ADMITTING STUDENTS TO SELECTIVE EDUCATION PROGRAMS:
MERIT, PROFILING, AND AFFIRMATIVE ACTION

Dario Cestau
Dennis Epple
Holger Sieg

Admitting Students to Selective Education Programs: Merit, Profiling, and Affirmative Action
Dario Cestau, Dennis Epple, and Holger Sieg
NBER Working Paper No. 21232
June 2015. Revised November 2015
JEL No. I2

## ABSTRACT

For decades, colleges and universities have struggled to increase participation of minority and disadvantaged students. Urban school districts confront a parallel challenge; minority and disadvantaged students are underrepresented in selective programs that use merit-based admission. In their referral and admission policies to such selective programs, school districts may potentially set different admission thresholds based on income and race (affirmative action), and they may potentially take account of differences in achievement relative to ability across race and income groups (profiling). We develop an econometric model that provides a unified treatment of affirmative action and profiling. Implementing the model for an urban district, we find profiling by race and income, and affirmative action for low-income students. Counterfactual analysis reveals that these policies achieve more than 80% of African American enrollment that could be attained by race-based affirmative action.

Dario Cestau
IE Business School
Madrid
Spain
Dario.Cestau@ie.edu

Dennis Epple
Tepper School of Business
Carnegie Mellon University
Posner Hall, Room 257B
Pittsburgh, PA  15213
and NBER
epple@cmu.edu

Holger Sieg
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104
and NBER
holgers@econ.upenn.edu

# 1 Introduction

For decades, colleges and universities have grappled with the challenge of increasing participation of minority and disadvantaged students. Differential admissions criteria designed to increase participation of minority and disadvantaged students inevitably collide with concerns for fairness toward those who are displaced. Courts, legislatures, and electorates have weighed in to define the acceptable limits of affirmative action in higher education.[1] While the spotlight has largely focused on higher education, school districts in large metropolitan areas confront the same challenge; minority and disadvantaged students are underrepresented in selective programs for gifted and talented students.[2] Efforts to increase participation of minority and disadvantaged students in these programs are no less controversial than their higher-education counterparts.[3]

School districts nationwide make a large investment in providing programs to serve gifted and talented students.[4] The National Association for Gifted Children estimates that there are approximately three million academically gifted children in grades K-12 in the U.S. - almost seven percent of the student population. As a proportion of

---

[1]The two most recent affirmative action rulings by the US Supreme Court are on Fisher v University of Texas (June 24, 2013) and Schuette v. Coalition to Defend Affirmative Action (April 22, 2014). We return to these in our policy analysis.

[2]To cite a recent example, on September 27, 2012, a complaint was filed with the U.S. Department of education claiming that the admissions policies of eight elite public high schools in New York City violate the 1964 Civil Rights Act, causing extreme under-representation of black and Hispanic students. (Baker, 2012).

[3]The New York Times dubbed the director of the city's gifted and talented program a "lightning rod for fury" when she announced changes to the admissions policies for the city's gifted programs (NYT, March 22, 2006). Approximately, two years later, the times reported: "When New York City set a uniform threshold for admission to public school gifted programs last fall, it was a crucial step in a prolonged effort to equalize access to programs that critics complained were dominated by white middle-class children whose parents knew how to navigate the system." The new policy has also met with intense criticism, and the director of the city's gifted and talented program has concluded: "We implemented the eligibility criteria, it didn't shake out that way and now we have to take another look at it." (NYT, June 19, 2008)

[4]The empirical evidence on the impact of gifted and talented programs on achievement is discussed in Abdulkadiroglu, Angrist, & Pathak (2011).

the student population, this is comparable to the percentage of high school students admitted to selective colleges and universities.[5] Most large public school districts in the U.S. operate selective programs to serve gifted and talented students.[6] Admission to these programs is competitive and merit-based, and often subject to state government requirements and guidelines. The purpose of this paper is to formulate and estimate a model to analyze the effects of profiling and affirmative action policies on participation of minority and disadvantaged students in gifted education. We believe this to be the first econometric model providing a unified treatment of affirmative action and profiling in urban schools that is based on an equilibrium framework that permits counterfactual analysis of policy.

A strong negative connotation is attached to the term profiling in some contexts, especially law enforcement. We emphasize that we use the term as it has traditionally been applied in the policy domain. For example, the U.S. Department of Labor uses a "Worker Profiling and Reemployment Services System"[7] to facilitate targeting of reemployment services to new claimants for unemployment benefits, and a widely cited report on OECD labor market policies (Martin, 1998) advocates profiling for the same purpose. A recent EU conference on 'Profiling systems for effective labour market integration' provides a definition that accords with our use of the term: "To target appropriate services, measures and programmes considered most suitable to meet the requirements of (the individual's) particular "profile" by statistics-based programme selection (European Commission, 2011)." As we discuss further below,

---

[5]Reardon, et. al. (2012) identify the 171 most selective colleges and universities and report that approximately 7% of high school students attend one of these selective institutions.

[6]International Baccalaureate programs also use stringent admission procedures. The most highly ranked public high schools in New York, Boston, and Philadelphia all use merit-based admission procedures. The referral and admission processes used for the gifted program in the district we study are widely used elsewhere. To provide just a few examples here, the state of Florida prescribes procedures that very closely parallel those used by our district, as does the Los Angeles Unified District, the second largest school district in the country, and the Atlanta district.

[7]http://www.doleta.gov/programs/wprs.cfm

the district we study is in a state in which the state government instructs school districts to consider the "profile of the person" in its decisions about gifted admission.

We develop a new model in which a school district faces a mandate to provide a special program to help highly able students reach their potential.[8] The district assigns a value for gifted program participation that is an increasing function of a student's ability. This value reflects the district's assessment of the educational gains obtained by the student taking into consideration the additional costs of operating the program. Since ability is not fully observed by the district, there is uncertainty about these gains. As a consequence, the district behaves in a Bayesian manner and relies on achievement scores to form beliefs about a student's ability. The school district can also obtain an additional signal about student's ability by conducting a costly IQ test. We interpret the existing state mandate to imply that students cannot be admitted to the program without taking an IQ test.[9] The objective function of the district is then given by the difference between the aggregate net benefits of the program and the aggregate costs of IQ testing. We characterize the optimal referral and admission policies that maximize the objective function of the district. One key property of our model is that optimal policies are cut-off strategies that are functions of the observed test scores and achievement measures.

Our baseline model considers the case of merit-based admission which arises when the objective function of the district is only a function of ability. However, the district must make a judgment of the extent to which the benefits of the program vary with other observed student characteristics such as poverty or race. We, therefore, consider two generalizations of our model. The first extension follows Becker (1957)

---

[8]Neal & Johnson (1996) show that black-white differences in premarket skill do account for a significant portion of the black- white earnings gap in the early 1990's.

[9]Ability is not only difficult to observe, but it is also non-verifiable. In contrast, IQ can be observed and verified by parents and teachers.

3

and models preference-based affirmative action. This case arises when the district assigns higher values to minority students holding ability constant.[10] Two primary arguments are generally advanced in support of affirmative action. One is that, by facilitating enrollment of a critical mass of minority students, affirmative action reduces the isolation otherwise experienced by minority students. The other is that the potential for interactions among diverse peers benefits both majority and minority students.[11] Affirmative action implies that the school district adopts lower referral and admission criteria for minority students.

Similar issues arise with respect to economic disadvantage. Hence, in addition to minority status, we consider whether a student is economically disadvantaged. Our measure of economic disadvantage is based on whether a student is eligible for free or reduced price lunch. Below, we use FRL to denote students eligible for free or reduced-price lunch.

A student who has, in some way, experienced hardship may underperform on achievement tests relative to his or her capability. By taking account of such empirically-grounded differences across demographic groups, a district may be better able to determine which students are most suited to admission to the gifted program. Hence, the second extension of our model allows for use of empirically grounded differ-

---

[10]Becker (1957) introduced the analysis of taste based discrimination into economics. Phelps (1972) and Arrow (1973) developed a theory of statistical discrimination. The effects of affirmative action in employment have been studied by Lundberg (1991), Coate & Loury (1993), and Moro & Norman (2003, 2004). Chung (2000) considers the relationship between role models and affirmative action. Affirmative action in higher education is considered by Chan & Eyster (2003) , Loury, Fryer, & Yuret (2008), and Epple, Romano, & Sieg (2006, 2008). Arcidiacono (2005) estimates a structural model to determine how affirmative action in admission and aid policies affects future earnings. Hickman (2010a, 2010b) develops and estimates a structural model of college admissions and compares the effects of alternative admissions policies on incentives for academic achievement, the racial achievement gap, and the racial college enrollment gap. Long (2007) provides a valuable summary of the legal status of affirmative action and a review of the evidence regarding the effects of affirmative action.

[11]These arguments are detailed in the University of Michigan brief in Grutter v. Bollinger: http://www.vpcomm.umich.edu/admissions/legal/grutter/UM-Grutter.pdf

ences in distributions of achievement, IQ, and ability for demographic groups defined by race and FRL status. We find that minority students have higher IQ conditional on achievement than non-minority students in the range of achievement relevant for referrals for IQ testing. This property then implies that the district can improve referral decisions by adopting a lower referral threshold for minority students. While this profiling based on differences in distributions across racial groups is beneficial to minority students, it is not preferential treatment.[12]

To gain additional insights into the quantitative properties of our model and demonstrate its practical relevance, we parametrize the model and develop a Maximum Likelihood Estimator. Our estimation approach acknowledges the fact that important variables are either latent (ability), partially latent (IQ scores), or measured with error (achievement). We estimate the parameters of the district's objective function from observed test scores, gifted referrals, and admissions by race and economic disadvantage (eligibility status for free or reduced-price lunch) of different demographic groups.

Our empirical analysis is based on a sample of three cohorts of elementary school students that entered first grade in the academic years 2003, 2004, and 2005. Our empirical analysis documents the divergence between the demographics of the district's student body relative to the demographics of those participating in the district's gifted program. We then implement our maximum likelihood estimator. We find that our model fits the data well and that estimated referral and admission thresholds are consistent with policies articulated by the district. We find profiling with respect to both race and subsidized lunch status. We find affirmative action in admission with

---

[12]Knowles, Persico, & Todd (2001) stimulated a large body of work on profiling in law enforcement. In higher education, research on profiling includes Loury et al. (2008), Epple et al. (2008), and Epple, Romano, Sarpca, and Sieg (2012). State policies mandating admission of a specified fraction of each high school to state universities, motivated in part to achieve racial balance through profiling, are studied by Long (2004a, 2004b) and Cullen, Long, & Reback (2012).

respect to subsidized lunch status. We do not find any evidence of affirmative action based on race.

A large proportion of African American students in our sample are eligible for free or reduced-price lunch. As a consequence, our estimated model implies that applying the FRL admission policies to all African American students, regardless of FRL status, would have a modest effect on increasing African American enrollment in the gifted program. Hence, in contrast to higher education[13], we find that referral and admission policies based on FRL status but not race can have a quite substantial effect in increasing minority student participation. Our model also implies that eliminating affirmative action for FRL students by standardizing referral and admission decisions at the level of non-FRL students would reduce the size of the gifted program by up to 50 percent. We thus conclude that adopting strict merit based referral and admission policies, would significantly alter the size and composition of the program.

