On Vertically Challenged and Horizontal Equity - Reassessing Anti-Discrimination Rules

Tomer Blumkin
Yoram Margalioth
Efraim Sadka

April 2004

IZA

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

On Vertically Challenged and Horizontal Equity – Reassessing Anti-Discrimination Rules

Tomer Blumkin
Ben-Gurion University

Yoram Margalioth
Tel Aviv University

Efraim Sadka
Tel Aviv University
and IZA Bonn

ABSTRACT

On Vertically Challenged and Horizontal Equity – Reassessing Anti-Discrimination Rules

We consider a model of prejudice-driven discrimination, where the advantaged 'tall' discriminate against the disadvantaged 'short'. We employ an egalitarian social welfare function to compare anti-discrimination legal rules with a non-discriminatory ('height-blind') income tax.

Corresponding author:

Efraim Sadka
The Eitan Berglas School of Economics
Tel-Aviv University
Tel-Aviv 69978
Israel
Email: sadka@post.tau.ac.il

1.   Introduction

Employer discrimination continues to play a role in generating different labor market outcomes by race and sex [Holzer and Neumark (2000)]. US Federal anti-discrimination laws[1] were at least partially successful in raising the relative earnings of blacks, especially in the southern states of the US during 1960-1975 period. However, the effect of such rules was more questionable since the 1980's [Donohue and Heckman (1991)]. The debate revolving around the effectiveness of anti-discrimination laws calls into attention the possibility of using complementary and/or alternative policy measures such as affirmative action and the tax and transfer system.

Proposals for using the tax and transfer system to redistribute wealth across individuals based on characteristics such as race, gender, national origin, religion, age, disability, etc., have been condemned on grounds of social meaning. Paying money to the discriminated individuals or paying money to employers to encourage ('bribe') them to employ individuals from those disadvantaged groups, are viewed as humiliating. 'No one wants to be told that his job is being subsidized, relative to the jobs of fellow workers; the polite way to describe the relationship is therefore to say that the recipient's civil rights are being protected against discrimination' [Kull (1994)]. Moreover, such transfers were viewed as equivalent to allowing whites to purchase permits to discriminate against blacks, therefore not a realistic policy tool, mentioned only to provoke discussion [Bell (1992)].

The manifested attitude of the legislator in favoring antidiscrimination laws over preferential tax treatment is guided by the principle of horizontal equity stating that 'those who are in all relevant senses identical should be treated identically' [Atkinson and Stiglitz

---

[1] U.S. Federal anti-discrimination laws include: Title VII of the Civil Rights Act of 1964, 42 U.S.C. §§ 2000e; The Age Discrimination in Employment Act (ADEA), 29 U.S.C §§ 621-634; The Americans with Disabilities Act (ADA), 42 USC§12101 et seq.; The Equal Pay Act of 1963, 29 U.S.C. § 206d; Pregnancy Discrimination Act, 42 USC§20; The Immigration Reform and Control Act of 1986 (IRCA), 8 USC§1324b.

(1980)]. This principle implies that individual characteristics, such as those mentioned above, should be ignored and all individuals should face the same tax schedule.

Indeed, anti-discrimination laws prevail over the tax and transfer system, because adjusting the tax base according to race, gender, national origin, etc., is viewed to be immoral. However, employing a simple model of discrimination, we nonetheless show that an appropriate tax and transfer system, namely, one that abides by the principle of horizontal equity, may dominate anti-discrimination laws, viewed from the disadvantaged point of view.

2.　　　A simple model of discrimination

We follow Becker (1971) seminal study in assuming that discrimination in the labor market is based on personal prejudice. Consider an economy with a population equally divided between 'tall' and 'short' individuals, who are otherwise identical.　We normalize the population of each type to unity. We assume that all employers are 'tall' agents. These 'tall' employers are also workers. 'Short' individuals are thus workers only. Vacancies (new jobs) are costless to create. We assume that production technology exhibits constant returns to scale. Denote the hourly productivity (hence wage rate) of a typical 'tall' worker by $\overline{w} > 0$. All individuals are equally productive, but, by virtue of discrimination, the productivity, hence the hourly wage rate of a typical 'short' worker, as perceived by a 'tall' employer, is given by $\underline{w} \equiv \overline{w} - k > 0$, where $k > 0$ measures the per-hour 'psychic' cost (in output terms) 'suffered' by a 'tall' employer who hires a 'short' worker.[2] Put differently, the discriminating 'tall' employer is willing to hire a 'short' employee only if she can pay her a wage rate ($\underline{w}$), which is below her productivity. All individuals share the same utility function given by:

