Redistribution and Affirmative Action

David Austen-Smith[1]

Department of Managerial Economics & Decision Sciences

Kellogg School of Management

Northwestern University

Evanston, IL 60208-2009

Michael Wallerstein

Department of Political Science

Yale University

New Haven, CT 06520-8301

January 2003. This revision, May 2006

[1]Corresponding author: Telephone: (847) 467-3496. Fax: (847) 467-1220. email: dasm@kellogg.northwestern.edu.

Abstract

The paper develops an integrated political economy model in which individuals are distinguished by earning ability and an ascriptive characteristic, race. 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 legislative blocs. 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.

JEL Classification: D78, H24, H53, J79

Key words: Affirmative action, Fiscal redistribution, Legislative bargaining

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. Alesina and Glaeser (2004) conclude that racial conflict is one of the most important reasons for the low level of redistribution in the US compared to Europe.

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 contrast, the more formal approach to the political economy of redistribution has largely ignored the role of race or divisions rooted in individuals' ascriptive characteristics. Assuming individuals differ only in the single dimension of wealth or income, the focus has been on how changes in income distribution or in the political franchise influence majority preferences over redistributive fiscal policy; Romer (1975) and Meltzer and Richard (1981) are the seminal contributions.[1] Likewise, most formal models of affirmative action (e.g. Lundberg, 1991; Foster and Vohra, 1992; Coate and Loury; 1993; Chung, 2000) have focused on the implications of affirmative action policies for labor market outcomes rather than on the political choice of affirmative action when alternative redistributive policies are also on the agenda.

In this paper, we study the effect of social cleavages on the politics and economics of redistribution due to the introduction of an additional dimension of potential redistribution; that is, affirmative action. Motivated in principle by perceptions of past and current injustice, affirmative action is a political decision designed to influence pre-wage labor market allocations; and fiscal redistribution is a political decision designed to influence post-wage allocations of income. It is reasonable, therefore, that individuals' induced preferences over the two sorts of policy should be related: affirmative action (differentially) affects the *ex ante* opportunities for both black and white workers to secure more lucrative employment which, in turn, affects their views regarding the appropriate level of *ex post* fiscal redistribution. Furthermore, the relative supply of more or less lucrative employment opportunites is itself likely to be influenced by both affirmative action and fiscal policy.

To explore the sorts of tradeoff above, we build a general equilibrium political economy model in which individuals differ in their exogenously given stock of human capital and some ascriptive characteristic. In our application, we suppose the ascriptive characteristic is race with blacks being the minority group; the racial gap in earnings is captured by the minority having lower human capital on average. Individuals are otherwise assumed both color-blind

[1] The most salient exceptions to one-dimensional models of redistribution are models of competition between special interests: see, for example, Grossman and Helpman (2001) and 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.

and identical with respect to their fundamental (self-interested) preferences. The policy space consists of two policies. The first is a standard redistributive fiscal policy with a proportional tax that is used to finance a uniform benefit to workers in low-wage jobs. The second policy is an affirmative action target that requires employers to fill a given share of their higher paying jobs with minority workers.

The affirmative action policy is motivated by the historical discrimination resulting in the different current distributions of human capital between whites and blacks. In equilibrium this difference implies that, despite color-blind hiring by employers, the share of better jobs going to blacks falls short of the share of blacks in the population. Individuals' preferences over the policy space, therefore, are derived through equilibrium behavior in the economy and both race and income affect these induced policy preferences. The economy is modeled as a simple labor market in which the distribution of job opportunities is endogenous and workers are randomly matched with firms. Policy decisions are the outcomes of a legislative bargaining process.

Before going on, it is worth emphasizing that we do not deny the potential importance of racial or ethnic differences in preferences for determining redistribution policies. Our purpose in assuming that individuals behave 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. A different approach is adopted in the articles most closely related, as far as we know, to this paper. In Roemer (1998) and Roemer and Lee (2004), racial or religious differences are built into voters' preferences. Roemer (1998) assumes that voters care about their (post-tax and transfer) income and about government policy along a non-economic dimension such as race or religion.[2] In the case of two-party competition, Roemer shows that the existence of a second, non-economic policy dimension may reduce the equilibrium level of redistribution via

[2]Roemer and Lee (2004) add an assumption that voters also altruistically care about aggregate inequality, with racially conservative voters attaching less weight to equality than racially liberal voters.

a "policy bundling effect." For example, the election may pit a party that favors redistribution and non-economic support for minorities against a party opposed to both. A low-income, white voter who think that minorities get more than they deserve may prefer a party that opposes redistribution but matches the voters' racial views, depending on the relative weights of the two dimensions in the voter's preferences.

Our approach differs from Roemer's in several ways. First, as mentioned above, we assume that voters have identical, self-interested preferences. Second, we focus on a different mechanism linking racial cleavages and redistributive politics. The political side of our model consists of a model of legislative politics rather than electoral competition. In our framework, the policy-bundling effect that drives Roemer's results is absent since all distinct combinations of derived preferences over redistributive taxation and affirmative action may be represented in the legislature. And a substantive argument for the legislative bargaining model is that, as we consider in more detail later (section 5), it captures the political strategies that established affirmative action at the Federal level during the first two years of the Nixon administration. Affirmative action had a number of advantages for the Nixon administration, the most important of which, according to insider accounts, was that the plan drove a wedge between two parts of the Democratic electoral coalition, blacks and the (almost exclusively white) unions.[3] And although there were only thirteen black representatives in the House in 1969, Congressional support for African-American interests was more widespread. Thus our focus is on the trade-off between redistribution by income and by race that occurs through a process of legislative bargaining in which both white and black representatives are critical players.

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 racially or ethnically divided society that cannot be addressed with existing models. Does the

[3]See Skrentny (1996) or Anderson (2004).

presence of policies that redistribute according to ascriptive characteristics reduce support for redistribution according to income even with race-blind preferences? Who benefits and who loses when the policy space is expanded to include policies that redistribute by ascriptive traits? What are the net redistributive implications of using both racial and fiscal, rather than only fiscal, policies? How do the policies selected in equilibrium change as the distributions of income within the majority and minority social groups become more similar?

In addition to a making a claim about how redistribution by race can reduce redistribution by income, we believe that our model captures the essence of the politics of affirmative action at the federal level with regard to the employment of minorities in jobs in the middle of the wage scale, a policy that reached its peak during the Nixon administration and has never been abandoned (Anderson 2004). Affirmative action with respect to higher education, which concerns access to jobs at the top of the wage scale, has a different set of potential beneficiaries and losers and hence a different political dynamic; the model below is not well-suited to address this aspect of contemporary affirmative action policy.[4] We discuss briefly the federal policy of affirmative action after developing our model of the economy and the political equilibrium.

2 The economy

In the standard competitive model of the labor market with complete contracts, affirmative action is pointless. If each worker receives a wage that is just equal to his or her best alternative, there is nothing to be gained from special treatment in hiring. For affirmative action to be of interest, some jobs must yield rents to those holding those jobs. There are a variety of reasons why some jobs might offer a premium above the competitive wage level, from the ability of unions to obtain higher wages via collective bargaining to employers'

[4] See Chan and Eyster (2002) for a model of political conflict over affirmative action with regard to higher education.

willingness to pay higher wages to reduce shirking. In this paper, we employ a model of the labor market in which the source of inefficiency is the holdup problem, whereby the workers obtain a share of the return on employers' investments via wage bargaining that occurs after the investment costs are sunk. We emphasize that we do not intend this paper to be a contribution to the large literature on the holdup problem and how it might be overcome.[5] The holdup model simply provides a convenient economic framework in which affirmative action policies can play a role. Other models that generate employment rents in the labor market, such as efficiency wage models, would yield similar results concerning the political equilibrium.

2.1 Demographics and the labor market

We consider a large static economy with a continuum of individuals, each of whom belongs to one of two ascriptively distinct groups, Whites (W) and Blacks (B); let $p < 1/2$ denote the share of Blacks in the population, the minority group. Individuals have one of two levels of human capital, hereafter called 'skill', $H \in \{0, h\}$, where $H = h > 0$ denotes a skilled individual, or worker, and $H = 0$ denotes an unskilled worker. Let θ_i be the share of skilled workers in group $i = W, B$ and let θ be the share of such workers in the population as a whole, $\theta = [p\theta_B + (1-p)\theta_W]$. A key assumption is that Blacks are disadvantaged in the labor market through an historical racial difference in average human capital due to past discrimination; thus, $\theta_B < \theta_W$.

There is a finite number of competitive firms, each producing a homogenous consumption good and employing a continuum of workers, with firms and workers being randomly matched. We focus throughout on a representative match between a firm and a member of each of the

[5] Grout (1984) was the first to discuss the hold-up problem in the context of the labor market, as far as we know. For more recent studies of the hold-up problem applied to the labor market, see MacLeod and Malcomson (1993), Agell and Lommerud (1997), Acemoglu and Shimer (1999) and Acemoglu (2001) among others.

four possible types (by group and skill) of individual.

Firms offer two types of job j, good jobs and bad jobs. All workers are equally productive in bad jobs, regardless of their skill. A worker's productivity in a good job at a given firm, however, depends both on the worker's skill and on the realization of an idiosyncratic, match-specific random variable, $x \in \mathbb{R}$. Thus different workers with a common skill level may exhibit varying productivities at any given firm. Specifically, let $y(H, j, x)$ be the marginal product of a representative worker with skill H in job j at a given firm with realized match-specific variable x; then

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

where the zero marginal productivity in a bad job is a normalization. On average, skilled workers are more productive in good jobs than unskilled workers, but the most productive unskilled worker may be more productive in a good job than the least productive skilled worker. Assume x is distributed according to a differentiable and strictly increasing cdf $F(x)$ having full support on $\mathbb{R}$ and continuous density $f(x)$. Assume further that F exhibits a strictly increasing hazard rate:

$$\frac{f(x)}{1 - F(x)} \text{ is strictly increasing in } x, \text{ all } x \in \mathbb{R}.$$

In effect, the assumption means that the probability density of any individual being more productive at any alternative firm, conditional on realizing a match x, is falling with x.

Bad jobs are costless to create and the labor market for such jobs is presumed competitive: firms make zero profit from workers in bad jobs and these workers earn their marginal product (zero). The same is not true, however, of workers in good jobs. We assume that a firm can create a good job only at cost $q > 0$ and the worker's wage is determined through bargaining with the firm. Furthermore, we suppose the firm's decision on whether to invest in a good job for a particular worker is made after observing the worker's productivity at the firm, but before bargaining over the worker's wage begins. Assume the outcome of such bargaining is defined by the generalized Nash bargaining solution.

Identifying the bargaining solution requires identifying the value of the worker's outside option. Assume the consequence of failing to agree in wage negotiations is that the worker obtains a bad job and the good job remains vacant. Let $b \geq 0$ denote the value of a bad job (that b might be strictly positive is justified below). Then, conditional on the firm having invested in creating a good job, the worker's gain from reaching agreement on wage w is $[w - b]$ and the firm's gain from the agreement is $[y - w]$. Hence, writing $w(y, b)$ for the wage of a worker in a good job with realized productivity y and outside option value b, we have

$$\begin{aligned} w(y,b) &= \arg\max (y-w)^{1-\beta}(w-b)^{\beta} \\ &= \beta y + (1-\beta) b \end{aligned} \tag{1}$$

where $\beta \in (0, 1)$ is the worker's bargaining power, assumed constant across all workers with an option on a good job, irrespective of their particular productivities. The corresponding profit for the firm from hiring a worker in a type j job, therefore, is given by

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

Substituting for $w(y, b)$ gives

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

as the profit earned from the creation of a good job if the worker is hired (and we suppress the subscript j on $\pi(y, b)$).

In the absence of any affirmative action policy, firms treat white and black workers identically, creating a good job for a worker with productivity $y = H + x$ if and only if it is profitable to do so; that is, if and only if $\pi(y, b) \geq 0$ or, equivalently,

$$H + x \geq \bar{y}(b) \equiv \frac{q}{1-\beta} + b. \tag{3}$$

The fraction of group $i = W, B$ with good jobs when there exists no formal affirmative action policy, denoted $\sigma_i(\phi, b)$, is given by

$$\sigma_i(\phi, b) = [1 - G_i(\bar{y}(b))]$$

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. Hence, the share of good jobs held by minority workers in the absence of affirmative action, $\alpha_\phi(b)$, is

$$\alpha_\phi(b) = \frac{p\sigma_B(\phi, b)}{p\sigma_B(\phi, b) + (1-p)\sigma_W(\phi, b)}.$$

By assumption, $\theta_B < \theta_W$ so $G_B(\bar{y}(b)) > G_W(\bar{y}(b))$ and, consequently, $\sigma_B(\phi, b) < \sigma_W(\phi, b)$. Therefore, lower average skill for the minority group implies the share of good jobs held by minority workers is less than the share of minority workers in the work force, $\alpha_\phi(b) < p$.