The rest of the paper is organized as follows. Section 2 develops the theoretical model of optimal referral and admission for a merit-based selective program. Section 3 discusses the implications of affirmative action and profiling within our model. Section 4 introduces a parameterization of the model and develops a Maximum Likelihood Estimator. Section 5 discusses our data set. Section 6 reports the empirical results of this study. Section 7 discusses the policy implications that can be drawn from this analysis. Section 8 offers some conclusions.

---

[13]In his review of affirmative action in colleges and universities, Long (2007) concludes that "...correlates of race are unlikely to substitute successfully for the consideration of race itself."

# 2 A Model of Merit-Based Referral and Admission

We consider the decision problem of a public school district that operates a selective educational program. Admission is competitive and merit-based. Ability is inherently difficult to observe and cannot be verified. Admission criteria are, therefore, based on prior academic achievement and a standardized aptitude or IQ test. Testing is costly. Hence, not all students are referred for testing. We model the choice of policy for referral of students for testing and the choice of policy for admission to the selective program.

There is a continuum of students that differ by ability, $b$. Let $q$ denote the student's score on an IQ test and $a$ performance on a prior achievement test. Let $f(a, b, q)$ be the joint density of achievement, academic ability, and IQ score in the population. Students may also differ by discrete characteristics such as race or low income status. We first present the analysis for a single discrete type and then extend to consideration of more than one such type.

**Assumption 1** *The density $f(a, b, q)$ is continuous on its support $(\underline{a}, \bar{a}) \times (0, \bar{b}) \times (0, \bar{q})$.*[14]

Gifted programs are motivated by the objective of targeting education to high-ability students to help them develop their capabilities. The value or value-added that the district attaches to having a student of type $b$ participate in the selective program is denoted by $v(b)$.

**Assumption 2**

*a) The value function $v(b)$ is continuous and differentiable. Moreover, it is monoton-*

---

[14]For notational simplicity, we assume in the following that the upper bounds are infinity and the lower bound for achievement is negative infinity.

*ically increasing in b, i.e. $v_b(b) \geq 0$.*

*b) The cost of testing a student a student is constant and equal to $c$.*[15]

The district initially observes $a$ and chooses a referral policy denoted by $\alpha(a)$. The testing procedure provides the student IQ score. $a$ and $q$ are both informative of students ability, hence the district grants access to the gifted program based on this information.

The district optimally chooses a referral policy and an admission policy. We make the following assumptions about the information revelation and hence the nature of the admission and referral process.

## Assumption 3

*a) The decision rule that determines who is referred for testing is a function of prior achievement:*

$$0 \leq \alpha(a) \leq 1 \tag{1}$$

*b) The decision rule that determines admission is a function of the IQ score and prior achievement:*

$$0 \leq \beta(a, q) \leq 1 \tag{2}$$

The district forms beliefs in a Bayesian way and chooses the optimal referral and admission policies to maximize the expected difference between benefits and costs.

---

[15]The IQ test employed by the district is an interactive test administered by a psychologist to an individual student. State law mandates that gifted status be determined by a certified district psychologist. Cost is typically cited as the reason for limiting the number of students who are tested. (See NYT, June 19, 2008)

The objective function of the district is then given by:

$$\int_0^\infty \int_0^\infty \int_0^\infty \alpha(a)\ \beta(a,q)\ v(b)\ f(a,b,q)\ da\ db\ dq$$
$$-\int_0^\infty \int_0^\infty \int_0^\infty \alpha(a)\ c\ f(a,b,q)\ da\ db\ dq \tag{3}$$

The first term captures the expected benefits of the program. The second term captures the costs due to IQ testing.[16]

We can solve the decision problem above using standard techniques of variational calculus. It is optimal for the district to use a cut-off strategy for both referral and admission policies.

**Proposition 1** *The optimal solution of our model can be characterized by thresholds $\bar{a}$ and $\bar{q}(a)$.*

All proofs are given in the Appendix.

As shown in the proof of Proposition 1, the district will only admit a student to the gifted program if he or she performs sufficiently well on the IQ score, i.e. if the expected value added to the program is non-negative:

$$V(a,q) \ = \ \int_0^\infty v(b)\ f(b|a,q)\ db \geq 0 \tag{4}$$

The district is indifferent if the equation above holds with equality. This condition implies a threshold function $\bar{q}(a)$ such that a student is admitted to the program if and only $q \geq \bar{q}(a)$.

---

[16]An alternative approach is to model the behavior of the district as wishing to allocate limited space in the gifted program based on fairness or efficiency considerations. These ideas are explored in Abdulkadiroglu (2005) and Kojima (2012).

Similarly, the optimality condition for referral for IQ testing can be written as:

$$W(a) = \int_{\bar{q}(a)}^{\infty} \int_{0}^{\infty} v(b) \ f(b, q|a) \ db \ dq \ \geq \ c \qquad (5)$$

The first term measures the benefits of the students that are referred. The second term captures the costs associated with testing. The district is indifferent if the equation above holds with equality. This condition then defines a threshold function $\bar{a}$ such that a student is referred to the program if and only $a \geq \bar{a}$.

Note that participation in the gifted programs is voluntary. Students who are admitted to the gifted program by the district can opt out of the gifted program and attend regular classes. Based on our conversations with members of the school district we know that the take-up rate among students is very high, easily exceeding 90 percent. As a consequence, while we do not observe take-up decisions by individual students, we feel quite confident that the take-up rate is sufficiently high as to make this issue moot. For applications where acceptance of admission is less routine, it would not be difficult to extend the model and the estimation procedure to allow for voluntary participation decisions as long as the researcher observes admission and attendance decisions.[17]

# 3   Affirmative Action and Profiling

Students from disadvantaged backgrounds tend to be underrepresented in selective programs. Factors such as hardship that may cause underperformance relative to ability on achievement tests may then be used in referral and admission decisions, i.e., profiling. Being a minority student or a student with FRL status are observable

---

[17]An appendix is available upon request which shows how to extend our model to account for a capacity constraint or additional operating costs that depend on the size of the program.

indicators that may be correlated with having experienced hardship. Another potential strategy to reduce underrepresentation in selective programs is affirmative action. Either profiling or affirmative action can give rise to different criteria for such students and can potentially be adopted in either admission or referral policies. In this section, we undertake analysis to provide an economic foundation for understanding and distinguishing affirmative action and profiling in referral and admission policies.

To study the implications of profiling and affirmative action, we extend our model to consider two discrete types. Let $m$ denote a student of a disadvantaged or minority type and $M$ denote the advantaged or majority type.[18] Where it does not cause confusion below, we use the term minority to refer to a student who is economically disadvantaged or a member of a racial minority or both.

Let $f_j(a, b, q)$ $j = m, M$ now denote the density of the each type and let $v_j(b)$ denote the value that the district assigns to type $j$ student given ability $b$.

## 3.1  Affirmative Action

Affirmative action arises in our model if preferential treatment is given to minority students. We formalize this concept of affirmative action as follows.

**Definition 1** *Preference based affirmative action arises if a school district assigns a higher value to students of type $m$ than students of type $M$, i.e if*

$$v_m(b) \geq v_M(b) \tag{6}$$

*where equation (6) holds with strict the inequality for a set of $b$ with positive measure.*

---

[18]In our empirical application we focus on African Americans and students that are eligible for free or reduced lunch.

Affirmative action then leads to lower referral and admission thresholds for minority students. We can formalize this result by considering the case in which there are no differences in the underlying distribution of types.

**Proposition 2** *If $f_m(a, b, q) = f_M(a, b, q)$, then preference-based affirmative action implies that referral and admission thresholds are lower for minorities, i.e. that $\bar{a}_m \leq \bar{a}_M$ and $\bar{q}_m(a) \leq \bar{q}_M(a)$.*

The intuition for this result is straightforward. A student is admitted to the gifted program if the expected benefit conditional on prior achievement and IQ is positive. Preference based affirmative action raises the benefits for minority students and as a consequence lowers the admission threshold. Lower admission thresholds also imply lower referral thresholds.

Thus far, we have cast the discussion in terms of differences between minority and majority students where minority status can be defined based on race (black vs non-black) or income (subsidized lunch vs regular lunch). However, it is also possible that heterogeneity in treatment arises for a variety of other reasons. For example, there may be systematic differences in treatment of students from different neighborhoods. For example, a district may use its gifted program as a mechanism to retain students from higher income neighborhoods in district schools. Even if not formally adopted by a district, similar variation may arise if, for example, a district defers to school principals to some degree in determining referral thresholds. In our empirical analysis below, we extend our model to test whether there is evidence of such policy variation across school types in the district we study.

## 3.2 Profiling

In practice, we observe significant differences in the distribution of observed outcomes between minority and majority students which suggests that $f_m(a, b, q) \neq f_M(a, b, q)$. As a consequence the district may want to adopt different referral and admission thresholds even if it does not engage in preference based affirmation action. For example, if, conditional on achievement, minorities perform better on intelligence tests that non-minorities, then the district would optimally take these differences into consideration when making referral decisions. We show later in the paper that some minority students perform better on IQ tests than majority students for some levels of school achievement, $(F_M(q|a) \geq F_m(q|a)$ for some values of $a)$. This in turn implies that for some levels of achievement minorities students are likely to have higher ability than majority students.

**Definition 2** *The district engages in profiling if it takes the differences in the distribution of ability conditional on achievement into consideration when making referral and admission decisions.*

To isolate the effects of profiling in referral, we have the following result:

**Proposition 3** *If $v_m(b) = v_M(b)$ and if $F_M(b|a, q) \geq F_m(b|a, q)$, then the district adopts a lower threshold for admission of minority relative to non-minority students in admission decisions, i.e. $\bar{q}_m(a) \leq \bar{q}_M(a)$.*

The intuition is as follows. Suppose, among students of a given ability, minority students perform worse than majority students on prior achievement tests. As a consequence, for given prior achievement, minority students are likely to be more able than majority students. Since the value function of the district depends on

ability, these students have higher expected benefits from gifted admission. This then translates into a lower referral threshold for minority than for majority students.

# 4    Identification and Estimation

Let $R$ denote a discrete random variable that is equal to one if the student is referred for testing and zero otherwise. Let $A$ denote a discrete random variable that is equal to one if the student is accepted into the selective program.

**Assumption 4** *Consider a random sample of students.*

1. *We observe the referral decision, denoted by $R$, for all students in the sample.*

2. *For students, who are referred for testing ($R = 1$), we observe $q$.*

3. *We observe the admission decision, denoted by $A$, for all students that are referred for testing.*

4. *We observe achievement with error, denoted by $\tilde{a}$.*

5. *We observe race and subsidized lunch status.*

6. *We do not observe $b$ for any student.*

We assume that $(a, b, q)$ are jointly normally distributed for each type $j$. Measured IQ equals ability plus a normally distributed error that is independent of $(a, q)$

$$q = b \ + \ \epsilon_q \tag{7}$$

14

The equation above implies that $E_j[b] = E_j[q]$ and $Var_j(b) < Var_j(q)$. Errors for achievement and ability are classical measurement errors and depend on the observed type $j$.

