Supplementary Appendices for Test-Optional Admissions

Wouter Dessein[*] Alex Frankel[†] Navin Kartik[‡]

June 6, 2025

[*]Columbia University, Graduate School of Business. Email: wd2179@columbia.edu.
[†]University of Chicago, Booth School of Business; Email: afrankel@chicagobooth.edu.
[‡]Columbia University, Department of Economics; Email: nkartik@columbia.edu.

The in-print appendices of the paper consist of Appendix A and Appendix B; hence, this document begins with Appendix C.

C. Details for Banning Affirmative Action

This section substantiates the discussion in Section IV.A of the paper.

C.1. A Model of Affirmative Action

There are two potentially observable non-test dimensions, $x = (x_0, x_1)$. Dimension x_0 is binary, with realizations in $\{r, g\}$ (red and green). Dimension x_1, which may represent some aggregate of GPA and/or extra-curricular achievement, takes continuous values in $\mathbb{R}$. Test scores are binary, with values normalized to 0 and 1.

The college and society have identical preferences over all factors except for the type dimension x_0. Society does not care about this dimension, but all else equal, the college wants to admit green types over red types.[1] Specifically, we assume that

$$\begin{aligned} u^s(x,t) &= x_1 + t, \\ u^c(x,t) &= x_1 + t + \beta \mathbb{1}_{x_0=g} - c, \end{aligned}$$

with $\beta > c > 0$, and $\mathbb{1}_{x_0=g}$ an indicator for green types. The parameter β is the bonus the college gives to green types over red types. The parameter c is not essential to our analysis, but it allows for the college and society to have different test-score bars for both red and green students. It can be interpreted as the (opportunity) cost for a college of admitting any student. We have normalized the analogous constant in society's utility to zero. The assumption $\beta > c > 0$ implies that the college has a lower test-score bar than society for green types and a higher one for red types. Note that the the college's ex-post utility is

$$u^*(x,t) = x_1 + t + \frac{\beta}{1+\delta}\mathbb{1}_{x_0=g} - \frac{c}{1+\delta}.$$

Let $x_0 = g$ with probability $q \in (0,1)$ and $x_0 = r$ with probability $1-q$. We assume that the distribution of test scores depends on x only through x_0: $\Pr(t = 1 | x = (x_0, x_1)) = p_{x_0} \in (0,1)$. Our primary interest is in the case of $p_r > p_g$, meaning that green types, which

[1] We could allow for society to have preferences over a student's x_0 dimension as well; what is important is that the college favors green types more than society does.

are favored by the college, have a worse distribution of test scores. This may correspond to green students being an underrepresented demographic group, for instance. But we also allow for the opposite case of $p_r < p_g$, in which the college's favored group has a better test score distribution. Here, green students may correspond to those from rich families, who have better access to test preparation, and are favored by the college because of donor considerations. If the green students correspond to legacy applicants, it may be that either $p_r < p_g$ or $p_r > p_g$.

We take x_1 to be independent of both x_0 and t. We also assume that x_1 is uniformly distributed over a large enough interval. Specifically, $x_1 \sim U[\underline{x}_1, \overline{x}_1]$, with $\underline{x}_1 < c - \beta - 1$ and $\overline{x}_1 > c$. The inequality on $\underline{x}_1$ guarantees that there are students with x_1 low enough that neither the college nor society wants to admit them, even if they are otherwise as desirable as possible ($x_0 = g$ and $t = 1$). The inequality on $\overline{x}_1$ guarantees that there are students with x_1 high enough that the college and society want to admit them even if they are otherwise as undesirable as possible ($x_0 = r$ and $t = 0$).

We will consider the college's choice over whether to be test mandatory or test blind in two observability regimes. First, we allow both dimensions of x to be observable, which we call *affirmative action allowed.* Then we consider only x_1 to be observable, with the dimension x_0 unobservable; we call this regime *affirmative action banned.* We interpret the switch from the first to the second regime as a policy change where society—which does not intrinsically care about x_0—bans the use of that dimension in admissions. This may represent a law or court decision forbidding the use of race or legacy status in admissions.[2]

C.2. Results

Affirmative action allowed. Consider first the case when affirmative action is allowed.

Under test mandatory, the college can choose a distinct threshold of x_1 above which to admit students at each (x_0, t) pair.[3] This threshold is determined by setting the ex-post utility to 0. Since the college favors green students, its x_1 threshold will be lower by $\beta/(1+\delta)$

[2] Note that we assume that when x_0 is unobservable to the college, it is also unobservable to society. While society does not value x_0 directly, the observability of x_0 to society could still matter for the calculation of the college's social costs. This is because, if society can observe x_0 but cannot observe test scores, then it would expect a different test score for green students (p_g) than red students (p_r). We assume that a law preventing the college from making inferences of this form also stop society from making/penalizing the college based on such inferences.

[3] Since we will be comparing test mandatory with test blind, it turns out to be convenient for our analysis to take the perspective of x_1 admissions thresholds rather than test score thresholds.

for green students than for red students at each score level t. From society's perspective, the college uses an x_1 threshold that is too low for green students and too high for red students—but crucially, the gap between society's preferred threshold and what the college uses does not vary with t.[4]

Under test blind, the college chooses an admissions threshold on dimension x_1 that depends on the student's type x_0 but not the test score t. However, x_0 is informative about t: the college and society evaluate students of type x_0 as if they have the expected test score $\mathbb{E}[t|x_0] = p_{x_0}$. If $p_r > p_g$, the college's preference for green students is countered by the fact that green students have lower test scores on average than red students. So the college will now use a lower x_1 threshold for green students than red students only if its preference parameter β is sufficiently large: specifically, if and only if $\beta/(1+\delta) > p_r - p_b$. Regardless, the gap between the college's chosen x_1 threshold and society's preferred threshold is the same as under test mandatory, for any test score t—that gap did not depend on the test score, and utilities are linear in the test score.

We can establish:

Proposition C.1. *If affirmative action is allowed, then the college prefers test mandatory to test blind.*

The reason is that going test blind leads to a set of students that the college prefers less, but in the current specification there is never a countervailing benefit of reducing disagreement cost. The latter point stems from two sources. First, as noted above, for any given x_0 type (and test score, under test mandatory), the gap between society's preferred x_1 threshold and what the college uses is independent of the regime, even though these thresholds do shift across regimes. Second, our assumption of a uniform distribution of x_1 means that the total disagreement cost for students of a given x_0 type (at a given test score, or averaging over test scores) only depends on the size of the gap.