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 (that is, $\beta > 0$), the number of good jobs created is less than the number that maximizes the joint income of employers and workers.[6] The additional output obtained from a marginal increase in the number of good jobs exceeds the cost of job creation: $\beta > 0$ implies $\bar{y}(b) > q$. Moreover, workers with $y = \bar{y}(b)$ obtain a discrete jump in pay of $w(\bar{y}(b), b) - b = [q\beta/(1-\beta)]$ in moving from a bad job to a good job, even though employers are indifferent between creating a good job or not when $y = \bar{y}(b)$. Thus, workers with productivity slightly below $\bar{y}(b)$ would gain from a policy that forces firms to create good jobs for them.

[6] In the context of our model, it is reasonable to ask why employers don't insist on negotiating the wage before investing in the creation of a good job, thereby avoiding the holdup problem. The assumption that wage bargaining takes place after the employer's investment is sunk can be seen as a reduced form version of a model where investments last for two or more periods. When investments last for more than one period, it is easy to show that workers receive rents and the number of good jobs is inefficiently low unless either (i) the initial wage contract covers the lifespan of the investment or (ii) employers can turn workers into co-investors by charging workers an upfront fee before investing in a good job.

2.2 Affirmative action and insurance

There are two types of policy in the society: social insurance and affirmative action. The social insurance policy provides a uniform transfer payment, $b \geq 0$ to all workers with bad jobs, financed by a budget-balancing flat tax $t \in [0, 1]$ on wages and welfare benefits (the assumption that benefits are taxed is purely for convenience; no result to follow depends upon it). Since workers with bad jobs receive zero wages, this benefit b defines the value of holding such a job. Affirmative action is modeled as a mandatory lower bound, $\alpha \in [0, p]$, on the proportion of good jobs filled by minority workers in any firm. If $\alpha \in [0, \alpha_\phi(b)]$ for any b, then the affirmative action policy is not binding; at the other extreme, if $\alpha = p$ the affirmative action policy requires firms to equalize the fraction of workers with good jobs in the two social groups.[7]

Typically, we expect $\alpha > \alpha_\phi(b)$ for any social insurance benefit b. In this case, affirmative action is binding and firms cannot use the same productivity threshold for both black and white workers when deciding whether or not to create a good job. Recall that each employer is large, in the sense of being randomly matched with a continuum of workers. Then a firm's optimal strategy with affirmative action policy α and benefit level b is to choose thresholds y_B and y_W to maximize expected profits

$$\max_{y_B, y_W} E[\pi|\alpha, b] = p \int_{y_B}^{\infty} \pi(y, b)\, dG_B(y) + (1 - p) \int_{y_W}^{\infty} \pi(y, b)\, dG_W(y)$$

[7] Holzer and Neumark (2000) observe that the enforcement of equal opportunity legislation has lead to affirmative action targets in practice, $\alpha > \alpha_\phi$. Since the underrepresentation of minorities in broad occupational categories is considered to constitute evidence of discrimination if it falls below numerical yardsticks set by the Equal Employment Opportunity Commission, firms that want to avoid discrimination claims would treat the EEOC yardsticks as constraints. With respect to the upper bound, $\alpha \leq p$, US courts have struck down affirmative action programs when the minority share of good jobs exceeded the minority share of the population at large: *Economist*, 10/4/03-10/10/03, p.30.

subject to the constraint that

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

where $\sigma_i(\alpha, b)$ denotes the share of group i workers with good jobs in the presence of an affirmative action policy α. Let $y_B(\alpha, b)$ and $y_W(\alpha, b)$ solve this problem; then all group i workers with $y \geq y_i(\alpha, b)$ are offered good jobs.

When the constraint (4) is binding, the first-order condition for a maximum can be written as two equations. The first equation replaces (3):

$$\alpha y_B(\alpha,b) + (1-\alpha) y_W(\alpha,b) = \bar{y}(b) \tag{5}$$

The second equation is simply the constraint, equation (4), written as an equality. When the constraint is not binding, that is, when $\alpha \leq \alpha_\phi(b)$ for any benefit level $b \geq 0$, equation (5) reduces to expression (3) with $y_B(\alpha, b) = y_W(\alpha, b) = \bar{y}(b)$ and the share of good jobs allocated to blacks being the *laissez faire* level, $\alpha_\phi(b)$. Hereafter, where there is no ambiguity, the arguments of $y_i(\alpha, b)$ and $\sigma_i(\alpha, b)$ are suppressed and we simply write y_i, σ_i etc.

Finally, it is useful to write the balanced budget constraint for financing the social insurance policy explicitly as

$$(1-\sigma)(1-t)b = tE[w|\alpha, b], \tag{6}$$

where $\sigma \equiv p\sigma_B + (1-p)\sigma_W$ is the share of the population with a good job (so receiving no benefit) and, from (1),

$$E[w|\alpha,b] = \beta\left[p\int_{y_B} y\, dG_B(y) + (1-p)\int_{y_W} y\, dG_W(y)\right] + (1-\beta)\sigma b \tag{7}$$

is the average wage at (α, b).

The policies (α, b) are fixed at the opening of the labor market. In the next section we identify equilibrium in the labor market conditional on these policies, with the description of exactly how the two policies are chosen deferred to a later section.

2.3 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).[8] Existence of a unique labor market equilibrium for any policy (α, b), insuring that individuals' induced preferences over policies are well defined, and two salient facts regarding the equilibrium, are established in Lemma 1. Proofs for all lemmas and propositions are collected in an Appendix.

Lemma 1 *(1) There exists a unique labor market equilibrium associated with every policy* (α, b)*.*

(2) If $\alpha > \alpha_\phi(b)$ *then* $y_W(\alpha, b) > \bar{y}(b) > y_B(\alpha, b)$*.*

(3) In equilibrium, $\alpha_\phi(b)$ *is strictly decreasing in* b*.*

Lemma 1 is fairly self-explanatory. If the affirmative constraint binds, then profit-maximizing firms adjust their hiring policy by setting distinct thresholds for creating good jobs for blacks and for whites; the threshold for blacks being necessarily lower than that for whites, with the weighted average of the two being exactly the *laissez faire* threshold, $\bar{y}(b)$. And the *laissez faire* share of good jobs allocated to blacks, $\alpha_\phi(b)$, is decreasing in b because higher benefit levels reduce the incentive of employers to create jobs, thus raising the threshold $\bar{y}(b)$, which affects blacks proportionately more than whites because of the differential distributions of human capital across the two groups. In other words, if affirmative action were not an available policy instrument, successful demands for higher levels of benefit to address income inequality result in a reduction in the share of higher paying jobs allocated to the minority group.

[8] 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.

The next result, Lemma 2, describes some useful comparative static properties of policy change.

Lemma 2 *(1)* $\partial y_W(\alpha, b)/\partial\alpha > 0$ *and, for all* $b \geq 0$, $\lim_{\alpha\downarrow\alpha_\phi(b)} \partial y_B(\alpha, b)/\partial\alpha < 0$;

(2) $\partial y_W(\alpha, b)/\partial b > 0; \partial y_B(\alpha, b)/\partial b > 0$;

(3) $\partial t(\alpha, b)/\partial b > 0$;

(4) $\lim_{\alpha\downarrow\alpha_\phi(b)} \partial E[w + \pi|\alpha, b]/\partial\alpha > 0$;

(5) $\lim_{\alpha\downarrow\alpha_\phi(b)} \partial t(\alpha, b)/\partial\alpha \leq 0$ *with strict inequality if and only if* $b > 0$.

The first three claims of Lemma 2 are intuitive. Lemma 2(1) says that affirmative action always induces firms to raise the threshold for hiring white workers and, at least for binding policies α close to the *laissez faire* level $\alpha_\phi(b)$, to lower the threshold for hiring minority workers.[9] For sufficiently onerous affirmative action mandates relative to the *laissez faire* level, and depending on details of the distribution function F, marginal increases in α may result in increases in the thresholds for both whites and blacks as firms reduce the overall level of good jobs in the economy. On the other hand, Lemma 2(2) states that both of these thresholds are strictly increasing functions of the benefit for any affirmative action policy α: an increase in benefits b raises the opportunity cost of a good job and thus the threat point for wage bargaining. As a result, profits fall and, as reported in Lemma 1, employers compensate by creating good jobs only for more productive workers which induces a fall in the *laissez faire* share of good jobs held by blacks. And Lemma 2(3) confirms that 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. First, the tax and benefit redistributes *ex post* from workers in good jobs, who pay taxes but don't receive the benefit,

[9] Holzer and Neumark (2000) report lower employment of white males (by 10-15 per cent) and the employment of minorities with lower qualifications (as defined by test scores or education) in establishments with affirmative action hiring policies.

to workers in bad jobs. Second, the tax and benefit redistributes *ex ante* from skilled workers who are likely to obtain good jobs to unskilled workers who are less likely to obtain good jobs. Third, fiscal policy redistributes income from employers to employees. Consequently, while benefit increases may raise the aggregate income received by workers provided the benefit is not too large, they surely reduce aggregate output. On the other hand, under the assumption that the distribution of match-specific productivities F exhibits a strictly increasing hazard rate, affirmative action policies in the neighborhood of the *laissez faire* level enhance aggregate output (Lemma 2(4)) and permit a lower tax-rate for any positive level of benefits (Lemma 2(5)). In view of Lemma 2(1), when $\alpha \approx \alpha_\phi(b)$ 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 at $\alpha \approx \alpha_\phi(b)$, aggregate output increases. Aggregate outcome may be reduced, however, with a sufficiently high affirmative action target as the acceptance thresholds y_W and y_B move further apart.

We now turn to consider the political choice of the two policy instruments.

3 Induced policy preferences

A policy is a pair $(\alpha, b) \in \mathbf{P} \equiv [0, p] \times [0, \infty)$. 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, let $c_{Hi}(\alpha, b)$ denote the expected consumption of an individual with skill $H \in \{0, h\}$ in group $i \in \{B, W\}$ at policy (α, b):

$$c_{Hi}(\alpha, b) = (1 - t) \left\{ F(y_i - H)b + \int_{y_i - H}^{\infty} w(H + x, b) dF(x) \right\}, \tag{8}$$

where y_i and t are defined by the labor market equilibrium at (α, b). The first term between the braces is the individual's income conditional on a bad job, and the second term is her expected income conditional on receiving a good job.

Although individuals' preferences are linear in consumption, Lemma 2 makes clear that consumption is nonlinear in policy. This greatly complicates explicit analysis. However, as $c_{Hi}(\alpha, b)$ is continuous on $\mathbf{P}$ and there is no loss of generality in assuming the maximal admissible benefit level is finite, there exists a global maximum for each individual type (H, i). And if this maximum is not a boundary point then it is generically unique; if the maximum is a point (α^*, b^*) with $\alpha^* \leq \alpha_\phi(b^*)$, however, definition of $\alpha_\phi(b^*)$ as the *laissez faire* share of good jobs going to blacks at the benefit level b implies that the individual's expected consumption is constant for all policies (α, b^*) such that $\alpha \leq \alpha_\phi(b^*)$. In this case, we simply report the most preferred policy as $(\alpha_\phi(b^*), b^*)$.

For each (H, i), let $(\alpha^*_{Hi}, b^*_{Hi}) \in \mathbf{P}$ denote the (H, i) type's most preferred policy pair. We are interested in identifying the relative locations of the ideal points across skill level and race. Because of the nonlinearities inherent in expected payoffs, additional restrictions on the parameters of the model are required to make some of the relevant comparisons. Essentially, these restrictions insure that fiscal policy is politically salient across low and high types of worker, and that affirmative action policy is politically salient across blacks and whites. If the relative expected return from a good rather than a bad job is too small (q or β too low) or if there is insufficient distinction between high and low skill types (h too small), then there is too little variation in policy preferences for policy to matter differentially across types. The restrictions, therefore, are to guarantee that good jobs are sufficiently attractive to induce variation in *ex ante* policy preferences and so insure that the policy choice problem is not trivial.

Lemma 3 *There exists a finite skill level* $\bar{h} > 0$ *and cost of good job creation* $\bar{q} > 0$ *such that* $h \geq \bar{h}$ *and* $q \geq \bar{q}$ *imply:*

(1) $[\alpha^*_{hB} > \alpha_\phi(0), b^*_{hB} = 0]$*;* $[\alpha^*_{0B} > \alpha_\phi(b^*_{0B}), b^*_{0B} > 0]$*.*

(2) $[\alpha^*_{hW} = \alpha_\phi(0), b^*_{hW} = 0]$*;* $[\alpha^*_{0W} = \alpha_\phi(b^*_{0W}), b^*_{0W} > 0]$*.*

Furthermore, firms' profits are maximal at $(\alpha_\phi(0), 0)$*.*

In sum, Lemma 3 says that blacks prefer higher levels of affirmative action to whites, and low skill types prefer both more fiscal redistribution than high skill types. And it is worth noting here that, while blacks always prefer some affirmative action policy to the *laissez faire* level at any benefit level, in general whites might prefer no affirmative action or some modicum thereof. Unlike blacks, whites face a tradeoff: at any *laissez faire* level $\alpha_\phi(b)$ with $b > 0$, whites benefit from the efficiency gains induced by small levels of affirmative action (Lemma 2(5)) but are hurt by the distributional losses that such action implies for the majority group (Lemma 2(1)). So it is possible that, at least for some strictly positive benefit levels and a sufficiently small difference between high and low skill levels, the efficiency gains dominate the distributional losses. However, for sufficiently high skill level h and cost of job creation q, the distributional costs dominate the efficiency gains. Hereafter, we assume $h \geq \bar{h}$ and $q \geq \bar{q}$ as required for Lemma 3.