The key assumptions here are that $b$ is linear in $q$ and that $\epsilon_q$ is independent of $b$. Both assumptions are important to disentangle latent ability from observed measured scores, and as a consequence study the differences in ability and scores across racial or socio-economic groups. To understand the limitations of these assumptions it is useful to recast the problem within the larger literature of latent factor models. We observe a number of different test scores and we do not need to aggregate these scores into one composite score. Instead, we could alternatively follow Carneiro, Hansen, & Heckman (2003) and assume that IQ scores and the test scores satisfy the following latent factor model:

$$
\begin{aligned}
q &= b + \epsilon_q \\
a_1 &= \beta_1 b + \epsilon_{a_1} \\
. \qquad . \\
a_k &= \beta_k b + \epsilon_{a_k}
\end{aligned}
$$

Moreover, we can relax the assumption that $b$ is a scalar and treat $b$ as a vector. This would allow us to distinguish between different types of skills (such as math versus reading skills). It is conceivable that the district could place different weights on the different ability components. It should be noted that most empirical studies that adopt the more general latent factor models follow Carneiro et al. (2003) and invoke the linearity and independence assumption. However, recent work by Cunha, Heckman, & Schenach (2010) show that the linearity assumption is not essential for

identification.

Moreover, $b$ is only identified up to a scaling factor, i.e. we need to normalize the coefficient of one of the ability measurements to be equal to one to identify the remaining parameters of the model. In our application it is natural to normalize the coefficient of the IQ score since IQ plays the largest role in the decision making of the district as we document in the paper.

Most of the test scores that are commonly used to measure ability are normed-reference scores that only allow us to rank students. These are, therefore, ordinal measures of ability. This then raises two additional question. First, we would like to know whether our model is identified if the school uses a monotonic transformation of $b$, denoted by $g(b)$. Note that we do not observed $g$. In that case the value of function of the school can be written as $V(g(b))$. It should be clear, that we cannot disentangle $V$ from $g$ without observing $g$. Hence our analysis focuses on the composite function:

$$v(b) \;=\; V(g(b))$$

Our formal identification proof only covers the case in which $v(b)$ is linear and thus should be interpreted as a first order Taylor approximation of the true underlying function $V(g(b))$. In all our work, we find that the linear specification of the value function performs well. Nonetheless, it would be desirable to extend our analysis to the non-linear case, but it is not a straightforward extension. Moreover, we can conduct our policy analysis purely based on our knowledge of $v(b)$, we do not need to know $V(\cdot)$ or $g(b)$.

Second, and maybe more importantly, the ordinality of ability raises a more fundamental question of how to model admission to merit based programs. If there is no inherent scale of test scores or if the comparison of test scores across cohorts are

difficult, it may make sense for the district to define admission thresholds not in terms of the level ability $b$, but in terms of the percentile rank of $b$. For example, the district may decide to admit all students of a given cohort that score in the top five percent of the ability distribution. Profiling then implies that we adjust the observed distributions for hardship, while affirmative action involves judgement about the desirability of admitting students that score in different percentiles from different socio-economic or racial groups. Our current framework assumes that the district makes decisions based on whether the expected level of ability exceeds a specified threshold. The district that we study is in a state that requires districts to provide gifted education to students above the 97.5 percentile of IQ (130 or higher). In this respect, the policy is in the spirit of the percentile ranking approach advocated by Barlevy & Neal (2012).

The district's value function for type $j$ is given by:

$$
\begin{aligned}
v_j(b) &= \zeta_{0j} + \zeta_{1j}\, b \\
&= \zeta_{1j}\, (b + \Delta_j)
\end{aligned}
\tag{8}
$$

where $\Delta_j = \zeta_{0j}/\zeta_{1j}$. In our application, we consider four types using race and FRL status to define types. Given this assumption, the admission rule can be written as:

$$
\int_0^\infty (b + \Delta_j)\, f_j(b|a,q)\, db \;\geq\; 0
$$

which implies that admission rule only depends on $\Delta_j$, but not on $\zeta_{1j}$. The referral rule is given by:

$$
\int_{\bar{q}(a)}^\infty \int_0^\infty \zeta_{1j}(b + \Delta_j)\, f_j(b,q|a)\, db\, dq \;\geq\; c
$$

The referral decision, therefore, also depends on $\zeta_1$. Given this parametrization, we can obtain a closed form solution of the admission policy and a simple characterization of the referral policy. As shown in Appendix D, we have the following result:

**Proposition 4** *Given the parametrization above, the optimal admission policy is given by the linear function $\bar{q}_j(a) = \tau_{0j} - \tau_{1j} a$ where:*

$$\tau_{0j} = \frac{-\frac{\zeta_{0j}}{\zeta_{1j}} - \mu_{b_j} + A_j \mu_{a_j} + Q_j \mu_{q_j}}{Q_j} \tag{9}$$

$$\tau_{1j} = \frac{A_j}{Q_j}$$

*where $A_j$ and $Q_j$ are known functions of the parameters of the model. A student of type $(j, a, q)$ is admitted to the program if and only if $q \geq \bar{q}_j(a)$.*

The proof of Proposition 4 also implies that the expected benefit of referral, $W_j(a)$, can be computed fairly efficiently by univariate integration. Moreover, there exists a unique $\bar{a}_j$, which is the solution to the equation $W_j(a) = c$ and can be efficiently computed.

With measurement error in achievement, the probability of observing $R = 0$ conditional on $j$ and $\tilde{a}$ is given by:

$$Pr\{R = 0 \,|j, \tilde{a}\} = \int_0^{\bar{a}_j} f_j(a|\tilde{a}) \, da \tag{10}$$

Similarly, the probability of observing $R = 1$ and $A = 1$ is:

$$Pr\{R = 1, A = 1|\, j, q, \tilde{a}\} = \int_{\bar{a}}^\infty 1\{q \geq \bar{q}_j(a)\} \ f_j(a|q, \tilde{a}) \, da \tag{11}$$

18

The probability of observing $R = 1$ and $A = 0$ is:

$$Pr\{R = 1, A = 0 | \ j, q, \tilde{a}\} = \int_{\bar{a}}^{\infty} 1\{q < \bar{q}_j(a)\} \ f_j(a|q, \tilde{a}) \ da \qquad (12)$$

Given the specification of our model above, we can prove that our model is identified.

**Proposition 5** *The parameters of our model are identified.*

A formal proof is provided in Appendix E. Here we provide the key intuition behind the proof. The parameters of the joint distribution of prior achievement and IQ are identified from the observed empirical distributions. While we do not observe IQ scores for the students who are not referred for testing, we can account for that truncation problem. The levels of the coefficients of the value function are thus identified from the observed conditional admission and referral probabilities.[19] The error variance of prior achievement is identified from the degree of misclassification observed in the data. If the error variance were zero, our model should perfectly explain the observed referral decisions once we condition on the observed prior achievement score. Differences between the model's referral and admission predictions and the observed outcomes are only due to measurement error in achievement.

Note that the proof above relies on the joint normality assumption. This is not an overly strong functional form assumption since our data are based on norm-referenced tests. Standardized achievement scores typically are normally distributed at the state level and IQ scores are normally distributed at the national level. Using multivariate normal distributions is computationally convenient since the conditional expectations of $E[b|a, q]$ can be analytically characterized as shown in Proposition 4. As a consequence, the threshold function for admission $q(a)$ and the cut-off point for referral $\bar{a}$

---

[19]Note that we treat testing costs, $c$, as observed by the econometrician.

can be computed. Extending this identification proof to models with for more flexible functional form assumptions is not trivial, but the basic intuition of the proof should carry over to the case in which the underlying distributions are not normal.[20]

The likelihood for a single observation is then given by:

$$
\begin{aligned}
L &= [Pr\{R = 0 | \, j, \tilde{a}\} \, f_j(\tilde{a})]^{(1-R)} \\
&\quad [Pr\{R = 1, A = 1 | \, j, q, \tilde{a}\} \, f_j(\tilde{a}, q)]^{R_q \, A} \\
&\quad [Pr\{R = 1, A = 0 | \, j, q, \tilde{a}\} \, f_j(\tilde{a}, q)]^{R \, (1-A)}
\end{aligned}
\tag{13}
$$

The likelihood for a sample of $N$ students is then a straightforward product of the terms above.

Evaluating the likelihood function is not straightforward. To evaluate the likelihood function above we need to numerically compute the optimal referral and admission policies for each type. The proof of Proposition 4 can be easily adapted to generate an algorithm that efficiently accomplishes this purpose. The only challenge here is to compute the optimal referral threshold which requires a line search algorithm (See equation (30) in the appendix). Once we have computed the optimal decision rules, it is relatively straightforward to evaluate the conditional probabilities that form the likelihood function. We then compute the maximum likelihood estimator using a standard nested fixed point argument. In the inner loop of the algorithm we evaluate the likelihood function for each parameter value. In the outer loop we use standard numerical optimization methods to find the argument that maximizes the likelihood function. It is straightforward to show that our estimator is $\sqrt{N}$ consistent and asymptotically normally distributed. We use a standard bootstrap algorithm to compute standard errors. Appendix F shows that our estimator works well in a

---

[20]Identification of latent factor models can be achieved without imposing strong functional form assumption as discussed, for example, in Schenach (2004) and Cunha et al. (2010).

Monte Carlo study.

# 5 Data

Our application focuses on a selective gifted program that is operated by a mid-sized urban district that prefers to stay anonymous. The district operates a Gifted Center that serves students in elementary and middle school. Gifted students in grades 1 through 8 participate in a one-day-per-week program at a designated location away from the student's home school. Students participate in programs designed to enhance creative problem solving and leadership skills and are offered specially designed instruction in math, science, literature, and a variety of other fields.

The district adheres to state regulations concerning gifted students and services. The state regulations outline a multifaceted approach used to identify whether a student is gifted and whether gifted education is needed. The state requires gifted status to be determined by a certified district psychologist. A mentally gifted student is defined as someone with an IQ score of at least 130. The regulation specifies that a student with IQ score below 130 may be admitted "... when other educational criteria in the *profile* of the person strongly indicates gifted ability (emphasis added)." The state guidelines provide for consideration of factors that may "mask" giftedness including "... gender or race bias, or socio/cultural deprivation..." Thus, the state gives districts discretion to employ profiling and affirmative action in referral and admission policies. Of course, the district is also subject to the U.S. constitution, which has been interpreted by the Supreme Court as requiring that affirmative action by race be a last resort in achieving diversity objectives.

The IQ tests that the district uses are the most widely used tests in the U.S. The tests differ primarily by student age. For the age range that we are studying,

21

the district uses the fourth edition of the Wechsler Intelligence Scale for Children (WISC4). The district administers this examination using standard WISC4 protocol, one-on-one, to a student by a psychologist trained in IQ testing. This test is designed for elementary school children and is administered orally so that IQ assessment is not heavily driven by the degree to which the student has developed reading and writing skills sufficient to comprehend and complete a written exam. The WISC4 has roughly 15 subtests designed to elicit a range of cognitive capabilities; each subtest in turn is comprised of a set of questions. The tester asks questions in a prescribed order and scores student responses.

