capital for minority workers implies the share of good jobs held by minority workers is less than the share of minority workers in the work force, or that $\alpha_0 < p$.

Because the cost of creating a good job is a sunk cost to the employer, there are match-specific rents over which employers and the prospective employees bargain. And since workers capture a share of the rents, the number of good jobs created is less than the number that maximizes the joint income of employers and workers. The additional output obtained from a marginal increase in the number of good jobs exceeds the cost of job creation; that is, $\beta > 0$ means $y_0 > q$. Moreover, workers with $y = y_0$ obtain a discrete jump in pay of $w(y_0) - b = [\beta/(1-\beta)]\, q$ in moving from a bad job to a good job, even though employers are indifferent between creating a good job or not when $y = y_0$. Thus, workers with productivity slightly below y_0 would gain from a policy that forces firms to create good jobs for them.

2.3 Affirmative action and insurance

Affirmative action entails setting a lower bound, $\alpha \in [\alpha_0, p]$, on the proportion of good jobs filled by minority workers. If the policy of affirmative action is binding, that is if $\alpha > \alpha_0$ or α is greater than the fraction of good jobs that would be filled by minority workers in the absence of affirmative action, then firms cannot use the same productivity threshold for both black and white workers when deciding whether or not to create a good job. If $\alpha = p$, the affirmative action policy requires firms to equalize the fraction of workers with good jobs in the two social groups.

Suppose each employer is large, in the sense of being matched with many workers. Suppose, in addition, that each employer is matched with a random sample of workers from both social groups. Employers' optimal strategy with affirmative action is to create a good job for all black workers with $y \geq y_B$ and for all white workers with $y \geq y_W$ where y_B and y_W solve the problem:

$$\max_{y_B, y_W} \Pi = p \int_{y_B}^{\infty} \pi(y)\, dG_B(y) + (1-p) \int_{y_W}^{\infty} \pi(y)\, dG_W(y)$$

subject to the constraint that

$$\frac{p\sigma_B}{p\sigma_B + (1-p)\sigma_W} \geq \alpha \tag{4}$$

When the constraint is binding, the first-order condition for a maximum can be written as two equations. The first equation replaces (3):

$$\alpha y_B + (1-\alpha) y_W = \frac{q}{1-\beta} + b \tag{5}$$

The second equation is simply the constraint, equation (4), written as an equality. When the constraint is not binding, $y_B = y_W = y_0$ and equation (5) reduces to equation (3).

The second policy instrument we consider is a social insurance policy that provides a uniform transfer payment, b, to all workers with bad jobs. Assume the social insurance policy is financed by a flat tax on wages and welfare benefits, t with $t \in [0, 1]$. The balanced budget constraint that tax revenues must equal welfare expenditures can be written

$$[1 - p\sigma_B - (1-p)\sigma_W]\,(1-t)b = tE(w) \tag{6}$$

where $1 - p\sigma_B - (1-p)\sigma_W$ is the share of the population receiving the welfare benefit and $E(w)$ is the average wage.

2.4 Labor market equilibrium

Conditional on policy (α, b) and on wages in good jobs being defined by (1), an equilibrium in the labor market is a triple (y_B, y_W, t) that solves the system of three equations, (4) (written as an equality), (5), and (6).[2] It is routine to check that, under the maintained assumptions, there exists a unique labor market equilibrium associated with every policy. Individuals' induced preferences over policies, therefore, are well defined. Before going on to examine such preferences in any detail, however, it is useful to identify some salient properties of the labor market equilibrium *per se.*[3]

Lemma 1 *(1)* $\partial y_W/\partial\alpha > 0 > \partial y_B/\partial\alpha$*;*

(2) $\partial y_W/\partial b \geq 1 \geq \partial y_B/\partial b > 0$ *with strict inequality for* $\alpha \in (\alpha_0, p)$*;*

(3) $\partial t/\partial b > 0$.

Affirmative action induces firms to raise the threshold for hiring white workers and to lower the threshold for hiring minority workers. On the other hand, both of these thresholds are strictly increasing functions of the benefit. Moreover, the threshold for whites increases at a greater rate than does the threshold for blacks when an affirmative action constraint forces employers to maintain a given proportion of minority employment in good jobs. Finally, the tax rate is a strictly increasing function of the benefit that must be financed.

An increase in the benefit increases redistribution in three ways. *Ex post,* the tax and benefit redistributes from workers in good jobs, who pay taxes but don't receive the benefit, to workers in bad jobs. *Ex ante,* the tax and benefit redistributes from high human capital

[2]As we show below, there is a monotonic relationship between b and t. Therefore, it makes no difference whether voters vote over t (the conventional approach) or over b. In our model, the mathematics is simplified by letting b be the policy instrument.

[3]Unless noted otherwise, technical arguments for the results to follow are confined to an Appendix. And we recall that throughout we maintain the assumption that the CDF F is approximately uniform over a suitable nontrivial interval; again, this is far from being a necessary restriction and is irrelevant for some of the claims.

workers who are likely to obtain good jobs to low human capital workers who are less likely to obtain good jobs. Third, the tax and benefits redistributes income from employers to employees. By raising the outside option of workers in good jobs, the benefit increases the wage that workers in good jobs are able to obtain through bargaining. Profits are reduced and the number of good jobs created declines as employers respond by increasing both y_B and y_W. Benefit increases reduce aggregate output, but benefit increases may raise the aggregate income received by workers provided the benefit is not too large. In contrast to redistribution through taxes and transfers, however, affirmative action policies may enhance aggregate output.

Proposition 2 *Suppose $b > 0$. There exist α_b, α_Y such that $\alpha_0 < \alpha_Y < \alpha_b < p$ and*

(1) $\partial t / \partial \alpha < 0$ for all $\alpha \in [\alpha_0, \alpha_b)$ and $lim_{\alpha \to p} \partial t / \partial \alpha > 0$;

(2) $\partial E\left(w + \pi(w)\right) / \partial \alpha > 0$ for all $\alpha \in [\alpha_0, \alpha_Y)$ and $lim_{\alpha \to p} \partial E\left(w + \pi(w)\right) / \partial \alpha < 0$.

