Online Appendices for Ban The Box? Information, Incentives, and Statistical Discrimination

John W. Patty*

Elizabeth Maggie Penn†

November 15, 2022

*Professor of Political Science and Quantitative Theory & Methods, Emory University. Email: *jwpatty@gmail.com*.

†Professor of Political Science and Quantitative Theory & Methods, Emory University. Email: *elizabeth.m.penn@gmail.com*.

A Proofs

Proposition 1 *The following table characterizes all equilibria in which positive qualification can be obtained. When multiple equilibria exist, they are strictly Pareto ranked.*

Equilibria when $p > p_E^*$

Parameters (c_L, w, ϕ_0, ϕ_1)	Equilibria
$w\phi_0 > c_L > w\phi_1$	FQE with $\chi^* = 1, \eta^* = 1$, MSE with $\chi^* = \chi_M(p), \eta^* = \eta_M$, ZQE with $\chi^* = 0, \eta^* = 0$,
$w\phi_0 > c_L$ and $w\phi_1 > c_L$	FQE with $\chi^* = 1, \eta^* = 1$
$w\phi_1 > c_L > w\phi_0$	MSE with $\chi^* = \chi_M(p), \eta^* = \eta_M$,
$c_L > w\phi_0$ and $c_L > w\phi_1$	ZQE with $\chi^* = 0, \eta^* = 0$

Equilibria when $p < p_E^*$

Parameters (c_L, w, ϕ_0, ϕ_1)	Equilibria
$c_L > w\phi_1$	ZQE with $\chi^* = 0, \eta^* = 0$
$w\phi_1 > c_L$	FQE with $\chi^* = 1, \eta^* = 0$

Proof: We proceed through the six regions identified in the statement of the proposition. For the first four cases, note that when $p > p_E^*$, as defined in (11), E receives a strictly positive payoff from hiring $\theta = 2$ if all low types have chosen qualification. When $p < p_E^*$ then E receives a strictly negative payoff from hiring $\theta = 2$ if all low types have chosen qualification, and E will consequently never hire if observing $\theta = 2$.

REGION 1. $p > p_E^*$ and $w\phi_0 > c_L > w\phi_1$. Because $p > p_E^*$, $\eta = 1$ is a unique best response to $\chi = 1$. As $\phi_0 w > c_L$, Equation (7) is satisfied and $\chi = 1$ is a unique best response to $\eta = 1$. Consequently, there is an FQE with $\eta^* = 1$ and $\chi^* = 1$ and no other pure strategy equilibrium with full qualification. It is straightforward to verify that when $p > p_E^*$ and

$w\phi_0 > c_L > w\phi_1$, then $\eta_M \in (0,1)$ and $\chi_M(p) \in (0,1)$, where η_M and $\chi_M(p)$ are characterized by Equation (9) and (10). Therefore there also exists a mixed strategy equilibrium for this parameter region. Finally, Equation 8 does not hold in this case, as $c_L > w\phi_1$ and $\chi = 0$ is a best response to $\eta = 0$. It follows that there also exists a zero qualification equilibrium in this region.

REGION 2. $p > p_E^*$ and $w\phi_0 > c_L$ and $w\phi_1 > c_L$. As in the above case, $\eta^* = 1$ is a unique best response to $\chi^* = 1$ and vice versa because $p > p_E^*$ and $\phi_0 w > c_L$. However in this case there is no MSE, because when $w\phi_1 > c_L$ and $w\phi_0 > c_L$, E can't choose a hiring strategy η to make W indifferent between qualification and no qualification. Regardless of E's hiring strategy, it is always strictly optimal for W to choose $q = 1$.