The district codes electronically the students IQ score, but not the scores on the subtests or the name of the psychologist administering the test. Thanks to excellent cooperation from the district, we were permitted to send a team of Ph.D. students to the district offices for a period of many days to pull each student's paper file, scan the student's IQ test report into a file, and, from these, code the sub-scores. We also coded the name of the psychologist who tested each student. We find some evidence of testers "giving the benefit of the doubt" to students who are very close to the threshold. For example, for most psychologists, the proportion of students scoring at the threshold is somewhat larger than the proportion scoring one or two points below the threshold. This appears to be modest in magnitude (on the order of 2 to 3 points), and is not systematically linked to student observables.

Another concern with matching our model to the data is that our model does not allow for private testing. Moreover, private testers may employ more lenient standards than public testers giving some opportunity to parents to buy access to the gifted program. Fortunately, our data set contains detailed information about the identity of person that administered the IQ test allowing us to differentiate between private and public testers. We find that there are only 24 students in the sample

for whom we observe private test scores. These are all students from higher income families and do not quality for free or reduced lunch. For all 24 students we observe one private test score. All of them are admitted to the program based on the private test. 22 out of the 24 students were also tested by a public psychologist. The average private test is 130.6 compared to 116.6 for the publicly administered test. Note that the average IQ of an accepted regular lunch student from a public test is 130.8. Of course, there may be a larger number of students who are privately tested, but it is reasonable to assume that we only observe private test scores when they are "above the bar." While there is some evidence that private testers are more lenient than public testers, we conclude that the number of students that are admitted into the gifted program based on private test score is sufficiently small that it does not raise serious doubts about our modeling and estimation approach.

Most referrals and admissions for the gifted program occur during elementary school in grades 1-3. We focus on this population in this paper. The sample consists of the cohort of students that were in $1^{st}$ grade in 2003/04, 2004/05, or 2005/06. We follow these students until the end of $3th$ grade.[21]

We start with a sample of 7,753 students. The district permits self-referral, and we retain those referrals in our sample.[22] We keep those students who were at the school district during the first three grades of school. This reduces our sample to 5,409 students. We drop 751 observations without lunch status data and that changed their lunch status over the three school grades. We eliminate 72 observations without achievement data. We drop 24 observations for students who were admitted based

---

[21]We find that the main qualitative and quantitative findings of this paper are robust to excluding any of the cohorts used in the analysis as well as including students that were referred for testing in the forth grade. Results are available upon request from the authors.

[22]Approximately 11.5% of referrals are self-referrals. Relative to their numbers in the district, self-referrals are proportionately higher among non-minority students who are not on free or reduced-price lunch. Hence, our conclusions with respect to referral and admission of minority students are certainly not driven by self-referrals.

on IQ scores provided by private testers after failing to gain admission by the IQ test administered by the district.[23] We eliminate 104 students marked as gifted at any time during elementary school with no history of IQ test data. Finally, we filter outliers by eliminating the five lowest achievement scores by cohort of referred students, and the top 5 achievement scores by cohort that were not referred, totaling 30 eliminated observations. The final sample size is 4,428 observations.

Achievement scores can be constructed based on a variety of observed scores. We observe the Oral Reading Fluency (ORF) test. We standardize ORF scores for every period in every grade and compute the average score for each student. In addition we have access to test scores on standardized tests that measure reading and math skills in 3rd grade. Finally, we observe an average GPA score for each student in the sample. As shown in Table 1 these measures are highly positively correlated.

Table 1: Correlation Matrix of Achievement Measures

|  | ORF scores | PSSA scores | Average GPA | Achievement |
|---|---|---|---|---|
| ORF scores | 1 | - | - | - |
| PSSA scores | 0.74 | 1 | - | - |
| Average GPA | 0.62 | 0.74 | 1 | - |
| Achievement | 0.89 | 0.95 | 0.82 | 1 |

The table above reports the correlation matrix between the components of the achievement measure (rows 1-3) and the achievement measure (row 4). The achievement measure is a weighted average of the students Oral Reading Fluency (ORF) tests, scores on standardized state tests in reading and math, and GPA average score, from $1^{st}$ grade through $3^{rd}$ grade. Achievement components and the achievement measure are standardized.

We construct the achievement measure used in the estimation of the model as a weighted average of the GPA, the scores on the standardized tests of math and reading, and the ORF scores.[24] While our achievement measure is unlikely to be

---

[23]These 24 individuals are all non-FRL students. The average of their scores on the district test was 116.6 and the average on the privately administered tests was 130.6.

[24]The weights are determined by regressing a referral indicator on the available achievement scores. We experimented with alternative weightings of achievement scores and found that our results are

the exact measure that is used by the district, we only require that our measure is sufficiently strongly correlated with the latent measure used by the district.

As previously noted, we use eligibility for free or reduced-price lunch as a measure of economic disadvantage for students. We also create a poverty indicator for each school which ranges between 0 and 2. To construct this variable, students eligible for free lunch are coded as 2, students eligible for reduced-price lunch are coded as 1, and regular lunch students are coded as zero. The school income indicator is then the school average of this variable over a ten year period. We use this variable to classify schools into two discrete categories, high- and low-income schools. As we discuss in detail in the next section, we can, therefore, test whether there are significant differences in referral and admission decisions for each type. Table 2 reports descriptive statistics for the full sample as well as the subsamples by FRL status and race.

Table 2: Sample Statistics by FRL status and Race

|  | All types | Sub-sample by Lunch status | | Sub-sample by Race | |
|---|---|---|---|---|---|
|  |  | Non-FRL | FRL | Non Black | Black |
| Sample size | 4428 | 961 | 3467 | 1872 | 2556 |
| FRL | 0.78 | 0 | 1 | 0.57 | 0.94 |
| Black | 0.58 | 0.16 | 0.70 | 0 | 1 |
| Achievement | 0.00 | 0.84 | -0.23 | 0.40 | -0.29 |
| School Income | 0.47 | 0.89 | 0.32 | 0.72 | 0.30 |
| Fraction Referred | 0.13 | 0.38 | 0.07 | 0.21 | 0.07 |
| Fraction Gifted | 0.07 | 0.22 | 0.02 | 0.13 | 0.02 |
| Ratio Gifted/Referred | 0.51 | 0.60 | 0.37 | 0.59 | 0.34 |

The table reports descriptive statistics of the full sample (column 1), the sample of students with regular lunch status (column 2), the sample of students with subsidized lunch status (column 3), the sample of non-black students (column 4), and the sample of black students (column 5). School income is the fraction of students in the sample that attend a high-income school as defined in the text.

robust.

We find, as with many other urban districts, that there is a high proportion of poor students and the majority of students are black. Achievement and school income are negatively correlated with FRL status and race. Table 1 also shows that a large majority of all African American students in our sample are FRL eligible while only slightly more than half of non-black students are FRL eligible.

Table 3 reports the demographics of the sample by referral and gifted status. We find that FRL and black students tend to be underrepresented among referred and gifted students. Not surprisingly, students referred for testing have higher achievement scores than students who are not referred. Similarly gifted students have significantly higher achievement and IQ scores than non-gifted students.

Table 3: Statistics for Non Referred, Referred and Gifted Students

|  | Sub-sample by referred status | | Sub-sample Referred by gifted status | |
|---|---|---|---|---|
|  | Non referred | Referred | Non Gifted | Gifted |
| FRL | 0.84 | 0.37 | 0.48 | 0.26 |
| Black | 0.62 | 0.30 | 0.41 | 0.20 |
| Achievement | -0.21 | 1.43 | 1.09 | 1.76 |
| IQ |  |  | 110 | 129 |

The table reports descriptive statistics of the sample of non-referred students (column 1), the sample of referred students (column 2), the sample of referred but non-gifted students (column 3), and the sample of referred gifted students (column 4).

To set the stage for our econometric analysis, we provide some additional evidence regarding the referral process. We partition the support of the achievement distribution into seven intervals. Table 4 reports the frequency of referral for each bin conditional on FRL status and race. We find that referral probabilities are monotonically increasing in the our observed measure of prior achievement. There are some pronounced differences in referrals probabilities by FRL status. There are much smaller differences in referral probabilities by race.

26

Table 4: Ratios of Referred Students by Achievement Bin

| Achievement | Regular | | FRL | |
|---|---|---|---|---|
| | total | % Referred | total | % Referred |
| <0 | 188 | 0.00 | 2092 | 0.00 |
| 0≤ <0.33 | 101 | 0.09 | 460 | 0.05 |
| 0.33≤ <0.66 | 111 | 0.14 | 354 | 0.09 |
| 0.66≤ <1.00 | 128 | 0.26 | 272 | 0.15 |
| 1.00≤ <1.33 | 122 | 0.42 | 152 | 0.30 |
| 1.33≤ <1.66 | 118 | 0.67 | 93 | 0.47 |
| 1.66≤ | 193 | 0.90 | 44 | 0.61 |
| | Non Black | | Black | |
| Achievement | total | % Referred | total | % Referred |
| <0 | 670 | 0.00 | 1610 | 0.00 |
| 0≤ <0.33 | 222 | 0.05 | 339 | 0.06 |
| 0.33≤ <0.66 | 227 | 0.09 | 238 | 0.12 |
| 0.66≤ <1.00 | 213 | 0.18 | 187 | 0.19 |
| 1.00≤ <1.33 | 170 | 0.34 | 104 | 0.38 |
| 1.33≤ <1.66 | 156 | 0.57 | 55 | 0.62 |
| 1.66≤ | 214 | 0.85 | 23 | 0.78 |

The table above reports the frequency of referral by achievement bins of four student sub-samples: Non-subsidized lunch students (Regular), subsidized lunch students (FRL), Non-black students (Non Black), and black students (Black).

To illustrate some of the properties of the admission process, we partition the support of the IO distribution into five intervals. Table 5 reports the frequency of admissions for each bin conditional on FRL status and race.

Table 5: Ratios of Gifted Students by IQ

| IQ score | Regular Lunch | | FRL | |
|---|---|---|---|---|
| | total | % Admitted | total | % Admitted |
| <110 | 44 | 0.00 | 88 | 0.00 |
| 110≤ <115 | 30 | 0.00 | 32 | 0.00 |
| 115≤ <120 | 51 | 0.20 | 38 | 0.58 |
| 120≤ <125 | 48 | 0.38 | 23 | 1.00 |
| 125≤ | 188 | 1.00 | 33 | 1.00 |
| | Non Black | | Black | |
| IQ score | total | % Admitted | total | % Admitted |
| <110 | 61 | 0.00 | 71 | 0.00 |
| 110≤ <115 | 39 | 0.00 | 23 | 0.00 |
| 115≤ <120 | 58 | 0.29 | 31 | 0.48 |
| 120≤ <125 | 51 | 0.51 | 20 | 0.75 |
| 125≤ | 191 | 1.00 | 30 | 1.00 |

The table reports the frequency of admissions by five IQ bins of four student sub-samples: Non-subsidized lunch students (Regular), subsidized lunch students (FRL), Non-black students (Non Black), and black students (Black).