Affirmative action banned. Now consider the case when affirmative action is banned.

Under test mandatory, the observed test score is informative about a student's type x_0. Specifically, since there are a fraction q of green types in the population and the probability of test score $t = 0$ for a student of type x_0 is $1 - p_{x_0}$, we compute the probability of a student

[4] The gap is $(\beta - c)/(1 + \delta)$ for green students and $c/(1 + \delta)$ for red students.

being green conditional on $t = 0$ as

$$P_g^0 := \Pr(x_0 = g | t = 0) = \frac{q}{q + (1-q)\frac{1-p_r}{1-p_g}}.$$

Analogously, conditional on $t = 1$, the probability of a green type is

$$P_g^1 := \Pr(x_0 = g | t = 1) = \frac{q}{q + (1-q)\frac{p_r}{p_g}}.$$

Let $\Delta := P_g^0 - P_g^1$ be the difference between these two quantities, i.e., a low test score implies a Δ higher probability of $x_0 = g$ than a high test score. Note that $\Delta > 0$ if $p_r > p_g$, whereas $\Delta < 0$ if $p_r < p_g$. Based on the inference of x_0 from t, the college's underlying utility gives a bonus of $\beta\Delta$ to students with low test scores relative to those with high scores. As a result, the college now values a high test score $1 - \beta\Delta$ units higher than a low score, whereas society still values it 1 unit higher. That is, unlike when affirmative action is allowed, the gap between society's preferred x_1 admissions threshold and what the college chooses now varies with the test score.[5] We impose the assumption that $\beta\Delta < 1$, so the college still prefers students with higher test scores.

There is now an avenue for test blind to help the college. Under test blind, since the college evaluates all students as having $\Pr(x_0 = g) = q$ and $\mathbb{E}[t] = qp_g + (1-q)p_r$, it is as if the college's utility from any student is $x_1 + \mathbb{E}[t] + q\beta - c$. Analogously, it is as if society's utility from any student is $x_1 + \mathbb{E}[t]$. If $c = q\beta$, which means the college and the society seek to admit the same number of students overall, then it is as if their utilities agree, and the college implements its preferred admissions policy—subject to being test blind and no affirmative action—at zero disagreement cost. More generally, the disagreement cost is always lower under test blind than test mandatory. Whether the reduced disagreement cost outweighs the allocative loss from being test blind depends on parameters, specifically the intensity of social pressure δ and the college's bonus to low-scoring students $\beta\Delta$.

Proposition C.2. *Suppose affirmative action is banned. If $(1+\delta)(2\beta\Delta - 1) \geq (\beta\Delta)^2$, then the college prefers test blind, and otherwise the college prefers test mandatory.*

[5] Absent affirmative action, it is as if the college's underlying utility from a student is $x_1 + t + \beta P_g^t - c$, and so the college's gain from a student with test score $t = 1$ over $t = 0$ is $1 + \beta P_g^1 - \beta P_g^0 = 1 - \beta\Delta$. Given its underlying utility, the college's ex-post utility from a student is $x_1 + t + \frac{\beta P_g^t - c}{1+\delta}$. The gap between the college's chosen x_1 admissions threshold with society's preference is the term $\frac{\beta P_g^t - c}{1+\delta}$, which varies with t so long as $P_g^0 \neq P_g^1$, or equivalently $\Delta \neq 0$.

Recall we assume $\beta\Delta < 1$. Proposition C.2 implies that if $\beta\Delta \leq 1/2$, the college always prefers test mandatory: the allocative losses ("admission mistakes") from not observing test scores are larger than those from simply implementing society's preferred decision rule and incurring no disagreement. When $\beta\Delta \in (1/2, 1)$, there is a trade-off, and test blind will be preferred if the intensity of social pressure, δ, is sufficiently large. The following corollary develops this and other comparative statics.

Corollary C.1. *Suppose that affirmative action is banned (x_0 is unobservable) and that a low test score is associated with $x_0 = g$ ($\Delta > 0$).*

1. *There is some $\beta^* \in \left(\frac{1}{2\Delta}, \frac{1}{\Delta}\right)$ such that the college prefers test mandatory when $\beta < \beta^*$ and prefers test blind when $\beta > \beta^*$.*
2. *There is some $\Delta^* \in \left(\frac{1}{2\beta}, \frac{1}{\beta}\right)$ such that the college prefers test mandatory when $\Delta < \Delta^*$ and prefers test blind when $\Delta > \Delta^*$.*
3. *If $\beta\Delta \leq 1/2$, then the college prefers test mandatory for all δ; if $\beta\Delta \in (1/2, 1)$, then there is some $\delta^* > 0$ such that the college prefers test mandatory when $\delta < \delta^*$ and prefers test blind when $\delta > \delta^*$.*

C.3. Society's Preferences

We now consider society's payoff under different affirmative action and testing regimes. Society's realized utility for an individual student is $Au^s(x, t)$, where the dummy variable A indicates whether the student is admitted. We assume that society's objective is to maximize its expected utility across the pool of applicants.

Proposition C.3. *Society's preferences over affirmative action and testing regimes are as follows:*

1. *Fixing the testing regime as mandatory or blind, society prefers banning affirmative action to allowing affirmative action.*
2. *Fixing affirmative action as banned or allowed, society prefers test mandatory to test blind.*
3. *Suppose society chooses the affirmative action regime and then the college chooses the testing regime. Then banning affirmative action can harm society. In particular, if $\beta\Delta \in (1/2, 1)$, there exist thresholds $0 < \underline{\delta} \leq \overline{\delta} < \infty$ such that (i) if affirmative action*

is banned, the college chooses test blind if $\delta > \underline{\delta}$, and (ii) society is harmed by banning affirmative action if $\delta > \overline{\delta}$, while it benefits if $\delta < \overline{\delta}$.[6]

The first two parts of the proposition are intuitive, since society does not want the admission decision to depend on whether a student is red or green (which suggests part 1) but does want the decision to depend on the test score (which suggests part 2). If society could choose both the testing and affirmative action regimes, it would ban affirmative action and choose test mandatory. However, part 3 of the proposition cautions that if society chooses the affirmative action regime and the college subsequently chooses the testing regime, society can be worse off by banning affirmative action. Specifically, when δ is large enough, banning affirmative action backfires because the college's response of going test optional results in a student pool that society likes less than under test mandatory and affirmative action allowed. Indeed, as δ gets arbitrarily large, society's payoff is arbitrarily close to society's first best when affirmative action is allowed and there is mandatory testing, while it is bounded away when affirmative is banned and the college goes test blind. But when δ is intermediate (between the thresholds $\underline{\delta}$ and $\overline{\delta}$ in Proposition C.3 part 3), society is better off by banning affirmative even though it results in the college going test blind.[7]

