ECONOMETRICA
JOURNAL OF THE ECONOMETRIC SOCIETY

Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings?

Author(s): Peter Arcidiacono

Source: *Econometrica*, Vol. 73, No. 5 (Sep., 2005), pp. 1477-1524
Published by: The Econometric Society
Stable URL: http://www.jstor.org/stable/3598881
Accessed: 06-10-2017 17:18 UTC

Econometrica, Vol. 73, No. 5 (September, 2005), 1477–1524

AFFIRMATIVE ACTION IN HIGHER EDUCATION: HOW DO ADMISSION AND FINANCIAL AID RULES AFFECT FUTURE EARNINGS?

BY PETER ARCIDIACONO[1]

This paper addresses how changing the admission and financial aid rules at colleges affects future earnings. I estimate a structural model of the following decisions by individuals: where to submit applications, which school to attend, and what field to study. The model also includes decisions by schools as to which students to accept and how much financial aid to offer. Simulating how black educational choices would change were they to face the white admission and aid rules shows that race-based advantages had little effect on earnings. However, removing race-based advantages does affect black educational outcomes. In particular, removing advantages in admissions substantially decreases the number of black students at top-tier schools, while removing advantages in financial aid causes a decrease in the number of blacks who attend college.

KEYWORDS: Dynamic discrete choice, returns to education, human capital, schooling decisions.

1. INTRODUCTION

SINCE 1996, THE USE OF RACE in the admissions decisions of public colleges and universities has been challenged through court cases in Michigan, Georgia, and Texas, through ballot initiatives in California and Washington, and through the governor's office in Florida. In addition to admissions policies, race-specific financial aid has also come under fire. Scholarships provided by the University of Maryland that were restricted to blacks alone were ruled to be unconstitutional in court.[2] Despite the considerable public debate on race-based advantages in higher education, there is little understanding of how these programs affect the future outcomes of their intended beneficiaries. This paper seeks to estimate the effects of removing race-based advantages in both admissions and financial aid on black earnings and educational choices.

To accomplish this, I estimate a structural model of the college decision-making process. In particular, I estimate a model of how individuals decide where to submit applications and, conditional on being accepted, in which college to enroll and what major to study. These educational decisions are then

[1]The author thanks Charlie Clotfelter, Jill Constantine, Mark Coppejans, Maria Ferreyra, Eric French, John Jones, Mike Keane, Levis Kochin, Rob McMillan, Marc Rysman, Jeff Smith, and seminar participants at the 2002 AEA Winter Meetings, Boston University, University of California–San Diego, CIRANO Conference on the Econometrics of Education, 2000 Cowles Conference on the Econometrics of Strategy and Decision Making, the NBER Higher Education Group, North Carolina State University, Queens University, Stanford University, Texas A&M University, University of Toronto, University of Washington, University of Western Ontario, University of Wisconsin, and Yale University. The paper benefited tremendously from the suggestions of the editor and the referees.

[2]See *http://chronicle.com/indepth/affirm/court.htm* for a review of recent court cases on affirmative action.

linked to future earnings. I also estimate the decisions by schools as to whether to admit a student and, conditional on admitting, how much financial aid to offer.

There are a number of complications with estimating the effect of affirmative action in higher education on earnings. The first complication arises from affirmative action in higher education not having a direct effect on earnings, but only an indirect effect through the college decision-making process. Affirmative action affects whether individuals are admitted or how much aid they are offered. Adjusting these admissions and financial aid rules indirectly affects future earnings by influencing where individuals apply to college and therefore where they attend college. Without understanding the process by which individuals decide where to apply, it is impossible to quantify the effects of affirmative action. For example, if blacks were subjected to the same admissions rules as whites, would they undo the effects of removing affirmative action by applying to a larger number of elite schools or would they decide not to apply to college at all? This paper is the first to structurally estimate how individuals decide where to submit applications by explicitly modeling individuals' expectations on admittance, financial aid, and future earnings.