---

[2]　There exists another form of discrimination, based on personal prejudice, which we choose to ignore for simplification reasons, namely, the one manifested by 'tall' employees against their 'short' counterparts. One possibility to warrant this is to assume that each firm is comprised of two workers – a 'tall' one who is also the owner/residual claimant and the 'short' worker. In such a case, the psychic cost, k, measure the overall disutility incurred by a 'tall' person, both as an employer and as a co-worker.

(1) $U = U(c, l)$,

where c denotes consumption (net of all 'psychic' costs), l denotes labor and the function U is concave, increasing with respect to c and decreasing with respect to l.[3] As is common in the optimal tax literature, denote by y the labor income of an individual, so that $l = y / w$. Thus, we re-write the utility function of an individual with a wage rate of w as:

(2) $V^w(c, y) = U(c, y / w)$.

Note that discrimination is the only source of heterogeneity (hence inequality) in the economy.

Both on moral grounds and in order to reduce the extent of earnings inequality in the economy, a social planner may naturally resort to impose anti-discrimination legal rules.

2.1   Anti-discrimination legal rules

Consider indeed the following simplified form of anti-discriminatory legislation in the spirit of Title VII of the 1964 Civil Rights Act. The law would require equal pay for equal workers and would also impose targets upon all employers to achieve mixing of 'tall' and 'short' in the workforce reflecting their shares in population. By virtue of the constant returns to scale assumption, one can assume, without loss of generality, that each typical firm is comprised of two workers, where compliance with equal opportunity implies that each firm will hire one 'short' worker and one 'tall' worker. The latter is also the owner of the firm (and its residual claimant). Equal pay cum equal opportunity implies, therefore, that the uniform wage rate across workers is given by $\hat{w} = (\overline{w} + \underline{w}) / 2$. The income of a 'tall' worker-employer is given by $\hat{w} \cdot \hat{l} + \hat{\pi}$, where $\hat{l}$ denotes her labor supply and $\hat{\pi}$ denotes the profit of her firm.[4] There is a

---

[3] This specification implicitly assumes that the utility of a 'tall' individual, $u(c'd, l)$, where c' denotes consumption and d - the 'psychic' cost associated with employing a 'short' individual (measured by the product of k and the number of working hours provided by the 'short' employee), takes the special form $u(c, l)$, where $c = c' - d$.

[4] With no loss in generality, we henceforth assume that the profits are not taxed (the implications of profit taxation are discussed in the appendix). Moreover, as the anti-discriminatory rule implies that the two types of

positive profit because both types of workers receive a wage rate $\hat{w}$, which is below their true common productivity $(\overline{w})$. Therefore, the profit is given by: $\hat{\pi} = 2 \cdot (\overline{w} - \hat{w}) \cdot \hat{l} = 2 \cdot (k/2) \cdot \hat{l} = k \cdot \hat{l}$. Note that this profit is exactly the total 'psychic' cost incurred by the 'tall' employer from having to employ the 'short' worker; denoted by d. Therefore, her consumption is equal to $\hat{w} \cdot \hat{l} + \hat{\pi} - d = \hat{w} \cdot \hat{l}$, that is her consumption is equal to her labor income. The same is true for the 'short' worker who makes no profit and suffers no 'psychic' costs (the latter follows as wage rates are equalized by law). For later purposes, denote the common consumption-labor income bundle by $(\hat{c}, \hat{y})$.

Anti-discriminatory policy of the form described above is not necessarily optimal, however. To show this, let us consider next a natural alternative remedy - a re-distributive income tax.