REGION 3. $p > p_E^*$ and $w\phi_1 > c_L > w\phi_0$. In this case, there does not exist a pure strategy equilibrium. If $\chi = 1$ then E optimally chooses $\eta = 1$, as $p > p_E^*$. However, Equation (7) does not hold; when E hires those receiving $\theta = 2$ "aggressively" (i.e. $\eta = 1$) and when ϕ_0 is sufficiently low, W is incentivized to not obtain qualification. However, if W obtains no qualification then E will not hire if observing $\theta = 2$. In this case there is only a mixed strategy equilibrium, and again it is straightforward to verify that when $p > p_E^*$ and $w\phi_1 > c_L > w\phi_0$ then $\eta_M \in (0,1)$ and $\chi_M(p) \in (0,1)$.

REGION 4. $c_L > w\phi_0$ and $c_L > w\phi_1$. In this case Equation 6 can never obtain for any value of η. It follows that $\chi^* = 0$ and $\eta^* = 0$ is the unique equilibrium.

REGION 5. $p < p_E^*$ and $c_L > w\phi_1$. In these remaining two cases E always sets $\eta^* = 0$, because $p < p_E^*$. When $c_L > w\phi_1$ then Equation (8) does not hold, and W sets $\chi^* = 0$.

REGION 6. $p < p_E^*$ and $w\phi_1 > c_L$. In this last case Equation (8) *does* hold, and W sets $\chi^* = 1$.

PARETO RANKING EQUILIBRIA IN REGION 1. We conclude by ranking the 3 equilibria in Region 1 ($w\phi_0 > c_L > w\phi_1$ and $p > p_E^*$) according to the Pareto principle. It is straightforward to show that the full qualification equilibrium with $\chi^* = 1, \eta^* = 1$ Pareto dominates the mixed strategy equilibrium. To see this, note that at an MSE, the low-cost worker must be indifferent between obtaining qualification and not, and the employer must be indifferent between hiring a worker with $\theta = 2$ and not, and so must receive an expected payoff of zero conditional on $\theta = 2$. However, when $p > p_E^*$, E receives a strictly positive payoff in the FQE from hiring a worker with $\theta = 2$. Moreover, at the FQE there is a higher probability a randomly drawn worker will receive a $\theta = 3$ (as there is a higher probability $q = 1$), and a lower probability that a randomly drawn worker will receive $\theta = 0$. Thus, E receives a strictly higher expected payoff in the FQE than in the MSE.

In the MSE, a low-cost worker receives a (positive) expected payoff of $(1 - \phi_o)w\cdot = \frac{(1-\phi_0)(c_L - w\phi_1)}{\phi_0 - \phi_1}$. In the FQE, W receives an expected payoff of $w - c_L$. The difference between these payoffs is

$$w - c_L - \frac{(1 - \phi_0)(c_L - w\phi_1)}{\phi_0 - \phi_1} = \frac{(1 - \phi_1)(\phi_0 w - c_L)}{\phi_0 - \phi_1},$$

which, by inspection, is strictly positive when $w\phi_0 > c_L > w\phi_1$. Therefore, a low-cost worker strictly prefers the FQE to the MSE when both equilibria exist and high-cost workers also strictly prefer the FQE to the MSE, because at the FQE the employer is hiring all workers who receive $\theta = 2$, which strictly benefits workers who are not qualified. Finally, note that both the FQE and MSE are strictly Pareto superior to the ZQE, in which both players receive a payoff of 0 with certainty. ∎

A Lemma. The following lemma is applied in the proof of Proposition 2.

Lemma 1 *Regardless of which group a worker belongs to, and whether the box is used or not, W's expected payoff from the potential mixed strategy equilibrium profile,* $(\chi^*, \eta^*) = (\chi_M(p), \eta_M)$, *is independent of the worker's realized cost of qualification,* $c \in \{c_L, c_H\}$, *and equal to the following:*

$$EU_W(MSE) \equiv \frac{(1 - \phi_0)(\phi_1 w - c_L)}{\phi_1 - \phi_0},$$

while W's conditional expected payoff in a full qualification equilibrium with conservative hiring, given $c \in \{c_L, c_H\}$, is:

$$EU_W(FQE \mid \eta^* = 0, c) \equiv \begin{cases} \phi_1 w - c_L & \textit{if } c = c_L \\ 0 & \textit{if } c = c_H. \end{cases}$$