Assuming there is at least some redistribution through taxes and transfers, expected aggregate income is strictly increased and the tax rate needed to finance a given benefit is strictly reduced by a marginal increase in affirmative action away from the laissez faire distribution of good jobs between blacks and whites, α_0. In this case, the increase in the number of good jobs filled by black workers exceeds the decline in the number of good jobs filled by white workers and, since white workers are being replaced by black workers with similar levels of productivity when $\alpha \approx \alpha_0$, aggregate income and aggregate earnings increase. This is an example of the theorem of the second best: in an economy where insiders' bargaining power reduces investment in good jobs below the optimal level, a policy that results in increased good job creation raises aggregate income. Of course, affirmative action policies need not result in job creation if $\alpha > \alpha_0$. Indeed, for α sufficiently large, increases in affirmative action reduce aggregate earnings and raise the tax rate needed to finance a given benefit. Moreover, because any increase in α above α_0 reduces profits, aggregate income must be decreasing in α if aggregate earnings are decreasing in α, although the converse is not true

$(\alpha_b > \alpha_Y)$. At $\alpha \approx \alpha_0$, the gain in aggregate earnings that follows from a small increase in α is greater than the loss in profits.

3 Induced policy preferences

A policy is a pair $(\alpha, b) \in \mathbf{P} \equiv [\alpha_0, p] \times [0, b_{\max}]$ where $b_{\max}$ is the benefit that maximizes $(1-t)b$. The policy is chosen prior to a workers' knowledge of x or of whether he or she will be offered a good job. In evaluating policy, however, individuals are assumed to anticipate correctly the consequences of their choice. Individual preferences over policy are induced by an understanding of the resultant labor market equilibrium: a policy (α, b) is preferred by some individual to an alternative policy (α', b') if and only if the individual's expected equilibrium consumption level under (α, b) is greater than that under (α', b'). Formally, an individual with human capital $H \in \{0, h\}$ in group $i \in \{B, W\}$ evaluates a policy (α, b) by

$$E\left[c_{Hi}(\alpha, b)\right] = (1-t)\left\{F(y_i - H)b + \int_{y_i - H}^{\infty} w(H, x, b)\, dF(x)\right\}, \tag{7}$$

where y_i and t are defined by the labor market equilibrium at (α, b).

Lemma 3 *(1) For all $(\alpha, b) \in \mathbf{P}$, $\partial E\left[c_{0i}(\alpha, b)\right]/\partial b > \partial E\left[c_{hi}(\alpha, b)\right]/\partial b$, $i = B, W$;*

(2) For all $b \in [0, b_{\max}]$, $\partial E\left[c_{HB}(\alpha_0, b)\right]/\partial b > \partial E\left[c_{HW}(\alpha_0, b)\right]/\partial b$, $H = 0, h$.

Consider, first, the effect of human capital on preferences regarding the tax rate. Voters with the high level of human capital pay a higher tax on average, and are less likely to receive the benefit b. By assumption, the distribution of x is approximately uniform and, therefore, high-type individuals are more tax averse than low-type individuals: Lemma 2(1) reflects this fact. It follows that unless both high and low types prefer $b = 0$ to any interior solution, voters with low levels of human capital prefer higher levels of taxation and spending than voters with high levels of human capital.

Race has counteracting effects on the support for welfare expenditures in the presence of affirmative action policies when the level of human capital is held constant. On the one hand, affirmative action policies provide some protection to minority workers with good jobs. The loss of good jobs may be greater for white workers than for black workers when employers reduce the number of good jobs in response to increases in the benefit in the presence of an affirmative action constraint. This effect works to increase black support for higher taxes and benefits. On the other hand, affirmative action raises the expected income of blacks for a given level of human capital, which reduces support for higher taxes and benefits. Under the maintained assumption on the distribution of x, however, there is a clear answer when the affirmative action policy is barely binding. This answer is given by Lemma 2(2): given $f(x)$ is approximately constant and $\alpha \approx \alpha_0$, black voters prefer higher levels of taxes and spending than white voters with equal levels of human capital (unless the level of human capital is sufficiently high that both black and white voters prefer zero taxes). In this respect, our model matches the survey evidence that the effect of race on support for welfare expenditures does not vanish when income is controlled for, even though we have assumed that voters have color-blind preferences (Kinder and Sanders 1996).[4]

The next lemma establishes how expected consumption is affected by changes in affirmative action.

Lemma 4 *There exists α_b such that $\alpha_0 < \alpha_b < p$ (as in Proposition 1) and $\underline{\beta} \in (0,1)$ such that, for all $(\alpha, b) \in \mathbf{P}$:*

(1) For $\beta \geq \underline{\beta}$, $\partial E\left[c_{HB}(\alpha,b)\right]/\partial\alpha > 0 > \partial E\left[c_{HW}(\alpha,b)\right]/\partial\alpha$, $H = 0, h$;

(2) $\partial E\left[c_{hi}(\alpha,b)\right]/\partial\alpha \gtreqless \partial E\left[c_{0i}(\alpha,b)\right]/\partial\alpha$ as $\alpha \lesseqgtr \alpha_b$, $i = B, W$.

Preferences with regard to affirmative action are straightforward. In general, members of the minority group favor increases in affirmative action while members of the majority group

[4] To be clear, we do not deny the possible existence of important differences in preferences of black and white voters in the US. Our point is that the racial differences in preferences over policies observed by survey researchers do not necessarily imply the existence of racial differences in the fundamental preferences from which policy preferences are derived.

are opposed, regardless of the level of human capital (Lemma 3(1)). Furthermore, high-type individuals care weakly more than low types about affirmative action policy in the neighborhood of α_0 (Lemma 3(2)). It is worth noting that if workers' bargaining power, β, is sufficiently high, then the qualitative properties reported in Lemma 3 apply over the entire policy space. The intuition here is that at high levels of β the impact on consumption of changes in the transfer b induced by changes in affirmative action, become dominated by the direct effect on earned income conditional on acquiring a good job. On the other hand, as β becomes negligible so does any advantage from holding a good job and all consumption derives from the transfer, b. In this case, even high-type white workers can prefer some moderate levels of affirmative action.

Lemmas 2 and 3 yield the following intuitive proposition.

Proposition 5 *For each $(H,i) \in \{0,h\}\times\{B,W\}$, let $I_{Hi} = (\alpha^*_{Hi}, b^*_{Hi})$ maximize $E[c_{Hi}(\alpha,b)]$ on $\mathbf{P}$. Assume low-type individuals strictly prefer at least some strictly positive benefit to zero benefit and that $\beta \geq \underline{\beta}$. Then I_{Hi} is unique and*

(1) $\alpha^*_{hB} = \alpha^*_{0B} = p$ and $\alpha^*_{0W} = \alpha^*_{hW} = 0$;

(2) $b^*_{B0} > b^*_{W0}$ for $\alpha \approx \alpha_0$, and $b^*_{i0} > b^*_{ih}$, $i = W, B$.