C.4. Proofs for Results on Banning Affirmative Action

As a preliminary observation, we can write the college's loss relative to first best as its allocative loss plus the cost of social pressure. At a given (x_0, t) pair of test scores and group memberships, the assumption of a uniform distribution over x_1 implies that the college's allocative loss depends only on the difference between the college's chosen x_1-cutoff for admission and the college's ideal x_1 cutoff. Specifically, let $f := \frac{1}{\overline{x}_1 - \underline{x}_1}$ be the (constant) density of the x_1 distribution on its support. If the college's chosen cutoff is r above its ideal cutoff, then its allocative loss on this (x_0, t) pair is

$$\int_0^r f x \mathrm{d}x = \frac{f}{2} r^2. \tag{C.1}$$

Society's (allocative) loss is given by the same formula, when the chosen cutoff is r above society's preferred cutoff.

[6] If $\beta\Delta \leq 1/2$, the college never goes test blind, and so, by part 1 of the proposition, society always benefits from banning affirmative action.

[7] It is possible that $\underline{\delta} = \overline{\delta}$, in which case whenever a ban on affirmative action leads to test optional, society is harmed by the affirmative-action ban.

Proof of Proposition C.1. Suppose that affirmative action is allowed. Here, there is no interaction between the college's decisions at different realizations of x_0. So, it suffices to show that test mandatory would be preferred to test blind for any fixed $x_0 = x_0'$ in $\{r, b\}$.

Fixing $x_0 = x_0'$, let $h := u^c(x_0', x_1, t) - u^s(x_0', x_1, t) = \beta \mathbb{1}_{x_0'=g} - c$ be the difference between the college's and society's utility for admitting a student of type $x_0 = x_0'$, which does not depend on x_1 or t. It then holds that $u^c(x_0', x_1, t) - u^*(x_0', x_1, t) = \frac{\delta}{1+\delta}h$, and that $u^*(x_0', x_1, t) - u^s(x_0', x_1, t) = \frac{1}{1+\delta}h$. Given its information, the college sets x_1 admissions cutoffs at the value of x_1 setting the expectation of $u^*(x_0', x_1, t)$ to 0. Note that the college's ideal x_1-cutoff for students in group $x_0 = x_0'$ with test score t is $-t - h$, whereas society's ideal x_1-cutoff is $-t$.

The college's loss under test mandatory. At (x_0', t), the college's chosen x_1-cutoff for admission is $\frac{\delta}{1+\delta}h$ above its ideal point, yielding allocative loss (from (C.1)) of

$$\frac{f}{2}\frac{h^2\delta^2}{(1+\delta)^2}. \tag{C.2}$$

Similarly, the college's chosen x_1-cutoff for admission is $-\frac{1}{1+\delta}h$ above society's ideal point, leading to an allocative loss for society of $\frac{f}{2}\frac{h^2}{(1+\delta)^2}$. The college then pays a social pressure cost equal to δ times that, or

$$\frac{f}{2}\frac{\delta h^2}{(1+\delta)^2}. \tag{C.3}$$

Both of these expressions are independent of t, meaning that these expressions also represent the college's losses averaged over test scores.

The college's loss under test mandatory, for students with $x_0 = x_0'$, is the sum of (C.2) and (C.3).

The college's loss under test blind. With unobservable test scores, the players evaluate students of type $x_0 = x_0'$ as if they have the expected test score of $p_{x_0'}$. The college's chosen x_1 cutoff for students of type $x_0 = x_0'$ sets $u^*(x_0', x_1, p_{x_0'})$ to 0, i.e., a cutoff of $x_1 = -p_{x_0'} - \frac{h}{1+\delta}$.

To calculate the college's allocative losses, we compare the college's chosen (test-independent) x_1 admissions cutoffs to its (test-dependent) ideal cutoffs. Recall that the college's ideal cutoff at test score t is $x_1 = -t - h$. So at $t = 1$, the college's chosen cutoff is $1 - p_{x_0'} + \frac{\delta}{1+\delta}h$ above its ideal point; at $t = 0$, the college's chosen cutoff is $-p_{x_0'} + \frac{\delta}{1+\delta}h$ above its ideal point. The college's expected allocative loss over test scores, once again plugging into (C.1), is therefore given by

$$p_{x_0'}\frac{f}{2}\left(1-p_{x_0'}+\frac{\delta}{1+\delta}h\right)^2+(1-p_{x_0'})\frac{f}{2}\left(-p_{x_0'}+\frac{\delta}{1+\delta}h\right)^2$$
$$=\frac{f}{2}\frac{h^2\delta^2}{(1+\delta)^2}+\frac{f}{2}p_{x_0'}(1-p_{x_0'}). \tag{C.4}$$

To calculate social costs, we compare the college's chosen x_1 admissions cutoff not to society's ideal cutoff, but to society's preferred cutoff given that test scores are not observed. Society's preferred x_1-cutoff is given by $-p_{x_0'}$. The chosen cutoff is $-\frac{h}{1+\delta}$ above society's preferred cutoff. We can now plug into (C.1) to calculate society's loss relative to its preferred cutoff (given its information) as $\frac{f}{2}\frac{h^2}{(1+\delta)^2}$. The college's social pressure cost is δ times that, or

$$\frac{f}{2}\frac{\delta h^2}{(1+\delta)^2}. \tag{C.5}$$

The college's loss under test blind, for students with $x_0 = x_0'$, is the sum of (C.4) and (C.5).

Comparison. Comparing expressions (C.3) and (C.5), the social pressure cost under test blind is identical to that under test mandatory. Comparing expressions (C.2) and (C.4), the allocative loss is higher under test blind. Hence, the college prefers test mandatory. □

Proof of Proposition C.2. Suppose that affirmative action is banned. Let $ET := \mathbb{E}[t] = qp_r + (1-q)p_g$ be the average test score in the population, i.e., the share with test score $t = 1$. Recall that $P_g^t = Pr(x_0 = g|t)$. We will now calculate the college's loss in each testing regime.