The upper panel of Table 5 breaks down the sample by free or reduced lunch status, the lower panel by race. We find that admission probabilities are increasing in IQ scores. For students who are not eligible for FRL, the probability of acceptance is equal to one if the IQ score is larger than 125. It drops to 38 percent if the IQ score is between 120 and 125. It is 20 percent for students with scores between 115 and 120. For students eligible for FRL, the admission probability is one for all students with scores larger than 120. It drops to 58 percent for students with IQ between 115 and 120. It is zero for student with IQ score below 115. These findings suggest that students on FRL face lower admission thresholds, an issue we study formally in our

empirical analysis.

Finally, we consider admission by race in the lower panel of Table 5. We find that there are some difference in the admission probabilities between black and non-black students in the interval between 115 and 125. These differences are substantially smaller that the differences we observe by FRL status.

# 6    Empirical Results

We begin by estimating our model allowing the fullest possible variation in policies across types with respect to both profiling and affirmative action. We designate this Model I. We then investigate the extent to which policy variation across types can be restricted without significantly reducing the fit of the model. It is useful to begin by summarizing the set of parameters to be estimated in Model I. The utility function of the district is linear and has two parameters for each type. With affirmative action varying across four types (race and FRL), there are then 8 parameters in the utility function. With full profiling, the distribution parameters differ across types. For each type, we have means and covariances for ability and achievement. This yields five parameters per type for a total of 20 parameters. In addition, there are eight measurement error variances for achievement and IQ by type. Hence, the most general specification, Model I, has 36 parameters.

Model specifications that impose common parameters in the underlying distributions of ability, IQ and achievement are clearly rejected by the data. Hence, we consider restrictions regarding the use of affirmative action while retaining type specific distributions of ability, achievement and measurement and, thus, allow for profiling. We consider four models that vary in the use of affirmative action. Model I allows for 8 different district utility parameters, thereby allowing for affirmative action by

FRL status and race. Model II allows for affirmative action by FRL status, but not race. This imposes 4 restrictions on the parameter space. In addition, it imposes the additional restriction that the slope parameters of the district utility function are the same, implying equal marginal benefit to ability across types. Model III allows for affirmative action by race, but not FRL status. This imposes 4 restrictions on the parameter space. In addition, it imposes the additional restriction that the slope parameters of the district utility function are the same, implying equal marginal benefit to ability across types. Finally, Model IV does not allow for affirmative action imposing 6 restrictions on the parameters of the utility function.

Table 6 summarizes these model specifications and reports the relevant log-likelihood statistics. Since Models II-IV are nested by Model I, we can use standard likelihood ratio tests to evaluate the different specifications. The results in Table 6 lead us to choose Model II as our preferred model specification. Model II imposes five restrictions relative to Model I. The difference in likelihood function values between these two models implies a chi-squared statistic of 1.6 with a p-value of 0.9. Models III and IV are very strongly rejected ($p < .001$) relative to either Model I or Model II.

Our choice of Model II carries the implication that preferences do not differ between black and non-black students. Profiling with respect to race enhances prediction of performance by race, but does not imply preferential treatment by race. Hence, we conclude that the district provides no preferential treatment by race. By contrast, we do find preferential treatment (affirmative action) with respect to FRL status, especially at the admission stage.

Next we discuss the estimation results in further detail. Table 7 reports goodness of fit statistics conditional on FRL status and race for the four models. Table 7 shows that the first two models all fit the data almost equally well. Model III is slightly

30

Table 6: Model Summary

| Model | Profiling | | Affirmative Action | | | | Log Likelihood | Free Param. |
|---|---|---|---|---|---|---|---|---|
| | | | Admission (Intercept) | | Referral (Slope) | | | |
| | FRL | Race | FRL | Race | FRL | Race | | |
| I | Yes | Yes | Yes | Yes | Yes | Yes | -8807.4 | 36 |
| II | Yes | Yes | Yes | No | No | No | -8808.2 | 31 |
| III | Yes | Yes | No | Yes | No | No | -8845.8 | 31 |
| IV | Yes | Yes | No | No | No | No | -8873.0 | 30 |

This table summarizes four model specifications and reports the relevant log-likelihood. Columns 1-2 indicate profiling by subsidized lunch status and race respectively. Columns 3-6 indicate affirmative action based on subsidized lunch status and race. Columns 3-4 signalize affirmative action through the intercept of the utility function while columns 5-6 pick out affirmative action through the slope of the utility function. The last column shows the number of free parameters estimated by specification.

Table 7: Goodness of Fit

| | Data | Model I | Model II | Model III | Model IV |
|---|---|---|---|---|---|
| | Non Black Students | | | | |
| Not Referred | 78.6 | 78.3 | 78.7 | 78.8 | 78.8 |
| Referred not Admitted | 8.9 | 9.6 | 9.2 | 9.5 | 9.2 |
| Referred Admitted | 12.5 | 12.1 | 12.1 | 11.8 | 12.0 |
| IQ Referred not Admitted | 111.8 | 112.8 | 112.6 | 112.6 | 110.5 |
| IQ Referred and Admitted | 130.0 | 130.4 | 130.2 | 130.3 | 128.5 |
| | Black Students | | | | |
| Not Referred | 93.2 | 93.1 | 93.1 | 93.2 | 93.4 |
| Referred not Admitted | 4.5 | 4.6 | 4.6 | 4.7 | 4.6 |
| Referred Admitted | 2.3 | 2.3 | 2.2 | 2.1 | 2.0 |
| IQ Referred not Admitted | 106.4 | 106.5 | 106.8 | 105.6 | 107.3 |
| IQ Referred and Admitted | 125.3 | 126.3 | 126.5 | 125.4 | 127.9 |
| | Non-FRL Lunch Students | | | | |
| Not Referred | 62.4 | 62.4 | 62.6 | 63.0 | 62.8 |
| Referred not Admitted | 15.1 | 16.0 | 16.2 | 16.2 | 15.7 |
| Referred Admitted | 22.5 | 21.6 | 21.2 | 20.9 | 21.5 |
| IQ Referred not Admitted | 112.8 | 113.4 | 113.5 | 111.7 | 110.2 |
| IQ Referred and Admitted | 130.8 | 131.4 | 131.4 | 130.1 | 128.8 |
| | FRL Students | | | | |
| Not Referred | 93.8 | 93.6 | 93.8 | 93.8 | 94.0 |
| Referred not Admitted | 3.9 | 4.2 | 3.9 | 4.1 | 4.0 |
| Referred Admitted | 2.2 | 2.2 | 2.3 | 2.1 | 2.0 |
| IQ Referred not Admitted | 106.1 | 107.0 | 106.4 | 107.6 | 108.1 |
| IQ Referred and Admitted | 124.2 | 124.3 | 124.4 | 127.2 | 127.3 |
| | All Students | | | | |
| Not Referred | 87.0 | 86.8 | 87.0 | 87.1 | 87.2 |
| Referred not Admitted | 6.3 | 6.7 | 6.6 | 6.7 | 6.5 |
| Referred Admitted | 6.6 | 6.4 | 6.4 | 6.2 | 6.3 |
| IQ Referred not Admitted | 109.6 | 110.3 | 110.2 | 109.7 | 109.2 |
| IQ Referred and Admitted | 129.1 | 129.5 | 129.4 | 129.3 | 128.4 |
| Black in Admitted | 20.4 | 20.5 | 20.2 | 19.5 | 18.8 |
| FRL in Admitted | 26.5 | 26.9 | 28.2 | 26.7 | 25.3 |

The table reports goodness of fit statistics conditional on subsidized lunch status (FRL) and race for the four models. The first column (Data) shows descriptive statistics of four sub-samples of students and the full sample. Columns 2-5 show the goodness of fit for each specification using simulated data from the parameter estimates.

worse. By contrast, the fit of the Model IV is significantly worse for black and FRL students.

We now turn to the parameter estimates. To save some space, Table 8 reports the estimates and estimated standard errors only for our preferred specification, Model II. Beginning with mean achievement estimates in the top panel, we see that students on FRL have lower achievement than their non-FRL, same-race counterparts. We see as well that non-black students have higher achievement than black students. The standard deviations of achievement exhibit some variation across race and FRL, with variance of non-white, non-FRL students being somewhat higher than for the other three student types.

Our estimates of the parameters of ability are shown in the second panel. The means of ability exhibit a striking contrast to the means of achievement. For students not on FRL, non-black students have a higher mean than black students, but the relative difference is substantially smaller than for achievement. The same holds true for mean ability estimates for those on FRL. The contrast between the variation in means of achievement by race and FRL to the variation in means of ability by race and FRL provides insight into our finding of significant scope for profiling by race and FRL. Other parameter estimates vary less across student types. The estimated correlations between achievement and ability are similar in magnitudes across types the four types, with the value regular lunch black students (.83) being somewhat lower than the other three groups. The standard deviations of ability are somewhat lower for non-FRL than for FRL students. Table 8 also reveals, not surprisingly, that there is substantial error in measurement of both ability and achievement.
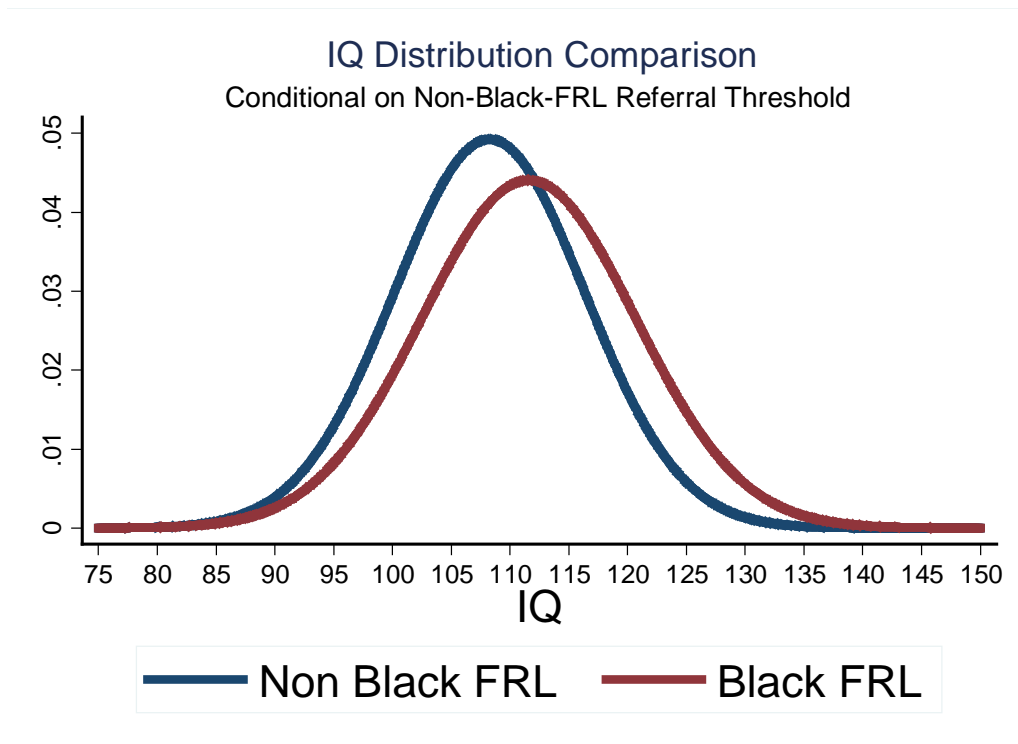
We provide in Figure 1 an illustration of how differences in distributions by type provide scope for profiling. Figure 1 plots the conditional distributions of IQ scores for a given value of achievement for black and non-black FRL students. The achievement