4 Legislative bargaining

We assume that legislators, when selecting between two alternatives, cast their ballot for the party that promises to implement the policy that generates the higher expected post-tax, post-transfer income for their constituents.

A complete model of the political process would include (at least) two stages. The first stage involves voters' choice of representatives while the second stage consists of representatives' choice of policy. In this paper we focus exclusively on the legislative policy decision stage. With regard to voters' choice, we simply assume the existence of blocs of representatives or parties, each of whom represents a distinct constituency. While a variety of divisions of the legislature might be considered, here we restrict our analysis to the case in which the legislature is divided into three groups: legislators who represent white workers with a high level of human capital ($\mathcal{H}$), legislators who represent white workers with a low level of human capital ($\mathcal{L}$), and legislators who represent black workers ($\mathcal{B}$). For purposes of exposition, we refer to such blocs of legislators as "parties", although we emphasize that critical feature of such parties for the model is as coherent voting blocs. Each party is assumed to act in a unified manner to maximize the welfare of its constituents. Because the policy is chosen before workers know what job they will be offered, all white workers with a given level of human capital are identical *ex ante*. Therefore, the objective functions for $\mathcal{H}$ and $\mathcal{L}$, respectively, are $u_{\mathcal{H}} = E(c_{hW})$ and $u_{\mathcal{L}} = E(c_{0W})$. Minority voters, however, include workers with both high and low levels of human capital. In this case, we assume the legislative group maximizes a weighted average of the consumption of its two types of constituents: $u_{\mathcal{B}} = (1-\lambda)E(c_{0B}) + \lambda E(c_{hB})$ for some $\lambda \in [0,1]$. The weight λ is a measure of the extent to which the minority party is concerned with low or high types of minority workers. Let $I_{\mathcal{H}} = (\alpha^*_{\mathcal{H}}, b^*_{\mathcal{H}}) \equiv I_{hW}$ and $I_{\mathcal{L}} = (\alpha^*_{\mathcal{L}}, b^*_{\mathcal{L}}) \equiv I_{0W}$ be the most preferred policy points for $\mathcal{H}$ and $\mathcal{L}$, respectively, and let $I_{\mathcal{B}} = (\alpha^*_{\mathcal{B}}, b^*_{\mathcal{B}}) \equiv \arg\max u_{\mathcal{B}}$ be the most preferred policy point for party $\mathcal{B}$; by Proposition 2, $I_{\mathcal{B}} \equiv [(1-\lambda)I_{0B} + \lambda I_{hB}]$.

To avoid a trivial solution to policy conflict, we assume that no single group has a majority of seats in the legislature. If the size of each legislative bloc reflects the relative size of each bloc's constituents, this implies that $p < 1/2$, $(1-p)\theta_W < 1/2$ and $(1-p)(1-\theta_W) < 1/2$. Thus any two of the three parties constitutes a majority. It is not hard to check that, as is usually the case with multidimensional policy spaces, the majority core is typically empty. Suppose β is sufficiently high that $\alpha^*_{hB} = \alpha^*_{0B} = p$ and $\alpha^*_{hW} = \alpha^*_{0W} = \alpha_0$. The coalition of $\mathcal{H}$ and $\mathcal{L}$ then strictly prefers the policy $(\alpha - \epsilon, b)$ for some $\epsilon > 0$ to any feasible pair $(\alpha, b) \in \mathbf{P}$ with $\alpha > \alpha_0$ while, for any pair (α_0, b) with $b \in (b^*_{\mathcal{H}}, b^*_{\mathcal{L}})$, there is a policy $(\alpha_0 + \epsilon, b \,\square\, \delta)$ for some $\epsilon \geq 0$ and $\delta > 0$ that is strictly preferred by either the coalition of $\mathcal{B}$ and $\mathcal{L}$ or the coalition of $\mathcal{B}$ and $\mathcal{H}$. The high white's ideal point cannot be in the core, since a majority prefer $(\alpha_0, b^*_{\mathcal{H}} + \delta)$ to $(\alpha_0, b^*_{\mathcal{H}})$. Therefore, either the core consists of the low white's ideal point $I_{\mathcal{L}} = (\alpha^*_{\mathcal{L}}, b^*_{\mathcal{L}})$ (a possible but unlikely case) or the core is empty.

In view of the general nonexistence of a majority core, we model the policy process as a legislative bargaining game. Specifically, we apply the by-now much used majoritarian version of the Rubinstein infinite horizon alternating offers model, introduced by Baron and Ferejohn (1989) and most recently generalized by Banks and Duggan (2000, 2001). In its simplest form, each party or bloc is associated with a probability of being selected to make a policy proposal (α, b) to the legislature in any period. If one (or both) of the non-proposing blocs accepts the proposal in some period then the proposed policy is implemented and bargaining ends; otherwise the process moves to the next decision period, a new proposer is randomly selected and the sequence repeats until some proposal is accepted.

The solution concept is a no-delay, stationary subgame perfect Nash equilibrium (hereafter, simply *equilibrium*).[5] An equilibrium consists of a (possibly degenerate) probability distribution ζ_j over a (possibly infinite) set of policies $P_j \subseteq \mathbf{P}$ that party j proposes whenever j is recognized to make a proposal, and an acceptance set, $A_j \subseteq \mathbf{P}$ that specifies the set of policies for which party j will vote if another party is the proposer. Let v_j be j's expected payoff at the beginning of the game. By stationarity, v_j is also j 's continuation value or its expected payoff after a proposal has been rejected. Finally, let $\rho_j \in (0,1)$ be party j's probability of being recognized to make a proposal. Then an equilibrium consists of a set of policy proposals and acceptance set for each $j = \mathcal{H}, \mathcal{L}, \mathcal{B}$ that satisfy the following conditions:

$$\begin{aligned} P_j &\subseteq \arg\max\{u_j(\alpha, b) \mid (\alpha, b) \in A_j \cup A_l\}, \\ A_j &= \{(\alpha, b) \mid u_j(\alpha, b) \geq v_j\} \end{aligned}$$

[5]Every stationary equilibrium is a no-delay equilibrium if the parties discount future periods at a discount rate of $\delta_i < 1$ (Banks and Duggan 2000, Theorem 1). When $\delta_i = 1$ for all i, the no-delay equilibria continue to exist, but many others exist as well. By considering only no-delay equilibria, we restrict attention to stationary equilibria that represent the limit of a sequence of stationary equilibria as $\delta_i \to 1$.

$$v_j = \sum_{k=\mathcal{H},\mathcal{L},\mathcal{B}} \rho_k \left[\int_{P_k} u_j(\alpha, b)\, d\zeta_k \right]$$

The first condition states that any policy a party proposes necessarily maximizes its constituents' welfare over the set of policies attracting majority support in the legislature, that is policies that lie within the acceptance set of either party k or party l. The second condition states that each party accepts any proposal that provides a higher payoff than the party's continuation value. The third condition states that in equilibrium the continuation value equals the expected value of the game. In a no-delay equilibrium, the first party to be recognized offers the best proposal it can for its voters from among the set of proposals that will be accepted and the game ends.