In each case, we will evaluate the college's allocative loss relative to a benchmark where the college must make decisions independently of the unobservable x_0 type. The college's ideal cutoff at test score t, given that it must pool together students across the two x_0 types, is $-t - \beta P_g^t + c$.

The college's loss under test mandatory. Society's ideal x_1-cutoff for admitting a student of with test score t is $-t$. The college's chosen cutoff, setting the expected ex post utility to 0, is $-t - \frac{1}{1+\delta}(\beta P_g^t - c)$.

To calculate the allocative loss, observe that the college's chosen cutoff is $\frac{\delta}{1+\delta}(\beta P_g^t - c)$ above its ideal cutoff at test score t. Plugging into (C.1), its allocative loss across the two test scores is given by

$$(1 - ET)\frac{f}{2}\left(\frac{\delta}{1+\delta}(\beta P_g^0 - c)\right)^2 + ET\frac{f}{2}\left(\frac{\delta}{1+\delta}(\beta P_g^1 - c)\right)^2. \tag{C.6}$$

To calculate the loss due to social pressure, observe that the chosen x_1-cutoff is $-\frac{1}{1+\delta}(\beta P_g^t - c)$ above society's preferred cutoff. The college's expected loss due to social pressure (plugging this difference into (C.1) for each test score, taking expectation over test scores to find society's loss, and then multiplying by δ) is therefore

$$\delta(1 - ET)\frac{f}{2}\left(\frac{1}{1+\delta}(\beta P_g^0 - c)\right)^2 + \delta ET\frac{f}{2}\left(\frac{1}{1+\delta}(\beta P_g^1 - c)\right)^2. \tag{C.7}$$

The college's total loss is (C.6) plus (C.7).

The college's loss under test blind. The average test score is ET, and so society's preferred x_1-cutoff is $-ET$. The college's chosen cutoff, setting the expected ex post utility to 0, is $-ET - \frac{1}{1+\delta}(\beta q - c)$, where q is the probability of $x_0 = g$.

Again, we calculate the college's allocative loss relative to its ideal point with observable t but unobservable x_0. At test score t, the chosen cutoff minus the ideal cutoff is

$$t - ET + \beta P_g^t - \frac{q\beta}{1+\delta} - \frac{c\delta}{1+\delta}$$

Plugging into (C.1) and taking the expectation across test scores, the college's allocative loss is given by

$$(1 - ET)\frac{f}{2}\left(-ET + \beta P_g^0 - \frac{q\beta}{1+\delta} - \frac{c\delta}{1+\delta}\right)^2 + ET\frac{f}{2}\left(1 - ET + \beta P_g^1 - \frac{q\beta}{1+\delta} - \frac{c\delta}{1+\delta}\right)^2. \tag{C.8}$$

The difference between the college's chosen cutoff and society's preferred cutoff is $-\frac{1}{1+\delta}(\beta q - c)$.

Plugging into (C.1) and multiplying by δ, the college's loss from social pressure is

$$\frac{f}{2}\frac{\delta(\beta q - c)^2}{(1+\delta)^2}. \tag{C.9}$$

The college's total loss is (C.8) plus (C.9).

Comparison. The net benefit of choosing test blind rather than test mandatory is given by the loss from test mandatory minus the loss from test blind, i.e.,

$$(\text{C.6}) + (\text{C.7}) - (\text{C.8}) - (\text{C.9}).$$

Substituting in $q = (ET)P_g^1 + (1-ET)P_g^0$ and $\Delta = P_g^0 - P_g^1$ and then simplifying, we can rewrite this net benefit as

$$\frac{f}{2}\frac{ET(1-ET)}{1+\delta}\left((1+\delta)(2\beta\Delta - 1) - (\beta\Delta)^2\right).$$

The above expression is weakly positive if and only if $(1+\delta)(2\beta\Delta - 1) \geq (\beta\Delta)^2$. □

Proof of Corollary C.1. Suppose that affirmative action is banned. Proposition C.2 establishes that the college prefers test blind to test mandatory if and only if

$$(1+\delta)(2\beta\Delta - 1) \geq (\beta\Delta)^2. \tag{C.10}$$

Recall we maintain the assumptions that $\beta > 0$, $\beta\Delta < 1$, and for this corollary, $\Delta > 0$. We prove each part of the corollary in turn:

1. Rewriting (C.10), the college prefers test blind if and only if

$$-\Delta^2\beta^2 + 2\Delta(1+\delta)\beta - (1+\delta) \geq 0.$$

   The LHS is a concave quadratic that is negative at $\beta = \frac{1}{2\Delta}$ (equal to $-1/4$) and positive at $\beta = \frac{1}{\Delta}$ (equal to δ). Hence, there exists $\beta^* \in (\frac{1}{2\Delta}, \frac{1}{\Delta})$ such that the college prefers test blind when $\beta > \beta^*$ and test mandatory when $\beta < \beta^*$. Using the quadratic formula, $\beta^* = \frac{1+\delta-\sqrt{\delta(1+\delta)}}{\Delta}$.

2. Since (C.10) is symmetric with respect to β and Δ, the argument of the previous part goes through unchanged after swapping β and Δ. We get $\Delta^* = \frac{1+\delta-\sqrt{\delta(1+\delta)}}{\beta}$.

3. If $\beta\Delta \in (0, 1/2)$, then the LHS of (C.10) is nonpositive and the RHS is strictly positive, implying that test mandatory is optimal.

   If $\beta\Delta > 1/2$, then we can rewrite (C.10) as $\delta \geq \frac{(1-\beta\Delta)^2}{2\beta\Delta-1}$, and hence the result holds for $\delta^* = \frac{(1-\beta\Delta)^2}{2\beta\Delta-1} > 0$. □

Proof of Proposition C.3. As in (C.1), at a given (x_0, t), society's loss relative to its first best when the college's chosen x_1-cutoff for admission is r above society's ideal cutoff is $\int_0^r fx\mathrm{d}x = \frac{f}{2}r^2$. Society's expected loss across all values of x_0 and t is equal to the expectation of $\frac{f}{2}r^2$ over the distribution of r, with r the difference between the chosen cutoff (which may depend on x_0 and t) and society's ideal cutoff (which depends only on t). Since the loss $\frac{f}{2}r^2$ is convex in r, mean-preserving spreads in the distribution of these cutoff differences make society worse off.