33

Table 8: Parameter Estimates: Model II

| | | | | |
|---|---|---|---|---|
| achievement distribution | $\mu_{a,rw}$ | 0.92 (0.04) | $\mu_{a,fw}$ | 0.00 (0.00) |
| | $\mu_{a,rb}$ | 0.45 (0.07) | $\mu_{a,fb}$ | -0.34 (0.02) |
| | $\sigma_{at,rw}$ | 0.98 (0.03) | $\sigma_{at,fw}$ | 0.91 (0.02) |
| | $\sigma_{at,rb}$ | 0.89 (0.05) | $\sigma_{at,fb}$ | 0.84 (0.01) |
| | $\sigma_{e,rw}$ | 0.52 (0.02) | $\sigma_{e,fw}$ | 0.57 (0.03) |
| | $\sigma_{e,rb}$ | 0.47 (0.05) | $\sigma_{e,fb}$ | 0.53 (0.02) |
| ability distribution | $\mu_{b,rw}$ | 111.52 (1.48) | $\mu_{b,fw}$ | 84.58 (4.25) |
| | $\mu_{b,rb}$ | 103.38 (4.50) | $\mu_{b,fb}$ | 73.74 (4.21) |
| | $\sigma_{b,rw}$ | 14.43 (1.25) | $\sigma_{b,fw}$ | 18.02 (2.26) |
| | $\sigma_{b,rb}$ | 15.81 (3.28) | $\sigma_{b,fb}$ | 20.24 (1.92) |
| | $\sigma_{q,rw}$ | 6.24 (0.41) | $\sigma_{q,fw}$ | 4.81 (0.90) |
| | $\sigma_{q,rb}$ | 6.13 (1.52) | $\sigma_{q,fb}$ | 3.47 (0.97) |
| correlations | $\rho_{ab,rw}$ | 0.90 (0.02) | $\rho_{ab,fw}$ | 0.93 (0.02) |
| | $\rho_{ab,rb}$ | 0.83 (0.05) | $\rho_{ab,fb}$ | 0.91 (0.02) |
| value function | $\Delta_{rw}$ | -121.44 (0.41) | $\Delta_{fw}$ | -115.67 (0.41) |
| | $\Delta_{rb}$ | -121.44 (0.41) | $\Delta_{fb}$ | -115.67 (0.41) |
| | $\zeta_{1,rw}$ | 0.19 (0.065) | $\zeta_{1,rb}$ | 0.19 (0.065) |
| | $\zeta_{1,fw}$ | 0.19 (0.065) | $\zeta_{1,fb}$ | 0.19 (0.065) |

The table reports parameter estimates for Model II. Bootstrapped standard errors from 50 replications are shown in parentheses.

This figure shows for Black and non-Black FRL students the distribution of IQ conditional on achievement after integrating out ability. Parameters are from Model II and achievement is set equal to .99.

level is set to 0.99 which is the referral threshold for non-black FRL students. Figure 1 illustrates our finding that, among relatively high-achieving FRL students, the IQ scores of black students are higher than that of non-black students, or stated differently, black students have lower achievement relative to ability than non-black students.

From the last panel of Table 8, we see that the intercept of the district utility function for FRL students is higher than for non-FRL students, implying affirmative action in favor of FRL students. The difference in intercepts is quite substantial, equivalent to adding approximately one third standard deviation to ability of FRL students. We established above in our comparison of Model IV to Model II that this difference is highly significant.

Finally, we have investigated whether different schools use different referral and admission policies. To conduct this analysis, we consider two sets of schools: high-income and low-income schools. We then allow different utility functions for the two school types, and hence different referral and admission thresholds. There are very few non-FRL students in poor schools, too few to permit meaningful testing of whether low-income schools treat non-FRL students differently than high-income schools. We, therefore, investigate whether preferences for gifted participation of FRL students differ between low-income and high-income schools by allowing the intercept of the district utility function to differ between high- and low-income schools. We obtain a difference in intercepts of 1.5, with low-income schools provide somewhat higher affirmative-action benefits to FRL students than high-income schools. This difference is relatively modest, equivalent to adding .1 standard deviation to ability of FRL student in low-income schools, and not significant by conventional standards (p=.09). Hence, we conclude that there is limited, if any, heterogeneity in treatment

36

of FRL students between high- and low-income schools.[25]

# 7  Policy Analysis

In 2003, the U.S. Supreme Court considered two affirmative action cases against the University of Michigan, naming then-president Lee Bollinger, as defendant. One, Grutter v. Bollinger (2003), challenged the admission policy of the law school. The other, Gratz v. Bollinger (2003), challenged the undergraduate admission policy. The court upheld the former, seeing it as narrowly construed and employing race only as a possible favorable factor, while striking down the latter as being too closely akin to a quota system. The court decision with respect to Michigan's undergraduate admission policy prompted many colleges and universities to change their admission policies to reduce or eliminate race as a distinct admission criterion.

The Gratz v. Bollinger decision prompted research investigating the extent to which race-blind policies, coupled with selection employing correlates of race, might preserve racial diversity in higher education.[26] Here we use race-blind to mean that criteria for student admission do not vary by race. These analyses suggest that, at the college level, banning affirmative action while permitting of use of correlates of race results in a very significant reduction in minority attendance, especially at top-tier institutions.

It is likely that the Gratz v. Bollinger decision also induced many school districts, possibly including the one that we study, to adopt race-blind referral and admission policies. Subsequent Supreme Court decisions directly focused on K-12 education reinforce the need for school districts develop race-blind policies. In 2007, the Court

---

[25]This sensitivity check is based on the 2004/05 cohort.
[26]See, for example, Chan & Eyster (2003), Epple et al. (2008) and Loury et al. (2008).

decided two cases involving K-12 education, one in Seattle (Parents Involved in Community Schools v. Seattle School District No. 1) and the other in Louisville (Meredith v. Jefferson County Board of Education). In those cases, the Court invalidated the districts' racial desegregation plans, ruling that students could not be assigned to schools solely to achieve racial balance. Not surprisingly, the Court was closely divided. Casting the pivotal vote, Justice Kennedy sided with the conservative wing of the Court to form a majority to ban use of race in determining school assignments. In his opinion, however, Justice Kennedy pointedly left the door open to use of correlates of race to achieve race-conscious objectives.[27] In 2006, voters in Michigan approved a ban on affirmative, which in turn was upheld in the April 2014 Supreme Court decision in Schuette v Coalition to Defend Affirmative Action. That decision has prompted increased focus on use of socioeconomic factors as an alternative to use of race in fostering affirmative action objectives.[28]

The following important policy question then arises: How closely can race-conscious but race-blind policies achieve the objectives of racial integration of gifted K-12 academic programs? Our model permits us to provide an analysis of the effect on racial diversity of affirmative action based on race. We can compare this approach to the effect of affirmative action based on a correlate of race—FRL status. We, therefore, consider two policy alternatives. Policy 1 employs affirmative action with respect to both FRL and race. This is accomplished by extending the admission threshold for FRL students to all African American Students. Policy 2 eliminates

---

[27]In Fisher v University of Texas (June 24, 2013), the Supreme Court strongly emphasized that placing any explicit weight on race in admissions should be a last resort, but, in its remand to the lower court, the Supreme Court again seemed to accept use of correlates of race to achieve race-conscious objectives: "The reviewing court must ultimately be satisfied that no workable race-neutral alternatives would produce the educational benefits of diversity."

[28]The President of the American Association of State Colleges and Universities commented that "Most of us have already started to look at other variables than race, especially first-generation students, and low-income students." (NYT, 4/22/2014).

Table 9: Policy Analysis

| | Data | Model II | Policy 1 | Policy 2 |
|---|---|---|---|---|
| | Non Black Students | | | |
| Not Referred | 78.6 | 78.7 | 78.7 | 81.0 |
| Referred not Admitted | 8.9 | 9.2 | 9.3 | 8.1 |
| Referred Admitted | 12.5 | 12.1 | 12.0 | 10.9 |
| IQ Referred not Admitted | 111.8 | 112.6 | 112.5 | 114.0 |
| IQ Referred and Admitted | 130.0 | 130.2 | 130.1 | 131.2 |
| | Black Students | | | |
| Not Referred | 93.2 | 93.1 | 92.1 | 95.6 |
| Referred not Admitted | 4.5 | 4.6 | 5.2 | 3.0 |
| Referred Admitted | 2.3 | 2.2 | 2.8 | 1.4 |
| IQ Referred not Admitted | 106.4 | 106.8 | 105.5 | 111.1 |
| IQ Referred and Admitted | 125.3 | 126.5 | 125.3 | 130.6 |
| | Non-FRL Lunch Students | | | |
| Not Referred | 62.4 | 62.6 | 59.9 | 62.6 |
| Referred not Admitted | 15.1 | 16.2 | 17.6 | 16.4 |
| Referred Admitted | 22.5 | 21.2 | 22.5 | 21.0 |
| IQ Referred not Admitted | 112.8 | 113.5 | 112.0 | 113.5 |
| IQ Referred and Admitted | 130.8 | 131.4 | 130.6 | 131.3 |
| | FRL Students | | | |
| Not Referred | 93.8 | 93.8 | 93.8 | 96.9 |
| Referred not Admitted | 3.9 | 3.9 | 3.9 | 2.1 |
| Referred Admitted | 2.2 | 2.3 | 2.3 | 1.1 |
| IQ Referred not Admitted | 106.1 | 106.4 | 106.3 | 112.0 |
| IQ Referred and Admitted | 124.2 | 124.4 | 124.5 | 129.8 |
| | All Students | | | |
| Not Referred | 87.0 | 87.0 | 86.4 | 89.4 |
| Referred not Admitted | 6.3 | 6.6 | 6.9 | 5.2 |
| Referred Admitted | 6.6 | 6.4 | 6.7 | 5.4 |
| IQ Referred not Admitted | 109.6 | 110.2 | 109.5 | 113.0 |
| IQ Referred and Admitted | 129.1 | 129.4 | 129.0 | 131.1 |
| Black in Admitted | 20.4 | 20.2 | 24.1 | 14.7 |
| FRL in Admitted | 26.5 | 28.2 | 27.0 | 15.5 |

 The table reports statistics conditional on subsidized lunch status (FRL) and race for the four models. The first column (Data) shows descriptive statistics of four sub-samples of students and the full sample. Columns 2 show the same statistics for the base line equilibrium, Columns 3-4 consider the policies discussed in detail in the text.

affirmative action entirely. Table 9 summarizes the impacts of these counterfactual policy experiments.