The most general equilibrium existence result for this game (at least, as far as we know) is due to Banks and Duggan (2000, 2001). However, their theorem assumes preferences are concave on the policy space. Unfortunately, although (as an example below illustrates) agents' induced preferences are indeed strictly concave on $\mathbf{P}$ for most parameterizations of the model, concavity is not a general property of induced preferences here. Moreover, this is not an artifact of the model *per se*. A high benefit reduces the importance of affirmative action, since the difference in consumption between workers in good and bad jobs declines as the benefit (and the tax rate) increase. Conversely, the lower the benefit, the greater the impact of affirmative action policies on workers' expected after-tax and transfer income. Consequently, the marginal rate of substitution between affirmative action and welfare benefit can be increasing and preferences over policies non-concave on the policy space.[6]

The difficulty here is essentially technical. So, rather than attempt to finesse complications with equilibrium existence due to nonconvexities in parties' induced preferences, we simply

[6]For a non-pathological example that generates non-concave preferences, let the benefit b be a universalistic payment to all workers rather than a benefit received only by workers in bad jobs. In this case, there is no deadweight cost of taxation, and the function $b(t)$ is linear. It is easy to check that no group has concave preferences. Moreover, the "problem" of non-concavity cannot be fixed in this example by assuming that workers are sufficiently risk averse.

assume in what follows that an equilibrium exists.[7] To show that the equilibrium concept is not vacuous, we present an example in which all parties' induced preferences are strictly concave on $\mathbf{P}$ and calculate the legislative bargaining equilibrium.

Assume x is uniformly distributed over the interval $[0,1]$, $q = h = 1/4$, $\beta = 1/2$, $p = 1/3$, $\theta_B = 1/5$ and $\theta_W = 1/2$. These parameter values imply that 80 per cent of minority workers and half of the majority workers have a low level of human capital and that the three groups are equal in size. We assume initially that the black party represents the 80 per cent of minority voters with less education, or $\lambda = 0$. Figure 1 illustrates the preferences of the three parties over welfare benefits and affirmative action and the equilibrium of the legislative bargaining game. The western and northern borders of the Pareto set are given by $b = 0$ and $\alpha = p$ respectively. The southern border is given by the function $\alpha_0(b)$ which represents the share of good jobs that the minority would receive without affirmative action. Note that $\alpha_0(b)$ declines as b increases. The ideal points of the three groups, $\mathcal{H}, \mathcal{L}, \mathcal{B}$, are denoted $I_{\mathcal{H}}$, $I_{\mathcal{L}}$ and $I_{\mathcal{B}}$ respectively. The figure illustrates the unique equilibrium for the case in which each group is recognized with equal probability. If recognized, the high-type whites propose $(\alpha_{\mathcal{H}}, b_{\mathcal{H}})$ with probability (.61) and receive the support of the low-type whites. With probability (.39) , the high-type whites propose $(\alpha'_{\mathcal{H}}, b'_{\mathcal{H}})$ and receive the support of the black legislators. The low-type whites, if recognized, propose $(\alpha_{\mathcal{L}}, b_{\mathcal{L}})$ with probability one and win the support of the black legislators. Finally, black legislators propose $(\alpha_{\mathcal{B}}, b_{\mathcal{B}})$ with probability one if recognized and win the support of high-type whites.

Figure 1 here

Table 1 records the numerical values of each group's proposal. The fourth line of Table 1 is the expected tax and affirmative action policy, as well as the expected levels of consumption

[7] One approach to existence in cases where preferences are not concave would be to allow parties to propose *lotteries* over policies rather than policy points. Then the relevant payoffs for the bargaining game are concave on the space of such lotteries and the Banks-Duggan theorem can be applied.

for each of the four groups. For comparative purposes, the fifth line of Table 1 shows the political equilibrium that exists absent the racial divide. When there is no racial divide, political conflict occurs over the single dimension of the tax rate. Since, in this example, $(1-\theta_B)p+(1-\theta_W)(1-p)=3/5$ of the population share the ideal point of poorly educated whites, the ideal point of the poorly educated majority would prevail. The bottom line presents a summary of the gains or losses to each group associated with the introduction of a second dimension of political conflict over affirmative action.

Table 1. The Equilibrium of the Legislative Bargaining Game when $\lambda = 0$

	Prob.	b_j	$t(b_j)$	α_j	$\alpha_0(b_j)$	c_{hW}	c_{0W}	c_{hB}	c_{0B}
$\mathcal{H}$'s proposal	.614	.042	.071	no aff.act.	.302				
	.386	.011	.018	.333					
$\mathcal{L}$'s proposal	1	.092	.158	no aff.act.	.296				
$\mathcal{B}$'s proposal	1	.118	.203	.333					
Expected Value		.080	.137	.314		.307	.194	.314	.201
$\mathcal{L}$'s Ideal Policy		.119	.206	no aff.act.	.291	.299	.199	.299	.199
Expected Gain						2.8%	-2.4%	5.1%	1.0%

Parameter values: $p = 1/3$, $\theta_B = 1/5$, $\theta_W = \beta = 1/2$, $q = h = 1/4$, $F(x) = x$ for $x \in [0, 1]$ and $\lambda = 0$. The term $\alpha_0(b_j)$ represents the share of minority workers with good jobs in the absence of affirmative action. Expected gain indicates the percentage gain or loss when one-dimensional conflict over t is replaced by two-dimensional conflict over b and α.

The example demonstrates that the equilibrium concept is not vacuous. In the case of a uniform distribution, the concavity property sufficient for equilibrium existence is satisfied for parameter values that cover most (but not all) of the feasible parameter space. While it is difficult to provide general conditions for the strict concavity of legislators' induced preferences, the central features of the equilibrium illustrated in Figure 1 and Table 1 are general characteristics of all equilibria of the legislative bargaining model. The impact of the addition of a second dimension of redistributive conflict on the policy outcome is summarized in the next proposition.