Part 1. Fix any testing regime. The distribution of cutoffs at each test score when affirmative action is allowed is a mean-preserving spread of the distribution when affirmative action is banned. Hence, society prefers banning affirmative action.

Part 2. First, suppose that affirmative action is allowed. Fix some type $x_0 = x_0'$, at which the college has a utility bonus of $h := u^c(x_0', x_1, t) - u^s(x_0', x_1, t) = \beta \mathbb{1}_{x_0'=g} - c$ relative to society. Under test mandatory, at each test score, the chosen x_1-cutoff is $\frac{h}{1+\delta}$ below society's ideal cutoff. Under test blind, at $t = 1$, the chosen cutoff is $1 - p_{x_0'} + \frac{h}{1+\delta}$ below society's cutoff; and at $t = 0$, the chosen cutoff is $-p_{x_0'} + \frac{h}{1+\delta}$ below society's cutoff. Hence, under test blind, at each type x_0', the distribution of society's cutoff minus the chosen cutoff is given by

$$\begin{cases} 1 - p_{x_0'} + \frac{h}{1+\delta} & \text{with probability } p_{x_0'} \\ -p_{x_0'} + \frac{h}{1+\delta} & \text{with probability } 1 - p_{x_0'}. \end{cases}$$

This distribution is a mean-preserving spread of the constant $\frac{h}{1+\delta}$. Hence, society is worse off under test blind for each realization x_0' of x_0, and so is worse off in expectation.

Next, suppose that affirmative action is banned. As also defined in the proof of Proposition C.2, we let $ET := \mathbb{E}[t] = qp_r + (1-q)p_g$ denote the average test score in the population, i.e., the share of students with test score $t = 1$. At test score t, the college's ideal x_1-cutoff is $-t - \beta P_g^t + c$ (recall $P_g^t = \Pr(x_0 = g|t)$), and society's ideal x_1-cutoff is $-t$.

Under test mandatory with affirmative action banned, the college's chosen x_1-cutoff is $\frac{1}{1+\delta}(\beta P_g^t - c)$ below society's ideal point at test score t. That is, a share ET of students have cutoffs $\frac{1}{1+\delta}(\beta P_g^1 - c)$ below society's ideal point, and a share $1-ET$ have cutoffs $\frac{1}{1+\delta}(\beta P_g^0 - c)$ below. Plugging in $q = (ET)P_g^1 + (1-ET)P_g^0$ and $\Delta = P_g^0 - P_g^1$, some algebra yields that the distribution of society's ideal cutoffs minus the chosen cutoffs is

$$\begin{cases} \frac{1}{1+\delta}(\beta q - c) - (1-ET)\frac{\beta\Delta}{1+\delta} & \text{with probability } ET \\ \frac{1}{1+\delta}(\beta q - c) + ET\frac{\beta\Delta}{1+\delta} & \text{with probability } 1-ET. \end{cases} \tag{C.11}$$

Under test blind with affirmative action banned, the college's chosen x_1 cutoff is $-ET - \frac{1}{1+\delta}(\beta q - c)$. This means that for the ET share of students with $t=1$, the chosen x_1-cutoff is $\frac{1}{1+\delta}(\beta q - c) - (1-ET)$ below society's ideal cutoff of -1; for the $1-ET$ share with $t=0$, the chosen cutoff is $\frac{1}{1+\delta}(\beta q - c) + ET$ below society's ideal cutoff of 0. That is, the distribution of society's ideal cutoffs minus the chosen cutoffs is

$$\begin{cases} \frac{1}{1+\delta}(\beta q - c) - (1-ET) & \text{with probability } ET \\ \frac{1}{1+\delta}(\beta q - c) + ET & \text{with probability } 1-ET. \end{cases} \tag{C.12}$$

Since $\beta\Delta < 1$ (by assumption) and $1+\delta > 1$, the distribution in (C.12) is a mean-preserving spread of that in (C.11). Hence, when affirmative action is banned, society prefers test mandatory to test blind.

Part 3. From Proposition C.2, if $(1+\delta)(2\beta\Delta - 1) < (\beta\Delta)^2$, then the college chooses test mandatory under an affirmative action ban. If $(1+\delta)(2\beta\Delta - 1) > (\beta\Delta)^2$, which implies $\beta\Delta > 1/2$, the college chooses test blind under an affirmative action ban.

So, when $\beta\Delta \in (0, 1/2]$, society prefers to ban affirmative action: it prefers test mandatory and no affirmative action to test mandatory with affirmative action (by part 1).

Now suppose that $\beta\Delta > 1/2$. Let $\underline{\delta} := \frac{(\beta\Delta)^2}{2\beta\Delta - 1} - 1$ be the solution to $(1+\delta)(2\beta\Delta - 1) = (\beta\Delta)^2$. For $\delta < \underline{\delta}$, the college chooses test mandatory, in which case society prefers to ban affirmative action. For $\delta > \underline{\delta}$, the college chooses test blind. In this case, we need to compare society's payoff of test mandatory with affirmative action versus test blind without affirmative action.

The distribution of chosen x_1-cutoffs minus society ideal cutoffs under test mandatory with affirmative action is

$$\begin{cases} \frac{\beta - c}{1+\delta} & \text{with probability } q \\ \frac{-c}{1+\delta} & \text{with probability } 1-q. \end{cases}$$

Society's corresponding payoff loss is

$$\frac{f}{2(1+\delta)^2}\left(c^2 - 2cq\beta + q\beta^2\right). \tag{C.13}$$

The distribution of cutoffs minus society ideal points under test blind without affirmative action is given by (C.12). Society's payoff loss is correspondingly

$$\frac{f}{2(1+\delta)^2}\left((\beta q - c)^2 + (1-ET)ET(1+\delta)^2\right) \tag{C.14}$$

with $ET = qp_g + (1-q)p_r$.

The sign of(C.14) minus (C.13) tells us whether society prefers test mandatory with affirmative action or test blind without affirmative action. The sign of that difference is the same as the sign of $ET(1-ET)(1+\delta)^2 - q(1-q)\beta^2$. This expression equals zero when δ equals

$$\delta' := \beta\sqrt{\frac{q(1-q)}{ET(1-ET)}} - 1.$$

When $\delta > \delta'$, society prefers test mandatory with affirmative action to test blind without affirmative action; when $\delta < \delta'$, the preference is reversed.

