Redistribution and Affirmative Action$^1$

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

Michael Wallerstein  
Department of Political Science  
Yale University  
New Haven, CT 06520

January 2003, This version October 2004

---

$^1$This paper is a revised version of an earlier working paper, “Redistribution in a Divided Society”. Among others, we are grateful to seminar participants at CEACS of the Juan March Foundation, Madrid, at UCLA, and at the PIEP Political Economy workshop, Harvard University, for constructive comments. 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.
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.
1 Introduction

Many scholars have observed that the politics of redistribution in the US is intertwined with the politics of race. Lipset and Bendix, writing in the 1950s, argued that the “social and economic cleavage” created by discrimination against blacks and Hispanics “diminishes the chances for the development of solidarity along class lines” (1959: 106). Myrdal (1960), Quadagno (1994) and, most recently, Gilens (1999) claim that racial animosity in the US is the single most important reason for the limited growth of welfare expenditures in the US relative to the nations of Western Europe. According to Quadagno (1994), political support for Johnson’s War on Poverty was undermined by the racial conflicts that erupted over job training and housing programs. Alesina, Baqir and Easterly (1999) find that localities in the US with high levels of racial fragmentation redistribute less and provide fewer public goods than localities that are racially homogeneous. The US is not the only democracy where redistributive politics has a racial dimension. At the time of this writing (autumn 2003), the Brazilian parliament is considering a proposed racial inequality law that would establish quotas for the employment of Afro-Brazilians in government and private-sector enterprises with more than 20 employees.

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. Likewise, formal models of affirmative action (e.g. Coate and Loury; 1993; Foster and Vohra, 1992; Lundberg, 1991; Chung, 2000) have, to the best of our knowledge, 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 of redistribution due to the introduction of additional dimensions of potential redistribution. As such, our model is neither one dimensional nor predicated on exogenous racial preferences. In particular, we develop a model where voters differ by their stock of human capital and by some ascriptive characteristic. 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 “good’ jobs with minority workers. Voters are assumed to be identical with respect to their fundamental (self-interested) preferences, but both race and wealth matter in determining voters’ derived preferences over policies. We then study how the expansion of the policy space to include affirmative action alters the expected outcome with respect to fiscal redistribution. To be clear about our research strategy, we do not deny the potential importance of racial or ethnic differences in preferences. Our purpose in assuming that voters vote in a color-blind fashion to maximize their post-tax and transfer

\[1\text{The most salient exceptions to one-dimensional models of redistribution are models of competition between special interests: see, for example, Grossman and Helpman (2001) and Dixit and Londregan (1998). The dynamics of political conflict among many narrow interests, however, is likely to be quite different from political conflict among a few, large social groups defined by non-economic criteria.}\]
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.

The articles most closely related to this paper, as far as we know, are the studies by Roe-
mer (1998) and Roemer and Lee (2003). In these papers, 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 law and order or the separation of church and state.\(^2\) 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 a “policy bundling effect.” For example, the election may pit a secular party that favors redistribution against a religious party opposed to redistribution. A low-income, religious voter may prefer the party opposed to redistribution, depending on the relative weights of the two dimensions in the voter’s preferences. The more voters care about the non-economic dimension relative to the redistributive dimension, the less the parties’ positions on redistribution influence the vote. In the limit, Roemer shows that the equilibrium tax rate approaches zero as the weight of the non-economic dimension increases, provided the mean wealth of voters with the median view on the non-economic dimension exceeds the mean wealth of the population.

Our approach differs from Roemer’s in several ways beyond a focus on affirmative action in particular. First, as mentioned above, we assume that voters have identical, self-interested preferences. Rather than assume that voters’ preferences differ with regard to a non-economic policy, we derive voters’ induced preferences over affirmative action and redistributive taxation from a model of the economy where voters are differentiated by human capital and by an ascriptive characteristic. 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

\(^2\)Roemer and Lee (2003) add an assumption that voters also altruistically care about aggregate inequality, with racially conservative voters attaching less weight to equality than racially liberal voters.
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. Our emphasis, instead, is on the tradeoff between redistribution by income and redistribution by race that occurs through a process of legislative bargaining.

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 presence of policies that redistribute according to ascriptive characteristics reduce support for redistribution according to income? Who benefits and who loses when the policy space is expanded to include policies that redistribute by ascriptive traits? How do the policies selected in equilibrium change as the distributions of income within the majority and minority social groups become more similar? We return to these questions after describing our model of the economy and our characterization of the political equilibrium.

2 A labor market with redistribution

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 worth fighting over, we need to start with a model of the labor market in which there is some inefficiency such that workers with “good” jobs receive rents. 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 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\textsuperscript{3}. The holdup model simply provides a convenient way to construct an economic framework in which affirmative action policies make sense. Other models that generate employment rents in the labor market, such as efficiency wage models, would yield similar results concerning the political equilibrium.

In this and the subsequent section respectively, we describe our model of equilibrium in the labor market and voters’ induced preferences over redistributive policies.

2.1 Demographics, human capital and jobs

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

Assume that workers have one of two levels of human capital, \( H = \{0, h\} \) where \( h > 0 \). The two levels might be interpreted, for example, as having a four year college degree or not. Thus, \( h \) refers to the additional human capital that educated workers have above the basic education that all workers share. Let the subscript \( W \) (for white) denote the majority group and the subscript \( B \) (for black) denote the minority group. Let \( \theta_i \) be the share of group \( i = W, B \) with the high level of human capital and let \( \theta \) be the share of the population with

\textsuperscript{3}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), Acemoglu and Shimer (1999) and Acemoglu (2001) among others.
the high level of human capital, or

\[ \theta = p\theta_B + (1 - p)\theta_W \]

Finally, we assume that blacks are disadvantaged in the labor market because they have less human capital on average than whites, or that \( \theta_B < \theta_W \). An alternative assumption would be that blacks are disadvantaged because they face discrimination in the job market. In this paper, we assume the racial gap in average earnings reflects a racial difference in average human capital.

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

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

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

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

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

The decision sequence in the polity is as follows:

1. The legislature chooses \( \alpha \) and \( b \) simultaneously.
2. Workers and employers are randomly matched and the match-specific component of productivity, \( x \), is revealed to both.
3. Employers decide whether to create a good job at a cost of \( q \) or a bad job at zero cost.
4. Workers and their employers bargain (individually) as necessary over the wage.

2.2 Wage bargaining and job creation

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

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

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

The firm’s profit is given by

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

Note that the cost \(q\) of creating a good job is a sunk cost when wage bargaining occurs. This assumption is a simple way of representing the common situation where wages are set for shorter periods of time than the life-span of the plant and equipment that must be purchased by the firm to usefully employ a worker in a good job. Further, while profits are assumed not to be taxed, profits are affected by taxes and transfers since taxes and transfers affect the wage that firms must pay.\(^4\) Once a good job is created, the firm’s gain from an agreement is \(y - w(y)\).