Proof: Any mixed strategy equilibrium is characterized by Equations 9 and 10, with group potential p varying depending on the group being considered (i.e. whether it is a subgroup with potential p_g or the set of all workers with potential $\overline{p}$). Note that E's mixed strategy, η_M^*, is not a function of group potential. E is simply making W indifferent between qualification and no qualification, and this indifference is solely dependent on costs to qualification, wages, and the testing technology, all of which are invariant to the presence or absence of the box.[1]

Any worker playing an MSE will receive an expected payoff of

$$\chi_M(w(\phi_1 + (1-\phi_1)\eta_M) - c) + w(1-\chi_M)(1-\phi_0)\eta_M,$$

which reduces to $\frac{(1-\phi_0)(\phi_1 w - c)}{\phi_1 - \phi_0}$ (*i.e.*, the group's potential, p, drops out of the equation). Therefore the MSE payoff to the worker is independent of the worker's group identity or the presence or absence of the box.

Finally, in any FQE with conservative hiring an unqualified worker receives a payoff of zero and a qualified worker receives a payoff of $\phi_1 w - c_L$. In this case expected payoffs again are independent of group identity. ∎

Proposition 2 *If the test is positively informative ($\phi_1 \geq \frac{c_L}{w} > \phi_0$), the groups are statistically distinct ($p_1 \geq p_E^* > p_2$), and population potential is low ($\overline{p} < p_E^*$), then W and E have opposed preferences over the box: E prefers that the box be present, W prefers that the box be banned.*

[1]Note that, conditional on W and E playing the MSE, W is indifferent about his or her cost of becoming qualified, c. This is because, in our setting, a worker with low costs of qualification is essentially choosing whether to "act like he or she must have a low cost of qualification" ($q = 1$) or "act like he or she might have had a high cost of qualification" ($q = 0$). The worker has a strict preference in equilibrium for a low cost of qualification only in an FQE.

Proof: When $\overline{p} < p_E^*$ then banning the box will generate an FQE with conservative hiring of all individuals. By Lemma 1 we know that W always prefers the MSE to the FQE with conservative hiring. By the supposition that $\phi_1 w - c_L > 0$ and $\phi_0 < \phi_1 \leq 1$, this follows from the fact that

$$\frac{(1-\phi_0)(\phi_1 w - c_L)}{\phi_1 - \phi_0} \geq \phi_1 w - c_L > 0.$$

Accordingly, both high and low-cost workers prefer the MSE, implying that workers in group 1 are made strictly worse off with the box, and workers in group 2 are indifferent about the box's presence.

The employer's payoff is affected by BTB solely through the change induced in group 1's qualification strategy behavior by BTB, because E was previously at a conservative hiring FQE with group 2 when the box was present. With the box, E received an expected payoff from hiring from group 1 equal to:

$$p_1((B-w)\chi_M(\phi_1 + (1-\phi_1)\eta_M) - w(1-\chi_M)(1-\phi_0)\eta_M) - w(1-p_1)(1-\phi_0)\eta_M.$$

This can be reduced to

$$EU_E(\text{MSE}|g=1) = \frac{(1-\phi_0)\phi_1(B-w)w}{B(1-\phi_1) + w(\phi_1 - \phi_0)}. \tag{1}$$

At the FQE, E's expected payoff from hiring from group 1 is

$$EU_E(\text{FQE}, \eta^* = 0|g=1) = p_1(\phi_1)(B-w).$$

Comparing these two payoffs we get that:

$$EU_E(\text{FQE}, \eta^* = 0|g=1) \geq EU_E(\text{MSE}|g=1)$$

when

$$\phi_1(B-w)\left(p_1 - \frac{w(1-\phi_0)}{B(1-\phi_1) + w(\phi_1 - \phi_0)}\right) \geq 0,$$

or

$$\phi_1(B-w)\left(p_1 - p_E^*\right) \geq 0.$$

Since we have supposed that $p_1 > p_E^*$, this inequality always holds. Therefore E receives a weakly higher payoff (strictly higher if $p_1 > p_E^*$) from banning the box in this case. ∎