4 Legislative policy choice

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 high skill white workers ($\mathcal{H}$), legislators who represent low skill white workers ($\mathcal{L}$), and legislators who represent black workers ($\mathcal{B}$).

Two comments on this choice of partition are worth making explicit. First, firms are not directly represented which fits with the secondary role of employers in the political conflict over affirmative action at the Federal level, as we discuss in section 5 below. Second, unlike white workers, we presume all black workers are represented by the same bloc of legislators, $\mathcal{B}$. In part this assumption is one of convenience; but more importantly it also reflects the empirical reality that while majority group parties have a history of development, with core constituents that are often separated by economic interests, the explicit representation of minority groups is a relatively recent phenomenon and as such, is better approximated as a single bloc.

For purposes of exposition, we refer to the three representative 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 of a given skill are identical *ex ante*. Therefore, the objective functions for $\mathcal{H}$ and $\mathcal{L}$, respectively, are $u_{\mathcal{H}} = c_{hW}(.)$ and $u_{\mathcal{L}} = c_{0W}(.)$. Minority voters, however, include both high and low skill workers. In this case, we assume the legislative group maximizes a weighted average of the consumption of its two types of constituents: $u_{\mathcal{B}} = \lambda c_{hB}(.) + (1-\lambda)c_{0B}(.)$ for some $\lambda \in [0,1]$. The weight λ is a measure of the extent to which the minority party is concerned with high or low skilled minority workers and we presume throughout that λ is not so high that low skill blacks have no incentive to remain in a party with high skill blacks (we pursue this issue a little further in the comparative statics section below). 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 $\mathcal{B}$.

4.1 Bargaining equilibrium

To avoid a trivial solution to policy conflict, we assume that no single group has a majority of seats in the legislature. If, for instance, the size or weight of each legislative bloc reflects the relative size of each bloc's constituents, then the weights of parties $\mathcal{B}$, $\mathcal{H}$ and $\mathcal{L}$ respectively are $p < 1/2$, $(1-p)\theta_W < 1/2$ 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 (generically) empty. By Lemma 3, if $b^*_{0i} > 0$ each i, the set of Pareto efficient policies in $\mathbf{P}$ for the three parties is (generically) two-dimensional and there is no majority core in $\mathbf{P}$ (Schofield 1984). If there is a core, it is at the low whites' ideal point; but as this is a knife-edge possibility at best, we ignore it hereafter.

Because there is generally no core policy, we model the policy process as a legislative bargaining game. Specifically, we apply the generalized version of the Baron and Ferejohn (1989) legislative bargaining model due to Banks and Duggan (2005). 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 stationary subgame perfect Nash equilibrium with no-delay (hereafter, simply *equilibrium*). 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 proposals by others that party j supports; along the equilibrium path, the first party to be recognized offers its best proposal from among the set of proposals that will be accepted and the game ends. 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:

$$\begin{aligned}
P_j &\subseteq \begin{cases} \arg\max \{u_j(\alpha, b) \mid (\alpha, b) \in (A_k \cup A_l)\} & \text{if } \sup [u_j(\alpha, b) : (\alpha, b) \in (A_k \cup A_l)] \geq v_j, \\ \mathbf{P} \backslash (A_k \cup A_l) \text{ otherwise} \end{cases} \\
A_j &= \{(\alpha, b) \mid u_j(\alpha, b) \geq v_j\} \\
v_j &= \sum_{k=\mathcal{H},\mathcal{L},\mathcal{B}} \rho_k \left[\int_{P_k} u_j(\alpha, b)\, d\zeta_k \right]
\end{aligned}$$

The first condition states that any policy a party j proposes must yield at least its constituents' continuation value v_j from having the policy rejected: either there is such a policy within the acceptance set of some party $k \neq j$ that weakly improves on v_j (in which case the second condition implies that such a proposal must be acceptable to j itself), or exactly the value v_j is achieved by proposing a policy that is sure to be rejected. The second condition states that each party accepts any proposal that provides a higher or equal payoff than the party's continuation value. The third condition states that in equilibrium the continuation value equals the expected value of the game.

The most general (no-delay) equilibrium existence result for this game (at least, as far as we know) is due to Banks and Duggan (2005, Theorem 1). However, their theorem assumes utilities are concave on the policy space. Unfortunately, concavity is not a general property of induced preferences here.[10] The difficulty introduced by nonconcave utilities

[10] The absence of concavity 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

is essentially technical: the exclusive role of concavity for insuring equilibrium existence is to insure the existence of nonempty acceptance sets A_j. So, rather than attempt to finesse complications with equilibrium existence due to nonconvexities in parties' induced preferences, we simply assume in what follows that an equilibrium exists. And to show that the equilibrium concept is not vacuous in this regard, we later present a nonpathological example in which all parties' induced preferences are strictly concave on $\mathbf{P}$ and calculate the legislative bargaining equilibrium.

Taking the existence of equilibrium (along with the sufficient conditions used for Lemma 3) as given, the impact of a second dimension of conflict over affirmative action on the legislative support for fiscal redistribution is summarized in our first proposition regarding equilibrium outcomes.

Proposition 1 *The expected level of fiscal redistribution when affirmative action is a policy variable is less than that most preferred by a majority of individuals when affirmative action is not a policy variable.*

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 exclusively advance the interests of the less educated members of the minority group as in our example below. Although the low-type black might prefer an even greater b than low-type whites in the absence of affirmative action, blacks and high whites now have the potential to lower the tax rate and raise affirmative action targets in a way that makes both better off. The possibility of splitting the coalition between low-type whites and blacks is sufficient to reduce the average

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 over some regions of the policy space.

transfer payment and to increase the average affirmative action target in equilibrium.[11]

To see the intuition for Proposition 1, suppose first that affirmative action is not an available policy instrument. Then it is easy to check that all low skill workers share the same ideal benefit level, defined by the low skill white's most preferred benefit level, b^*_{0W}. Furthermore, b^*_{0W} is strictly greater than the high skill workers' most preferred level and low skill workers constitute a strict majority of the population. Therefore b^*_{0W} is a majority core allocation and this uniquely defines the equilibrium decision on social insurance policy.[12] Given the assumption of only two skill levels in the economy, the proposition is then immediate if, when affirmative action is an available policy instrument, the most preferred benefit level for the Black party, $b^*_{\mathcal{B}}$, is no greater than that for low skill whites, i.e. if $b^*_{\mathcal{B}} \leq b^*_{0W} = b^*_{\mathcal{L}}$. Now suppose that the most preferred benefits level for the Black party when affirmative action is an available policy is strictly greater than that for low skill whites, i.e. if $b^*_{\mathcal{B}} > b^*_{\mathcal{L}}$. Then it is *a priori* possible for the equilibrium benefit level to rise relative to the level when affirmative action is unavailable. But any proposal for such an outcome can be surely blocked by the high white party $\mathcal{H}$ proposing the low white party's ideal point, $(\alpha^*_{\mathcal{L}}, b^*_{\mathcal{L}})$. It follows that no equilibrium policy pair can have a benefit level (and thus a tax-rate) higher than $b^*_{\mathcal{L}} = b^*_{0W}$.

A limitation of the result is that the argument uses the assumption of only two skill levels in the economy. If there were a middle skill level such that, conditional on affirmative action being unavailable, individuals with this level of human capital prefer a benefits level between those most preferred by the high and the low skill workers, then the proposition is not so transparent. When the affirmative action policy is not available, the ideal level of fiscal redistribution for the middle skilled workers is the core allocation and the unique bargaining

[11] See Levy (2004) for a model that captures the same phenomenon of the rich reducing their tax bill by playing off divisions among poor voters with regard to paying for school.

[12] Strictly speaking, if parties are impatient and increasingly discount payoffs with the time taken to reach a legislative decision, the final outcome can be slightly less than b^*_{0W} depending on which party makes the first proposal.

outcome. As in the two-skills model, making affirmative action policy available admits a richer set of legislative coalitions and, in principle, there is scope for a coalition between low skill white workers and the Black party to implement higher benefits in exchange for high levels of affirmative action. Whether or not such an outcome is sustainable in equilibrium, however, depends more subtly on the possible counter-coalitions in the legislature, as well as on details of the various individuals' induced preferences over the policy space. Nevertheless, the logic of the proposition, that being able to redistribute income through affecting job prospects induces a legislative substitution away from a reliance on redistribution through social insurance, seems compelling.

The second proposition summarizes the impact of adding of a second dimension of redistributive conflict on the expected post-tax, post-transfer equilibrium income of each group of voters.

Proposition 2 *The possibility of redistribution through affirmative action (1) raises the expected income of both high whites and of blacks as a group (where the incomes of high and low type blacks are aggregated with weights $\lambda \geq 0$ and $(1-\lambda)$ respectively), and (2) lowers the expected income of low whites.*

In expected value, the high-type whites gain more from a low tax and benefit level than they lose from a higher affirmative action target while blacks, in aggregate, gain more from the affirmative action target than they lose from a lower benefit. The big losers, in comparison to the political equilibrium in the absence of affirmative action, are the low-type whites who are made worse off by both affirmative action and lower benefits.

Propositions 1 and 2 can be illustrated with the following example, that also serves to show the existence of bargaining equilibria here. It is worth observing that the example is similar to the equilibria generated by other parameterizations. Existence is not a knife-edge property of the model. 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 the example are general characteristics of all equilibria of the legislative bargaining model.

Example 1 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 are low skill 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 low skills, 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 $\alpha_\phi(b)$ which represents the share of good jobs that the minority would receive without affirmative action. Note that $\alpha_\phi(b)$ declines as b increases (see Lemma 1). 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: $\rho_j = 1/3$, all j. 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 some numerical details of the equilibrium in this example. For comparative purposes, Table 1 also show the expected value of the benefit, the expected tax rate, and the expectied share of good job held by minorities associated with the legislative equilibrium. When there is no racial divide, political conflict occurs over the single dimension of the tax rate. Since, in the 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 is a core point. Finally, we report the share of goods jobs held by minorities for the *laissez faire* policy of $\alpha = \alpha_\phi(0)$ and $b = 0$.

Table 1 here

Using the equilibrium proposals reported in Table 1 above, the expected net percentage gain in consumption from the equilibrium outcome when affirmative action is a policy variable relative to when it is not can be calculated. Doing this yields, as predicted by Proposition 2, that the expected consumption (1) of high whites *increases* by 2.8%; (2) of low whites *decreases* by 2.8%; (3) of high blacks *increases* by 5.1%; and (4) of low blacks *increases* by 1.0%. Thus, the largest winners from the presence of redistributive policies along racial lines are high skill workers, especially high skill minority workers. High skill minority workers gain both from affirmative action and the tax reduction, while high skill majority workers benefit from the tax reduction. Even though black legislators were assumed to be representatives *exclusively* of black workers with the low skills ($\lambda = 0$), 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 low skill majority who lose from affirmative action and lose again from the reduction in the average amount of redistribution along income lines.

Proposition 2 and the numerical illustration raise a question about the effect of affirmative action on the inequality of post-tax, post-transfer income. In the setting in which the only policy dimension is fiscal redistribution, then the distribution of post-tax and transfer income is unequivocally more equal than the laissez faire distribution. To get some idea of the impact of choosing affirmative action along with fiscal redistribution on income inequality, we computed the Gini coefficient for the economy described in Example 1. Under *laissez faire*, the Gini is 0.482; in the case that only fiscal policy is available and there is no affirmative action constraint, the low white workers' most preferred tax-policy prevails and the resulting

Gini falls to 0.347. Introducing affirmative action attenuates the extent to which income is redistributed: the expected (equilibrium) Gini from the legislative bargaining process is 0.392, with the high white party's proposal to the black party involving the least equalization and the black party's proposal to the low white party involving the greatest equalization. Nevertheless, although there is less equalization when affirmative action is subject to policy choice than when only fiscal redistribution is feasible, black workers benefit more as a group (at the expense of low white workers) when affirmative action is a policy issue than when it is not.