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

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

as the wage offered in good jobs and

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

as the profit earned from the creation of a good job, where \(\beta \in (0,1)\) represents the worker’s share of the joint gains (and we suppress the subscript \(j\) on \(\pi(y)\)).

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

\[
y \geq y_0(b) \equiv \frac{q}{1-\beta} + b.  \tag{3}\]

\(^4\) Taxing profits at the same rate as wages has virtually no consequences for the qualitative results to follow. This is not true under the plausible assumption that profits are taxed at a different rate; but this greatly complicates the political decision-making problem. So, for now at least, we simply assume these complications away by having only wages be taxed.
The fraction of group $i = W, B$ with good jobs, denoted $\sigma_i$, is given by

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

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 $y$. Without affirmative action, the share of good jobs held by minority workers, denoted $\alpha_0(b)$, is given by

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

By assumption, $\theta_B < \theta_W$ so $G_B(y_0(b)) > G_W(y_0(b))$ and, consequently, $\sigma_B(b) < \sigma_W(b)$. Therefore, lower average human capital for minority workers implies the share of good jobs held by minority workers is less than the share of minority workers in the workforce, or that $\alpha_0(b) < p$. Hereafter, where there is no ambiguity, we leave the dependency of $\alpha_0$, $y_0$ and $\sigma_i$ on $b$ implicit.

Because the cost of creating a good job is a sunk cost to the employer, there are match-specific rents over which employers and the prospective employees bargain. And since workers capture a share of the rents, the number of good jobs created is less than the number that maximizes the joint income of employers and workers. The additional output obtained from a marginal increase in the number of good jobs exceeds the cost of job creation; that is, $\beta > 0$ implies $y_0 > q$. Moreover, workers with $y = y_0$ obtain a discrete jump in pay of $w(y_0) - b = [\beta/(1 - \beta)] q$ in moving from a bad job to a good job, even though employers are

---

5In 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. Our assumption that wage bargaining takes place after the employer’s investment is sunk is a reduced form version of a slightly more complicated model where investments last for two (or more) periods while wage contracts only last for a single period. In this slightly more complicated model, 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.
indifferent between creating a good job or not when $y = y_0$. Thus, workers with productivity slightly below $y_0$ would gain from a policy that forces firms to create good jobs for them.

### 2.3 Affirmative action and insurance

Affirmative action entails setting a lower bound, $\alpha \in [0, p]$, on the proportion of good jobs filled by minority workers. If the policy of affirmative action is binding, that is if $\alpha > \alpha_0(b)$ or $\alpha$ if is greater than the fraction of good jobs that would be filled by minority workers in the absence of affirmative action, then firms cannot use the same productivity threshold for both black and white workers when deciding whether or not to create a good job. If $\alpha = p$, the affirmative action policy requires firms to equalize the fraction of workers with good jobs in the two social groups.\(^6\)

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

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

subject to the constraint that

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

\(^6\)Holzer and Neumark (2000) observe that the enforcement of equal opportunity legislation has lead to affirmative action targets in practice, $\alpha > \alpha_0$. 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 (EEOC), 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.
When the constraint is binding, the first-order condition for a maximum can be written as two equations. The first equation replaces (3):

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

(5)

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

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

$$[1 - p\sigma_B - (1 - p)\sigma_W] (1 - t)b = tE(w)$$

(6)

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

2.4 Labor market equilibrium

Conditional on policy \((\alpha, b)\) 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).\(^7\) It is routine to check that, under the maintained assumptions, there exists a unique labor market equilibrium associated with every policy. Individuals’ induced preferences over policies, therefore, are well defined. Before going on to examine such preferences in any detail, however, it is useful to identify some salient properties of the labor market equilibrium per se.\(^8\)

\(^7\)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.

\(^8\)Unless noted otherwise, technical arguments for the results to follow are confined to an Appendix. And we recall that throughout we maintain the assumption that the CDF \(F\) is approximately uniform over a suitable
Lemma 1  (1) $\partial y_W/\partial \alpha > 0 > \partial y_B/\partial \alpha$;

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

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

Affirmative action induces firms to raise the threshold for hiring white workers and to lower the threshold for hiring minority workers.\(^9\) On the other hand, both of these thresholds are strictly increasing functions of the benefit. Moreover, the threshold for whites increases at a greater rate than does the threshold for blacks when an affirmative action constraint forces employers to maintain a given proportion of minority employment in good jobs. Finally, the tax rate is a strictly increasing function of the benefit that must be financed.

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

Proposition 1 Suppose $b > 0$. There exist $\alpha_b, \alpha_Y$ such that $\alpha_0 < \alpha_Y < \alpha_b < p$ and

\(^9\)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.
(1) $\partial t / \partial \alpha < 0$ for all $\alpha \in [0, \alpha_b)$ and $\lim_{\alpha \to p} \partial t / \partial \alpha > 0$;
(2) $\partial E (w + \pi(w)) / \partial \alpha > 0$ for all $\alpha \in [0, \alpha_Y)$ and $\lim_{\alpha \to p} \partial E (w + \pi(w)) / \partial \alpha < 0$.

Assuming there is at least some redistribution through taxes and transfers, expected aggregate income is strictly increased and the tax rate needed to finance a given benefit is strictly reduced by a marginal increase in affirmative action away from the laissez faire distribution of good jobs between blacks and whites, $\alpha_0$. In this case, the increase in the number of good jobs filled by black workers exceeds the decline in the number of good jobs filled by white workers and, since white workers are being replaced by black workers with similar levels of productivity when $\alpha \approx \alpha_0$, aggregate income and aggregate earnings increase.\textsuperscript{10} This is an example of the theorem of the second best: in an economy where insiders’ bargaining power reduces investment in good jobs below the optimal level, a policy that results in increased good job creation raises aggregate income. Of course, affirmative action policies need not result in job creation if $\alpha > \alpha_0$. Indeed, for $\alpha$ sufficiently large, increases in affirmative action reduce aggregate earnings and raise the tax rate needed to finance a given benefit. Moreover, because any increase in $\alpha$ above $\alpha_0$ reduces profits, aggregate income must be decreasing in $\alpha$ if aggregate earnings are decreasing in $\alpha$, although the converse is not true ($\alpha_0 > \alpha_Y$). At $\alpha \approx \alpha_0$, the gain in aggregate earnings that follows from a small increase in $\alpha$ is greater than the loss in profits.