Proposition 3 *If the test is positively informative ($\phi_1 \geq \frac{c_L}{w} > \phi_0$), the groups are statistically distinct ($p_1 \geq p_E^* > p_2$), and population potential is high ($\overline{p} > p_E^*$), then BTB is Pareto dominant, strictly benefiting E and group 2 workers, and leaving the payoffs of group 1 workers unchanged.*

Proof: Note that workers from group 1 are indifferent about banning the box when $\overline{p} > p_E^*$. First E's mixed equilibrium strategy, η_M, is unchanged regardless of whether the box is present or not. Second, with or without the box, workers in the advantaged group play an MSE with E. We demonstrate the result by considering each groups of workers in turn, followed by the employer.

WORKERS IN THE ADVANTAGED GROUP. The equilibrium probability that a low-cost worker becomes qualified when the box is banned, $\chi_M(\overline{p})$, is higher than it is for workers from group 1 when the box is present. This is because the population's potential, $\overline{p}$, is less than p_1 and therefore these workers must become qualified at a higher rate in order to keep E indifferent in the absence of the box when considering whether to hire a worker who received a test score of $\theta = 2$. However, this higher rate of qualification by workers from group 1 has no effect on their equilibrium expected payoffs because these workers are indifferent between qualification and no qualification in equilibrium regardless of the box's presence:

$$\begin{aligned} EU_W(\text{MSE} \mid g = 1) &= p_1 \cdot \left(\chi_M(\overline{p})(\phi_1 w + (1 - \phi_1)\eta_M - c_L) + (1 - \chi_M(\overline{p}))(1 - \phi_0)\eta_M w\right) \\ &\quad +(1 - p_1)(1 - \phi_0)\eta_M w, \\ &= p_1 \cdot \left((1 - \phi_0)\eta_M w\right) + (1 - p_1)(1 - \phi_0)\eta_M w, \\ &= (1 - \phi_0)\eta_M w. \end{aligned}$$

WORKERS IN THE DISADVANTAGED GROUP. Turning to workers in group 2, Lemma 1 implies that both low-and high-cost workers in group 2 receive a strictly higher payoff in the MSE. This is the Pareto efficient equilibrium if the box is banned, implying that workers from group 2 strictly benefit from banning the box.

THE EMPLOYER. The employer strictly prefers to ban the box in this setting. His or her payoff from banning the box in the MSE is equivalent to his or her payoff from the workers that send $\theta = 3$; this is because for E to mix conditional on $\theta = 2$, E must be receiving an expected payoff of zero conditional on $\theta = 2$. Consequently, E's expected payoff with the box banned is

$$\begin{aligned} EU_E(\text{MSE}) &= (B - w)\left(\gamma p_1 \chi_M(\overline{p})\phi_1 + (1-\gamma)p_2\chi_M(\overline{p})\phi_1\right), \\ &= (B - w)\cdot \overline{p} \cdot \chi_M(\overline{p})\phi_1, \\ &= (B - w)\phi_1 p_E^*. \end{aligned}$$

In the presence of the box, the employer's expected payoff in the Pareto efficient equilibrium in this case is

$$\begin{aligned} EU_E(\text{FQE}) &= (B - w)\left(\gamma p_1 \chi_M(p_1)\phi_1 + (1-\gamma)p_2\phi_1\right), \\ &= (B - w)\left(\gamma p_1 \frac{w(1-\phi_0)}{p_1(B(1-\phi_1) + w(\phi_1 - \phi_0))}\phi_1 + (1-\gamma)p_2\phi_1\right), \\ &= (B - w)\phi_1\left(\gamma p_E^* + (1-\gamma)p_2\right). \end{aligned}$$

Accordingly, by the supposition that $p_2 < p_E^*$, it follows that $\gamma p_E^* + (1-\gamma)p_2 < p_E^*$, implying that $EU_E(\text{MSE}) > EU_E(\text{FQE})$, so that E strictly benefits from BTB.