Proposition 6 *Assume induced preferences over the policy space are concave, that $\beta \geq \underline{\beta}$ and that the majority core is empty. Then the possibility of redistribution through affirmative action reduces the expected redistribution through the tax and transfer policy.*

Proof Let (α_j, b_j) be the policy proposed by $j = \mathcal{H}, \mathcal{L}, \mathcal{B}$ in equilibrium and $(\alpha^*_{\mathcal{L}}, b^*_{\mathcal{L}})$ be the ideal point of low whites. Suppose the proposition is false. Then the expected equilibrium of the bargaining game must consist of $E\alpha_j \geq \alpha^*_{\mathcal{L}} = \alpha_0$ and $Eb_j \geq b^*_{\mathcal{L}}$ with $(E\alpha_j, Eb_j) \neq (\alpha^*_{\mathcal{L}}, b^*_{\mathcal{L}})$. By virtue of concavity, $c_{hW}(E\alpha_j, Eb_j) \geq Ec_{hW}(\alpha_j, b_j)$. Moreover, $c_{hW}(\alpha^*_{\mathcal{L}}, b^*_{\mathcal{L}}) > c_{hW}(E\alpha_j, Eb_j)$ since c_{hW} is declining in both α and b (for $b \geq b^*_{\mathcal{L}}$). Therefore, $c_{hW}(\alpha^*_{\mathcal{L}}, b^*_{\mathcal{L}}) > Ec_{hW}(\alpha_j, b_j) = v_{\mathcal{H}}$. But high whites can obtain $c_{hW}(\alpha^*_{\mathcal{L}}, b^*_{\mathcal{L}})$ with certainty by proposing $(\alpha^*_{\mathcal{L}}, b^*_{\mathcal{L}})$ if selected to be the proposer (a proposal low whites would certainly accept) and rejecting all proposals such that $c_{hW}(\alpha, b) < c_{hW}(\alpha^*_{\mathcal{L}}, b^*_{\mathcal{L}})$. If high whites can obtain $c_{hW}(\alpha^*_{\mathcal{L}}, b^*_{\mathcal{L}})$ with certainty, the expected payoff to high whites cannot be lower that $c_{hW}(\alpha^*_{\mathcal{L}}, b^*_{\mathcal{L}})$ in equilibrium, or $v_{\mathcal{H}} \geq c_{hW}(\alpha^*_{\mathcal{L}}, b^*_{\mathcal{L}})$, which contradicts the claim that $c_{hW}(\alpha^*_{\mathcal{L}}, b^*_{\mathcal{L}}) > v_{\mathcal{H}}$. $\square$

The addition of redistributive conflict along ethnic or racial lines always reduces the equilibrium amount of redistribution that occurs along income lines. The addition of a racial or ethnic dimension of redistribution reduces the amount of redistribution according to income, even when (i) the racial minority is poorer than the majority on average and (ii) the legislators who represent the minority group advance the interests of the less educated members of the minority group as in our example above. The potential alliance of blacks and high whites to lower the tax rate and raise affirmative action targets to make both better off is sufficient to reduce the average transfer payment and to increase the average affirmative action target.

Proposition 7 *If an equilibrium exists, then redistribution through affirmative action (1) raises the expected income of both high whites and blacks, and (2) lowers the expected income of low whites.*

Proof Suppose part (1) was false. Then either high whites or blacks could do better by proposing the low white's ideal point, which would certainly be accepted. Note that if high blacks are better off with the low white's ideal point, then all blacks are better off with the low white's ideal point. But then an equilibrium could not exist by the same reasoning as in the proof of Proposition 3. Part (2) is an immediate corollary of Proposition 3. Since the equilibrium with affirmative action results, on average, in a tax rate that is less than what low whites prefer and a binding affirmative action target, low whites are worse off. □

As Table 1 illustrates, the largest winners from the presence of redistributive policies along racial lines are workers with high levels of human capital. Highly educated minority workers gain both from affirmative action and the tax reduction, while highly educated majority workers benefit from the tax reduction. Even though black legislators were assumed to be pure representatives of black workers with the low level of education, such workers gain much less, with the gains from affirmative action partly offset by the loss from the lower benefit. The losers are members of the majority who lose from affirmative action and lose again from the reduction in the average amount of redistribution along income lines with the addition of a second dimension.

In addition to existence, Banks and Duggan (2000) prove that if the equilibrium is unique, as in our example, then the equilibrium is a continuous function of the recognition probabilities and parameters of the legislators' utility functions, so justifying comparative static exercises.[8] From the political perspective, an important comparative static concerns the impact of how interests are represented in the legislature on policy outcomes. Of particular

[8] In the case of multiple equilibria, Banks and Duggan (2000, Theorem 3) prove that the set of solutions is upper hemicontinuous in the recognition probabilities and utility parameters.

concern here are the consequences of shifting the balance of influence (λ) between high and low types within the minority ($\mathcal{B}$) party. Table 2 presents the unique legislative bargaining equilibrium when the weight given to high-type blacks in the black party is increased to equal the share of high types within the black population as a whole, that is, $\lambda = .2$ in the example.

Table 2. The Equilibrium of the Legislative Bargaining Game when $\lambda = 1/5$

	Prob.	b_j	$t(b_j)$	α_j	$\alpha_0(b_j)$	c_{hW}	c_{0W}	c_{hB}	c_{0B}
$\mathcal{H}$'s proposal	.427	.025	.041	no aff.act.	.304				
	.573	0	0	.326					
$\mathcal{L}$'s proposal	.137	.110	.189	.315					
	.863	.067	.113	no aff.act.	.299				
$\mathcal{B}$'s proposal	1	.074	.126	.333					
Expected Value		.053	.089	.317		.314	.192	.322	.200
$\mathcal{L}$'s Ideal Policy		.119	.206	no aff.act.	.291	.299	.199	.299	.199
Expected Gain						5.1%	-3.7%	7.8%	0.4%

Parameter values: $p = 1/3$, $\lambda = \theta_B = 1/5$, $\theta_W = \beta = 1/2$, $q = h = 1/4$, and $F(x) = x$ for $x \in [0, 1]$.

As compared to Table 1, where no weight is given to the high-type blacks, an increase in the weight of highly educated blacks in the black party $\mathcal{B}$ increases the average affirmative action target and reduces the average level of fiscal redistribution. Intuitively, the shift of intra-party weight to high types within $\mathcal{B}$ enables high-type blacks and high-type whites to reach more profitable compromises than they otherwise could when low-type blacks exerted more control over party bargaining. This in turn improves $\mathcal{H}$'s bargaining power relative to $\mathcal{L}$. Relative to Table 1, all high types benefit from an increase in λ and all low types do worse. With $\lambda = 0.2$, however, it is still the case that low-type blacks are better off, albeit only marginally, relative to the equilibrium policy outcome under one-dimensional politics, that is, where the market alone determines the allocation of good jobs. This is not true for all values of λ: if the weight given to high types by minority legislators exceeds the share of high types in the minority population, low type blacks can be worse off with affirmative action than if taxes and transfers are the only redistributive policy.