Finally, let $\overline{\delta} := \max\{\underline{\delta}, \delta'\}$. We now see that (i) when $\delta > \underline{\delta}$, the college chooses test blind if affirmative action is banned; and (ii) taking into account the college's response in choosing its testing regime, society prefers to ban affirmative action if $\delta < \overline{\delta}$, and prefers to allow affirmative action if $\delta > \overline{\delta}$. $\square$

D. Competition Examples

This section provides three numerical examples to substantiate the discussion in Section IV.B about competition.

All three examples have two colleges. There is a single observable, and thus we omit the dependence of all variables on x. At this observable, students have test scores uniformly distributed between 0 and 100. Two identical colleges have underlying utility $u^c(t) = t - \underline{t}^c$; society has utility $u^s(t) = t - \underline{t}^s$; and the colleges place a weight $\delta = 1$ on social pressure, implying ex-post utilities $u^*(t) = t - \underline{t}^*$ with $\underline{t}^* = (\underline{t}^c + \underline{t}^s)/2$. A test-optional college is restricted to impute $\tau = 50$, the average test score, for nonsubmitters. If admitted by both colleges, a student chooses betwen them uniformly at random.

Example D.1 (Complements due to adverse selection)**.** Let $\underline{t}^c = 5$ and $\underline{t}^s = 25$, implying $\underline{t}^* = 15$. Here, society is more selective than the college. Our calculations below will establish strategic complementarity. That is, when a college's competitor is test mandatory, the college prefers to be test mandatory; but when a college's competitor is test optional, the college prefers to be test optional.

A test-mandatory college admits students with $t > \underline{t}^* = 15$ and rejects students with $t \leq 15$. Assume that a test-optional college admits nonsubmitters with $t \in [0, 50]$, which is optimal so long as the yield-weighted average test score of nonsubmitters is above $\underline{t}^* = 15$, as will be the case. The test-optional college also admits the submitting students with $t > 50$.

A test-mandatory college facing another test-mandatory college has yield of 1/2 for all of the students it admits, since the other college makes identical admission decisions. The college then gets underlying utility of $\frac{1}{2}\int_{15}^{100} \frac{1}{100}(t-5)\mathrm{d}t = 22.3125$, social pressure costs of $\frac{1}{2}\int_{15}^{25} \frac{1}{100}(25-t)\mathrm{t} = .25$, and a net payoff of $\simeq 22.1$.

A test-mandatory college facing a test-optional college also has yield of 1/2 for all of the students it admits, because the other college admits all students. So its payoff is also $\simeq 22.1$.

A test-optional college facing another test-optional college has yield of 1/2 for all students. The yield-weighted average test score for nonsubmitters is just the unweighted expectation of 25. The college's underlying utility from admitting every student is $\frac{1}{2}\int_{0}^{100} \frac{1}{100}(t-5)\mathrm{d}t = 22.5$, and social pressure costs are 0. So its payoff is 22.5.

Finally, a test-optional college facing a test-mandatory college has a yield of 1/2 for students with $t > 15$, and a yield of 1 for students with $t \leq 15$. The yield-weighted average test score for nonsubmitters is 20.9615.[8] The college's underlying utility from admitting every student is 22.5, and social pressure costs are $(25 - 20.9615) \cdot \frac{1}{100} \cdot (1 \cdot (15-0) + \frac{1}{2}(50-15)) = 1.3125$. Its payoff is $\simeq 21.2$.

We see that when a college's competitor is test mandatory, the college prefers to be test mandatory ($22.1 > 21.2$). When a college's competitor is test optional, the college prefers to be test optional ($22.5 > 22.1$). We also see that a test-optional college prefers its competitor to be test-optional ($22.5 > 21.2$).

Example D.2 (Substitutes due to adverse selection)**.** Let $\underline{t}^c = 50$ and $\underline{t}^s = 20$, implying $\underline{t}^* = 35$. Here, the college is more selective than society. Our calculations below will establish

[8] The average test score between 0 and 15 is 7.5; the average test score between 15 and 50 is 32.5; and the weighted average, putting a weight of 1/2 on test scores between 15 and 50, is $(7.5 \cdot (15-0) + 1/2 \cdot 32.5 \cdot (50-15))/((15-0) + 1/2 \cdot (50-15))$.

strategic substitutability. That is, when a college's competitor is test mandatory, the college prefers to be test optional; but when a college's competitor is test optional, the college prefers to be test mandatory.

A college that is test mandatory admits students with $t > \underline{t}^* = 35$ and rejects students with $t \leq 35$. Assume that a test-optional college rejects nonsubmitters with $t \in [0, 50]$, which is optimal so long as the yield-weighted average test score of nonsubmitters is below $\underline{t}^* = 35$, as will be the case. The test-optional college also admits the submitting students with $t > 50$.

A test-mandatory college facing another test-mandatory college has yield of 1/2 for students with $t > 35$ and yield of 1 for students with $t \leq 35$, since the other college makes identical admission decisions. The college then gets underlying utility of $\frac{1}{2}\int_{35}^{100} \frac{1}{100}(t-50)\mathrm{d}t = 5.6875$, social pressure costs of $\int_{20}^{35} \frac{1}{100}(t-20)\mathrm{d}t = 1.125$, and a net payoff of $\simeq 4.6$.

A test-mandatory college facing a test-optional college has a yield of 1/2 for students with $t > 50$ and a yield of 1 for students with $t \leq 50$. So it gets underlying utility of $\int_{35}^{50} \frac{1}{100}(t-50)\mathrm{d}t + \frac{1}{2}\int_{50}^{100} \frac{1}{100}(t-50)\mathrm{d}t = 5.125$, social pressure costs of $\int_{20}^{35} \frac{1}{100}(t-20)\mathrm{d}t = 1.125$, and a net payoff of 4.

A test-optional college facing another test-optional college has a yield of 1/2 for students with $t > 50$ and yield of 1 for students with $t \leq 50$. The yield-weighted average test score for nonsubmitters is just the unweighted expectation of 25. The college's underlying utility from rejecting nonsubmitters and accepting submitters with $t > 50$ is $\frac{1}{2}\int_{50}^{100} \frac{1}{100}(t-50)\mathrm{d}t = 6.25$, and social pressure costs are $\frac{1}{100}(25-20)(50-0) = 2.5$. So its payoff is 3.75.