Our model predicts that adoption of Policy 1 would increase gifted enrollment of African American students in this cohort from 20.2 to 24.1 percent. By contrast, under Policy 2, enrollment of African American students would drop to 14.7 percent. Thus, eliminating affirmative action would reduce African American student enrollment by approximately 39% relative to a policy of affirmative action for African American students. The current policy, providing no affirmative action for African American students but retaining affirmative action for FRL students reduces African American student enrollment by 16% relative to Policy 1. Thus, affirmative action with respect to FRL is extremely important in preserving enrollment of African American students in the district's gifted program.

# 8   Conclusions

U.S. educational institutions at all levels invest extensive resources in providing advanced programs to help high-ability students reach their potential. The challenge that arises in deciding admission to these programs is that ability is unobserved. Obtaining measures of ability that are not affected by family resources and vagaries of circumstance has proven to be extraordinarily challenging. Despite extensive research and design efforts, measures of ability such as IQ exhibit systematic variation across demographic groups. Educational institutions confront the challenge of using available evidence to assess suitability of students for admission to selective programs, distinguishing ability from influences attributable to family resources and circumstance. While not uncontroversial, differential treatment based on income has gained a considerable measure of acceptance. Race-based criteria have proven to be much

more controversial.

The Supreme Court has endorsed a holistic approach, one that considers a broad set of factors in decisions about the admission of a student to selective programs and institutions, i.e., the Court encourages profiling. Going somewhat farther in his decision for the majority in Gratz versus Bollinger, Justice Kennedy explicitly declined to reject use of non-race factors to achieve race-conscious ends, a view he reaffirmed in a concurring opinion in the Seattle and Louisville cases. It is natural to wonder whether it is possible to formalize the distinctions being made by the court, and, if so, whether it is possible to test empirically whether a policy does or does not conform to court guidelines. In this light, we have pursued four objectives in this paper. One objective is to develop a framework sufficiently broad not only to encompasses merit, profiling, and affirmative action, but also to permit statistical testing of the roles that these factors play in practice. A second is to demonstrate, via econometric analysis, the practical relevance of this framework in application. A third is to extend the analysis of these policy issues to an important domain that has been understudied, selective programs for high-ability students in primary and secondary education. A fourth objective is to make a substantive contribution to policy by statistical testing of the extent of use of profiling and affirmative action in an operational setting, and by counterfactuals that investigate the consequences, either of curtailing or extending, the use of profiling or affirmative action.

We believe our application succeeds in these objectives. Of particular interest are the findings of our counterfactual analysis of race-based affirmative action. Long (2007), in his review of the evidence with respect to race-based affirmative action in colleges and universities, concludes, and we agree, that: "The pressures—and in some cases, legal requirements—to replace affirmative action are very real. All the same, the stark evidence of the inefficacy of replacement programs remains." By contrast,

for the central school district we study, we find that profiling by race and income, coupled with affirmative action by income, can achieve more than 80 percent of the level of African American enrollment that could be achieved by race-based affirmative action.

Two questions naturally arise: Why the sharp contrast to findings with respect to elimination of race-based affirmative action in our setting as compared to findings for colleges and universities? To what extent is it likely to generalize to other central city school districts? The answer to the first question is that middle- and upper-income households with children tend to locate in suburban school districts, leaving relatively low-income households in the central city districts. African American households are disproportionately represented among low-income households. Hence, in central city districts, policies that courts have thus far found admissible (profiling by race and income and affirmative action by income) can go a long way toward achieving outcomes that would be achieved by adding race-based affirmative action to the district's decision criteria. The answer to the second question is that our findings are indeed likely to generalize to other central city school districts; the propensity of middle- and high-income households with children to locate in suburban districts is a phenomenon observed in every large metropolitan area in the United States.

The methods developed in this paper offer much scope for future research. Our framework may also be applicable in other settings in which costly screening procedures are used with the goal of increasing minority participation.[29] For example, costly search for qualified minority applicants as an alternative or complement to affirmative action is quite common in post-secondary and post-graduate education (e.g., minority recruitment fairs hosted by colleges). Minority job fairs often require par-

---

[29]We are grateful to an anonymous referee for pointing out the potential relevance of our framework in these domains.

ticipating candidates to submit a specified set of materials that are then reviewed by participating employers as they decide which candidates will be invited for interviews.

# References

Abdulkadiroglu, A. (2005). College admission with affirmative action. *International Journal of Game Theory, 33*, 535–49.

Abdulkadiroglu, A., Angrist, J., & Pathak, P. (2011). The elite illusion: achievement effects at boston and new york exam schools. NBER Working Paper 17264.

Anwar, S. & Fang, H. (2006). An alternative test for racial prejudice in motor vehicle searches: theory and evidence. *American Economic Review, 96 (1)*, 127–151.

Arcidiacono, P. (2005). Affirmative action in higher education: how do admission and financial aid rules affect future earnings?. *Econometrica, 73 (5)*, 1477–1524.

Arrow, K. (1973). The theory of discrimination. In Ashenfelter, O. (Ed.), *Discrimination in Labor Markets*. Princeton University Press.

Baker, A. (2012). Charges of bias in admission test policy at eight elite public high schools. *New York Times, Sept 22*.

Barlevy, G. & Neal, D. (2012). Pay for percentile. *American Economic Review, 102 (5)*, 1805–31.

Becker, G. (1957). *The Economics of Discrimination*. Chicago University Press.

Carneiro, P., Hansen, K., & Heckman, J. (2003). Estimating distributions of treatment effects with an application to the returns of schooling and measurement of the effects of uncertainty on college choice. *International Economic Review, 44(2)*, 361–422.

Chan, J. & Eyster, E. (2003). Does banning affirmative action harm college quality?. *American Economic Review, 93 (3)*, 858–873.

Chung, K. (2000). Role models and arguments for affirmative action. *American Economic Review, 90*, 640–48.

Coate, S. & Loury, G. (1993). Will affirmative action policies eliminate negative stereotypes?. *American Economic Review, 83*, 1220–40.

Commission, E. (2011). Dialogue conference: profiling systems for effective labour market integration. Executive Summary, Brussels.

Court, U. S. (2003a). Gratz and versus and bollinger, 539 u.s. 244..

Court, U. S. (2003b). Grutter versus bollinger, 539 u.s. 306..

Court, U. S. (2003c). Meredith v. jefferson county board of education, 548, us 938,..

Court, U. S. (2007). Parents involved in community schools v. seattle school district no. 1, 551 u.s. 701..

Cullen, J., Long, M., & Reback, R. (2012). Jockeying for position: strategic high school choice under texas' top ten percent plan. *Journal of Public Economics, forthcoming.*

Cunha, F., Heckman, J., & Schenach, S. (2010). Estimating the technology of cognitive and non-cognitive skill formation. *Econometrica, 78 (3)*, 883–931.

Epple, D., Romano, R., & Sieg, H. (2006). Admission, tuition, and financial aid policies in the market for higher education. *Econometrica, 74(4)*, 885–928.

Epple, D., Romano, R., & Sieg, H. (2008). Diversity and affirmative action in higher education. *Journal of Public Economic Theory, 10 (4)*, 475–501.

Hickman, B. (2010a). Effort, race gaps, and affirmative action: a game-theortic analysis of college admissions. Working Paper.

Hickman, B. (2010b). Effort, race gaps, and affirmative action: a structural policy analysis of u.s. college admissions. Working Paper.

Knowles, J., Persico, N., & Todd, P. (2001). Racial bias in motor vehicle searches: theory and evidence. *Journal of Political Economy, 109*, 203–229.

Kojima, F. (2012). School choice: impossibilities for afformative action. *Games and Economic Behavior, 75*, 685–93.

Long, M. (2004a). College applications and the effect of affirmative action. *Journal of Econometrics, 121*, 319–42.

Long, M. (2004b). Race and college admissions: an alternative to affirmative action. *Review of Economics and Statistics, 86 (4)*, 1020–33.

Long, M. (2007). Affirmative action and its alternatives in public universities: what do we know?. *Public Administration Review, 67 (2)*, 315–30.

Loury, G., Fryer, R., & Yuret, T. (2008). An economic analysis of color-blind affirmative action. *Journal of Law, Economics, and Organization, 24(2)*, 319–55.

45

Lundberg, S. (1991). The enforcement of equal opportunity laws under imperfect information: affirmative action and alternatives. *The Quarterly Journal of Economics*, *106*(1), 309–26.

Martin, J. (1998). What works among active labour market policies: evidence from oecd countries' experiences. OECD Labour Market and Social Policy Occasional Papers, No. 35.

Moro, A. & Norman, P. (2003). Affirmative action in a competitive economy. *Journal of Public Economics*, *87*, 567–94.

Moro, A. & Norman, P. (2004). A general equilibrium model of statistical discrimination. *Journal of Economic Theory*, *114*, 1–30.

Neal, D. & Johnson, W. (1996). The role of pre-market factors in black-white wage differences. *Journal of Political Economy*, *104(5)*, 869–895.

Phelps, E. (1972). The statistical theory of racism and sexism. *American Economic Review*, *62*, 659–661.

Reardon, S., Baker, R., & Klasik, D. (2012). Race, income, and enrollment patterns in highly selective colleges, 1982-2004. Stanford University Working Paper.

Schenach, S. (2004). Estimation of nonlinear models with measurement error. *Econometrica*, *72 (1)*, 33–75.

# A    Proof of Proposition 1

Optimal admission for referral policies for IQ testing satisfies:

$$\alpha(a) \begin{pmatrix} = 1 \\ \in [0,1] \\ = 0 \end{pmatrix} \text{ as } \int_0^\infty \int_0^\infty \Big( \beta(a,q)\, v(b) - c \Big) f(b,q|a)\, db\, dq \begin{pmatrix} > 0 \\ = 0 \\ < 0 \end{pmatrix} \qquad (14)$$

Optimal admission satisfies:

$$\beta(a,q) \begin{pmatrix} = 1 \\ \in [0,1] \\ = 0 \end{pmatrix} \text{ as } \int v(b)\, f(b|a,q)\, db \begin{pmatrix} > 0 \\ = 0 \\ < 0 \end{pmatrix} \qquad (15)$$

Rearranging the FOC gives the result.

# B    Proof of Proposition 2

If $v_m(b) \geq v_M(b)$ and $f_m(a,b,q) = f_M(a,b,q) = f(a,b,q)$, then we have

$$
\begin{aligned}
V_m(a,q) &= \int_0^\infty \int_0^\infty v_m(b)\, f(b|a,q)\, db && (16) \\
&\geq \int_0^\infty \int_0^\infty v_M(b)\, f(b|a,q)\, db \\
&= V_M(a,q)
\end{aligned}
$$

47

and hence $\bar{q}_m(a) \leq \bar{q}_M(a)$. Similarly

$$
\begin{aligned}
W_m(a) &= \int_{\bar{q}_m(a)}^{\infty} \int_0^{\infty} v_m(b) \ f(b, q|a) \ db \ dq & (17) \\
&\geq \int_{\bar{q}_M(a)}^{\infty} \int_0^{\infty} v_m(b) \ f(b, q|a) \ db \ dq \\
&\geq \int_{\bar{q}_M(a)}^{\infty} \int_0^{\infty} v_M(b) \ f(b, q|a) \ db \ dq \\
&= W_M(a)
\end{aligned}
$$

which proves the result.