\textsuperscript{10}Formally, given $F$ uniform, we have

$$
\frac{d\sigma}{d\alpha} = p \frac{d\sigma_B}{d\alpha} + (1 - p) \frac{d\sigma_W}{d\alpha},
$$

$$
= -[p \frac{dy_B}{d\alpha} + (1 - p) \frac{dy_W}{d\alpha}] f
$$

where $dy_B/d\alpha < 0$ and $dy_W/d\alpha > 0$. Since the affirmative action constraint binds, at $\alpha = \alpha_0$ we have $y_B = y_W = y_0$ so $\alpha y_B + (1 - \alpha) y_W = y_0$ and therefore

$$
\alpha \frac{dy_B}{d\alpha} + (1 - \alpha) \frac{dy_W}{d\alpha} = 0.
$$

Hence, $[d\sigma/d\alpha]_{\alpha=\alpha_0} > 0$ if $p > \alpha_0$, which is true.
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 \((\alpha, b)\) is preferred by some individual to an alternative policy \((\alpha', b')\) if and only if the individual’s expected equilibrium consumption level under \((\alpha, b)\) is greater than that under \((\alpha', b')\). Formally, an individual with human capital \(H \in \{0, h\}\) in group \(i \in \{B, W\}\) evaluates a policy \((\alpha, b)\) by

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

where \(y_i\) and \(t\) are defined by the labor market equilibrium at \((\alpha, b)\). Let \(b_{max}(\alpha)\) denote the benefit that maximizes \((1 - t)b\) given \(\alpha\).

**Lemma 2** (1) For all \((\alpha, b) \in [0, p] \times [0, b_{max}(\alpha)]\), \(\partial E [c_{0i}(\alpha, b)] / \partial b > \partial E [c_{hi}(\alpha, b)] / \partial b\), \(i = B, W\);

(2) There exists \(\epsilon > 0\) such that, for all \(\alpha < \alpha_0(0) + \epsilon\) and for all \(b \in [0, b_{max}(\alpha)]\), \(\partial E [c_{HB}(\alpha, b)] / \partial b > \partial E [c_{HW}(\alpha, b)] / \partial b\), \(H = 0, h\).

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

Race has counteracting effects on the support for welfare expenditures in the presence of affirmative action policies when the level of human capital is held constant. On the one
hand, affirmative action policies provide some protection to minority workers with good jobs. The loss of good jobs is greater for white workers than for black workers when employers reduce the number of good jobs in response to increases in the benefit in the presence of an affirmative action constraint. This effect works to increase black support for higher taxes and benefits. On the other hand, affirmative action raises the expected income of blacks for a given level of human capital, which reduces support for higher taxes and benefits. Under the maintained assumption on the distribution of $x$, however, there is a clear answer when the affirmative action policy is barely binding. This answer is given by Lemma 2(2): given $f(x)$ is constant and $\alpha \approx \alpha_0$, black voters prefer (weakly) higher levels of taxes and spending than white voters with equal levels of human capital. In this respect, our model matches the survey evidence that the effect of race on support for welfare expenditures does not vanish when income is controlled for, even though we have assumed that voters have color-blind preferences (Kinder and Sanders 1996).11

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

**Lemma 3** There exists $\alpha_b$ such that $\alpha_0 < \alpha_b < p$ (as in Proposition 1) and $\beta \in (0, 1)$ such that, for all $(\alpha, b) \in P$:

1. For $\beta \geq \beta^*$, $\partial E[c_{HB}(\alpha, b)] / \partial \alpha > 0 > \partial E[c_{HW}(\alpha, b)] / \partial \alpha$, $H = 0, h$;
2. $\partial E[c_{hi}(\alpha, b)] / \partial \alpha \underset{i < b}{\leq} \partial E[c_{0i}(\alpha, b)] / \partial \alpha$ as $\alpha \underset{\geq}{\approx} \alpha_B, \ i = B, W$.

Preferences with regard to affirmative action are straightforward. Provided workers’ bargaining power is sufficiently high, members of the minority group favor increases in affirmative action while members of the majority group are opposed, regardless of the level of human

---

11To be clear, we do not deny the possible existence of important differences in preferences of black and white voters in the US. Our point is that the racial differences in preferences over policies observed by survey researchers do not necessarily imply the existence of racial differences in the fundamental preferences from which policy preferences are derived.
capital (Lemma 3(1)). The intuition here is that, if $\beta$ is not too low, the direct effect of affirmative action on the likelihood of acquiring a good job dominates its indirect effect on the tax-rate required to finance a given transfer. On the other hand, as $\beta$ becomes negligible so does any advantage from holding a good job and all that matters for workers’ consumption is the tax and transfer. In this case, all workers can prefer a moderate level of affirmative action. Furthermore, in the neighborhood of $\alpha_0$, high-type blacks care more about affirmative action than low-type blacks, while high-type whites are less opposed to affirmative action than low-type whites (Lemma 3(2)).

Because the affirmative action constraint on firms’ hiring decisions is binding, a tightening of the constraint must cause profits to decline (if not, the constraint was not binding). Hence, expected profits are strictly decreasing in affirmative action on the interval $\alpha \in [\alpha_0, p]$ and, by Lemma 1(3) and the definition of $\pi(\cdot)$, they are also decreasing in $t \in [0, 1]$. Together, these remarks and Lemmas 2(1) and 3 prove the following proposition.

**Proposition 2** For each $(H, i) \in \{0, h\} \times \{B, W\}$, let $I_{Hi} = (\alpha_{Hi}^*, b_{Hi}^*)$ maximize $E[c_{Hi}(\alpha, b)]$ on $P$. Assume low-type individuals strictly prefer at least some strictly positive benefit to zero benefit and that $\beta \geq \underline{\beta}$. Then $I_{Hi}$ is unique and

1. $\alpha_{hB}^* = \alpha_{0B}^* = p$ and $\alpha_{0W}^* = \alpha_{hW}^* = 0$;
2. $b_{0i}^* > b_{hi}^*$, $i = W, B$.

Furthermore, a firm’s most preferred policy is laissez faire, $(0, 0)$.

Although nothing in general can be said about the relative location of the majority’s and minority’s most preferred benefit levels at their respective ideal points, it is worth recalling that if the only policy instrument is fiscal redistribution, so $\alpha \equiv \alpha_0(b)$ at every $b$, then Lemma 2(2) implies that high (respectively, low) type blacks prefer weakly higher levels of benefit and taxation than do high (respectively, low) type whites.
4 Legislative bargaining

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

A complete model of the political process would include (at least) two stages. The first stage involves voters’ choice of representatives while the second stage consists of representatives’ choice of policy. In this paper we focus exclusively on the legislative policy decision stage. With regard to voters’ choice, we simply assume the existence of blocs of representatives or parties, each of whom represents a distinct constituency. While a variety of divisions of the legislature might be considered, here we restrict our analysis to the case in which the legislature is divided into three groups: legislators who represent white workers with a high level of human capital ($H$), legislators who represent white workers with a low level of human capital ($L$), and legislators who represent black workers ($B$).