4.2 Comparative statics

In addition to existence, Banks and Duggan (2005, Theorem 3) prove that if the equilibrium is unique, as in our example, then it is a continuous function of the recognition probabilities and parameters of the legislators' utility functions, so justifying comparative static exercises.[13] And note that if the only political issue is fiscal policy (so the share of blacks with good jobs is invariably defined by the *laissez faire* level), then under our assumptions, there is surely a core allocation at the low white's most preferred benefit level conditional on no affirmative action policy.

4.2.1 Social variation

One motivation for affirmative action in divided societies derives from an intuition that increases in the accessibility of good jobs and in the numbers of minority individuals holding such jobs, provide both indirect (through role models) and direct (through expected economic returns) incentives for increased investment in human capital formation among minorities (e.g. Foster and Vohra, 1992; Coate and Loury, 1993; Chung, 2000). For this and for many

[13] Although their argument assumes concavity of the utility functions, such concavity is used only to insure equilibrium existence. Given existence, the continuity (strictly speaking, upper hemicontinuity) property depends essentially on the continuity of preferences and compactness of the set of equilibrium proposals.

other reasons, the educational gap between majority and minority groups is not static. The question of how the equilibrium changes with the distribution of human capital, therefore, is of interest. Proposition 3 states affirmative action and direct fiscal redistribution vary with the share of high blacks in the population (θ_B).

Proposition 3 *The political equilibrium approaches the majority core of the policy space where fiscal redistribution is the only dimension as $\theta_B \to \theta_W$.*

As the average level of human capital between white and blacks approaches equality, affirmative action drops out as a separate policy dimension, given the constraint that affirmative action target cannot exceed the share of minorities in the population. With both white and black workers receiving good jobs in the same proportion or as $\alpha_\phi(b) \to p$ for all b, redistribution politics is exclusively concerned with taxes and benefits.

We cannot prove that the convergence to the one-dimensional majority core is monotonic, but monotonic convergence seems to be the typical case. Table 2 continues with the basic parameterization of Example 1 (in which x is uniformly distributed over the interval $[0, 1]$, $q = h = 1/4$, $\beta = 1/2$, $p = 1/3$, and $\theta_W = 1/2$), but varies θ_B between 1/5 and 1/2. Choosing $\lambda = 0$ implies the increases in the proportion of high blacks have no effect on their influence in the legislative bargaining process. As the proportion of high-type blacks among the black population rises, so too does the share of good jobs going to minorities and the expected equilibrium benefit and tax rates. Not surprisingly, given that an increasing share of good jobs are allocated to minorities independently of affirmative action (because highly educated blacks constitute an increasing proportion of the labor force and firms are color-blind), there is less pressure on affirmative action targets. Moreover, because $\lambda = 0$ and high blacks have no influence on the legislative bargaining process, the black party becomes increasingly allied with the low white party in the legislature as the importance of affirmative action diminishes.

Table 2 here

4.2.2 Economic variation

Another comparative static concerns the degree of economic efficiency (q). So far we have assumed q sufficiently large to induce politically salient variations in workers' preferences over policy.[14] As q falls the market inefficiency due to employers' investment costs being borne prior to any wage bargaining becomes less important and the corresponding wage-premium for securing a good job declines. Intuitively, then, the rationale and demand for affirmative action policy weaken with reductions in q. Exactly this intuition is contained in Proposition 4.

Proposition 4 *The political equilibrium approaches the majority core of the policy space where fiscal redistribution is the only dimension as* $q \to 0$*.*

As q goes to zero, the difference between good jobs and bad jobs is reduced and affirmative action loses its importance. Again, the political equilibrium approaches to majority core when fiscal redistribution is the only dimension.

Table 4 continued the same parametrization as Example 1 (with $\lambda = 0$), letting q go from $q = 1/4$ to $q = 0$. The interesting feature to note in Table 4 is that, as q falls, the share of good jobs held by minorities goes up somewhat, reflecting a willingness of employers to create more good jobs in total, but the equilibrium benefit (and tax) rate goes up considerably. With a falling marginal return to obtaining a good job, increasing weight is placed by the black party on fiscal redistribution relative to affirmative action.

Table 3 here

[14] It is worth noting that increases in the technical cost parameter q have a qualitatively identical impact on equilibrium outcomes as increases in the relative bargaining strength of workers, β.

4.2.3 Political variation

From the political perspective, an important issue involves the impact of how interests are represented in the legislature on policy outcomes. And of particular concern here are the consequences of shifting the balance of influence (λ) between high and low types within the black party. We are unable to obtain analytical results which respect changes in λ, but the numerical results are intuitive. Given the parameterization of Example 1 with x is uniformly distributed over the interval $[0, 1]$, $q = h = 1/4$, $\beta = 1/2$, $p = 1/3$, $\theta_W = 1/2$, and $\theta_B = 1/5$, Table 4 describes the unique legislative bargaining equilibrium as the weight given to high-type blacks in the black party, λ, is increased from zero (as in Table 1) to the share of high types within the black population as a whole, $\lambda = \theta_B = 1/5$, and beyond to $\lambda = 1/3$. The last four columns of the table report the expected percentage gain or loss of high-type whites (HW), low-type whites (LW), high-type blacks (HB) and low-type blacks (LB) of the expected equilibrium policy outcome with affirmative action relative to the equilibrium without affirmative action.

Table 4 here

It is apparent from the table that an increase in the weight of highly educated blacks in the black party $\mathcal{B}$ increases the average affirmative action target (minority share of good jobs) 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 $\mathcal{L}$. Consequently, all high types benefit from an increase in λ and all low types do worse. In particular, once the relative influence of high blacks increases sufficiently beyond their share of the black population at large, low-type blacks are left strictly worse off with affirmative action than if taxes and transfers were the only redistributive policy, that is, where the market alone determines the allocation of good jobs.

Proposition 2 states that affirmative action can be expected to improve the lot of black workers *as a group* relative to when only fiscal policy is at issue when black incomes are aggregated using the weights λ and $1 - \lambda$ for high and low types respectively. On the other hand, Table 4 makes clear that the distribution of such gains between high and low type black workers, however, need not leave low type blacks better off as a subgroup. If enough weight is given to the high blacks (λ high) to leave the low blacks with strictly less in equilibrium than they would receive at the low whites party's ideal point, $I_{\mathcal{L}} = (\alpha_\phi(.12), .12)$, then representatives of low blacks prefer to defect from $\mathcal{B}$ and vote with $\mathcal{L}$ to insure $I_{\mathcal{L}}$ as the legislative decision. In other words, for the parameterization of Table 4, $\lambda = .3$ provides an upper-bound on the politically sustainable representation of high blacks within the black party. It is worth noting that no such lower bound is apparent: high blacks benefit disproportionately even in the case $\lambda = 0$.

5 Discussion

In this paper, we explore the consequences of ethnic or racial divisions for redistributive policy choice in a world devoid of prejudice. There is no suggestion here that racial prejudice is in fact irrelevant, only that it seems sensible to identify what happens in the absence of racial prejudice. Results derived in a prejudice-free setting provide a clear illustration of the impact of the introduction of additional dimensions of potential redistribution on the amount of redistribution that occurs in equilibrium.

The motivation for redistribution along racial lines in the model is an inefficiency in the labor market creating rents for those holding good jobs, coupled with an exogenously (historically) given difference in the distributions of human capital across races. When racial divisions lead to demands for redistributive policies along racial lines via affirmative action, we show that legislative policy bargaining implies that the amount of redistribution along income lines is less on average than would exist were racial divisions absent. When affirmative action is on the agenda, redistribution along racial lines partly replaces redistribution along income lines and total redistribution, as measured by the change of the Gini coefficient, declines. 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 the low skilled majority (white) workers.

An advantage of the legislative bargaining model is that it captures the political strategies that lead to the adoption of affirmative action at the Federal level during the first two years of the Nixon administration.[15] The term "affirmative action" applied to race was first used in the Kennedy administration, but the meaning was to insure that minority employees were treated "without regard to race, creed, color and national origin" by the federal government or by government contractors. Title VII of the 1964 Civil Rights Act make it unlawful for any employer with more than 25 employees to discriminate according to race, religion, sex and national origin and gave the courts the power to enforce the act. When, in 1968, the Labor Department issued a regulation that federal construction contractors were required to have affirmative actions programs that had "specific goals and timetable", the goals were voluntary. After the US Comptroller General (who works for Congress) declared "goals and timetable" to be in violation of the intent of Title VII that all employees be treated in a race-blind manner, the attempt to encourage federal contractors to hire more minorities came to a halt until Nixon came to office.

Affirmative action was advocated by the Nixon administration with far more vigor than displayed by the previous Democratic administrations. Soon after Nixon's inauguration, Labor Secretary George Schultz presented a plan of affirmative action whereby the Office of Federal Contract Compliance would establish a target range for bidders on Federal construc-

[15] The story is told in two books: Skentny (1996) and Anderson (2004). Our account follows Anderson closely. Holzer and Neumark (2000) also present a brief history and an discussion of how affirmative action worked in practice.

tion projects that was related to the share of minority workers within a given area, plus timetables for progress toward the target. Affirmative action had a number of advantages for the Nixon administration. It was "consistent with the spirit of self-reliance," in Schultz' words (Anderson 2004, 116). Unlike fiscal redistribution, affirmative action required no government money. But what was most important, according to insider accounts, is that the plan drove a wedge between two parts of the Democratic coalition, black and the unions. In the words of John Ehrlichmen, Nixon's aide, "The NAACP wanted a tougher requirement; the unions hated the whole thing. Before long, the AFL-CIO and the NAACP were locked in combat over one of the passionate issues of the day and the Nixon administration was located in the sweet and the reasonable middle" (Anderson 2004, 120).

In 1969, the Comptroller General again declared affirmative action to be in violation of Title VII of the 1964 Civil Rights Acts. In the Senate, Democratic Senator Sam Ervin introduced a rider to an appropriations bill stating that Comptroller General, not the executive branch, should determine who could receive federal contracts and a second rider asserting that Title VII was the nation's only law regarding employment. In the House, where the key legslative vote was held, opposition to the riders and support for affirmative action was led by the Republican minority leader, Gerald Ford. In December 1969, the riders were defeated, first in the House and then in the Senate. In the House, affirmative action was supported by Republicans by 3-1, while a majority of Democratic representatives voted in opposition. In early 1970, Labor Secretary Schutz expanded affirmative action to include all businesses, not just construction companies, that obtained federal contracts of more the $50,000 and who employed more that 50 employers to submit minority hiring plans with goals, targets and timetables.

Nixon's strategy introduced a new policy dimension that split black workers and white union-members, establishing a Federal program of race-based hiring that has never been revoked. While the Office of Contact Compliance policed the hiring practices of federal contractors, the Equal Employment Opportunity Commission and, ultimately, the Department of Justice policed the rest of the labor market. Large firms, in particular, found that the best defense against the possibility of a costly lawsuit was to implement affirmative action targets based on the share of minorities in the local population. Affirmative action was effective in raising the employment of minorities. Holzer and Neumark (2000) estimate that affirmative action plans lowered the hiring of white men by 10-15 percent in favor of minority men and women.

It is worth noting that firms did not appear to play an active role in lobbying either for or against affirmative action. We came across no instances in which firms used affirmative action to lower their wage costs by replacing white workers with cheaper black workers, although that doesn't mean it never happened. Initially, most firms defended their hiring practices and their existing work force against the demands of minorities for increased employment. But once affirmative action was in place, firms found it easy to live with. Employers realized the advantages of having one federal law to comply with rather than a multiplicity of state regulations. Most importantly, employers found affirmative action to have small effects on profits. As Holzer and Neumark (2000) summarize the economic literature on the effects of affirmative action programs on efficiency, there is evidence that affirmative action resulted in the hiring of minorities with lower educational credentials than their white coworkers, but much less evidence that minorities recruited in affirmative action programs were less productive. In sum, our neglect of employers as a political actor does not appear to be an important lack in the politics of affirmative action.

There are several fairly obvious ways it would be desirable to extend the framework suggested here, short of explicitly including racial preferences. Among these, three seem especially salient. First, as observed in the text, it is important to explore the implications of allowing more than two skill levels for workers. Second, a natural extension is to expand the set of policies considered to include education. Education, however, is inherently a dynamic problem. At any moment, 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.[16] And third, it is a commonplace in the contemporary political economy literature that details of legislative and party structures are important for understanding policy outcomes. Legislative bargaining is likely to differ in parliamentary system with proportional representation where the government typically consists of a coalitions of parties and party discipline is high. A general theory of the impact of introducing a second dimension of redistribution by race, religion or language to redistribution by income requires considering of the full range of political institutions that shape political conflict.