2.2 An income tax

Unlike the standard optimal income tax framework, where the source of inequality derives from unobserved variation in innate earning abilities, the 'height' characteristic is both observable and immutable. Thus, one could in principle implement a 'height' sensitive system. In such a case, each 'short' individual would be eligible for a transfer from the state, for example, through a refundable credit.[5] This transfer would be financed by a lump-sum adjustment to the income tax. Such a policy would constitute a differential lump-sum system, and would entail no deadweight loss. Being first best efficient, notwithstanding, such a system stands in sharp contrast to the notion of horizontal equity. We thus turn to examine a second-best 'height-blind' income tax system. We will show that even such a restricted income tax policy may dominate anti-discrimination rules.

---

workers would earn the same income in equilibrium, there would be no scope for a labor income tax as a complementary policy measure to the legal rule, in order to attain enhanced re-distribution.

[5] Or alternatively, the tax subsidy could be given to the employers of 'short' individuals.

As is common in the income tax literature, we describe the tax system as a set of gross labor income-net labor income bundles (y, z), where z denotes net labor income. A tax system has to satisfy the self-selection constraints, which state that each type of worker has no incentive to mimic the other type. Denote the (y, z)-bundles of the 'tall' and the 'short' type by $(\bar{y}, \bar{z})$ and $(\underline{y}, \underline{z})$, respectively. A 'tall' person must be as well-off as with the 'short' person's bundle. The consumption (net of 'psychic' costs) of a 'tall' person is given by $\bar{c} = \bar{z} + \pi - d = \bar{z}$, where the latter follows because $\pi = d$, by virtue of perfect competition. Similarly, the consumption of a 'short' person is given by $\underline{c} = \underline{z}$. The self-selection constraints are therefore given as follows:

(3) $\qquad V^{\bar{w}}(\bar{c}, \bar{y}) \geq V^{\bar{w}}(\underline{c}, \underline{y})$,

(4) $\qquad V^{\underline{w}}(\underline{c}, \underline{y}) \geq V^{\underline{w}}(\bar{c}, \bar{y})$.

The social planner is seeking to maximize some welfare measure, which strikes a balance between efficiency and equity considerations. To put things in the sharpest relief, we choose to focus on the *Rawlsian* planner case, exhibiting the highest degree of inequality (discrimination) aversion, seeking to maximize the well being of the least well-off individual ('short' agent in our case).[6] This seems to be a natural choice, reflecting perhaps the anti-discriminatory legislation key premise, that discrimination-driven inequality is categorically socially illegitimate. Following the conventional practice in the public finance literature, psychic costs of the form incurred by 'tall' employers in our setting, reflecting aversion towards 'short' employees should not count.[7] Therefore, these 'psychic' costs are 'laundered' out of the social welfare measure (Note, however, that our results hold also when no such 'laundering' out is done; see our discussion in section 2.3 below.). Accordingly, the social

[6] The main results of the paper can be derived under other welfare measures.

[7] For instance, the public finance literature on externalities may justify a tax on alcohol on the grounds that alcohol consumption may cause drunkenness, which is a real cost to society (generating violence, car accidents, etc.); but not on the grounds that some people may incur "psychic" costs because they view alcohol consumption as intrinsically immoral (say, on religious or other grounds).

planner adds to the consumption of a 'tall' employer the amount d. The social welfare function is therefore given by:

(5) $W = \min[V^{\bar{w}}(\bar{c} + d, \bar{y}), V^{\underline{w}}(\underline{c}, \underline{y})]$.

The *Rawlsian* social planner is thus seeking to maximize the social welfare function given in (5), subject to the self-selection constraints (3) and (4) and a revenue constraint:

(6) $(\bar{y} - \bar{c}) + (\underline{y} - \underline{c}) \geq 0$.

As is shown in the literature [see, e.g., Stiglitz (1982) and Balcer and Sadka (1982)], only one of the two self-selection constraints is binding, namely, the self-selection constraint for the 'tall' type [given by inequality (3)]. Zero marginal tax rate (efficiency) at the top income and strictly positive marginal tax rate at the bottom income are also shown in the literature [see, e.g., Balcer and Sadka (1982)].

2.3   Welfare comparison

The turn next to show that an income tax system may deliver a better social outcome than a legal anti-discrimination rule of the form described above. Because our social welfare function is *Rawlsian*, we in fact show that an optimal tax can even make the disadvantaged themselves better off, compared to the anti-discrimination legal rule.