Finally, a test-optional college facing a test-mandatory college has a yield of 1/2 for students with $t > 35$, and a yield of 1 for students with $t \leq 35$. The yield-weighted average test score for nonsubmitters is 21.9118.[9] The college's underlying utility from rejecting nonsubmitters and accepting submitters is 6.25, as above, and social pressure costs are $(21.9118 - 20) \cdot \frac{1}{100} \cdot (1 \cdot (35-0) + 1/2 \cdot (50-35)) = .812515$. Its payoff is $\simeq 5.4$.

We see that when a college's competitor is test mandatory, the college prefers to be test optional at this observable ($5.4 > 4.6$). When a college's competitor is test optional, the college prefers to be test mandatory ($4 > 3.75$). We also see that a test-optional college prefers its competitor to be test-mandatory ($5.4 > 3.75$).

[9] The average test score between 0 and 35 is 17.5; the average test score between 35 and 50 is 42.5; and the weighted average, putting a weight of 1/2 on test scores between 35 and 50, is $(17.5 \cdot (35-0) + 1/2 \cdot 42.5 \cdot (50-35))/((35-0) + 1/2 \cdot (50-35))$.

Example D.3 (Substitutes due to cherry picking). Let $\underline{t}^c = 39$ and $\underline{t}^s = 25$, implying $\underline{t}^* = 32$. Here, the college is again more selective than society. As in Example D.2, our calculations below will establish strategic substitutability, but the mechanism now owes to "cherry-picking" rather than adverse selection.

A college that is test mandatory admits students with $t > \underline{t}^* = 32$ and rejects students with $t \leq 32$. As in Example D.2, a test-optional college rejects nonsubmitters with $t \in [0, 50]$ and admits the submitting students with $t > 50$.

In contrast to Example D.2, the payoff of a test-optional college is now independent of its competitor's testing regime: regardless of whether the competitor is test mandatory or test optional, the college gets a yield of 1/2 for the submitters that it admits, and it faces no social pressure costs for the nonsubmitters that it rejects. (Society's bar is 25; and the yield-weighted average test score is 25 when the competitor is test optional, and it is below 25 when the competitor is test mandatory.) The test-optional college's payoff is thus $\frac{1}{2}\int_{50}^{100} \frac{1}{100}(t-39)\mathrm{d}t = 9$.

A test-mandatory college facing a test-mandatory competitor has a yield of 1/2 for students with $t > 32$ and a yield of 1 for students with $t \leq 32$, since the other college makes identical admission decisions. The college then gets underlying utility of $\frac{1}{2}\int_{32}^{100} \frac{1}{100}(t-40)\mathrm{d}t = 9.18$, social pressure costs of $\int_{25}^{32} \frac{1}{100}(t-25)\mathrm{d}t = 0.245$, and a net payoff of 8.935.

A test-mandatory college facing a test-optional competitor has a yield of 1/2 for students with $t > 50$ and yield of 1 for students with $t \leq 50$. So it gets underlying utility of $\int_{32}^{50} \frac{1}{100}(t-39)\mathrm{d}t + \frac{1}{2}\int_{50}^{100} \frac{1}{100}(t-39)\mathrm{d}t = 9.36$, social pressure costs of $\int_{25}^{32} \frac{1}{100}(t-25)\mathrm{d}t = 0.245$, and a net payoff of 9.115.

We see that, as in Example D.2, when a college's competitor is test mandatory, this college prefers to be test optional ($9 > 8.935$). When a college's competitor is test optional, this college prefers to be test mandatory ($9.115 > 9$). We also see that a test-mandatory college prefers its competitor to be test-optional ($9.115 > 8.935$), because that allows it to cherry-pick the students with $t \in (32, 50)$—whom it wants to admit, on average—without competition.

E. Connection to Bayesian Persuasion

In this section, we show that our college's payoff can be transformed into a familiar setup of Bayesian persuasion (Kamenica and Gentzkow, 2011). That is, we can view the college (sender) as having an indirect utility function over society's (receiver's) belief about the test score (state of the world).

However, we cannot simply apply standard Bayesian persuasion tools because there is a restricted set of information structures available. Instead of generating an arbitrary experiment about the test scores at each observable, our college can choose only an imputation level τ. This imputation then determines the rest of the information structure: all test scores below τ are pooled together, and all test scores above τ are revealed perfectly. (This information structure is sometimes called "lower censorship" in the persuasion literature; see Remark E.1 below.)

The transformation. Fixing some observable and omitting that for notational convenience, we can write our college's underlying utility as $u^c(t) = -\underline{t}^c + t$ and society's utility as $u^s(t) = -\underline{t}^s + t$ for a student with test score t. (At a fixed observable, we can normalize both parties' "weights" on the test score to 1.) For a student for whom the available information induces posterior belief $\mathbb{E}[t] = t^s$, the college's payoff from making admission decision $A \in \{0, 1\}$ can be written as

$$U^c(t^s, A) = Au^c(t^s) - \delta \cdot \begin{cases} -u^s(t^s) & \text{if } A = 1 \text{ and } t^s < \underline{t}^s \\ u^s(t^s) & \text{if } A = 0 \text{ and } t^s > \underline{t}^s \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 1 establishes that for a fixed information structure, the college's admission decision on the equilibrium path is made as if it maximizes the ex-post utility, $u^*(t) = u^c(t)/(1+\delta) + \delta u^s(t)/(1 + \delta) = -\underline{t}^* + t$, for $\underline{t}^* = (\underline{t}^c + \delta\underline{t}^s)/(1 + \delta)$. The college accepts students with an expected test score t^s above $\underline{t}^*$, and rejects students with an expected test score t^s below $\underline{t}^*$.

We now present an analogous result for the college's choice of information. Consider some arbitrary set of possible information structures from which the college may choose, and take as given that the college will use this information to make ex-post optimal admission decisions. We will find that the college chooses the information that maximizes the expectation of an indirect utility function $\tilde{u}(t^s)$. As we will see, however, this new utility function $\tilde{u}$ will be distinct from the ex-post utility u^*.[10]

[10] The ex-post utility $u^*(t)$ is linear in t, meaning that its expectation is the same—the value at the mean test score—for all information structures at a given distribution. However, the new indirect utility function $\tilde{u}(t^s)$ will not be linear.