[16]See Roemer (2004) for an analysis of a model of the political choice of taxes, transfers and investment 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

For convenience, recall from (3) in the text that

$$\bar{y}(b) = \left[\frac{q}{1-\beta} + b\right]$$

is the *laissez faire* productivity threshold for a firm to create a good job. Also, for each $i = B, W$,

$$\begin{aligned} \sigma_i(\alpha, b) &= 1 - G_i\left(y_i(\alpha, b)\right) \\ &= 1 - \left[\left(1-\theta_i\right) F\left(y_i(\alpha, b)\right) + \theta_i F\left(y_i(\alpha, b) - h\right)\right] \end{aligned}$$

is the share of group i workers with good jobs; let $g_i(y) = ((1-\theta_i)\, f\,(y) + \theta_i f\,(y-h)$ be the density $dG_i(y)$. Recall that $\sigma(\alpha, b) = p\sigma_B(\alpha, b) + (1-p)\sigma_W(\alpha, b)$ is the share of the population with good jobs and, from equation (7),

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

is the average wage at (α, b). From the budget constraint (6),

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

is the budget-balancing tax rate. The *laissez faire* share of good jobs held by blacks at any benefit level b is

$$\alpha_\phi(b) = \frac{p\sigma_B(\phi, b)}{\sigma(\phi, b)} < p.$$

Because F has full support on $\mathbb{R}$, b finite implies $\alpha_\phi(b) > 0$. Hereafter, where there is no ambiguity we take the dependency of y_i, σ_i, etc on (α, b) as understood. Similarly, we occasionally write $E(w) \equiv E[w|\alpha, b]$ and $E(\pi) \equiv E[\pi|\alpha, b]$.

The following fact is used repeatedly in the arguments to follow. For any $x \in \mathbb{R}$, define the difference

$$Z(x) = \left[\frac{g_B\left(x\right)}{\left[1-G_B\left(x\right)\right]} - \frac{g_W\left(x\right)}{\left[1-G_W\left(x\right)\right]}\right];$$

then we have

Lemma 0 *For any $x \in \mathbb{R}$ and any $h > 0$, $Z(x) > 0$ iff F has a strictly increasing hazard rate at x.*

Proof Write

$$\begin{aligned} Z(x) &= \frac{g_B(x)}{[1 - G_B(x)]} - \frac{g_W(x)}{[1 - G_W(x)]} \\ &= \frac{g_B(x)[1 - G_W(x)] - g_W(x)[1 - G_B(x)]}{[1 - G_B(x)][1 - G_W(x)]}. \end{aligned}$$

Since the denominator is positive, the sign of $Z(x)$ is the same as the sign of the numerator. Expanding the numerator we obtain

$$\begin{aligned} & g_B(x)[1 - G_W(x)] - g_W(x)[1 - G_B(x)] \\ = \; & [\theta_B f(x-h) + (1-\theta_B) f(x)][1 - \theta_W F(x-h) - (1-\theta_W) F(x)] \\ & - [\theta_W f(x-h) + (1-\theta_W) f(x)][1 - \theta_B F(x-h) - (1-\theta_B) F(x)] \\ = \; & (\theta_W - \theta_B)[f(x)(1 - F(x-h)) - f(x-h)(1 - F(x))]. \end{aligned}$$

Since $(\theta_W - \theta_B) > 0$ we have

$$Z(x) > 0 \Leftrightarrow \frac{f(x)}{1 - F(x)} > \frac{f(x-h)}{1 - F(x-h)},$$

which proves the lemma.$\square$

Proof of Lemma 1 *(1) To show the existence and uniqueness of the labor market equilibrium.* It is immediate from a firm's profit-maximizing problem that, for any $b \geq 0$ and any $\alpha < \alpha_\phi(b)$, the solution involves setting $y_i(\alpha, b) = \bar{y}(b)$, $i = B, W$, with the share of good jobs allocated to blacks being $\alpha_\phi(b)$. Moreover, $\bar{y}(b)$ is finite for any finite level of benefit. Hence, if the affirmative action constraint is not binding, that is, if $\alpha \leq \alpha_\phi(b)$, then the labor market equilibrium clearly exists and is uniquely defined by $y_i(\alpha, b) = \bar{y}(b)$, $i = B, W$, and (9). So assume $\alpha > \alpha_\phi(b)$ and, without loss of generality, suppose b is finite.

Fix $(\alpha, b) \in \mathbf{P}$, $\alpha > \alpha_\phi(b)$ and b finite. Equilibrium in the labor market is characterized by a system of two equations that jointly determine y_B and y_W and a balanced budget constraint (9) that determines t. Given the affirmative action constraint binds, equations (5) and (4) in the text can be written, respectively, as

$$\begin{aligned} \alpha y_B + (1-\alpha) y_W - \bar{y}(b) &= 0 \\ p(1-\alpha)\sigma_B - \alpha(1-p)\sigma_W &= 0. \end{aligned} \tag{10}$$

Writing $Y(y_W) \equiv \frac{1}{\alpha}\left[\bar{y}(b) - (1-\alpha)y_W\right]$ and solving for y_B from the first equation of (10) yields $y_B = Y(y_W)$. Substituting for y_B into the second equation and collecting terms yields

$$\begin{aligned} \frac{p(1-\alpha)}{(1-p)\alpha} &= \frac{1-(1-\theta_W)F(y_W) - \theta_W F(y_W - h)}{1-(1-\theta_B)F(Y(y_W)) - \theta_B F(Y(y_W)-h)} \\ &= \frac{\sigma_W(y_W)}{\sigma_B(Y(y_W))}. \end{aligned}$$

Since the affirmative action constraint is binding, $\alpha > \alpha_\phi(b)$. By assumption, $\frac{1}{2} > p \geq \alpha$ and so $\alpha > \alpha_\phi(b) > 0$ implies

$$\frac{p(1-\alpha)}{(1-p)\alpha} \in [1, \frac{p(1-\alpha_\phi(b))}{(1-p)\alpha_\phi(b)}].$$

On the other hand, we have that for any $b \in [0, \infty)$,

$$\lim_{y_W \uparrow +\infty} \frac{\sigma_W(y_W)}{\sigma_B(Y(y_W))} = 0 < \lim_{y_W \downarrow -\infty} \frac{\sigma_W(y_W)}{\sigma_B(Y(y_W))} = \infty.$$

Moreover, since F is continuous and strictly increasing, $\sigma_W(y_W)/\sigma_B(Y(y_W))$ is a strictly decreasing function of y_W. Hence, there exists a unique y_W and $y_B = Y(y_W)$ solving (10). And given b, y_W and y_B, the tax rate is uniquely defined by (9). This proves the existence and uniqueness of a labor market equilibrium.

(2) To show $y_B(\alpha, b) < \bar{y}(b) < y_W(\alpha, b)$. Suppose $\alpha \in (\alpha_\phi(b), p]$. If $y_B(\alpha, b) \geq y_W(\alpha, b)$ then the first equation of (10) implies $y_W(\alpha, b) \leq \bar{y}(b) \leq y_B(\alpha, b)$ and so

$$\begin{aligned} \sigma_B(\alpha, b) &= 1 - (1-\theta_B) F(y_B(\alpha, b)) - \theta_B F(y_B(\alpha, b) - h) \\ &\leq 1 - (1-\theta_B) F(\bar{y}(b)) - \theta_B F(\bar{y}(b) - h) \end{aligned}$$

and

$$\begin{aligned}\sigma_W(\alpha, b) &= 1-(1-\theta_W)\,F\left(y_W(\alpha,b)\right)-\theta_W F\left(y_W(\alpha,b)-h\right)\\ &\geq 1-(1-\theta_W)\,F\left(\bar{y}(b)\right)-\theta_W F\left(\bar{y}(b)-h\right).\end{aligned}$$

Therefore

$$\frac{p\sigma_B(\alpha,b)}{\sigma(\alpha,b)} \leq \frac{p\sigma_B(\phi,b)}{\sigma(\phi,b)} = \alpha_\phi(b)$$

which contradicts $\alpha > \alpha_\phi(b)$. Hence $\alpha \in (0,1)$ and (10) imply $y_B(\alpha,b) < \bar{y}(b) < y_W(\alpha,b)$.

(3) To show $\alpha_\phi(b)$ is strictly decreasing in b. Because $\sigma_B(\phi,b) > 0$, we can write

$$\alpha_\phi(b) = \frac{p\sigma_B(\phi,b)}{\sigma(\phi,b)} = \frac{p}{p+(1-p)\left(\frac{\sigma_W(\phi,b)}{\sigma_B(\phi,b)}\right)}.$$

And

$$\begin{aligned}&sign\frac{d}{db}\left(\frac{\sigma_W(\phi,b)}{\sigma_B(\phi,b)}\right)\\ =\ & sign\left[\frac{1}{\sigma_W(\phi,b)}\frac{d\sigma_W(\phi,b)}{db}-\frac{1}{\sigma_B(\phi,b)}\frac{d\sigma_B(\phi,b)}{db}\right]\\ =\ & sign\left[\frac{g_B(\bar{y}(b))}{1-G_B(\bar{y}(b))}-\frac{g_W(\bar{y}(b))}{1-G_W(\bar{y}(b))}\right]\frac{d\bar{y}(b)}{db}.\end{aligned}$$

By assumption, F exhibits a strictly increasing hazard rate everywhere so Lemma 0 implies the term in square brackets is strictly positive. And since $d\bar{y}(b)/db = 1$, $\frac{d}{db}\left(\frac{\sigma_W(\phi,b)}{\sigma_B(\phi,b)}\right) > 0$. Hence, $\alpha_\phi(b)$ is strictly decreasing in b as required.□

Write the system of equations (10) as

$$\begin{pmatrix}\psi_1\left(y_B, y_W\right)\\ \psi_2\left(y_B, y_W\right)\end{pmatrix} = 0$$

and define Ψ as

$$\begin{aligned}\Psi &\equiv \begin{vmatrix}\partial\psi_1/\partial y_B & \partial\psi_1/\partial y_W\\ \partial\psi_2/\partial y_B & \partial\psi_2/\partial y_W\end{vmatrix}\\ &= \alpha^2(1-p)g_W\left(y_W\right)+(1-\alpha)^2 p g_B\left(y_B\right) > 0\end{aligned}$$

Proof of Lemma 2 *(1)* Write $\Delta \equiv (y_W - y_B)$; by Lemma 1, $\Delta \geq 0$. Differentiate (10) with respect to α to obtain

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

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

Since $\Delta \to 0$ as $\alpha \to \alpha_\phi(b)$, it is clear that for all $b \geq 0$,

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

and, for all $\alpha \in [\alpha_\phi(b), p]$,

$$\frac{\partial y_W(\alpha, b)}{\partial \alpha} > 0.$$

(2) Differentiating (10) with respect to b, one obtains

$$\frac{\partial y_B(\alpha, b)}{\partial b} = \frac{\alpha(1-p)g_W(y_W)}{\Psi} > 0 \tag{13}$$

$$\frac{\partial y_W(\alpha, b)}{\partial b} = \frac{(1-\alpha)pg_B(y_B)}{\Psi} > 0. \tag{14}$$

(3) Differentiating (9) with respect to $b \geq 0$ yields

$$\begin{aligned}
\frac{\partial t}{\partial b} &= \frac{1}{(1-\sigma)b + E(w)} \left[\frac{\partial\left((1-\sigma)b\right)}{\partial b} - t\frac{\partial\left((1-\sigma)b + E(w)\right)}{\partial b} \right] \\
&= \frac{1}{(1-\sigma)b + E(w)} \left[(1-t)\frac{\partial\left((1-\sigma)b\right)}{\partial b} - t\frac{\partial\left(E(w)\right)}{\partial b} \right] \\
&= \frac{1}{(1-\sigma)b + E(w)} \left\{ \left[(1-t)(1-\sigma) - t(1-\beta)\sigma\right] + A_1 \right\},
\end{aligned} \tag{15}$$

where

$$A_1 \equiv \left[(1-t)b + tw(y_B)\right] pg_B(y_B)\frac{\partial y_B}{\partial b} + \left[(1-t)b + tw(y_W)\right](1-p)g_W(y_W)\frac{\partial y_W}{\partial b}$$

and $w(y_i) = \beta y_i + (1-\beta)b$ is the wage of the marginal individual from group $i \in \{B, W\}$ with a good job. By *(2)* above, $A_1 \geq 0$ with equality if and only if $b = 0$. Subtracting $t(1-\beta)\sigma b$ from both sides of the balanced budget constraint (6) in the text, we obtain

$$\left[(1-t)(1-\sigma) - t(1-\beta)\sigma\right] b = \left[E(w) - (1-\beta)\sigma b\right] t$$

Since $E(w) - (1-\beta)\sigma b > 0$ the left-hand side of the equality must be positive for $b > 0$ and zero otherwise. Hence, $\partial t/\partial b > 0$ for $b > 0$. If $b = 0$, $\partial t/\partial b = (1-\sigma)/E(w) > 0$. Thus, $\partial t/\partial b > 0$ for $b \geq 0$ as required.