Proposition 1: For sufficiently high levels of discrimination, that is, high 'psychic' costs, the anti-discriminatory legal rule is dominated by the *Rawlsian* optimal tax system.

Proof: See the appendix. QED

The rationale for the result is as follows. Suppose for a moment that the psychic costs were not laundered out. In this case, the marginal product of a 'short' worker is indeed lower than that of a 'tall' one from a social point of view, because of the negative externality exerted on the 'tall' employers, and Pareto-efficiency would indeed require paying them different wage rates. In this case, an anti-discrimination rule forces equal wages for workers with unequal marginal productivities from a social point of view. This causes inefficiency, but at the same time entails a benefit of achieving full equality. An income tax, on the other hand,

causes no leisure-consumption distortion at the top (for the 'tall' worker), but can never achieve full equality. Now, as the degree of discrimination rises, the distortion caused by anti-discrimination legislation magnifies, as the equal wage rate paid to both 'short' and 'tall' workers shifts further away from their social productivity levels (recall that the wage is set at the level of the average social productivity). At the same time, the leisure-consumption distortion of the income tax, which is confined to the bottom only (the 'short' worker), weakens, as the social productivity of the 'short' worker declines and becomes less relevant. As discrimination becomes more and more pronounced, the efficiency advantage of the income tax turn to dominate the egalitarian advantage of the anti-discrimination rule. Now, suppose that we launder out the psychic cost, as we indeed do. Nothing of the above argument changes, because the psychic costs affect the 'tall' individual whose welfare is ignored by our extremely egalitarian *Rawlsian* social welfare function.[8]

3.   Conclusion

We employed a model of discrimination based on pure prejudice à la Becker. We show that anti-discrimination legal rules may be inferior for the disadvantaged themselves relative to a non-discriminating tax and transfer system. This is true when focusing on ways to correct for the adverse consequences of discrimination. Nevertheless, one may argue that education is the best way to deal with the phenomenon of prejudice-driven discrimination. Such a policy is aimed at eradicating in the long-run prejudice itself, instead of correcting for its adverse consequences [see, for instance, Donohue (1998)]. In this respect, anti-discrimination rules could play a role in changing prejudiced attitudes.

---

[8]   Naturally, by virtue of continuity, the dominance of the income tax over anti-discrimination legislation when the latter is sufficiently high prevails even when the social welfare is not necessarily *Rawlsian*, but still exhibits a high enough degree of inequality (discrimination) aversion.

References

Atkinson, A. and J. Stiglitz (1980) "Lectures on Public Economics", London: McGraw-Hill

Balcer, Y. and E. Sadka (1982) "Horizontal Equity, Income Taxation and Self Selection with and Application to Income Tax Credit", Journal of Public Economics, 19, 291-309

Becker, G. (1971) "The Economics of Discrimination", Chicago: University of Chicago Press

Bell, D. (1992) "Faces at the Bottom of the Well: The Permanence of Racism", New York, Basic Books

Donohue, J. (1998) "Discrimination in Employment", The New Palgrave Dictionary of Economics and the Law 615, Peter Newman Edition

Donohue and Heckman (1991) "Continuous Versus Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Blacks", Journal of Economic Literature, 29, 1603-1643

Holzer and Neumark (2000) "Assessing Affirmative Action", Journal of Economic Literature, 38, 483-568

Kull, A. (1994) "The Discrimination Shibboleth", 31 San Diego L. Rev. 195

Mirrlees, J. (1971) "An Exploration in the Theory of Optimum Income Taxation", Review of Economic Studies, 38, 175-208

Stiglitz, J. (1982) "Self Selection and Pareto Efficient Taxation", Journal of Public Economics, 17, 213-240

Appendix

A1.   Proof of Proposition 1

Denote the welfare measures for the optimal tax system and the legal rule, respectively, by W^{Tax} and W^{Legal}. Maintaining the notation used in the text and employing the self-selection constraint in (3), it follows that:

(A1) $W^{Tax} = V^{\underline{w}}(\underline{c},\underline{y})$,

(A2) $W^{Legal} = V^{\hat{w}}(\hat{c},\hat{y})$.