Thus, relative to the expected payoff from the Pareto efficient equilibrium with the box present, the expected payoff from the Pareto efficient equilibrium with BTB is

1. identical for workers from the advantaged group,

2. strictly higher for workers from the disadvantaged group, and

3. strictly higher for the employer.

Accordingly, BTB is Pareto dominant in this case, as was to be shown. ∎

Proposition 4 *When the test is negatively informative ($\phi_0 > \frac{c_L}{w} > \phi_1$), the groups are statistically distinct ($p_1 \geq p_E^* > p_2$), and population potential is low ($\overline{p} < p_E^*$), BTB is Pareto inefficient.*

Proof: When $\overline{p} < p_E^*$, the inequality in Equation 5 fails to hold and $\eta^*(\varnothing) = 0$; E hires conservatively from the group at large. As Inequality 6 doesn't hold when $\frac{c_L}{w} > \phi_1$, it follows that $\chi^*(1) = \chi^*(2) = 0$ and the effect of banning the box is to shut the labor market down entirely. No worker obtains qualification, and no worker is hired. This leaves payoffs for workers in group 2 unchanged. E and workers in group 1 are strictly worse off than they were with the box. As described above, with the box both E and workers from group 1 received a strictly positive expected payoff. ∎

Proposition 5 *When the test is negatively informative ($\phi_0 > \frac{c_L}{w} > \phi_1$), the groups are statistically distinct ($p_1 \geq p_E^* > p_2$), and population potential is high ($\overline{p} \geq p_E^*$), BTB strictly benefits group 2 workers and leaves the payoffs of group 1 workers unchanged.*

Proof: By satisfaction of Equation 5, E hires aggressively from the group at large when $\overline{p} \geq p_E^*$. And by satisfaction of Equation 7, all workers obtain qualification and an FQE exists with $\eta^*(\varnothing) = \chi^*(1) = \chi^*(2) = 1$. The payoff to members of group 1 at this equilibrium is identical to their payoff when the box was present. However every member of group 2 is strictly better off in expectation. With the box, all members of group 2 received a payoff of zero. Without the box, high cost individuals in group 2 receive an expected payoff of $w(1 - \phi_0) > 0$ and low-cost individuals receive a payoff of $w - c_L > 0$. ∎

Proposition 6 *When the test is negatively informative ($\phi_0 > \frac{c_L}{w} > \phi_1$), the groups are statistically distinct ($p_1 \geq p_E^* > p_2$), and population potential is high ($\overline{p} \geq p_E^*$), BTB is Pareto dominant if*

$$p_2 \in \left[\frac{w(1 - \phi_0)}{B - w\phi_0}, p_E^*\right),$$

and E is hurt by BTB if

$$p_2 < \frac{w(1 - \phi_0)}{B - w\phi_0}.$$

Proof: Banning the box strictly benefits E when E's expected payoff from hiring members of group 2 aggressively is positive, but when it is not sequentially rational for E to hire group 2 aggressively. This implies that $p_E^* > p_2$ but

$$(B - w) - w(1 - p_2)(1 - \phi_0) \geq 0.$$

This latter inequality is satisfied when

$$p_2 \geq \frac{w(1 - \phi_0)}{B - w\phi_0}. \tag{2}$$

Thus, when $p_2 \in \left[\frac{w(1-\phi_0)}{B-w\phi_0}, p_E^*\right)$ banning the box Pareto dominates the box, leaving members of group 2 strictly better off; members of group 1 indifferent; and E weakly better off. When $p_2 < \frac{w(1-\phi_0)}{B-w\phi_0}$ E is strictly made worse off by the box, as E's receives a negative payoff from hiring from group 2. ∎

B When the Box Has No Effect in Equilibrium (Section 5.1)