(4) In equilibrium, (1) and (2) imply

$$E[w+\pi|\alpha,b] = p\int_{y_B}(y-q)\,dG_B(y) + (1-p)\int_{y_W}(y-q)\,dG_W(y).$$

Taking the derivative with respect to α yields

$$\frac{\partial E[w+\pi|\alpha,b]}{\partial\alpha} = -\left\{p(y_B-q)g_B(y_B)\frac{\partial y_B}{\partial\alpha} + (1-p)(y_W-q)g_W(y_W)\frac{\partial y_W}{\partial\alpha}\right\}.$$

Taking limits then gives

$$\begin{aligned}\lim_{\alpha\downarrow\alpha_\phi(b)}\frac{\partial E[w+\pi|\alpha,b]}{\partial\alpha} &= -(\bar{y}-q)\left[pg_B(\bar{y})\frac{\partial y_B}{\partial\alpha} + (1-p)g_W(\bar{y})\frac{\partial y_W}{\partial\alpha}\right]_{\alpha=\alpha_\phi(b)} \\ &= (\bar{y}-q)\frac{\partial\sigma(\alpha_\phi(b),b)}{\partial\alpha}. \qquad (16)\end{aligned}$$

Since $\bar{y} > q$, expected income is increasing at $\alpha_\phi(b)$ if and only if the share of good jobs in the economy as a whole is increasing at $\alpha_\phi(b)$. So consider the effect of α on σ at $\alpha_\phi(b)$:

$$\frac{\partial\sigma(\alpha,b)}{\partial\alpha} = -\left[pg_B(y_B)\frac{\partial y_B(\alpha,b)}{\partial\alpha} + (1-p)g_w(y_W)\frac{\partial y_W(\alpha,b)}{\partial\alpha}\right]. \qquad (17)$$

Substituting for $\partial y_i/\partial\alpha$, $i = B, W$, and taking limits gives

$$\begin{aligned}\lim_{\alpha\downarrow\alpha_\phi(b)}\frac{\partial\sigma(\alpha_\phi(b),b)}{\partial\alpha} &= \left[p(1-\alpha_\phi(b))g_B(\bar{y}) - (1-p)\alpha_\phi(b)g_W(\bar{y})\right]\frac{\sigma}{\Psi} \\ &= \left[(1-G_W(\bar{y}))g_B(\bar{y}) - (1-G_B(\bar{y}))g_W(\bar{y})\right]\frac{p(1-p)}{\Psi} \\ &= \frac{p(1-p)(1-G_W(\bar{y}))(1-G_B(\bar{y}))}{\Psi}Z(\bar{y}) \\ &> 0, \qquad (18)\end{aligned}$$

where the second equality follows on substituting for $\alpha_\phi(b) = p\sigma_B(\phi,b)/\sigma(\phi,b)$ and collecting terms; and the inequality follows from Lemma 0 and the assumption that F has a strictly increasing hazard rate. Thus, $\lim_{\alpha\downarrow\alpha_\phi(b)}\partial E[w+\pi|\alpha,b]/\partial\alpha > 0$ as claimed.

(5) Using (9), we have for all $b \geq 0$,

$$\frac{\partial t}{\partial \alpha} = -\frac{b}{\left[(1-\sigma)b+E(w)\right]^2}\left[E(w)\frac{\partial \sigma}{\partial \alpha}+(1-\sigma)\frac{\partial E\left(w\right)}{\partial \alpha}\right]. \tag{19}$$

By definition of $\alpha_\phi(b)$, $\partial E[\pi|\alpha_\phi(b), b]/\partial\alpha = 0$. Hence,

$$\left.\frac{\partial E[w|\alpha,b]}{\partial \alpha}\right|_{\alpha=\alpha_\phi(b)} = \left.\frac{\partial E[w+\pi|\alpha,b]}{\partial \alpha}\right|_{\alpha=\alpha_\phi(b)}.$$

Taking limits, substituting from (16) and collecting terms yields,

$$\lim_{\alpha\downarrow\alpha_\phi(b)}\frac{\partial t}{\partial \alpha} = -\frac{\left[(1-\sigma)\left(\bar{y}-q\right)+E(w)\right]b}{\left[(1-\sigma)b+E(w)\right]^2}\frac{\partial \sigma(\alpha_\phi(b),b)}{\partial \alpha}. \tag{20}$$

Therefore, by (18), $\lim_{\alpha\downarrow\alpha_\phi(b)}\partial t/\partial\alpha \leq 0$ with inequality strict for all $b > 0$.□

The expected consumption of an individual with human capital $H \in \{0, h\}$ in group $i \in \{B, W\}$ is given by equation (8) in the text; substituting for w, this expression can be rewritten:

$$c_{Hi}(\alpha,b) = (1-t)\left\{\left[1-\beta\left(1-F(y_i-H)\right)\right]b+\beta\int_{y_i-H}^{\infty}(x+H)\,dF(x)\right\} \tag{21}$$

Proof of Lemma 3 Differentiating (21) with respect to α and b yields,

$$\frac{\partial c_{Hi}\left(\alpha,b\right)}{\partial \alpha} = -\left(\frac{c_{Hi}}{1-t}\right)\frac{\partial t}{\partial \alpha}-(1-t)\beta\left(y_i-b\right)f\left(y_i-H\right)\frac{\partial y_i}{\partial \alpha}. \tag{22}$$

$$\begin{aligned}\frac{\partial c_{Hi}}{\partial b} &= -\left(\frac{c_{Hi}}{1-t}\right)\frac{\partial t}{\partial b}+(1-t)\left[1-\beta\left(1-F(y_i-H)\right)\right] \\ &\quad -(1-t)\beta\left(y_i-b\right)f\left(y_i-H\right)\frac{\partial y_i}{\partial b}.\end{aligned} \tag{23}$$

The first two claims of the lemma follow from steps *(1)* through *(5)* below.

(1) To show $\alpha^*_{HB} > \alpha_\phi(b^*_{HB})$, $H \in \{0, h\}$. Set $i = B$ in (22) and fix $b \geq 0$. By Lemma 2(1), $\lim_{\alpha\downarrow\alpha_\phi(b)}\partial y_B/\partial\alpha < 0$ and, by Lemma 2(5), $\lim_{\alpha\downarrow\alpha_\phi(b)}\partial t/\partial\alpha \leq 0$. Further, $\lim_{\alpha\downarrow\alpha_\phi(b)}(y_B - b) = (\bar{y} - b) > 0$. Hence, for all $b \geq 0$, $\lim_{\alpha\downarrow\alpha_\phi(b)}\partial c_{HB}\left(\alpha, b\right)/\partial\alpha > 0$ which suffices to show $\alpha^*_{HB} > \alpha_\phi(b^*_{HB})$, $H \in \{0, h\}$.

(2) To show there exists finite $\bar{h}_{1i}$ and $\bar{q}_1$ such that $h \geq \bar{h}_{1i}$ and $q \geq \bar{q}_1$ implies $b^*_{hi} = 0$, $i \in \{B, W\}$. To prove $b^*_{hi} = 0$, note that (15) implies,

$$\lim_{b \to 0} \frac{1}{1 - t(\alpha_\phi(b), b)} \frac{\partial t(\alpha_\phi(b), b)}{\partial b} = \left.\frac{1 - \sigma}{E(w)}\right|_{(\alpha_\phi(0), 0)} \tag{24}$$

in which case

$$\begin{aligned} & \lim_{b \to 0} \frac{\partial c_{hi}(\alpha_\phi(b), b)}{\partial b} \\ = & -\left[c_{hi} \frac{1 - \sigma}{E(w)} - [1 - \beta(1 - F(\bar{y} - h))] + \beta \bar{y} f(\bar{y} - h) \frac{\partial y_i}{\partial b}\right]_{(\alpha_\phi(0), 0)}. \end{aligned}$$

By Lemma 2(2), the last term inside square brackets is nonnegative for all h. So, fixing the policy $(\alpha_\phi(0), 0)$, define

$$\begin{aligned} \Phi(h) &= -c_{hi} \frac{1 - \sigma}{E(w)} + 1 - \beta(1 - F(\bar{y} - h)) \\ &= -\frac{[\theta F(\bar{y} - h) + (1 - \theta) F(\bar{y})] \int_{\bar{y} - h} (x + h)\, dF(x)}{\theta \int_{\bar{y} - h} (x + h)\, dF(x) + (1 - \theta) \int_{\bar{y}} x\, dF(x)} + 1 - \beta(1 - F(\bar{y} - h)) \end{aligned}$$

which is independent of i. Since (given $b = 0$) $\bar{y} = q/(1 - \beta)$ is constant and $\lim_{h \to \infty} F(\bar{y} - h) = 0$,

$$\begin{aligned} \lim_{h \to \infty} \left(c_{hi} \frac{1 - \sigma}{E(w)}\right) &= \lim_{h \to \infty} \frac{(1 - \theta) F(\bar{y}) \left[\int_{\bar{y} - h}^\infty x f(x) dx + h \int_{\bar{y} - h}^\infty f(x) dx\right]}{(1 - \theta) \int_{\bar{y}} x\, dF(x) + \theta \left[\int_{\bar{y} - h}^\infty x f(x) dx + h \int_{\bar{y} - h}^\infty f(x) dx\right]} \\ &= \frac{(1 - \theta) F(\bar{y})}{\theta}. \end{aligned}$$

Hence

$$\begin{aligned} \lim_{h \to \infty} \Phi(h) &= [1 - \beta] - \frac{(1 - \theta) F(\bar{y})}{\theta} \\ &< 0 \Leftrightarrow F(\bar{y}) > \frac{\theta}{(1 - \theta)} (1 - \beta). \end{aligned}$$

Therefore, because $\lim_{q \to \infty} F(\bar{y}) = 1$ and $\theta < 1/2$, there exists $\bar{h}_{1i} > 0$ and $\bar{q}_1$ such that $h \geq \bar{h}_{1i}$ and $q \geq \bar{q}_1$ implies $\lim_{b \to 0} \partial c_{hi}(\alpha_\phi(b), b)/\partial b \leq 0$, $i = B, W$, in which case $b^*_{hi} = 0$, $i = B, W$.

(3) To show there exists finite $\bar{h}_{2i}$ and $\bar{q}_{2i}$ such that $h \geq \bar{h}_{2i}$ and $q \geq \bar{q}_{2i}$ implies $b^*_{0i} > 0$, $i \in \{B, W\}$. Evaluating (23) for $i = W$ at $(\alpha_\phi(0), 0)$ yields

$$\begin{aligned}\lim_{h\to\infty} \frac{\partial c_{0W}\left(\alpha_\phi(0),0\right)}{\partial b} &= \lim_{h\to\infty}\left[-c_{0W}\frac{1-\sigma}{E(w)} + \left[1-\beta\left(1-F(\bar{y})\right)\right] - \beta\bar{y}f\left(\bar{y}\right)\frac{\partial y_W\left(\alpha_\phi(0),0\right)}{\partial b}\right] \\ &= \lim_{h\to\infty}\left[-K(h) + \left[1-\beta\left(1-F(\bar{y})\right)\right] - \beta\bar{y}f\left(\bar{y}\right)\frac{p(1-\alpha_\phi(0))g_W\left(\bar{y}\right)}{\Psi}\right],\end{aligned}$$

where

$$K(h) \equiv \frac{\left[\theta F\left(\bar{y}-h\right) + (1-\theta)F\left(\bar{y}\right)\right]\int_{\bar{y}}^{\infty} xf(x)dx}{(1-\theta)\int_{\bar{y}} x\, dF(x) + \theta\left[\int_{\bar{y}-h}^{\infty} xf(x)dx + h\int_{\bar{y}-h}^{\infty} f(x)dx\right]},$$

and we have substituted for $\partial y_W/\partial b$ from (14) and for $(\partial t/\partial b)/(1-t)$ from (24). Therefore

$$\lim_{h\to\infty} \frac{\partial c_{0W}\left(\alpha_\phi(0),0\right)}{\partial b} = \left[1-\beta\left(1-F(\bar{y})\right)\right] - \beta\bar{y}f\left(\bar{y}\right)\frac{p(1-\alpha_\phi(0))(1-\theta_W)f(\bar{y})}{\Psi}$$

Hence

$$\lim_{h\to\infty} \frac{\partial c_{0W}\left(\alpha_\phi(0),0\right)}{\partial b} > 0 \Leftrightarrow F(\bar{y}) > \frac{1}{\beta\Psi}\left[\beta\bar{y}f\left(\bar{y}\right)^2 p(1-\alpha_\phi)(1-\theta_W) - (1-\beta)\Psi\right]$$

Now $\lim_{q\to\infty} F(\bar{y}) = 1$ and, by L'Hôpital's Rule and the assumption of that F has a strictly increasing hazard rate, $\lim_{q\to\infty} \bar{y}f(\bar{y})^2 = 0$.[17] Hence the right-hand side of the last inequality

[17] To see this, write

$$\bar{y}f(\bar{y})^2 = \frac{\bar{y}}{1/f(\bar{y})^2}.$$

Both numerator and denominator go to infinity with q or, equivalently, with $\bar{y}$ and L'Hôpital's Rule implies

$$\lim_{\bar{y}\to\infty} \bar{y}f(\bar{y})^2 = \lim_{\bar{y}\to\infty} -\frac{f(\bar{y})^2}{2f'(\bar{y})}$$

if this latter limit exists. Since F has a strictly increasing hazard rate, at every $x > 0$ we have

$$\frac{d}{dx}\frac{f(x)}{1-F(x)} = \frac{(1-F(x))f'(x) + f(x)^2}{\left[1-F(x)\right]^2} > 0$$

so that, for all $x > 0$,

$$1 - F(x) > -\frac{f(x)^2}{f'(x)}.$$

And since $\lim_{x\to\infty}[1-F(x)] = 0$ and $f'(x) \leq 0$ for sufficiently large x,

$$\lim_{x\to\infty} -\frac{f(x)^2}{f'(x)} = 0.$$

is negative for q sufficiently large. Therefore, $\lim_{q\to\infty} F(\bar{y}) = 1$ implies there exists $\bar{h}_{2W} > 0$ and $\bar{q}_{2W}$ such that $h \geq \bar{h}_{2W}$ and $q \geq \bar{q}_{2W}$ implies $\lim_{b\to 0} \partial c_{0W}(\alpha_\phi(b), b)/\partial b > 0$. The same argument for $i = B$ likewise yields that there exists $\bar{h}_{2B} > 0$ and $\bar{q}_{2B}$ such that $h \geq \bar{h}_{2B}$ and $q \geq \bar{q}_{2B}$ implies $\lim_{b\to 0} \partial c_{0B}(\alpha_\phi(b), b)/\partial b > 0$. In each case, therefore, we obtain $b^*_{0i} > 0$, $i \in \{B, W\}$.