Two comments on this choice of partition are worth making explicit. First, firms and employers are not directly represented, yet they are, as we discuss in the concluding section, important to the development of affirmative action policy in the US. Nevertheless, in this paper we make the analytically convenient simplification that the bloc $H$ speaks also for firms: at least on the policy space $P$, firms and high whites share an identical most preferred policy although, as suggested by Proposition 2, the indifference map on $P$ of firms and of high whites do not coincide everywhere. Second, unlike white workers, we presume all black workers are represented by the same bloc of legislators, $B$. 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.\footnote{Exploring the implications of alternative patterns of legislative representation is a topic for future research.} Having said this, it is
worth noting two things. First, we compare the implications of the bargaining process with the three groups as assumed with those of a majority rule competition between two groups comprising, respectively, high types and low types. While the details of the equilibrium would differ, none of our results concerning the comparison of the equilibrium with and without affirmative action depend on assuming one rather than two groups of legislators representing minority voters. Second, we consider the stability of the black workers’ group $B$ with respect to the relative weights of high and low type blacks in the $B$-bloc’s decision-making process. In particular, we show that so long as the relative weight given to the interests of high type blacks in the $B$ objective function is not “too large”, then low type of black workers are on average strictly better off with $B$ than if they defected to either of the white groups.\textsuperscript{13}

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 with a given level of human capital are identical \textit{ex ante}. Therefore, the objective functions for $H$ and $L$, respectively, are $u_H = E(c_{hW})$ and $u_L = E(c_{0W})$. Minority voters, however, include workers with both high and low levels of human capital. In this case, we assume the legislative group maximizes a weighted average of the consumption of its two types of constituents: $u_B = (1-\lambda)E(c_{0B}) + \lambda E(c_{hB})$ for some $\lambda \in [0,1]$. The weight $\lambda$ is a measure of the extent to which the minority party is concerned with low or high types of minority workers. Let $I_H = (\alpha_{H}^*,b_{H}^*) \equiv I_{hW}$ and $I_L = (\alpha_{L}^*,b_{L}^*) \equiv I_{0W}$ be the most preferred policy points for $H$ and $L$, respectively, and let $I_B = (\alpha_{B}^*,b_{B}^*) \equiv \arg\max u_B$ be the most preferred policy point for party $B$. By Lemmas 2(1) and 3(1), $I_B \equiv [(1 - \mu(\lambda))I_{0B} + \mu(\lambda)I_{hB}]$, where $\mu(0) = 0$, $\mu(1) = 1$,and $\mu' > 0$.\textsuperscript{14}

\textsuperscript{13}High type black voters are always better off supporting $B$ than defecting to either of the white groups, even when black legislators care only about low type black voters.

\textsuperscript{14}Lemma 3(1) asserts that both high and low type blacks strictly prefer a maximal affirmative action target
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 the size of each legislative bloc reflects the relative size of each bloc’s constituents, this implies that $p < 1/2$, $(1 - p)\theta_W < 1/2$ and $(1 - p)(1 - \theta_W) < 1/2$.

Thus any two of the three parties constitutes a majority. It is not hard to check that, as is usually the case with multidimensional policy spaces, the majority core is typically empty.

Suppose $\beta$ is sufficiently high that $\alpha_{hB}^* = \alpha_{0B}^* = p$ and $\alpha_{hW}^* = \alpha_{0W}^* = 0$. The coalition of $\mathcal{H}$ and $\mathcal{L}$ then strictly prefers the policy $(\alpha - \epsilon, b)$ for some $\epsilon > 0$ to any feasible pair $(\alpha, b) \in \mathcal{P}$ with $\alpha > \alpha_0$ while, for any pair $(\alpha, b)$ with $\alpha \leq \alpha_0$ and $b \in (b_{hL}^*, b_{L}^*)$, there is a policy $(\alpha_0 + \epsilon, b \pm \delta)$ for some $\epsilon \geq 0$ and $\delta > 0$ that is strictly preferred by either the coalition of $\mathcal{B}$ and $\mathcal{L}$ or the coalition of $\mathcal{B}$ and $\mathcal{H}$. The high white’s ideal point cannot be in the core, since a majority prefer $(\alpha_0, b_{hL}^* + \delta)$ to $(\alpha_0, b_{hL}^*)$. Therefore, either the core consists of the low white’s ideal point $I_L = (\alpha_L^*, b_L^*)$ (a possible but unlikely case) or the core is empty.

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

The solution concept is a no-delay, stationary subgame perfect Nash equilibrium (here-$\alpha = p$) whatever the level of fiscal redistribution, $b$. Therefore the Black party is divided only in respect of its members’ most preferred benefits level. By Lemma 2(1), the low types strictly prefer higher benefit levels than high types at any level of affirmative action. It follows that the Black party’s ideal point must be a convex combination of the ideal points of the low and high types.
An equilibrium consists of a (possibly degenerate) probability distribution $\zeta_j$ over a (possibly infinite) set of policies $P_j \subseteq P$ that party $j$ proposes whenever $j$ is recognized to make a proposal, and an acceptance set, $A_j \subseteq P$ that specifies the set of policies for which party $j$ will vote if another party is the proposer. Let $v_j$ be $j$'s expected payoff at the beginning of the game. By stationarity, $v_j$ is also $j$'s continuation value or its expected payoff after a proposal has been rejected. Finally, let $\rho_j \in (0, 1)$ be party $j$'s probability of being recognized to make a proposal. Then an equilibrium consists of a set of policy proposals and acceptance set for each $j = H, L, B$ that satisfy the following conditions:

$$P_j \subseteq \begin{cases} \arg \max \{u_j(\alpha, b) \mid (\alpha, b) \in A_j \cup A_l \} & \text{if } \sup[u_j(\alpha, b) : (\alpha, b) \in A_j \cup A_l] \geq v_j \\ P \setminus [A_j \cup A_l] & \text{otherwise} \end{cases}$$

$$A_j = \{ (\alpha, b) \mid u_j(\alpha, b) \geq v_j \}$$

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

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

The most general equilibrium existence result for this game (at least, as far as we know) is due to Banks and Duggan (2000, 2001). However, their theorem assumes preferences are

---

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

20
concave on the policy space. Unfortunately, concavity is not a general property of induced preferences here, although (as an example below illustrates) agents’ induced preferences are indeed strictly concave on $P$ for most parameterizations of the model. 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 as the benefit (and the tax rate) increase. Conversely, the lower the benefit, the greater the impact of affirmative action policies on workers’ expected after-tax and transfer income. Consequently, the marginal rate of substitution between affirmative action and welfare benefit can be increasing and preferences over policies non-concave on the policy space.\footnote{For a non-pathological example that generates non-concave preferences, let the benefit $b$ be a universalistic payment to all workers rather than a benefit received only by workers in bad jobs. In this case, there is no deadweight cost of taxation, and the function $b(t)$ is linear. It is easy to check that no group has concave preferences. Moreover, the “problem” of non-concavity cannot be fixed in this example by assuming that workers are sufficiently risk averse.}

The difficulty here is essentially technical. So, rather than attempt to finesse complications with equilibrium existence due to nonconvexities in parties’ induced preferences, we simply assume in what follows that an equilibrium exists. To show that the equilibrium concept is not vacuous, we later present an example in which all parties’ induced preferences are strictly concave on $P$ and calculate the legislative bargaining equilibrium. Taking the existence of equilibrium as given, the impact of the addition of a second dimension of redistributive conflict on the policy outcome, reflected in Table 1, is summarized in the next proposition.