Statistically Non-Distinct Group Potentials. The logic behind Corollary 1 is relatively straight-forward. Because $\overline{p}$ is a convex combination of p_1 and p_2, both groups' potentials being greater than p_E^* implies that $\overline{p}$ is also greater than p_E^*. In this case, E will use the same hiring strategy for each group if group identity is observed, and E will also use this same strategy in the event that group identity is not observed ($g = \varnothing$). Consequently, if both groups have high potentials ($p_1 > p_2 > p_E^*$), then BTB can have no effect on equilibrium behavior when comparing the Pareto optimal equilibrium in each case.[2] The same logic follows if both groups have low potential ($p_E^* > p_1 > p_2$): in this case, the employer will always use a conservative hiring strategy in equilibrium, *regardless of whether the box is present or not*. Note that, as we do throughout, Corollary 1 restricts our comparisons to Pareto efficient equilibria. This focus separates our analysis from that provided by Coate and Loury (1993) (and many other models of Arrovian statistical discrimination), because in that model, the causal mechanism for discrimination operates through the role of a worker's group membership as an equilibrium selection device.

[2]As mentioned above, we are focusing on Pareto efficient equilibria throughout so that the presence or absence of the box does not have an effect on outcomes merely as an equilibrium selection device.

Uninformative Testing Structures. When the test is uninformative, its result is so noisy that it is not in any worker's interest to invest in qualification. The most straightforward example of this scenario would be when ϕ_1 and ϕ_0 both approach zero. In the limit, every applicant would receive a test score of 2 regardless of qualification status, and no applicant would choose to become qualified.

Corollaries 1 and 2 separately indicate the theoretical limits of BTB as a policy tool for ameliorating discrimination in hiring and, more fundamentally, illustrate the "informational foundations" of BTB's impact (or lack thereof) on equilibrium qualification and hiring. Corollary 1 highlights that BTB can have an impact on statistical discrimination only if the employer's beliefs about the two groups are statistically distinct, implying that the employer would treat workers from the two groups differently conditional on a garbled test result even if the employer believes that workers from both groups were obtaining qualification whenever qualification is not strictly dominated. Corollary 2 clarifies that BTB can have an effect only if the testing structure is sufficiently precise.

Taken together, the two results indicate that "coarsening" the employer's information by obscuring an applicant's group membership can have an impact on outcomes only if the employer's information about applicants — encompassing both his or her prior beliefs about the groups' potentials and his or her interim information about the applicant in question's true qualification — is sufficiently rich. Put another way, Corollary 1 states that the box can have an impact only if the employer might treat workers from the two groups differently even if they are using the same strategy to obtain qualification, and Corollary 2 states that the box can have impact only if the employer's information about any given worker's qualification is sufficiently precise for the employer to actually condition upon the test result when making his or her hiring decision.

C Incorporating Other Sources of Discrimination

In this appendix we present two examples to demonstrate that our qualitative result that BTB can represent a Pareto improvement over "the box" extends to two different settings in which group potential is identical across the two groups. The first is a setting of taste-based discrimination, in which the benefit to the employer of hiring qualified individuals from each group differs, so that $B_1 > B_2$. The second is a setting in which the testing structure is noisier for one group than another.

Example 1 (BTB with Taste-based Discrimination.) Modify the model as presented in the body of the article in two ways as follows. First, assume that group potential is identical across the two groups ($p_1 = p_2 = p$). Second, assume that the employer receives a higher payoff from hiring a qualified member of group 1 than group 2 ($B_1 > B_2$). This implies that the employer has an intrinsic **taste for discrimination**. FInally, to keep the analysis simple, let $p = \frac{1}{3}$, $B_1 = 2$, $B_2 = \frac{19}{12}$, $w = 1$, $c_L = \frac{1}{2}$, $\phi_0 = \frac{3}{4}$, and $\phi_1 = \frac{1}{4}$.

Suppose $\chi_i = 1$, so that all low-cost individuals from group i attain qualification. By Equation (5) the employer will only hire an individual receiving a test score of $\theta = 2$ when

$$\Pr[q = 1|\theta = 2] = \frac{(1-\phi_1)p}{(1-\phi_1)p + (1-\phi_0)(1-p)} \geq \frac{w}{B_i},$$

or when

$$B_i \geq \frac{5}{3}.$$

Consequently, E will hire from group 1 aggressively, and from group 2 conservatively.