(4) To show $\alpha^*_{hW} = \alpha_\phi(b^*_{hW})$. By step *(2)*, $b^*_{hW} = 0$ and, by Lemma 2(5), we have $\lim_{b\to 0} \partial t(\alpha_\phi(b), b)/\partial\alpha = 0$. Finally, by Lemma 2(1), $\partial y_W/\partial\alpha > 0$. Hence, (22) implies $\lim_{b\to 0} \partial c_{hW}(\alpha_\phi(b), b)/\partial\alpha < 0$, proving $\alpha^*_{hW} = \alpha_\phi(0)$ as required.

(5) To show there exists finite $\bar{h}_3$ *and* $\bar{q}_3$ *such that* $h \geq \bar{h}_3$ *and* $q \geq \bar{q}_3$ *implies* $\alpha^*_{0W} = \alpha_\phi(b^*_{0W})$. Using (22), we have that for any $b \geq 0$,

$$\frac{\partial c_{0W}(\alpha, b)}{\partial\alpha} = -\left(\frac{c_{0W}}{1-t}\right)\frac{\partial t\,(\alpha, b)}{\partial\alpha} - (1-t)\beta\,(\bar{y} - b)\, f\,(\bar{y})\,\frac{\partial y_W\,(\alpha, b)}{\partial\alpha}$$

When evaluated at $\alpha = \alpha_\phi(b)$, we get

$$\begin{aligned}
&\lim_{\alpha\downarrow\alpha_\phi(b)} \frac{\partial c_{0W}\,(\alpha, b)}{\partial\alpha} \\
= \quad & -\left(\frac{c_{0W}}{1-t}\right)\frac{\partial t\,(\alpha_\phi(b), b)}{\partial\alpha} - (1-t)\left(\frac{\beta q}{1-\beta}\right) f\,(\bar{y})\,\frac{p\,[1 - G_B\,(\bar{y})]}{\Psi} \\
= \quad & \left\{\frac{c_{0W}}{(1-t)}\left[\frac{t\,[(\beta q/(1-\beta)) + b] + (1-t)b}{[(1-\sigma)b + E(w)]}\right]\frac{\partial\sigma\,(\alpha_\phi(b), b)}{\partial\alpha}\right. \\
& \left. -(1-t)\left(\frac{\beta q}{1-\beta}\right)[1 - G_B\,(\bar{y})]\, f\,(\bar{y})\,\frac{p}{\Psi}\right\} \\
= \quad & \left\{\frac{c_{0W}}{(1-t)}\frac{[(t\beta q/(1-\beta)) + b]\,(1-p)}{[(1-\sigma)b + E(w)]}\left[(1 - G_W\,(\bar{y}))\left(\theta_B\frac{f\,(\bar{y} - h)}{f\,(\bar{y})} + (1-\theta_B)\right)\right.\right. \\
& \left. - (1 - G_B\,(\bar{y}))\left(\theta_W\frac{f\,(\bar{y} - h)}{f\,(\bar{y})} + (1-\theta_W)\right)\right] \\
& \left. -(1-t)\left(\frac{\beta q}{1-\beta}\right)[1 - G_B\,(\bar{y})]\right\}\frac{p}{\Psi} f\,(\bar{y}) \\
= \quad & \left\{\frac{c_{0W}}{1-t}\frac{[(t\beta/(1-\beta)) + (b/q)]\,(1-p)}{[(1-\sigma)b + E(w)]}\left[\left(\frac{1 - G_W\,(\bar{y})}{1 - G_B\,(\bar{y})}\right)\left(\theta_B\frac{f\,(\bar{y} - h)}{f\,(\bar{y})} + (1-\theta_B)\right)\right.\right. \\
& \left.\left. - \left(\theta_W\frac{f\,(\bar{y} - h)}{f\,(\bar{y})} + (1-\theta_W)\right)\right] - (1-t)\left(\frac{\beta}{1-\beta}\right)\right\}\frac{p\,(1 - G_B\,(\bar{y}))}{\Psi} q f\,(\bar{y})
\end{aligned}$$

where we have substituted for $\partial y_W(\alpha, b)/\partial\alpha$, $\partial t(\alpha_\phi(b), b)/\partial\alpha$ and $\partial\sigma(\alpha_\phi(b), b)/\partial\alpha$ from (12), (20) and (18), respectively, and collected terms. All of the terms outside the curly brackets are positive. Therefore,

$$\text{sign}\,\frac{\partial c_{0W}(\alpha_\phi(b), b)}{\partial\alpha} = \text{sign}\left[K_1 \frac{f(\bar{y}-h)}{f(\bar{y})} + K_2\right]$$

where

$$\begin{aligned} K_1 &= \frac{c_{0W}}{1-t}\frac{[(t\beta/(1-\beta)) + (b/q)](1-p)}{[(1-\sigma)b + E(w)]}\left[\left(\frac{1-G_W(\bar{y})}{1-G_B(\bar{y})}\right)\theta_B - \theta_W\right] \\ K_2 &= \frac{c_{0W}}{1-t}\frac{[(t\beta/(1-\beta)) + (b/q)](1-p)}{[(1-\sigma)b + E(w)]}(\theta_w - \theta_B) - (1-t)\left(\frac{\beta}{1-\beta}\right) \end{aligned}$$

Consider the term

$$\begin{aligned} &\left[\left(\frac{1-G_W(\bar{y})}{1-G_B(\bar{y})}\right)\theta_B - \theta_W\right] = \frac{\theta_B(1-G_W(\bar{y})) - \theta_W(1-G_B(\bar{y}))}{(1-G_B(\bar{y}))} \\ = &\ [\theta_B[1-\theta_W F(\bar{y}-h) - (1-\theta_W)F(\bar{y})] \\ &- \theta_W[1-\theta_B F(\bar{y}-h) - (1-\theta_B)F(\bar{y})]]\frac{1}{(1-G_B(\bar{y}))} \\ = &\ \frac{(\theta_B-\theta_W) - (\theta_B-\theta_W)F(\bar{y})}{(1-G_B(\bar{y}))} \\ = &\ \frac{(\theta_B-\theta_W)\,(1-F(\bar{y}))}{(1-G_B(\bar{y}))} < 0 \end{aligned}$$

Therefore $K_1 < 0$.

Fixing b and letting $q \to \infty$ (so $\bar{y} \to \infty$) and $h \to \infty$ in such a way to insure $\bar{y} - h$ is constant, we obtain

$$\begin{aligned} &\text{sign}\lim_{\substack{q,h\to\infty \\ \bar{y}-h \text{ constant}}} \frac{\partial c_{0W}(\alpha_\phi(b), b)}{\partial\alpha} = \text{sign}\lim_{\substack{q,h\to\infty \\ \bar{y}-h \text{ constant}}}\left[K_1\frac{f(\bar{y}-h)}{f(\bar{y})} + K_2\right] \\ = &\ \text{sign}\left[\frac{(1-p)tb}{[(1-\sigma)b+E(w)]}\frac{(1-F(\bar{y}))}{f(\bar{y})}\frac{f(\bar{y}-h)}{(1-F(\bar{y}-h))}\frac{\theta_B-\theta_W}{\theta_B}\right. \\ &\left. + \frac{(1-p)tb}{[(1-\sigma)b+E(w)]}(\theta_W-\theta_B) - (1-t)\right]\frac{\beta}{1-\beta} \end{aligned}$$

where we observe $\lim_{q\to\infty} c_{0W} = (1-t)b$ and, given $\bar{y}-h$ constant, $\lim_{q\to\infty}(1-G_B(\bar{y})) = \theta_B(1-F(\bar{y}-h))$.

By the argument above for $K_1 < 0$, the assumption that F has a strictly increasing hazard rate and $\bar{y} - h$ constant, $\lim K_1 \leq 0$. Hence the first term in square brackets is nonpositive, in which case

$$\lim_{\substack{q,h\to\infty \\ \bar{y}-h \text{ constant}}} \frac{\partial c_{0W}\left(\alpha_\phi(b), b\right)}{\partial \alpha} < 0$$

if

$$\left[\frac{(1-p)tb}{(1-\sigma)b + E(w)}\left(\theta_W - \theta_B\right) - (1-t)\right] < 0.$$

Using (6), this inequality can be written

$$\begin{aligned}
\left[\frac{(1-p)(1-\sigma)b^2}{(1-\sigma)b + E(w)}\left(\theta_W - \theta_B\right) - E(w)\right] &< 0 \Leftrightarrow \\
(1-p)(1-\sigma)b^2\left(\theta_W - \theta_B\right) &< E(w)\left[(1-\sigma)b + E(w)\right] \Leftrightarrow \\
(1-p)\left(\theta_W - \theta_B\right) &< \frac{E(w)}{b}\frac{\left[(1-\sigma)b + E(w)\right]}{(1-\sigma)b}.
\end{aligned}$$

The left hand side of the last inequality must be less than $1/2$ since $(1-p)\theta_W < 1/2$ and $\theta_B \geq 0$; and the right hand side of this inequality must be greater that one since $b \leq E(w)$, benefits must be paid out of wages, and $b = b^*_{0W} > 0$. Therefore, there is some $\bar{q}_3 < \infty$ and some $\bar{h}_3 < \infty$ such that $q > \bar{q}_3$ and $h > \bar{h}_3$ implies

$$\frac{\partial c_{0W}\left(\alpha_\phi\left(b^*_{0W}\right), b^*_{0W}\right)}{\partial \alpha} < 0.$$

The claim follows.

To complete the argument for the first two claims of the lemma, let $\bar{h} = \max\{\bar{h}_{1i}\, \bar{h}_{2i}, \bar{h}_3\}_{i\in\{B,W\}}$ and $\bar{q} = \max\{\bar{q}_1, \bar{q}_{2i}, \bar{q}_3\}_{i\in\{B,W\}}$.

(6) To show profits are maximal at $(\alpha_\phi(0), 0)$. Immediate from (2) and definition of $\alpha_\phi(0)$ as the outcome of profit-maximizing choice of the threshold $\bar{y}(0)$.□

The argument for Proposition 1 uses the following claim. Let $\alpha^*_{Hi}(b)$ denote (H,i)'s most preferred level of affirmative action conditional on b.

Lemma 4 *there exists finite* $\bar{h}_4$ *and* $\bar{q}_4$ *such that* $h \geq \bar{h}_4$ *and* $q \geq \bar{q}_4$ *implies* $\alpha^*_{hW}(b^*_{0W}) = \alpha_\phi(b^*_{0W})$.