**Proposition 3** Assume induced preferences over the policy space are concave, that $\beta \geq \beta_\cap$ and that the majority core is empty. Then the possibility of redistribution through affirmative action reduces the expected redistribution through the tax and transfer policy.

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

The addition of redistributive conflict along ethnic or racial lines always reduces the equilibrium amount of redistribution that occurs along income lines. The addition of a racial or ethnic dimension of redistribution reduces the amount of redistribution according to income, even when (i) the racial minority is poorer than the majority on average and (ii) the legislators who represent the minority group exclusively advance the interests of the less educated members of the minority group as in our example below. The potential alliance of blacks and high whites to lower the tax rate and raise affirmative action targets to make both better off is sufficient to reduce the average transfer payment and to increase the average affirmative action target. Note that the proof of Proposition 3 applies equally well in the case of an equilibrium with four voting blocs where low blacks and high blacks act independently, provided that the two parties of less-well educated voters constitute a majority. The impact of the addition of a second dimension of redistributive conflict on the expected post-tax, post-transfer income of each group of voters is summarized in the next proposition.

**Proposition 4** If an equilibrium exists, then 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 $\lambda$ and $(1 - \lambda)$ respectively), and (2) lowers the expected income of low whites.
Proof Suppose part (1) was false. Then either high whites or blacks could do better by proposing the low white’s ideal point, which would certainly be accepted. But then an equilibrium could not exist by the same reasoning as in the proof of Proposition 3. Part (2) is an immediate corollary of Proposition 3. Since the equilibrium with affirmative action results, on average, in a tax rate that is less than what low whites prefer and a binding affirmative action target, low whites are worse off. □

Propositions 3 and 4 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 here is not a knife-edge property of the 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 have a low level of human capital and that the three groups are equal in size. We assume initially that the black party represents the 80 per cent of minority voters with less education, or $\lambda = 0$. Figure 1 illustrates the preferences of the three parties over welfare benefits and affirmative action and the equilibrium of the legislative bargaining game.\textsuperscript{17} The western and northern borders of the Pareto set are given by $b = 0$ and $\alpha = p$ respectively. The southern border is given by the function $\alpha_0(b)$ which represents the share of good jobs that the minority would receive without affirmative action. Note that $\alpha_0(b)$ declines as $b$ increases. The ideal points of the three groups, $\mathcal{H}, \mathcal{L}, \mathcal{B}$, are denoted $I_{\mathcal{H}}$, $I_{\mathcal{L}}$ and $I_{\mathcal{B}}$ respectively. The figure illustrates the unique equilibrium for the case in which each group is recognized with equal probability. If recognized, the high-type whites propose $(\alpha_{\mathcal{H}}, b_{\mathcal{H}})$ with probability (.61) and receive the support of the low-type whites.

\textsuperscript{17}Since the impact of choosing any $\alpha \leq \bar{\alpha}_0(b)$ is that the \textit{de facto} allocation of good jobs at $b$ is $\alpha_0(b)$, there is no loss in generality in normalizing the whites’ most preferred policies to $\alpha_0(b_{\mathcal{L}})$ and $\alpha_0(b_{\mathcal{H}})$, respectively, rather than zero, and Figure 1 is drawn to reflect this fact.
With probability (.39), the high-type whites propose \((\alpha_H', b_H')\) and receive the support of the black legislators. The low-type whites, if recognized, propose \((\alpha_L, b_L)\) with probability one and win the support of the black legislators. Finally, black legislators propose \((\alpha_B, b_B)\) 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 under the assumption that the probability any party \(j\) is recognized to make the proposal for any period is \(\rho_j = 1/3\).

Table 1 here

For comparative purposes, the last line of Table 1 shows the political equilibrium that exists absent the racial divide. When there is no racial divide, political conflict occurs over the single dimension of the tax rate. Since, in 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 would prevail. \(\square\)

Example 1 demonstrates that the equilibrium concept is not vacuous. In the case of a uniform distribution, the concavity property sufficient for equilibrium existence is satisfied for parameter values that cover most (but not all) of the feasible parameter space. While it is difficult to provide general conditions for the strict concavity of legislators’ induced preferences, the central features of the equilibrium illustrated in Figure 1 and Table 1 are general characteristics of all equilibria of the legislative bargaining model.

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 4, that the expected consumption (1) of high whites increases by 2.8%; (2) of low whites decreases by 2.4%; (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 workers with high levels of human capital. Highly educated minority workers gain both from affirmative action and the tax reduction, while highly educated majority workers benefit from the tax reduction. Even though black legislators were assumed to be pure representatives of black workers with the low level of education, such workers gain much less, with the gains from affirmative action partly offset by the loss from the lower benefit. The losers are members of the majority who lose from affirmative action and lose again from the reduction in the average amount of redistribution along income lines with the addition of a second dimension.

Proposition 4 and the numerical illustration raise a question about the effect of affirmative action on the inequality of post-tax, post-transfer income. In the canonical 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. But although the direct impact of fiscal redistribution in the current model is no different, the impact of affirmative action is equivocal. Proposition 1(2) suggests affirmative action can have different redistributive effects depending on how aggressive is the policy relative to the laissez faire situation: at low levels of $\alpha$, more high blacks acquire good jobs than are lost by high whites, but this is not true for higher levels of $\alpha$. 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 (2000) 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. Unfortunately, analytical comparative statics are unavailable; providing an explicit and tractable general solution to the system of simultaneous equations jointly characterizing the bargaining equilibrium is not feasible here.\footnote{For instance, an equilibrium can involve mixed strategy proposals by all three parties over each of the two policy dimensions. In addition, these proposals must satisfy a set of endogenously determined acceptability conditions and reflect subsequent labor market equilibrium behavior.} Instead, we report a set of numerical comparative static calculations for variations of Example 1. There are three sort of variation of interest in the model: political, social and economic. Consider these in turn. In every case, we assume (as for Example 1) that $x$ is uniformly distributed over the interval $[0, 1]$, set

$$h = 1/4, \ \beta = 1/2, \ p = 1/3, \ \theta_W = 1/2$$

and suppose $\rho_j = 1/3$ for all $j \in \{B, H, L\}$; values for the remaining parameters, $\lambda$, $\theta_B$ and $q$, are specified as needed below.