To define $\tilde{u}$, we separately consider two cases. First, suppose the college is more selective than society: $\underline{t}^s < \underline{t}^* < \underline{t}^c$. In this case, when $t^s \leq \underline{t}^s$, the payoff $U^c(t^s, A)$ is zero because the applicant is rejected ($A = 0$) and there is no disagreement cost. When $\underline{t}^s < t^s \leq \underline{t}^*$, the applicant is rejected but the college pays a disagreement cost of $\delta u^s(t^s) = \delta(t - \underline{t}^s)$. Finally, when $t^s > \underline{t}^*$, the applicant is accepted ($A = 1$) and there is no disagreement cost, yielding $U^c(t^s, A) = u^c(t^s) = -\underline{t}^c + t$. Putting this all together, it holds that $U^c(t^s, A) = \tilde{u}(t^s)$ for

$$\tilde{u}(t^s) := \begin{cases} 0 & \text{if } t^s \leq \underline{t}^s \\ -\delta \cdot (t^s - \underline{t}^s) & \text{if } \underline{t}^s < t^s \leq \underline{t}^* \\ -\underline{t}^c + t^s & \text{if } t > \underline{t}^*. \end{cases} \tag{College More Selective}$$

Next, suppose the college is less selective than society: $\underline{t}^c < \underline{t}^* < \underline{t}^s$. In this case, when $t^s \leq \underline{t}^*$, the payoff $U^c(t^s, A)$ is zero because the applicant is rejected ($A = 0$) and there is no disagreement cost. When $\underline{t}^* < t^s \leq \underline{t}^s$, the applicant is accepted ($A = 1$), generating underlying payoff $u^c(t^s) = -\underline{t}^c + t^s$; but the college also pays a disagreement cost of $-\delta u^s(t^s) = \delta(t - \underline{t}^s)$. Finally, when $t^s > \underline{t}^*$, the applicant is accepted ($A = 1$) and there is no disagreement cost, yielding $U^c(t^s, A) = u^c(t^s) = -\underline{t}^c + t$. Putting this all together, it holds that $U^c(t^s, A) = \tilde{u}(t^s)$ for

$$\tilde{u}(t^s) := \begin{cases} 0 & \text{if } t^s \leq \underline{t}^* \\ (1+\delta)(t^s - \underline{t}^*) & \text{if } \underline{t}^* < t^s \leq \underline{t}^s \\ -\underline{t}^c + t^s & \text{if } t^s > \underline{t}^*. \end{cases} \tag{College Less Selective}$$

Figure E.1 illustrates $\tilde{u}$. Notice that in both cases, $\tilde{u}$ is neither globally convex nor globally concave.

Remark E.1. As mentioned above, the college can only use lower-censorship information information structures: pool scores below a threshold τ and reveal scores above τ.

When the college is less selective than society (Figure E.1b), lower censorship can be suboptimal in the class of all information structures. Specifically, as shown in Proposition 2 (case 2), the optimal admission policy for the college can entail setting $\tau = \underline{t}^*$ and accepting all students with scores $t > \tau$. The college then bears a disagreement cost for all those high scores. When $\mathbb{E}[t|t > \underline{t}^*] < \underline{t}^s$, the college would be strictly better off by instead pooling those scores above $\tau = \underline{t}^*$, now accepting the same set of students at no disagreement cost.

When the college is more selective than society (Figure E.1a), on the other hand, we know from Proposition 3 that it is optimal for the college to set $\tau \in [\underline{t}^*, \underline{t}^c]$ and again accept all students with scores $t > \tau$. By since $\tilde{u}$ is now linear in this region, there is never a (strict) benefit of pooling these high scores.

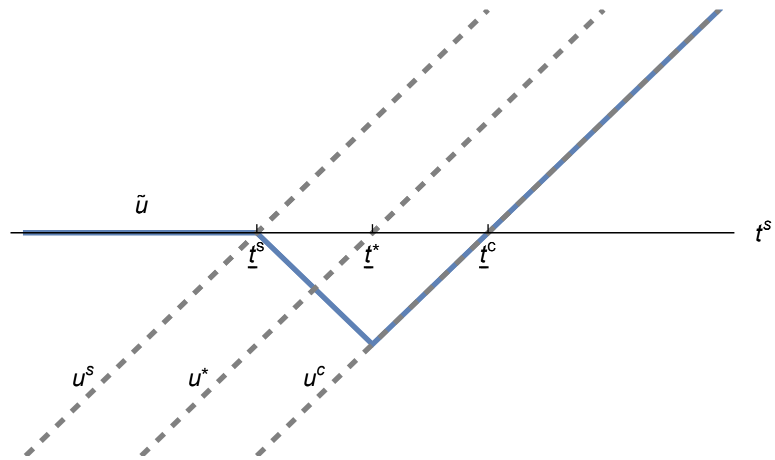
In fact, the results of Kolotilin, Mylovanov, and Zapechelnyuk (2022) suggest that lower-censorship information structures are optimal in the class of all information structures when the college is more selective than society; whereas when the college is less selective, "upper censorship" (pooling scores above a threshold and revealing them below) is optimal. We say "suggest" rather than "imply" because formally those authors' smoothness assumptions on the sender's indirect utility preclude the kinks in our function $\tilde{u}$.

References

Kamenica, E. and M. Gentzkow (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

Kolotilin, A., T. Mylovanov, and A. Zapechelnyuk (2022): "Censorship as Optimal Persuasion," *Theoretical Economics*, 17, 561–585.

(a) College more selective



(b) College less selective

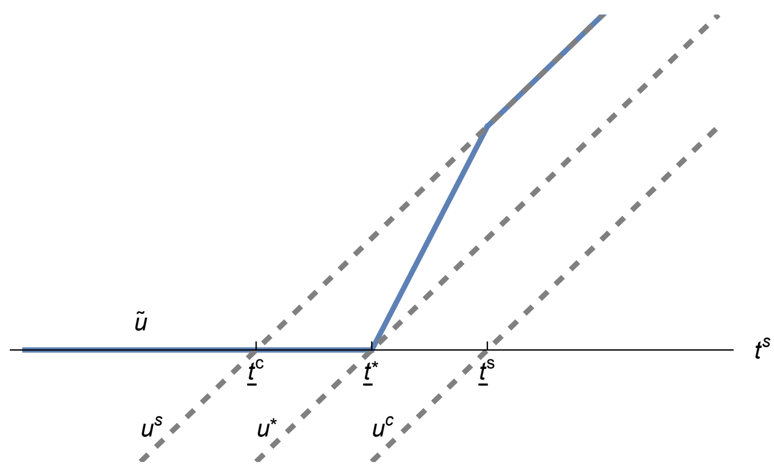


Figure E.1 – The indirect utility function $\tilde{u}(t^s)$ for Bayesian persuasion.