By Equations (7) and (8), group 1 will attain full qualification, and group 2 will attain no qualification. However, if the box is banned then E hires aggressively from the group at large when $\gamma B_1 + (1-\gamma)B_2 \geq \frac{5}{3}$, or when $\gamma > \frac{1}{5}$. If all individuals attain qualification – which they will in this example when E hires aggressively, by Equation (7) – then E will receive a payoff from group 2 members equal to

$$p(B_2 - w) - w(1-p)(1-\phi_0) = \frac{1}{3}\left(\frac{19}{12} - 1\right) - \frac{2}{3}\left(\frac{1}{4}\right) = \frac{1}{8} > 0.$$

Thus E strictly prefers to ban the box. While group 1 is indifferent about the box, members of group 2 strictly prefer to ban the box. With the box, all group 2 individuals receive a payoff of 0 in equilibrium. When the box is banned, low cost individuals in group 2 receive a certain payoff of $w - c_L = \frac{1}{2}$, and high cost individuals receive an expected payoff of $w(1-\phi_0) = \frac{1}{4}$. $\triangle$

Example 2 (BTB with Differentially Accurate Tests.) Modify the model as presented in the body of the article in two ways as follows. First, assume that group potential is identical across the two groups ($p_1 = p_2 = p$). Second, the testing structure is noisier for group 2 than for group 1, with $\phi_1(i)$ and $\phi_0(i)$ denoting the informativeness of the test for group i. When $\phi_0(1) \neq \phi_0(2)$ and/or $\phi_1(1) \neq \phi_1(2)$, the employer's posterior beliefs about qualification for any muddled test result ($\theta = 2$) will be sensitive to the employee's group membership even if the groups make identical qualification decisions. FInally, to keep the analysis simple, let $\phi_0(1) = \frac{3}{4}$, $\phi_0(2) = \frac{9}{16}$, and $\phi_1(1) = \phi_1(2) = \frac{1}{4}$. All other parameters are the same across groups, with $p = \frac{1}{3}$, $B = 2$, $w = 1$, and $c_L = \frac{1}{2}$.

Using our results for group 1 from the previous example, when E can observe group identity then E hires from group 1 aggressively, and members of group 1 attain full qualification. However, again by Equation (5) the employer will only hire an individual receiving a test score of $\theta = 2$ when

$$\Pr[q = 1|\theta = 2] = \frac{(1 - \phi_1(i))p}{(1 - \phi_1(i))p + (1 - \phi_0(i))(1 - p)} \geq \frac{w}{B_i},$$

or when

$$\phi_0(i) \geq \frac{5}{8}.$$

Again, E will hire from group 2 conservatively, and no one in group 2 will attain qualification (again, by Equation (8)).

If the box is banned then E hires aggressively from the group at large when $\gamma\phi_0(1) + (1 - \gamma)\phi_0(2) \geq \frac{5}{8}$, or when $\gamma > \frac{1}{3}$. If all individuals attain qualification – which they will in this example when E hires aggressively, as Equation (7) is satisfied for both groups – then E will receive a payoff from group 2 members equal to

$$p(B - w) - w(1 - p)(1 - \phi_0(2)) = \frac{1}{3}(2 - 1) - \frac{2}{3}\left(\frac{7}{16}\right) = \frac{1}{24} > 0.$$

Thus E strictly prefers to ban the box. While group 1 is indifferent about the box, members of group 2 also strictly prefer to ban the box. With the box, all group 2 individuals receive a payoff of 0 in equilibrium. When the box is banned, low cost individuals in group 2 receive a certain payoff of $w - c_L = \frac{1}{2}$, and high cost individuals receive an expected payoff of $w(1 - \phi_0) = \frac{7}{16}$. △

References

Coate, Stephen and Glenn C Loury. 1993. "Will Affirmative-action Policies Eliminate Negative Stereotypes?" *The American Economic Review* pp. 1220–1240.