Although we expect the comparative static illustrated in comparing Tables 1 and 2 reflects a general property of the model, we have so far been unable to prove such a conjecture.

5 Conclusion

In this paper, we have explored the consequences of ethnic or racial divisions for redistributive policy choice. When racial divisions lead to demands for redistributive policies along racial lines via affirmative action, we show that the bargaining equilibrium implies that the amount of redistribution along income lines is less on average that would exist were racial divisions absent. Redistribution along racial lines partly replaces redistribution along income lines in equilibrium. We also show that the expansion of the dimensions of redistribution benefits both highly educated members of the majority (who gain from lower taxation) as well as members of the minority (who gain from the affirmative action policies). The losers are members of the poorly educated members of the majority.

Not surprisingly, the model points to the importance of both political and economic factors for understanding the nature and consequences of inequality and redistributive policy in divided societies. Given the prominence of differences in human capital in our model, a natural extension is to expand the set of policies considered to include education. Education, however, is inherently a dynamic problem. At any moment in time, the distribution of human capital is relatively fixed. Over time, as new generations receive schooling, the distribution of human capital reflects investments in education as well as the distribution of human capital in previous periods. This in turn gives rise to dynamics in the politics of redistribution as earlier policies affect the distribution of resources and political interests in later periods.[9]

[9] See Roemer (2003) for an analysis of a model of the political choice of taxes, transfers and investment

As with the distribution of human capital in the model, the details of legislative and party structures are important for the results: the policy prediction in the absence of a minority party or caucus, for example, is radically different from that with such an explicit and independent minority representation. Thus a second extension to the model is look more deeply at the impact of alternative assumptions on legislative decision making and party composition. In particular, we are interested in identifying which forms of representation best promote the interests of the various subgroups within the population, both in the short run and the longer term.

in education in a dynamic setting where the current distribution of human capital reflects past investment in education and the past distribution of human capital.

6 Appendix

Equilibrium in the labor market is characterized by a system of two equations that jointly determine y_B and y_W and as functions of α and b, and a balanced budget constraint that determines t. Equations (4)-(5) in the text can be written as

$$\begin{aligned} \alpha y_B + (1-\alpha) y_W - \left[\frac{q}{1-\beta} + b\right] &= 0 \\ p(1-\alpha)\sigma_B - \alpha(1-p)\sigma_W &= 0 \end{aligned} \tag{8}$$

where $\sigma_i = [1 - G_i(y_i)] = [1 - (1-\theta_i)F(y_i) - \theta_i F(y_i - h)]$. If F is uniform, this is a system of linear equations in y_B and y_W with a unique solution. The tax rate is determined by the budget constraint, equation (6), which can be written as

$$t = \frac{(1-\sigma)b}{(1-\sigma)b + E(w)}. \tag{9}$$

where

$$\sigma = p\sigma_B + (1-p)\sigma_W$$

is the share of workers with good jobs and

$$E(w) = \beta\left[p\int_{y_B}^{\infty} y\, dG_B(y) + (1-p)\int_{y_W}^{\infty} y\, dG_W(y)\right] + (1-\beta)\sigma b \tag{10}$$

is the average wage. We denote the minority share of good jobs in the absence of affirmative action by $\alpha_0 < p$.

In what follows, we will frequently make use of the function

$$\Psi(\alpha) \equiv \left[\alpha^2(1-p) + (1-\alpha)^2 p\right] f > 0.$$

In addition, we will have occasion to use the following facts regarding y_W and y_B

$$\begin{aligned} y_B(\alpha_0, b) &= y_W(\alpha_0, b) = y_0 = \frac{q}{1-\beta} + b \\ y_B(p, b) &= y_0 - (1-p)h(\theta_W - \theta_B) \\ y_W(p, b) &= y_0 + ph(\theta_W - \theta_B) \end{aligned}$$

Proof of Lemma 1 (1) Differentiate (8) with respect to α, using f constant and writing $\Delta \equiv (y_W - y_B)$ to obtain

$$\frac{\partial y_B(\alpha,b)}{\partial \alpha} = \frac{\alpha(1-p)f\Delta - (1-\alpha)\sigma}{\Psi}; \tag{11}$$

$$\frac{\partial y_W(\alpha,b)}{\partial \alpha} = \frac{p(1-\alpha)f\Delta + \alpha\sigma}{\Psi}. \tag{12}$$

Since $\Delta \to 0$ as $\alpha \to \alpha_0$, it is clear that

$$\lim_{\alpha \to \alpha_0} \frac{\partial y_B(\alpha,b)}{\partial \alpha} = -\frac{(1-\alpha)\sigma}{\Psi} < 0$$

and

$$\lim_{\alpha \to \alpha_0} \frac{\partial y_W(\alpha,b)}{\partial \alpha} = \frac{\alpha\sigma}{\Psi} > 0.$$

So the result surely holds for $\alpha \approx \alpha_0$. Furthermore, since the affirmative action constraint is binding, $\theta_W > \theta_B$ implies $\Delta \geq 0$ for all $\alpha \in [\alpha_0, p]$; hence $\partial y_W/\partial\alpha > 0$.

>From (11), we have $\partial y_B/\partial\alpha < 0 \Leftrightarrow \alpha(1-p)f\Delta - (1-\alpha)\sigma < 0$. Note that, since F is uniform, we can write $\sigma = 1 - [py_B + (1-p)y_W - \theta h]\, f = 1 - (y_W - p\Delta - \theta h)\, f$. Therefore, we have

$$\alpha(1-p)f\Delta - (1-\alpha)\sigma = (\alpha - p)f\Delta - (1-\alpha)\left[1 - (y_W - \theta h)f\right]$$

Since $(\alpha - p)f\Delta \leq 0$ for $\alpha \leq p$, a sufficient condition for $\partial y_B/\partial\alpha < 0$ is that $[1 - (y_W - \theta h)f] > 0$. Since y_W is strictly increasing in α, $y_W(\alpha,b) \leq y_W(p,b) = y_0 + (\theta_W - \theta_B)\, ph$. Hence

$$\begin{aligned} [1 - (y_W - \theta h)f] &\geq [1 - (y_0 + (\theta_W - \theta_B)ph - \theta h)\, f] \\ &= [1 - F(y_0)] + [2p\theta_B + (1-2p)\theta_W]\, hf > 0 \end{aligned}$$

since $p < 1/2$.