4.2.1 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 ($\lambda$) between high and low types within the minority ($B$) party. Given the parameterization of Example 1 with $\theta_B = 1/5$ and $q = 1/4$,
Table 2 describes the unique legislative bargaining equilibrium as the weight given to high-type blacks in the black party, \( \lambda \), 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 \( \lambda = 1/2 \).

Table 2 here

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.

It is apparent from the table that an increase in the weight of highly educated blacks in the black party \( B \) 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 \( B \) 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 \( H \)'s bargaining power relative to \( L \). Consequently, all high types benefit from an increase in \( \lambda \) 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 4 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 \( \lambda \) and \( 1 - \lambda \) for high and low types respectively. On the other hand, Table 2 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 (\( \lambda \) high) to leave the low blacks with strictly less in equilibrium than they would receive at the low white party's (effective) ideal point, \( I_L = (\alpha_0, .12) \),
then representatives of low blacks prefer to defect from $B$ and vote with $L$ to insure $I_L$ as the legislative decision. In other words, for the parameterization of Table 2, $\lambda = .3$ 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 $\lambda = 0$. Similarly, at a more aggregate level, the extent to which post-policy incomes are equalized, as measured by changes in the Gini coefficient, is inversely related to $\lambda$. The (expected value of the) Gini coefficient with $\lambda = 0$ is, as reported earlier, 0.392; as $\lambda$ rises to 0.5, the Gini coefficient rises to 0.458. Thus while there is always some equalization of incomes relative to the laissez faire benchmark, the more are high blacks weighted in the evaluation of policy by the black party, the less leveling occurs.

4.2.2 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 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. Specifically, we ask how affirmative action and direct fiscal redistribution vary with the share of high blacks in the population ($\theta_B$).

As before, assume the basic parameterization of Example 1 with $\lambda = 0$ and $q = 1/4$. Choosing $\lambda = 0$ helps comparison with the benchmark Example 1 and, for this particular comparative static, is the conservative assumption, as increases in the proportion of high blacks have no effect on their influence in the legislative bargaining process; the consequences of allowing $\theta_B$ and $\lambda$ to covary are considered shortly.
Table 3 describes the (again unique) legislative equilibrium as $\theta_B$ varies between $1/5$ (as in Example 1) and $\theta_W = 1/2$.

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.

It is worth noting that, at $\theta_B = \theta_W = 1/2$, the laissez faire labor market equilibrium sets $\alpha_0 = p$ and so affirmative action is redundant. At $\theta_B = \theta_W$, that is, the effective policy space becomes one-dimensional, involving only fiscal policy. Furthermore, the argument for Lemma 2(2) (with $\alpha_0 = p$) implies that, in the limit, all low-type workers have a common ideal benefit, $b_B^* = b_L^* > 0$. Therefore, because $\lambda = 0$ and low types are a strict majority in the population, there is a core allocation given by the low-types’ (effective) most preferred policy, $(\alpha_0(b_L^*), b_L^*)$. And, as predicted by Banks and Duggan (2000, Theorems 3 and 5), it is this allocation, $(p, b_L^*) = (.333, .14)$, to which the sequence of equilibrium bargaining outcomes listed in Table 3 converges.

Although the relative importance of high-type black workers on the black party, $\lambda$, has so far been treated parametrically, it is plausible that the relative weights of the two types of black worker within the party reflect their relative weights in the black population as a whole; that is, it is reasonable to assume $\lambda = \theta_B$. Adopting this alternative assumption and, ceteris paribus, recomputing Table 3 yields Table 4.
As before, at $\theta_B = \theta_W = 1/2$ we have $\alpha_0 = p$ and affirmative action becomes irrelevant. In this case, however, although all low-type workers again have the same preferences over the benefit policy, this policy (with $\alpha_0$) is not the core allocation. Because $\lambda = \theta_B > 0$, Proposition 2 implies that the black party’s most preferred benefit rate at $\alpha_0 = p$ is a convex combination of the ideal rates of the low- and high-type black workers,

$$b^*_B = [\mu(\lambda)b^*_{hB} + (1 - \mu(\lambda))b^*_{0B}] = [\mu(\theta_B)b^*_{hB} + (1 - \mu(\theta_B))b^*_{0B}],$$

where $b^*_{hB} \in (0, b^*_{0B})$ is the most preferred level of benefits for the high blacks and $b^*_{0B} = b^*_L$. It follows that $(p, b^*_B) = (0.333, 0.05)$, is the core allocation to which the sequence of bargaining equilibria reported in Table 4 converges.

It is apparent from Table 4 that the expected level of benefits is roughly constant in equilibrium for every value of $\theta_B$, but the tax-rate financing this benefit varies with the share of good jobs going to minorities; the total number of good jobs created in the economy rises with $\theta_B$ permitting the tax-rate to fall as $\theta_B$ goes up without reducing the benefit rate. As the average level of human capital among black workers approaches that among white workers, the room for affirmative action vanishes and the bargaining solution converges to the policy ideal of the group with median income. Tables 3 and 4 reflect different ways such convergence might occur depending on the political influence of high blacks in the bargaining process.

### 4.2.3 Economic variation

The final comparative static concerns the degree of economic efficiency ($q$). 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.

\[\text{It is worth noting that increases in the technical cost parameter } q \text{ have a qualitatively identical impact on equilibrium outcomes as increases in the relative bargaining strength of workers, } \beta.\]
with reductions in $q$. Exactly this intuition shows up in Table 5, where $\lambda = 0$ and $\theta_B = 1/5$.

Table 5 here

The interesting feature to note in Table 5 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.

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 absent such a sociological complication. 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 that 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 members of
the poorly educated members of the majority.

The political comparative static exercises (especially Tables 2 and 4) suggest the legisla-
tive structure of minority representation is consequential; the greater the weight of blacks
with high human capital the more we expect to see redistribution through affirmative action
rather than through fiscal policy. In the equilibrium, the typical coalition supporting positive
affirmative action comprises the high-type whites and blacks, the one conceding on affirma-
tive action in exchange for the other limiting demands for fiscal redistribution. *Prima facie,*
this coalitional prediction seems at odds with casual empiricism; at least in the USA, it is
precisely the party most closely associated with high-income whites that is most resistant
to affirmative action policy. In a study of the history and evolution of affirmative action
in the US, however, David Skrentny (1996:5) writes that “[t]hough civil rights and African-
American groups may have supported affirmative action as the preferred civil rights measure
since at least the 1970s, the policy is largely the construction of white male elites who tra-
ditionally have dominated government and business.” For example, it was the Department
of Labor of the Nixon administration that initiated the policy of requiring employers with
federal contracts to meet numerical goals and timetables for minority employment.20 Thus
the political reality does not completely mesh with the current political rhetoric. And an
important piece of the story Skrentny relates is that a motivation behind the legislative ini-
tiatives of the “white male elites” was in large part due to a desire to suppress the threat
of growing African-American civil disobedience in response to extreme economic inequities.
Although the model studied above is unable to capture the complexities of this interaction,
it nevertheless reflects the incentives of high whites to bargain through policy with the racial
minority.