The second complication comes from the self-selection inherent in the educational process. This self-selection may take many forms. First, the returns to college may differ across the abilities of individuals, with those who have the highest returns being the most likely to take part in the "treatment" of attending college (see Card (2001), Heckman and Vytlacil (1998), and Heckman, Tobias, and Vytlacil (2000)). Second, the choices individuals make while in college may affect the treatment of attending college as well. For example, attending college and choosing to major in the natural sciences may result in higher earnings than choosing to major in the humanities. Finally, higher ability individuals may find college less difficult and therefore may be more likely to attend college, even if the returns to ability do not depend on whether one has a college degree.[3] I capture all of these sources of heterogeneous treatment effects through explicitly modeling the choice of major and allowing the monetary returns to different majors to vary with college quality and observed and unobserved ability. This is particularly important in evaluating affirmative action programs because the marginal individual who attends college may have very different expected returns than the average individual who attends college.

By estimating the structural model and appropriately accounting for selection, the admissions and financial aid rules are linked to where individuals submit applications, which school they attend, and what field they study. It is then possible to track how these decisions would change given a change in the admissions and financial aid rules. Earnings expectations under the different

[3]Many of these points are discussed in Altonji (1993).

rules can then be calculated before individuals submit applications by assigning probabilities of applying and attending particular schools under different rules and calculating the associated expected earnings for each of the possible education paths.

Simulating the effects of removing preferential treatment for blacks in admissions and in financial aid shows surprisingly little effect on black male earnings despite blacks enjoying much larger premiums for attending college than their white counterparts. The small effects on expected earnings from removing black advantages in financial aid occur because those individuals who are at the margin of attending are also the ones who have the lowest treatment effect; their abilities are relatively better suited to the noncollege market and they are likely to choose majors with low premiums. On the admissions side, preferential treatment for blacks occurs only at top-tier schools. Removing the preferential treatment in admissions has little effect on earnings because the return to college quality is small and those blacks affected by the policy are most likely to attend college regardless of whether affirmative action is in place.

While the effects of affirmative action in higher education on expected earnings are small, removing affirmative action programs would have effects on the distribution of blacks at top-tier schools and the percentage of blacks attending college. Although removing affirmative action in admissions has a very small effect on the overall college attendance rate, I find that the percentage of black students falls dramatically at top-tier schools. For example, the percentage of black males attending colleges with average SAT scores above 1,200 falls by over 40%. In contrast, removing advantages in financial aid does not affect the distribution of blacks at the top schools as much as removing admissions advantages, but does have a larger effect on the college attendance rates. Even when controlling for unobserved ability, the parameter estimates imply an over 5% drop in the college attendance rates of black males had they faced the white financial aid rules.

Whereas the study of the returns to college has a rich history in labor economics, it is important to understand how this paper builds on the previous literature. Although this is the first study to quantify the effects of affirmative action in higher education by structurally modeling all aspects of the college decision-making process, many papers have estimated different parts of the full model. Three papers in particular have reduced form versions of many parts of the education process. Manski and Wise's (1983) book *College Choice in America* provides a series of chapters on college application,[4] admissions, financial aid, and enrollment. Bowen and Bok (1998) contains perhaps the most comprehensive description of the correlations between race and education in top-tier schools. Most relevant to the work here is their documentation of the black advantage in admissions at top-tier schools and the finding that higher earnings are correlated with attending higher quality colleges. Brewer, Eide,

[4]See Venti and Wise (1982) for the first paper on college application choice.

and Goldhaber (1999) estimate reduced form application, admissions, and enrollment rules. They document a larger role of affirmative action in the early seventies than in the early nineties. None of these papers models the links between the various parts of the college decision-making process. For example, estimation of advantages in admissions is not tied to expected future earnings or the choice of college. Furthermore, these papers do not control for selection on unobservables.