(2) Differentiating (8) with respect to b, one obtains

$$\frac{\partial y_B(\alpha,b)}{\partial b} = \frac{\alpha(1-p)f}{\Psi} > 0$$

$$\frac{\partial y_W(\alpha,b)}{\partial b} = \frac{(1-\alpha)pf}{\Psi} > 0$$

To see that $\partial y_W/\partial b \geq 1 \geq \partial y_B/\partial b$, observe that

$$\Psi - \alpha(1-p)f = (1-\alpha)(p-\alpha)f \geq 0$$

while

$$\Psi - (1-\alpha)pf = -\alpha(p-\alpha)f \leq 0.$$

(3) Differentiating (9) with respect to b yields

$$\frac{\partial t}{\partial b} = \frac{[(1-\sigma)(1-t) - t(1-\beta)\sigma]\,\Psi + p(1-p)f\,[(1-t)b + t\beta\,(y_0+b)]}{[E(w) + (1-\sigma)b]\,\Psi} \tag{13}$$

Clearly $\partial t/\partial b > 0$ if $(1-\sigma)(1-t) - t(1-\beta)\sigma > 0$. Rearranging (9), we can write

$$b = \frac{t\,[E(w) - (1-\beta)\sigma b]}{(1-\sigma)(1-t) - t(1-\beta)\sigma}$$

Equation (10) implies $t\,[E(w) - (1-\beta)\sigma b] \geq 0$. Hence, the denominator must be positive (since $0 \leq b < \infty$). Therefore, we conclude $(1-\sigma)(1-t) - t(1-\beta)\sigma > 0$ which implies $\partial t/\partial b > 0$. $\square$

Proof of Proposition 1 Suppose $b > 0$ and consider the first claim, (1). Differentiate (9) with respect to α to obtain

$$\frac{\partial t(\alpha,b)}{\partial \alpha} = -\frac{b}{[(1-\sigma)b + E(w)]^2}\left\{E(w)\frac{\partial \sigma}{\partial \alpha} + (1-\sigma)\frac{\partial E(w)}{\partial \alpha}\right\}. \tag{14}$$

From the definition of σ and equations (11) and (12), we have

$$\begin{aligned}\frac{\partial \sigma}{\partial \alpha} &= -f\left[p\frac{\partial y_B}{\partial \alpha} + (1-p)\frac{\partial y_W}{\partial \alpha}\right] \\ &= \frac{f}{\Psi}\left[(p-\alpha)\sigma - p(1-p)f\Delta\right]. \end{aligned} \tag{15}$$

Note that $\partial\sigma/\partial\alpha = (p-\alpha)\sigma f/\Psi > 0$ when $\alpha = \alpha_0$ while $\partial\sigma/\partial\alpha = -p(1-p)f\Delta/\Psi < 0$ when $\alpha = p$. Differentiating (10)yields

$$\begin{aligned}\frac{\partial E(w)}{\partial \alpha} &= -f\left[w\,(y_B)\,p\frac{\partial y_B}{\partial \alpha} + w\,(y_W)\,(1-p)\frac{\partial y_W}{\partial \alpha}\right] \\ &= w\,(y_B)\,\frac{\partial \sigma}{\partial \alpha} - f\,[w\,(y_W) - w\,(y_B)]\,(1-p)\frac{\partial y_W}{\partial \alpha} \end{aligned} \tag{16}$$

where $w(y_i) = \beta y_i + (1-\beta)b$ is the wage received by the least productive member of group i who is given a good job. When $\alpha = \alpha_0$, then $w(y_B) = w(y_W) = w(y_0)$ and $\partial E(w)/\partial\alpha = w(y_0)\,\partial\sigma/\partial\alpha > 0$. When $\alpha = p$, $\partial E(w)/\partial\alpha < 0$ since $\partial\sigma/\partial\alpha < 0$ when $\alpha = p$ and $\partial y_W/\partial\alpha > 0$. In sum, both $\partial\sigma/\partial\alpha$ and $\partial E(w)/\partial\alpha$ are positive (negative) when $\alpha = \alpha_0$ ($\alpha = p$). Equation (14) then implies that $\partial t/\partial\alpha < 0$ ($\partial t/\partial\alpha > 0$) when $\alpha = \alpha_0$ ($\alpha = p$).

Consider the second claim, (2). By (1) and (2), we have

$$E(w+\pi) = \left[p\int_{y_B}^{\infty}(y-q)\ dG_B(y) + (1-p)\int_{y_W}^{\infty}(y-q)\ dG_W(y)\right]$$

Differentiating and collecting terms yields

$$\frac{\partial E(w+\pi)}{\partial\alpha} = f\left[p\frac{\partial y_B}{\partial\alpha}(q-y_B) + (1-p)\frac{\partial y_W}{\partial\alpha}(q-y_W)\right]. \tag{17}$$

Evaluating (17) at $\alpha = \alpha_0$ produces

$$\lim_{\alpha\to\alpha_0}\frac{\partial E(w+\pi)}{\partial\alpha} = (y_0-q)\frac{\partial\sigma(\alpha_0,b)}{\partial\alpha} = \left[\left(\frac{\beta}{1-\beta}\right)q + b\right]\frac{\partial\sigma(\alpha_0,b)}{\partial\alpha} > 0$$

since $\partial\sigma(\alpha_0,b)/\partial\alpha > 0$.

To see that $\partial E(w+\pi)/\partial\alpha < 0$ when $\alpha = p$, note that $\partial E(\pi)/\partial\alpha < 0$ for all $\alpha \in [\alpha_0, p]$. When the affirmative action constraint is binding, a tightening of the constraint must cause profits to decline. If not, the constraint was not binding. Since $\partial E(w)/\partial\alpha < 0$, when $\alpha = p$, it follows that $\partial E(w+\pi)/\partial\alpha < 0$ when $\alpha = p$.