# C  Proof of Proposition 3

If $v_m(b) = v_M(b)$ and $F_m(b|a, q) \leq F_M(b|a, q)$, then we have

$$
\begin{aligned}
V_m(a, q) &= \int_0^{\infty} \int_0^{\infty} v(b) \ f_m(b|a, q) \ db & (18) \\
&\geq \int_0^{\infty} \int_0^{\infty} v(b) \ f_M(b|a, q) \ db \\
&= V_M(a, q)
\end{aligned}
$$

where the inequality follows from the fact that $v(b)$ is monotonically increasing in $b$.
Hence $q_m(a) \leq q_M(a)$.

# D Proof of Proposition 4

The expected value of admitting a student type $(a, q)$ is given by:

$$V(a, q) = \int_0^\infty \int_0^\infty v(b) \, f(b|a, q) \, db = \zeta_0 + \zeta_1 \, \mu_{b/a,q} \tag{19}$$

The covariance matrix is given by:

$$\Sigma_{b \, a \, q} = \begin{bmatrix} \sigma_b^2 & \sigma_{ab} & \sigma_b^2 \\ \sigma_{ab} & \sigma_a^2 & \sigma_{ab} \\ \sigma_b^2 & \sigma_{ab} & \sigma_b^2 + \sigma_q^2 \end{bmatrix} \tag{20}$$

Hence the conditional expectation of $b$ given $(a, q)$ is given by the following equation:

$$
\begin{aligned}
\mu_{b/a,q} &= \mu_b + \begin{bmatrix} \sigma_{ab} & \sigma_b^2 \end{bmatrix} \begin{bmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 + \sigma_q^2 \end{bmatrix}^{-1} \begin{bmatrix} a - \mu_a \\ q - \mu_q \end{bmatrix} \\
&= \mu_b - (A\mu_a + Q\mu_q) + Aa + Qq
\end{aligned} \tag{21}
$$

where the constants $A$ and $Q$ are defined as:

$$
\begin{aligned}
A &= \frac{\sigma_{ab}(\sigma_b^2 + \sigma_q^2) - \sigma_b^2 \sigma_{ab}}{\sigma_a^2(\sigma_b^2 + \sigma_q^2) - \sigma_{ab}^2} \\
Q &= \frac{-\sigma_{ab}^2 + \sigma_b^2 \sigma_a^2}{\sigma_a^2(\sigma_b^2 + \sigma_q^2) - \sigma_{ab}^2}
\end{aligned} \tag{22}
$$

The school is indifferent between admitting and not-admitting a student to the gifted program if the expected value added of the program is zero:

$$0 = V(a, q) = \zeta_0 + \zeta_1 \left( \mu_b - (A\mu_a + Q\mu_q) + Aa + Qq \right) \tag{23}$$

49

Equation (23) defines a linear function $\bar{q}(a)$ which fully characterizes the admission policy $\beta(a, q)$:

$$\bar{q}(a) \;=\; \tau_0 \;-\; \tau_1 \, a \tag{24}$$

where:

$$\tau_0 \;=\; \frac{-\frac{\zeta_0}{\zeta_1} \;-\; \mu_b \;+\; (A\mu_a \;+\; Q\mu_q)}{Q}$$

$$\tau_1 \;=\; \frac{A}{Q} \tag{25}$$

Substituting the optimal admission rule into the expected value of referring a student type $(a)$ for testing, we obtain:

$$W(a) = \int_0^\infty \int_{\bar{q}(a)}^\infty v(b) \, f(b, q|a) \, dq \, db \tag{26}$$

Now $f(b, q|a) = f(b|a) \, f_\varepsilon(q - b)$. The mean and variance of the condition distribution of $b$ given $a$ are given by

$$\mu_{b/a} \;=\; \mu_b \;+\; \frac{\sigma_{ab}}{\sigma_a^2}(a - \mu_a) \tag{27}$$

and

$$\sigma_{b/a}^2 \;=\; \sigma_b^2 \;-\; \frac{\sigma_{ab}^2}{\sigma_a^2} \tag{28}$$

Hence, we have:

$$v(b) \, f(b, q/a) = (\zeta_0 + \zeta_1 \, b) \, \frac{1}{\sqrt{2\pi}\sigma_{b|a}} e^{-\frac{(b - mu_{b|a})^2}{2\sigma_{b|a}^2}} \, \frac{1}{\sqrt{2\pi}\sigma_q} e^{-\frac{(q - b)^2}{2\sigma_q^2}} \tag{29}$$

Substituting (29) into (26), we obtain:

$$W(a) \;=\; \int_{-\infty}^{\infty} [1 - \Phi(\frac{\tau_0 \; - \; \tau_1 a \; - \; b}{\sigma_\epsilon})] \, (\zeta_0 + \zeta_1 \, b) \, \frac{1}{\sqrt{2\pi}\sigma_{b|a}} e^{-\frac{(b-mu_{b|a})^2}{2\sigma_{b|a}^2}} \, db \quad (30)$$

$\bar{a}$ is then the unique solution to the equation $W(a) = c$ . A student is referred if and only if $a \geq \bar{a}$. Letting the structural parameters of the underlying distribution of $(a, b, q)$ and the parameters of the utility function to be functions of $j$ then proves the result.

# E   Proof of Proposition 5

Consider the model with one type.[30] Achievement is measured with error:

$$\tilde{a} \;=\; a + \epsilon_a \quad (31)$$

First, consider the case in which the variance of the measurement error in achievement, denoted by $\sigma_{\epsilon_a}^2$, is known. In that case, $\mu_a, \sigma_a^2$ are identified from the observed distribution of achievement.

Next consider the distribution of IQ condition on being referred and conditional on $\tilde{a}$. This distribution is given by:

$$f(q|R = 1, \tilde{a}) \;=\; \frac{\int_{\bar{a}} f(a, q|\tilde{a}) \, da}{\int \int_{\bar{a}} f(a, q|\tilde{a}) \, da \, dq} \quad (32)$$

The left hand side of the equation is non-parametrically identified. Under the normality assumption, $f(a, q|\tilde{a})$ is also normally distributed. Hence, the right hand side is a function of the three structural parameters $(\mu_q, \sigma_q^2, \sigma_{aq})$ and the nuisance parameter

---

[30]The identification argument easily extends to the case of multiple types.

$\bar{a}$. All of these parameters are thus identified under the joint normality assumption (as well as more flexible functional forms).

Next, we assume, that

$$q = b + \epsilon_q \qquad (33)$$

Hence $\mu_b = \mu_q$ and $\sigma_{ab} = \sigma_{aq}$ are identified as well. That leaves us with three parameters: $\sigma_b^2 = \sigma_q^2 - \sigma_{\epsilon_q}^2$, and the parameters of the value function $\zeta_0$ and $\zeta_1$.

Consider the probability of being admitted to the gifted program, which is given by:

$$Pr\{A = 1|\tilde{a}, q\} = \int 1\{\tau_0 + \tau_1 a \leq q\} \, f(a|\tilde{a}, q) \, da \qquad (34)$$

where the functional forms of $\tau_0$ and $\tau_1$ are given in the proof of Proposition 4. The left hand side of the equation above is non-parametrically identified. Consider the right hand side which the same probability predicted by our model. First, note that the joint distribution $f(a|\tilde{a}, q)$ has already been identified. Second note that $\tau_1$ is function of only $\sigma_b^2$. $\tau_0$ is a function of the ratio $\Delta = \zeta_0/\zeta_1$ and $\sigma_b^2$, given that all other parameters that affect $\tau_0$ and $\tau_1$ have been identified. We can thus conclude that the ratio $\Delta = \zeta_0/\zeta_1$ and $\sigma_b^2$ are identified of the conditional admission probabilities.

Next, consider the referral probability, which is given by:

$$Pr\{R = 1|\tilde{a}, q\} = \int_{\bar{a}} f(a|\tilde{a}) \, da \qquad (35)$$

where $\bar{a}$ is the unique solution to

$$\int_{-\infty}^{\infty}[1 - \Phi(\frac{\tau_0 - \tau_1 a - b}{\sigma_{\epsilon_q}})]\,(\zeta_1\,(b + \Delta))\,\frac{1}{\sqrt{2\pi}\sigma_{b|a}}e^{-\frac{(b-mu_{b|a})^2}{2\sigma_{b|a}^2}}\,db\,-\,c\,=0 \qquad (36)$$

Given that $c$ is a known parameter, it is straight forward to show that $\bar{a} = \bar{a}(\zeta_1)$. We thus conclude that $\zeta_1$ is identified of the observed referral probabilities.

Finally, note that the error variance of prior achievement is identified by the dependence of $Pr\{R = 1|\tilde{a}\}$ and $Pr\{A = 1|\tilde{a}, q\}$ on $\sigma_{\epsilon_a}^2$. If the measurement error variance is zero, then there is a deterministic relationship between referral and admission decisions and observed achievement and IQ scores. As the amount of measurement error increases, the correlation between $\tilde{a}$ and $R$ and $A$ decreases. If the measurement error variance is infinite, the referral probability is not a function of the observed achievement

# F   A Monte Carlo Study

We conducted a Monte Carlo study to get some additional insights into the small sample properties of our ML estimator. We generated 50 samples with sizes equal to 5000 observations. Table 10 summarizes the findings of the Monte Carlo exercise. It reports the true parameter values under which the data were generated, the mean of the estimates, the standard deviation of the estimates and the sample mean of the estimated asymptotic standard errors.

We find that the small sample bias of most parameters are small.

Table 10: Monte Carlo Results

|  | True Value | Guess | Mean | Std Dev. | Asymp. Errors |
|---|---|---|---|---|---|
| $\mu_a$ | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 |
| $\mu_b$ | 100.00 | 105.00 | 99.97 | 1.09 | 0.99 |
| $\sigma_b$ | 13.00 | 12.00 | 13.06 | 0.60 | 0.58 |
| $\sigma_{IQ\ error}$ | 7.00 | 7.50 | 6.99 | 0.32 | 0.28 |
| $\sigma_{\widetilde{a}\ error}$ | 0.35 | 0.40 | 0.35 | 0.01 | 0.01 |
| $\sigma_{\widetilde{a}}$ | 1.00 | 1.00 | 1.00 | 0.01 | 0.01 |
| $\rho_{a,b}$ | 0.65 | 0.70 | 0.65 | 0.03 | 0.03 |
| $\Delta = \frac{\zeta_0}{\zeta_1}$ | -116.67 | -100.00 | -116.68 | 0.23 | 0.24 |
| $\zeta_1$ | 300 | 250 | 299.48 | 36.16 | 36.14 |

Simulations: 50. Simulations size: 7500

The table summarizes the findings of an identification exercise using a Monte Carlo study. It reports the true parameter values under which the data were generated, the mean of the estimates, the standard deviation of the estimates and the sample mean of the estimated asymptotic standard errors.