While the works discussed above have examined the general trends across a variety of college education decisions, many papers have focused on one aspect of the market for higher education. On the school side, while the literature is sparse, Kane (1998) documents black advantages in admissions at top-tier schools, while Kane and Spizman (1994) document similar advantages in financial aid across all schools. On the demand side, Fuller, Manski, and Wise (1982) and Brewer and Ehrenberg (1999) estimate multinomial logit models of the choice of college, with the latter also estimating the returns to college type and controlling for selection using the methodology developed in Lee (1983). Light and Strayer (2002) examine the decision to enroll and graduate from colleges of different qualities. They pay particular attention to race and find that, conditional on the same observed and unobserved characteristics, blacks are more likely to attend colleges of all quality levels. Both Berger (1988) and Arcidiacono (2004) model the choice of major and find large earnings differences for particular majors, even after controlling for selection.[5] Many studies have estimated the returns to college quality, with mixed evidence on how important college quality is to future earnings.[6] Particularly relevant is the work by Dale and Krueger (2002), who find much smaller effects of college quality on earnings when using information about what schools individuals were rejected at to control for unobserved ability. The same sort of variation is used to control for selection in this paper, but in the context of a structural model. This paper links much of the above literature by explicitly modeling each of the relevant decisions, which then makes policy analysis possible.

In addition to the detailed work on specific aspects of college education, there has been a vast literature on the returns to years of schooling.[7] Most relevant to the work here are the dynamic, structural models of Cameron and Heckman (1998, 2001) and Keane and Wolpin (1997, 2000, 2001). These papers look at much longer time horizons, trading off the details of the college education process for a more explicit modeling of the year-by-year decisions as to whether to further one's education.[8] The latter three papers estimate

[5]Grogger and Eide (1995), James et al. (1989), and Loury and Garman (1995) also document substantial earnings differences across majors.

[6]Daniel, Black, and Smith (1997), James et al. (1989), and Loury and Garman (1995) find strong positive effects of college quality.

[7]See Card (1999) for a review.

[8]One of the benefits of modeling the year-to-year transitions is that years of schooling are measured more accurately. By essentially allowing individuals only one choice about whether to

expectations of future utility, taking into account the option values of each of the possible decisions. In calculating these expectations, researchers face a trade-off between the correlation structure of the unobservable preferences and having closed form expressions for the expectations of future utility. One of the advantages of assuming a generalized extreme value (GEV) distribution for unobservable preferences is that, under certain conditions on the evolution of the state space,[9] closed form expressions exist for the expectations of future utility. The trade-off is the very restrictive correlation structure of the unobserved preferences, where the unobservable preferences usually take on either a multinomial logit or a nested logit form. With the nested logit, unobservable preferences within a nest share a common component but there is no correlation across nests. This paper applies a GEV framework developed in the industrial organization literature (Bresnahan, Stern, and Trajtenberg (1997)) that allows the unobservable preferences to be correlated across multiple nests. In this paper, unobservable preferences are correlated across both schools and majors while still having closed form expressions on the expectations of future utility.

The rest of the paper proceeds as follows. Section 2 presents the model and estimation strategy. Section 3 discusses the data. Results are presented in Section 4 with a discussion of how well the model matches the data given in Section 5. Policy simulations are examined in Section 6. Section 7 provides some concluding remarks as well as ideas for future research. Additional simulations are contained in the Appendix, which is available on the *Econometrica* website.

2. THE MODEL AND ESTIMATION STRATEGY

In this section I present a model of how individuals decide where to submit applications, where to attend college (conditional on being accepted), and what field to study. The model has four stages, which are outlined below.

Stage 1: Individuals choose where to submit applications.

Stage 2: Schools make admissions and financial aid decisions.

Stage 3: Conditional on the offered financial aid and acceptance set, individuals decide which school to attend and what field to study. Individuals may also choose to opt out of school altogether and enter the labor market.

Stage 4: All individuals enter the labor market.

Since decisions made in stage 1 are conditional on expectations of what will happen in the future, discussion of the model begins with stage 4 and works backward to stage 1.

There are two types of parameters for dynamic discrete choice models: transition parameters (γ's), which affect the probability of being in particular

advance one's education, those who drop out early or attend community colleges are mixed in with those who did not attend college at all.

[9]See Rust (1987, 1994).

states, and preference parameters (α's), which affect the utility of particular choices at particular states. Since transition or preference parameters appear in each stage of the model, to avoid confusion I subscript parameters and variables for each stage. Namely, parameters and variables for the labor market are subscripted by w, for the choice of college and major by c, for admissions by a, for financial aid by f, and for applications by s. Individual subscripts are suppressed.