We first turn to examine the case of extreme discrimination in which $\underline{w}$ is sufficiently small relative to $\overline{w}$ ($\underline{w} \to 0$, that is 'short' individuals are virtually excluded from the labor market), and show that the legal rule attains a lower level of welfare than the optimal tax system. To see this note that when $\underline{w} \to 0$, the optimal tax system involves no distortions. The standard result implies efficiency at the top (zero marginal tax rate on the top-earner – in our case the 'tall'-earner), and the fact that the 'short' workers are perceived to be un-productive by their 'tall' employers implies that they are crowded out of the market and obtain their entire income as a transfer from the government. Thus, we obtain a differential lump-sum tax and transfer system, which is superior to the distorting legal rule granting a wage subsidy to the 'short' workers and levying a wage tax on 'tall' ones. To see this formally, consider a standard two-type economy optimal tax framework (where differences in wage rates simply reflect differences in earning abilities). Consider the following two-part tariff tax and transfer system comprised of a wage tax/subsidy imposed on the high type and low type, respectively, denoted by $\overline{t}$ and $\underline{t}$, and a lump-sum tax/subsidy denoted correspondingly by $\overline{T}$ and $\underline{T}$. Let $V(w,t,T)$ denote the indirect utility derived by and agent of type *w*, faced by a two-part tariff (t, T).[a1] Suppose that the social planner has perfect information and is seeking to maximize a *Rawlsian* welfare measure. The maximization program would then be the following:

[a1]   Formally, $V(w,t,T) = \max\{U(c,l) + \lambda[l(w-t) - T - c]\}$, where λ denotes the *lagrange* multiplier.

$\max_{\bar{t},\underline{t},\bar{T},\underline{T}} \{\min[V(\underline{w},\underline{t},\underline{T}),V(\overline{w},\overline{t},\overline{T})]\}$

subject to:

$\overline{T} + \underline{T} + \overline{t}\cdot\overline{l} + \underline{t}\cdot\underline{l} \geq 0$,

where $\overline{l}$ and $\underline{l}$ denote the labor supply of the high type and low type worker, respectively. It is easy to show that in the optimal solution the revenue constraint is binding. Moreover, in the optimal solution it necessarily holds that $V(\overline{w},\overline{t},\overline{T}) = V(\underline{w},\underline{t},\underline{T})$. Re-writing the constrained maximization one thus obtains:

(A3) $\max_{\bar{t},\underline{t},\bar{T},\underline{T}} V(\underline{w},\underline{t},\underline{T})$

subject to:

$\overline{T} + \underline{T} + \overline{t}\cdot\overline{l} + \underline{t}\cdot\underline{l} = 0$
$V(\overline{w},\overline{t},\overline{T}) = V(\underline{w},\underline{t},\underline{T})$

Formulating the *Lagrangean*, we obtain:

$\max_{\bar{t},\underline{t},\bar{T},\underline{T},\mu,\eta} \{V(\underline{w},\underline{t},\underline{T}) + \mu[\overline{T} + \underline{T} + \overline{t}\cdot\overline{l} + \underline{t}\cdot\underline{l}] + \eta[V(\overline{w},\overline{t},\overline{T}) - V(\underline{w},\underline{t},\underline{T})]\}$,

with μ and η denoting the multipliers. Formulating the first order conditions we obtain (lower bar and upper bar refer correspondingly to high type and low type):

(A4) $\mu\cdot\overline{l} + \mu\cdot\overline{t}\cdot\partial\overline{l}/\partial\overline{t} - \eta\cdot\overline{\lambda}\cdot\overline{l} = 0$

(A5) $-\underline{\lambda}\cdot\underline{l}\cdot(1-\eta) + \mu\cdot\underline{l} + \mu\cdot\underline{t}\cdot\partial\underline{l}/\partial\underline{t} = 0$