20As Skrentny writes, “Perhaps the greatest irony of all in the story of affirmative action is that this
controversial model of justice owes its most advanced and explicit race-based formulation to a Republican
An important feature of our model is the effort to root political incentives and opportunities in the economy. All individuals are presumed rational utility-maximizers, unencumbered by racial prejudice. Legislative preferences over policy are induced from equilibrium behavior in the labor market, permitting a direct analysis of both the economic implications of political decisions and the political consequences of parametric variations in the economy. Certainly it is the case, as observed in Lemma 1 above, that affirmative action increases the minority, at the expense of majority, share of good jobs in the economy (Holzer and Neumark, 2000:506); on the other hand, “the use of production and cost function estimates to infer affirmative action effects on productivity has so far generated inconclusive [empirical] results” (Holzer and Neumark, 2000:535). Our (theoretical) results suggest that the direction of changes in efficiency due to affirmative action depend on just how binding a constraint is the policy. Proposition 1 states that economic efficiency is improved under mandated levels of affirmative action only slightly more restrictive than those implied by unfettered profit-maximization, but the converse is true for sufficiently high mandated levels of minority recruitment. To the extent that the laissez faire share of minorities with good jobs varies by industry or even firm, therefore, equivocal empirical estimates might reasonably be expected.

There several fairly obvious ways it is desirable to extend the framework suggested here, short of explicitly including racial preferences. Among these, two seem especially salient. First, given the prominence of differences in human capital in our framework, 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.21

21 See Roemer (2003) 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
And this motivates the second salient extension. It is a commonplace in the contemporary political economy literature that details of legislative and party structures are important for understanding policy outcomes. The policy prediction in the absence of a minority party or caucus derived in this paper, for example, is radically different from that with such an explicit and independent minority representation. Thus it is important to look more deeply at the impact of alternative assumptions on legislative decision making and party composition. In particular, we are interested in identifying the sorts of representation that best promote the interests of various subgroups within the population, both in the short run and in the longer term.
Appendix

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

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

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

\[
t = \frac{(1 - \sigma)b}{(1 - \sigma)b + E(w)}.
\]

where

\[
\sigma = p\sigma_B + (1 - p)\sigma_W
\]
is the share of workers with good jobs and

\[
E(w) = \beta \left[ p \int_{y_B}^{\infty} y \, dG_B(y) + (1 - p) \int_{y_W}^{\infty} y \, dG_W(y) \right] + (1 - \beta)\sigma b
\]
is the average wage. We denote the minority share of good jobs in the absence of affirmative
action by \( \alpha_0(b) < p. \)

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

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

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

\[
\begin{align*}
y_B(\alpha_0, b) &= y_W(\alpha_0, b) = y_0 = \frac{q}{1 - \beta} + b \\
y_B(p, b) &= y_0 - (1 - p)h(\theta_W - \theta_B) \\
y_W(p, b) &= y_0 + ph(\theta_W - \theta_B)
\end{align*}
\]
Proof of Lemma 1

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

$$
\begin{align*}
\frac{\partial y_B(\alpha, b)}{\partial \alpha} &= \frac{\alpha(1-p)f\Delta - (1-\alpha)\sigma}{\Psi}; \\
\frac{\partial y_W(\alpha, b)}{\partial \alpha} &= \frac{p(1-\alpha)f\Delta + \alpha\sigma}{\Psi}.
\end{align*}
$$

(11) (12)

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

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

and

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

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

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

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

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

$$
[1 - (y_W - \theta h)f] \geq [1 - (y_0 + (\theta_W - \theta_B)ph - \theta h)f]
$$

$$
= [1 - F(y_0)] + [2p\theta_B + (1 - 2p)\theta_W] hf > 0
$$

since $p < 1/2$.

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

$$
\begin{align*}
\frac{\partial y_B(\alpha, b)}{\partial b} &= \frac{\alpha(1-p)f}{\Psi} > 0 \\
\frac{\partial y_W(\alpha, b)}{\partial b} &= \frac{(1-\alpha)pf}{\Psi} > 0
\end{align*}
$$
To see that \( \partial y_W / \partial b \geq 1 \geq \partial y_B / \partial b \), observe that

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

while

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

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

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

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

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

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

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

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

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

\[
\frac{\partial \sigma}{\partial \alpha} = -f \left[ \frac{\partial y_B}{\partial \alpha} + (1 - p) \frac{\partial y_W}{\partial \alpha} \right]
\]

\[
= \frac{f}{\Psi} [(p - \alpha)\sigma - p(1 - p)f\Delta].
\]

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

\[
\frac{\partial E(w)}{\partial \alpha} = -f \left[ w(y_B) p \frac{\partial y_B}{\partial \alpha} + w(y_W)(1 - p) \frac{\partial y_W}{\partial \alpha} \right]
\]

\[
= w(y_B) \frac{\partial \sigma}{\partial \alpha} - f \left[ w(y_W) - w(y_B) \right] (1 - p) \frac{\partial y_W}{\partial \alpha}
\]

37
where \( w(y_i) = \beta y_i + (1 - \beta)b \) is the wage received by the least productive member of group \( i \) who is given a good job. When \( \alpha = \alpha_0 \), then \( w(y_B) = w(y_W) = w(y_0) \) and \( \partial E(w)/\partial \alpha = w(y_0) \partial \sigma/\partial \alpha > 0 \). When \( \alpha = p \), \( \partial E(w)/\partial \alpha < 0 \) since \( \partial \sigma/\partial \alpha < 0 \) when \( \alpha = p \) and \( \partial y_W/\partial \alpha > 0 \). In sum, both \( \partial \sigma/\partial \alpha \) and \( \partial E(w)/\partial \alpha \) are positive (negative) when \( \alpha = \alpha_0 \) (\( \alpha = p \)). Equation (14) then implies that \( \partial t/\partial \alpha < 0 \) (\( \partial t/\partial \alpha > 0 \)) when \( \alpha = \alpha_0 \) (\( \alpha = p \)).