Proof Consider the high type whites' choice of α^*_{hW} at the point $\alpha_\phi(b^*_{0W})$:

$$\frac{\partial c_{hW}(\alpha_\phi(b^*_{0W}), b^*_{0W})}{\partial \alpha} = -\left(\frac{c_{hW}}{1-t}\right)\frac{\partial t(\alpha_\phi(b^*_{0W}), b^*_{0W})}{\partial \alpha} - (1-t)\beta(\bar{y} - b^*_{0W}) f(\bar{y} - h)\frac{\partial y_W}{\partial \alpha}$$

Evaluating $\partial t(\alpha_\phi(b^*_{0W}), b^*_{0W})/\partial\alpha$ yields

$$\frac{\partial t(\alpha_\phi(b^*_{0W}), b^*_{0W})}{\partial \alpha} = \frac{[(1-\sigma)(\bar{y}-q) + E(w)]\, b^*_{0W}}{\left[(1-\sigma)b^*_{0W} + E(w)\right]^2}\frac{\partial \sigma(\alpha_\phi(b^*_{0W}), b^*_{0W})}{\partial \alpha}$$

Finally, evaluating $\partial\sigma(\alpha_\phi(b^*_{0W}), b^*_{0W})/\partial\alpha$ when $q \to \infty$ and $h \to \infty$ to insure $(\bar{y} - h)$ is constant, yields

$$\begin{aligned}
\lim_{q\to\infty}\frac{\partial \sigma(\alpha_\phi(b^*_{0W}), b^*_{0W})}{\partial \alpha} &= \frac{p(1-p)}{\Psi}\left[(1 - G_W(\bar{y}))\, g_B(\bar{y}) - (1 - G_B(\bar{y}))\, g_W(\bar{y})\right] \\
&= \frac{p(1-p)\theta_W\theta_B}{\Psi}\left[(1 - F(\bar{y}-h))\, f(\bar{y}-h) - (1 - F(\bar{y}-h))\, f(\bar{y}-h)\right] \\
&= 0.
\end{aligned}$$

Since, by L'Hôpital's rule,

$$\lim_{\bar{y}\to\infty}\bar{y}\frac{\partial\sigma}{\partial\alpha} = \lim_{q\to\infty} q\frac{\partial\sigma}{\partial\alpha} = 0$$

we have

$$\lim_{q\to\infty}\frac{\partial t(\alpha_\phi(b^*_{0W}), b^*_{0W})}{\partial \alpha} = 0.$$

And therefore $\partial y_W/\partial\alpha > 0$ implies

$$\lim_{q\to\infty}\frac{\partial c_{hW}(\alpha_\phi(b^*_{0W}), b^*_{0W})}{\partial \alpha} < 0.$$

Hence, $\partial c_{hW}(\alpha_\phi(b^*_{0W}), b^*_{0W}) < 0$ for $q > \bar{q}_4$ and $h > \bar{h}_4$.□

Without loss of generality, assume $\bar{h} \geq \bar{h}_4$ and $\bar{q} \geq \bar{q}_4$.

Proof of Proposition 1 If affirmative action is not a political decision variable, the share of good jobs held by blacks in the economy is identically $\alpha_\phi(b)$ for all $b \geq 0$. In this case, the thresholds y_i satisfy $y_i \equiv \bar{y}(b)$, $i = B, W$, and all low human capital types have identical preferences over fiscal policy. So, because the coalition of all low types is a strict majority, $(1-\theta) > 1/2$, the majority most preferred fiscal policy when affirmative action is not a decision variable, say b_0^*, solves

$$\left.\frac{\partial c_{0i}(\alpha_\phi(b_0^*), b_0^*)}{\partial b}\right|_{\alpha \equiv \alpha_\phi(b), \forall b \geq 0} = 0, \ i \in \{B, W\}.$$

By similar reasoning to that for Lemma 3[18], $b_0^* > 0$; and that b_0^* is finite with $t(\alpha_\phi(b_0^*), b_0^*) < 1$ follows directly from (23) and $\partial t/\partial b > 0$ all $b \geq 0$. By Lemma 3 and the maintained assumptions, $\alpha_{\mathcal{L}}^* = \alpha_\phi(b_{\mathcal{L}}^*)$ for $b_{\mathcal{L}}^* = b_{0W}^* > 0$, in which case $b_0^* = b_{\mathcal{L}}^*$. Hence the core outcome when affirmative action is not a policy variable coincides with the low white party's ideal point when affirmative action is a policy choice. Now let affirmative action be a decision variable.

By Lemma 3 and Lemma 4, for sufficiently high q and h, $b_{hi}^* = 0$, $i = B, W$, and

$$\alpha_{\mathcal{H}}^*(b_{0W}^*) = \alpha_{\mathcal{L}}^*(b_{0W}^*) = \alpha_\phi(b_{0W}^*) < \alpha_{\mathcal{B}}^*(b_{0W}^*),$$

where $\alpha_j^*(b)$ is party j's most preferred level of affirmative action conditional on b. Therefore, if $b_{0W}^* \geq b_{\mathcal{B}}^*$ then surely $Eb_j < b_{\mathcal{L}}^* = b_0^*$, where (α_j, b_j) is the policy proposed by $j = \mathcal{H}, \mathcal{L}, \mathcal{B}$ in equilibrium and the expectation is with respect to the party first selected to propose a policy. Suppose $b_{0W}^* < b_{\mathcal{B}}^*$. Because $\nabla c_{\mathcal{H}}(\alpha_\phi(b), b) < 0$ for all $b \leq b_{\mathcal{L}}^*$, $\alpha_{\mathcal{L}}^* < \alpha_{\mathcal{B}}^*(b_{\mathcal{L}}^*)$ and $0 < b_{\mathcal{L}}^* < b_{\mathcal{B}}^*$ imply $c_{\mathcal{H}}(\alpha_{\mathcal{L}}^*, b_{\mathcal{L}}^*) > c_{\mathcal{H}}(\alpha_{\mathcal{B}}^*, b_{\mathcal{B}}^*)$. Moreover, 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 accepting all proposals such that $c_{hW}(\alpha, b) \geq c_{hW}(\alpha_{\mathcal{L}}^*, b_{\mathcal{L}}^*)$ (a proposal low whites would certainly make). If blacks proposed a policy that entailed lower utility for both the high and the low white party than the low whites' ideal point, $I_{\mathcal{L}} = (\alpha_{\mathcal{L}}^*, b_{\mathcal{L}}^*)$, both white parties would reject the black party's proposal and wait until $I_{\mathcal{L}} = (\alpha_{\mathcal{L}}^*, b_{\mathcal{L}}^*)$ or better was proposed. Hence, the high white party's equilibrium continuation value in the bargaining

[18] Specifically, step *(3)* of the proof above.

game, $v_{\mathcal{H}}$, must be at least $c_{\mathcal{H}}(\alpha^*_{\mathcal{L}}, b^*_{\mathcal{L}})$, in which case there is some $b' \leq b^*_{\mathcal{L}}$ such that $v_{\mathcal{H}} = c_{\mathcal{H}}(\alpha_\phi(b'), b') \geq c_{\mathcal{H}}(\alpha^*_{\mathcal{L}}, b^*_{\mathcal{L}})$. Therefore, if $c_{\mathcal{H}}(\alpha, b) = c_{\mathcal{H}}(\alpha_\phi(b'), b')$, then $\nabla c_{\mathcal{H}}(\alpha_\phi(b), b) < 0$ for all $b < b^*_{\mathcal{L}}$ implies $\alpha > \alpha_\phi(b)$; if $b = b^*_{\mathcal{L}}$, then we have $\alpha = \alpha_\phi(b)$; and, finally, the party's indifference curve $c_{\mathcal{H}}(\alpha, b) = v_{\mathcal{H}}$ is downward sloping in (b, α)-space. Consequently, because $v_{\mathcal{H}}$ is the high whites expected equilibrium payoff, it must be that $Eb_j < b^*_{\mathcal{L}}$ as required.$\Box$

Proof of Proposition 2 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. But then an equilibrium could not exist by the same reasoning as in the proof of Proposition 1. Part (2) is an immediate corollary of Proposition 1. 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.$\Box$

Proof of Propositions 3 and 4 Both of these results follow from the continuity result on the equilibrium correspondence proved in Banks and Duggan (2005, Theorem 3) and the convergence of interests between low-type whites and blacks, who jointly constitute a majority of the population.$\Box$

Acknowledgments

This paper is a revised version of an earlier working paper, "Redistribution in a Divided Society". We are grateful to the editor and two anonymous referees for very constructive comments and, among others, to seminar participants at CEACS of the Juan March Foundation, Madrid, at UCLA, and at the PIEP Political Economy workshop, Harvard University. We are also grateful to the NSF (SES-0212049) and the MacArthur Foundation for financial support. All responsibility for the content of the paper resides with the authors.

Michael Wallerstein died from brain cancer in January 2006. I am truly grateful for his friendship over many years and for all that I learned from him during that time.

References

Acemoglu, D. 2001. Good Jobs versus Bad Jobs. Journal of Labor Economics, 19: 1-21.

Acemoglu, D. and R. Shimer. 1999. Holdups and Efficiency with Search Frictions. International Economic Review, 40: 827-49.

Agell, J. and K. E. Lommerud. 1997. Minimum Wages and the Incentives for Skill Formation. Journal of Public Economics, 64: 25-40.

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

Alesina, A. and E. Glaeser. 2004. Fighting Poverty in the US and Europe. Oxford: Oxford University Press.

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

Anderson, T. 2004. The Pursuit of Fairness: A History of Affirmative Actions. Oxford: Oxford University Press.

Banks, J. S., and J. Duggan 2005, A General Bargaining Model of Legislative Policy-Making. Forthcoming in the Quarterly Journal of Political Science.

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.

Chan, J. and E. Eyster. 2002. Admission Impossible: Self-interest and Affirmative Action. Working Paper, The Johns Hopkins University.

Chung, K-S 2000, Role Models and Arguments for Affirmative Action. American Economic Review, 99, 640-48.

Coate, S. and G. Loury 1993, Will Affirmative Action Policies Eliminate Negative Stereotypes? American Economic Review, 83, 1220-40.

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

Foster, D. and R. Vohra 1992, An Economic Argument for Affirmative Action. Rationality and Society, 4(2), 176-188.

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

Grossman, G. and E. Helpman. 2001. Special Interest Politics. Cambridge: MIT Press.

Grout, P. 1984. Investment and Wages in the Absence of Binding Contracts: A Nash Bargaining Approach. Econometrica, 52: 449-60.

Holzer, H. and D. Neumark. 2000. Assessing Affirmative Action. Journal of Economic Literature, 38: 483-568.

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.

Levy, G. 2004. Public Education for the Minority, Private Education for the Majority. Working Paper, London School of Economics.

Lundberg, S. J. 1991. The enforcement of equal opportunity laws under imperfect information: affirmative action and alternatives. Quarterly Journal of Economics, 106(1), 309–326.

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

MacLeod, W. B. and J. Malcomson. 1993. Investments, Holdups, and the Form of Market Contracts. American Economic Review, 83: 811-837

Meltzer, A. H. and S. F. Richard 1981, 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. 2004. The Democratic Dynamics of Educational Investment and Income Distribution. In E. J. Mansfield and R. Sisson, eds., The Evolution of Political Knowledge: Democracy, Autonomy, and Conflict in Comparative and International Politics, Columbus: University of Ohio Press.

Roemer, J. E. and W. Lee. 2004. Racialism and Redistribution in the United States: A Solution to the Problem of American Exceptionalism. Working Paper, Yale University.

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

Schofield, N. 1984. Classification Theorem for Smooth Social Choice on a Manifold. Social Choice and Welfare, 1(3), 187-210.

Skrentny, J. D. 1996, The Ironies of Affirmative Action. Chicago: Chicago University Press.

Table 1: Equilibrium for Example 1

Equilibrium with Affirmative Action	Prob.	Affirmative Action Target	Benefit	Tax Rate	Minority Share of Good Jobs	Gini
$\mathcal{H}$'s proposal	.61	none	.04	.07	.301	.448
	.39	.333	.01	.02	.333	.470
$\mathcal{L}$'s proposal	1	none	.09	.16	.300	.364
$\mathcal{B}$'s proposal	1	.332	.12	.20	.330	.351
Expected Value with Affirmative Action			.08	.14	.314	.392
Equilibrium without Affirmative Action			.12	.21	.298	.347
Laissez-faire			0	0	.306	.482

Table 2: Comparative static on θ_B

θ_B	Minority Share of Good Jobs	Average Benefit	Average Tax
.2	.314	.08	.14
.25	.318	.09	.15
.3	.321	.09	.15
.35	.325	.10	.16
.4	.328	.11	.18
.45	.331	.12	.19
.495	.333	.14	.22

Table 3: Comparative static on q

q	Minority Share of Good Jobs	Average Benefit	Average Tax
.25	.314	.08	.14
.22	.314	.12	.17
.19	.315	.15	.19
.16	.316	.19	.21
.13	.317	.23	.22

Table 4: Comparative static on λ

	Minority Share	Average	Average	Percentage Net Gain,				Average
λ	of Good Jobs	Benefit	Tax	HW	LW	HB	LB	Gini
0	.314	.08	.14	2.8	-2.8	5.1	1.0	.392
.1	.316	.06	.11	4.2	-3.2	6.8	0.8	.410
.2	.317	.05	.09	5.1	-3.7	7.8	0.4	.422
.3	.317	.04	.07	5.8	-4.2	8.5	0	.433

FIGURE 1
Legislative bargaining equilibrium