(A6) $\mu - \overline{\lambda}\cdot\eta = 0$

(A7) $\mu - \underline{\lambda}\cdot(1-\eta) = 0$

Substituting (A7) and (A6) into (A5) and (A4), correspondingly, it follows that in the optimum, $\overline{t} = \underline{t} = 0$. However, note that for the case $\underline{w} = 0$, the constrained maximization in (A3) is equivalent to the optimal tax problem formulated in the text [maximization of (5) subject to the two binding constraints (3) and (6)], as $\underline{y} = 0$, hence $V^{\overline{w}}(\underline{c},\underline{y}) = V^{\underline{w}}(\underline{c},\underline{y})$, and the optimal tax system indeed involves no distortions and attains full equalization. Moreover, the outcome of the legal rule is equivalent to the result of the maximization program in (A3) if

we restrict ourselves to the use of a differential wage tax/subsidy. By virtue of the maximization results, the latter is however dominated by the differential lump-sum system, and the result follows. By virtue of continuity, the result extends to sufficiently small levels of $\underline{w}$.[a2] This concludes the proof. QED

A2. Profit Taxation

In the main body of the paper we assumed that profits were not taxed. In what follows we relax this assumption and turn to examine the implications of profit taxation on the optimal tax system. Suppose then that a proportional tax, denoted by t, is levied on profits. Obviously, psychic costs are not tax deductible. The constant returns-to-scale assumption implies that the tax burden is fully shifted to the workers. Moreover, the entire tax burden is shifted to the 'short' workers in equilibrium. To see this, suppose by negation, that part of the tax burden is shifted to 'tall' workers. Thus, the equilibrium wage rate of a typical 'tall' worker falls short of its productivity level $\overline{w}$. However, this obviously can not be equilibrium, for a typical entrepreneur (a 'tall' person) would like to attract, in this case, a 'tall' worker by offering her slightly more than the ostensible equilibrium wage rate, but less than her productivity level. It follows that a profit tax is equivalent to a wage tax imposed on the 'short' worker.[a3] By the same token, a profit subsidy is equivalent to a wage subsidy given to 'short' workers only. Relying on the latter equivalence, we show in what follows that a labor income tax cum profit subsidy entails a higher level of social welfare than anti-discrimination legal rules for any

---

[a2]   The range of parameters for which the optimal income tax system dominates the anti-discriminatory legal rule may be quite large. For instance, when the utility function is given by $u(c,l) = c - l^2/2$, then one can show that for any wage-ratio $\overline{w}/\underline{w} > 2$, the optimal tax system attains a higher welfare measure than the legal rule entails.

[a3]   To see this formally, note that when a proportional tax t is imposed on profits, the wage rate of a typical 'short' worker will adjust downwards until the point where the (non-deductible) 'psychic' costs will be just offset by net profits. In such a case, the wage rate of a 'short' worker will be given by $\underline{w} = \overline{w} - k/(1-t)$. The implicit wage tax will thus be given by $s = t \cdot k/(1-t)$.

degree of discrimination. Note that with anti-discrimination rules, a profit tax/subsidy is shifted to both types, because wages are forced to be equal by legislation. Therefore, a profit tax/subsidy is not useful in this case. We state and prove the following proposition:

Proposition A1: When profit taxation is allowed, the optimum tax system attains a strictly higher level of welfare than the one entailed by the anti-discriminatory legislation, for any degree of discrimination.

Proof: The proof is by means of an equivalence result. Consider the following tax system. Let s denote the wage subsidy provided to 'short' workers and let t denote a proportional labor income tax levied on all workers. Set s and t such that the following two conditions hold:

(A8) $\quad (\underline{w}+s)\cdot(1-t)=(\underline{w}+\overline{w})/2$,

(A9) $\quad \overline{w}\cdot(1-t)=(\underline{w}+\overline{w})/2$.

As both types of workers expect the same net wage rate per hour, they both choose the same workload. One can verify that the fiscal system is balanced, namely, the total tax revenues are equal to the amount of subsidy paid to the 'short' worker. To see that, let the common amount of labor chosen by both types be given by l. Then, we need to show that:

(A10) $\quad l\cdot[(\underline{w}+s)+\overline{w}]\cdot t=l\cdot s$.

However, by virtue of (A8) and (A9) it follows that $\overline{w}=\underline{w}+s$ and $t=1/2\cdot(\overline{w}-\underline{w})/\overline{w}$. Substitution into (A10) establishes the result. It follows that the tax system given by equations (A8) and (A9) is equivalent to the anti-discriminatory legal rule. However this system is not chosen optimally, and is therefore dominated by the optimum tax system (in the optimum, for example, the marginal tax rate on the high level of income should be set to zero). This concludes the proof. QED