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

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

Differentiating and collecting terms yields

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

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

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

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

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

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

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

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

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

\[
\frac{\partial E[c_{Hi}(\alpha, b)]}{\partial b} = - \frac{E[c_{Hi}(\alpha, b)]}{1 - t} \frac{\partial t}{\partial b} + (1 - t) [1 - \beta (1 - F(y_i - H))] - (1 - t) \beta (y_i - b) f \frac{\partial y_i}{\partial b} \\
\equiv -T_1(H,i) \frac{\partial t}{\partial b} + T_2(H,i) - T_3(i) \frac{\partial y_i}{\partial b}. \quad (19)
\]

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

(2) If \(\alpha = \alpha_0\) and \(f\) is a constant, then \(T_1(H,B) = T_1(H,W), T_2(H,B) = T_2(H,W)\) and \(T_3(B) = T_3(W)\). By Lemma 1(2), \(\partial y_W / \partial b > \partial y_B / \partial b > 0\) and \(\partial E[c_{HW}(\alpha_0, b)] / \partial b > \partial E[c_{HW}(\alpha_0, b)] / \partial b\). The result now follows by continuity of the derivatives in \(\alpha\). \(\square\)

**Proof of Lemma 3** Differentiating (18) with respect to \(\alpha\) yields

\[
\frac{\partial E[c_{Hi}(\alpha, b)]}{\partial \alpha} = - \frac{E[c_{Hi}(\alpha, b)]}{1 - t} \frac{\partial t}{\partial \alpha} - (1 - t) \beta (y_i - b) f \frac{\partial y_i}{\partial \alpha} \\
\equiv -T_1(H,i) \frac{\partial t}{\partial \alpha} - T_3(i) \frac{\partial y_i}{\partial \alpha}. \quad (20)
\]

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


Table 1: Equilibrium for Example 1

<table>
<thead>
<tr>
<th>Equilibrium with Affirmative Action</th>
<th>Affirmative</th>
<th>Minority Share</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prob.</td>
<td>Action Target</td>
</tr>
<tr>
<td>H's proposal</td>
<td>.61</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td>.39</td>
<td>.333</td>
</tr>
<tr>
<td>L's proposal</td>
<td>1</td>
<td>none</td>
</tr>
<tr>
<td>B's proposal</td>
<td>1</td>
<td>.332</td>
</tr>
<tr>
<td>Expected Value</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Equilibrium without Affirmative Action</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.12</td>
</tr>
</tbody>
</table>

| Laissez-faire                         | 0          | 0            | .306     | .482      |

Table 2: Comparative static on λ

<table>
<thead>
<tr>
<th>λ</th>
<th>Minority Share of Good Jobs</th>
<th>Average Benefit</th>
<th>Average Tax</th>
<th>Percentage Net Gain, Average HW</th>
<th>LW</th>
<th>HB</th>
<th>LB</th>
<th>Gini</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.314</td>
<td>.08</td>
<td>.14</td>
<td>2.8</td>
<td>-2.8</td>
<td>5.1</td>
<td>1.0</td>
<td>.392</td>
</tr>
<tr>
<td>.1</td>
<td>.316</td>
<td>.06</td>
<td>.11</td>
<td>4.2</td>
<td>-3.2</td>
<td>6.8</td>
<td>0.8</td>
<td>.410</td>
</tr>
<tr>
<td>.2</td>
<td>.317</td>
<td>.05</td>
<td>.09</td>
<td>5.1</td>
<td>-3.7</td>
<td>7.8</td>
<td>0.4</td>
<td>.422</td>
</tr>
<tr>
<td>.3</td>
<td>.317</td>
<td>.04</td>
<td>.07</td>
<td>5.8</td>
<td>-4.2</td>
<td>8.5</td>
<td>0</td>
<td>.433</td>
</tr>
<tr>
<td>.4</td>
<td>.317</td>
<td>.03</td>
<td>.06</td>
<td>6.5</td>
<td>-4.6</td>
<td>9.0</td>
<td>-0.8</td>
<td>.443</td>
</tr>
<tr>
<td>.5</td>
<td>.317</td>
<td>.02</td>
<td>.04</td>
<td>7.3</td>
<td>-5.6</td>
<td>10.0</td>
<td>-1.7</td>
<td>.458</td>
</tr>
</tbody>
</table>
Table 3: Comparative static on $\theta_B$; $\lambda \equiv 0$

<table>
<thead>
<tr>
<th>$\theta_B$</th>
<th>Minority Share of Good Jobs</th>
<th>Average Benefit</th>
<th>Average Tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>.2</td>
<td>.314</td>
<td>.08</td>
<td>.14</td>
</tr>
<tr>
<td>.25</td>
<td>.318</td>
<td>.09</td>
<td>.15</td>
</tr>
<tr>
<td>.3</td>
<td>.321</td>
<td>.09</td>
<td>.15</td>
</tr>
<tr>
<td>.35</td>
<td>.325</td>
<td>.10</td>
<td>.16</td>
</tr>
<tr>
<td>.4</td>
<td>.328</td>
<td>.11</td>
<td>.18</td>
</tr>
<tr>
<td>.45</td>
<td>.331</td>
<td>.12</td>
<td>.19</td>
</tr>
<tr>
<td>.495</td>
<td>.333</td>
<td>.14</td>
<td>.22</td>
</tr>
</tbody>
</table>

Table 4: Comparative static on $\theta_B$; $\lambda \equiv \theta_B$

<table>
<thead>
<tr>
<th>$\lambda \equiv \theta_B$</th>
<th>Minority Share of Good Jobs</th>
<th>Average Benefit</th>
<th>Average Tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>.2</td>
<td>.317</td>
<td>.05</td>
<td>.09</td>
</tr>
<tr>
<td>.25</td>
<td>.322</td>
<td>.05</td>
<td>.09</td>
</tr>
<tr>
<td>.3</td>
<td>.327</td>
<td>.05</td>
<td>.08</td>
</tr>
<tr>
<td>.35</td>
<td>.331</td>
<td>.05</td>
<td>.08</td>
</tr>
<tr>
<td>.4</td>
<td>.333</td>
<td>.05</td>
<td>.08</td>
</tr>
<tr>
<td>.45</td>
<td>.333</td>
<td>.05</td>
<td>.08</td>
</tr>
<tr>
<td>.495</td>
<td>.333</td>
<td>.05</td>
<td>.07</td>
</tr>
</tbody>
</table>
Table 5: Comparative static on $q$

<table>
<thead>
<tr>
<th>$q$</th>
<th>Minority Share of Good Jobs</th>
<th>Average Benefit</th>
<th>Average Tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>.25</td>
<td>.314</td>
<td>.08</td>
<td>.14</td>
</tr>
<tr>
<td>.22</td>
<td>.314</td>
<td>.12</td>
<td>.17</td>
</tr>
<tr>
<td>.19</td>
<td>.315</td>
<td>.15</td>
<td>.19</td>
</tr>
<tr>
<td>.16</td>
<td>.316</td>
<td>.19</td>
<td>.21</td>
</tr>
<tr>
<td>.13</td>
<td>.317</td>
<td>.23</td>
<td>.22</td>
</tr>
</tbody>
</table>
\[(\alpha_L, b_L) \quad (\alpha_B, b_B) \quad (\alpha_H, b_H) \quad (\alpha_{\mathcal{H}}, b_{\mathcal{H}}) \quad \alpha_0(0) \quad \alpha_0(b) \quad I_B \quad I_L \quad 0 \quad p \quad b\]