It remains to show $\alpha_Y < \alpha_b$. Let α_w be the smallest value of $\alpha > \alpha_0$ such that $\partial E(w)/\partial\alpha = 0$. Since $\partial E(\pi)/\partial\alpha < 0$, it follows that $\partial E(w+\pi)/\partial\alpha < 0$ when $\partial E(w)/\partial\alpha = 0$. Therefore, $\alpha_Y < \alpha_w$. Equation (15) implies that $\partial\sigma/\partial\alpha > 0$ when $\partial E(w)/\partial\alpha = 0$. Equation (14) indicates that $\partial t/\partial\alpha < 0$ when $\partial\sigma/\partial\alpha > 0$ and $\partial E(w)/\partial\alpha = 0$. Therefore, $\alpha_b > \alpha_w$ and the claim that $\alpha_Y < \alpha_b$ is confirmed. $\square$

The expected consumption of an individual with human capital $H \in \{0, h\}$ in group $i \in \{B, W\}$ is given by equation (7) in the text; substituting for w, this expression can be rewritten:

$$E\left[c_{Hi}(\alpha,b)\right]=(1-t)\left\{\left[1-\beta\left(1-F(y_i-H)\right)\right]b+\beta\int_{y_i-H}^{\infty}(x+H)\,dF(x)\right\} \tag{18}$$

Proof of Lemma 2 (1) Differentiate (18) with respect to b to obtain (using f constant)

$$\begin{aligned}\frac{\partial E\left[c_{Hi}(\alpha,b)\right]}{\partial b} &= -\frac{E\left[c_{Hi}(\alpha,b)\right]}{1-t}\frac{\partial t}{\partial b}+(1-t)\left\{\left[1-\beta\left(1-F(y_i-H)\right)\right]-\beta(y_i-b)f\frac{\partial y_i}{\partial b}\right\}\\ &\equiv -T_1(H,i)\frac{\partial t}{\partial b}+T_2(H,i)-T_3(i)\frac{\partial y_i}{\partial b}. \end{aligned} \tag{19}$$

We have $0<T_1(0,i)<T_1(h,i)$ and $0<T_2(h,i)<T_2(0,i)$ since F is a CDF. The third term, $T_3(i)(\partial y_i/\partial b)$, is independent of H. By Lemma 1(2), $\partial t/\partial b>0$ for all b; hence, $\partial E\left[c_{0i}(\alpha,b)\right]/\partial b>\partial E\left[c_{hi}(\alpha,b)\right]/\partial b$ for all (α,b).

(2) If $\alpha=\alpha_0$ and f is a constant, then $T_1(H,B)=T_1(H,W)$, $T_2(H,B)=T_2(H,W)$ and $T_3(B)=T_3(W)$. By Lemma 1(2), $\partial y_W/\partial b>\partial y_B/\partial b>0$ and therefore $\partial E\left[c_{HB}(\alpha_0,b)\right]/\partial b>\partial E\left[c_{HW}(\alpha_0,b)\right]/\partial b$ as required. □

Proof of Lemma 3 Differentiating (18) with respect to α yields

$$\begin{aligned}\frac{\partial E\left[c_{Hi}(\alpha,b)\right]}{\partial \alpha} &= -\frac{E\left[c_{Hi}(\alpha,b)\right]}{1-t}\frac{\partial t}{\partial \alpha}-(1-t)\beta(y_i-b)f\frac{\partial y_i}{\partial \alpha}\\ &\equiv -T_1(H,i)\frac{\partial t}{\partial \alpha}-T_3(i)\frac{\partial y_i}{\partial \alpha}.\end{aligned} \tag{20}$$

Now, both $T_1(H,i)>0$ and $T_3(i)>0$, each $i\in\{B,W\}$. The first claims then follow from $\partial y_W/\partial\alpha>0>\partial y_B/\partial\alpha$ (Lemma (1)) and from the fact that $y_i-b=q/(1-\beta)$. Thus, the second term dominates if β is sufficiently close to one. The second claim follows from Proposition (1) and $T_1(h,i)>T_1(0,i)$. □

7 References

Alesina, A., Baqir and Easterly 1999, Public Goods and Ethnic Divisions, *Quarterly Journal of Economics,* 114(4): 1243-1284.

Alesina, A., E. Glaeser and B. Sacerdote 2001, Why Doesn't the US have a European-Style Welfare State. paper presented at Brookings Panel on Economic Activity, Sept. 7.

Alesina, A. and La Ferrara 2000, Participation in Heterogeneous Communities, *Quarterly Journal of Economics,* 115(3): 847-904.

Altonji, J. and R. Blank, 1999, Race and Gender in the Labor Market. In O. Ashenfelter and D. Card (eds) *Handbook of Labor Economics, volume 3* , New York: Elsevier

Banks, J. S., and J. Duggan 2000, Bargaining in a Model of Social Choice, *American Political Science Review,* 94, 73-88.

Baron, D. 1991, A Spatial Bargaining Theory of Government Formation in Parliamentary Systems, *American Political Science Review,* 85, 137-164.

Baron, D. and J. Ferejohn 1989, Bargaining in Legislatures, *American Political Science Review*, 83, 1181-1206.

Bartolini, Stefano. 2000. *The Political Mobilization of the European Left 1860-1980.* Cambridge: Cambridge University Press.

Dixit, A. and J. Londregan 1998, Ideology, Tactics and Efficiency in Redistributive Politics, *Quarterly Journal of Economics*, 113, 497–529.

Gelbach, J. and L. Pritchet 1997, Indicator Targeting in a Political Equilibrium: Leakier can be Better, Washington DC: World Bank Policy Research Working Paper 1523.

Gilens, M. 1999, *Why Americans Hate Welfare*, Chicago: University of Chicago Press.

Grossman, G. and E. Helpman 1999, Competing for Endorsements, *American Economic Review,* 89(3), 501-524

Kinder, D. and L. Sanders 1996, *Divided by Color: Racial Politics and Democratic Ideals*, Chicago: University of Chicago Press.

Lipset, S. M. and R. Bendix 1959, *Social Mobility in Industrial Society,* Berkeley: University of California Press.

Luttmer, E. 2001, Group Loyalty and the Taste for Redistribution, *Journal of Political Economy,* 89(5): 500-528.

Meltzer, A. H. and S. F. Richards1981, A Rational Theory of the Size of Government, *Journal of Political Economy*, 89(5), 914-27.

Myrdal, G. 1960, *Beyond the Welfare State,* New Haven: Yale University Press.

Quadagno, J. 1994, *The Color of Welfare,* Oxford: Oxford University Press.

Roberts, K. W. S. 1977. Voting over Income Tax Schedules. *Journal of Public Economics*, 8, 329–340.

Roemer, J. E. 1998, Why the Poor don't Expropriate the Rich, *Journal of Public Economics*, 70(3), 399-424.

Roemer, J. E. 2003, Does Democracy Engender Equality, in Edward J. Mansfield and Richard Sisson, eds., *Political Knowledge and Public Interest,* Columbus: University of Ohio Press.

Romer, T. 1975, Individual Welfare, Majority Voting, and the Properties of a Linear Income Tax. *Journal of Public Economics*, 14, 163–185.

Figure 1: Legislative Equilibrium
When $\lambda = 0$