Throughout, the discussion will be conducted as though all the errors in the various stages are independent of one another and, hence, each stage could be estimated separately—essentially assuming away the selection problem. This assumption will be relaxed later in the paper through the use of mixture distributions. Mixture distributions allow for the various stages to be connected through an individual's unobserved "type," controlling for the dynamic selection that occurs in the model. These unobserved types then affect the intercepts in each stage of the estimation as well as the expectations on future utility in both the application decision and the choice of college and major. The use of the mixture distribution is discussed in more detail in Section 2.6.

2.1. *Stage 4: The Labor Market and the Utility of Working*

Once individuals enter the work force, they make no other educational decisions: the labor market is an absorbing state. Individuals then receive utility only through earnings.[10] Earnings are a function of ability, A, where A is individual specific. I assume that the human capital gains for attending the jth college operate through the average ability of the students at the college, $\overline{A}_j$. In some majors individuals may acquire more human capital than in other majors, leading to earnings differentials across majors. Heterogeneity in these earnings differentials may also exist because the amount of human accumulation an individual obtains in a particular major may depend on their ability. Log earnings t years after high school are then given by

$$\ln(W_{jkt}) = \gamma_{wk1} + \gamma_{wk2}A + \gamma_{wk3}\overline{A}_j + \gamma_{wk4}X_w + g_{wkt} + \epsilon_{wt}, \tag{1}$$

where X_w is a vector of other characteristics that may affect earnings, k indicates major, and g_{wkt} represents how earnings grow over times. Should the individual not choose one of the college options, $j = k = 0$ and the variables that characterize the college are set to zero. The shocks (the ϵ_{wt}'s) are assumed to be distributed $N(0, \sigma_w^2)$.

Note that this model explicitly incorporates comparative advantage. Particular majors may have low wage intercepts, but high returns to ability. Similarly,

[10]Nonmonetary benefits in the work force for particular schooling paths will not be separately identified from the utility of those paths while in college. Hence, for ease of exposition, I speak of these only in terms of the utility in college.

other majors or the no-college option may have high intercepts, but low returns to ability.[11]

The data set I use to estimate the model is a short panel of high school graduates from a particular cohort. Accurately estimating growth rates for particular majors far out into the life cycle is not possible. Instead, I estimate the log earnings equation with year indicator variables interacted with sex and whether the individual choose one of the college options. Since I estimate the model using only one cohort, the coefficients on these year indicator variables will be a mixture of the returns to experience, age, and overall growth of the economy. The corresponding growth rates are then estimated only for years in which we actually have wage observations.

The expected utility of being in the work force is given by the log of the expected present value of lifetime earnings,[12]

$$u_{wjk} = \alpha_w \log\left(E_w\left[\sum_{t=t'}^{T} \beta^{t-t'} P_{kt} W_{jkt} \right] \right), \tag{2}$$

where T is the retirement date, t' is the year the individual enters the work force, and β is the discount factor. The probability of working in a particular year is given by P_{kt}. The expectation is then taken with respect to future labor force participation and shocks to earnings. I assume that, conditional on sex and major, all individuals have the same expectations regarding future labor

[11]As will be discussed in the section on unobserved heterogeneity, comparative advantage across schooling options will also be present because certain "types" of individuals will see higher returns in one major, but lower returns in another.

[12]This expression for the indirect utility function falls out of utility maximization when (i) period utility is given by log period consumption, (ii) there is perfect insurance on the lifetime earnings stream after one leaves college, and (iii) the discount factor is common across individuals and is equal to the market discount factor. Normalizing the price of consumption to 1, the individual's maximization problem is given by

$$\max_{c_t} \sum_{t=0}^{T} \beta^t \ln(c_t) \quad \text{s.t.}$$

$$\sum_{t=0}^{T} \beta^t c_t = E_w\left(\sum_{t=0}^{T} \beta^t P_t W_t \right).$$

The first-order conditions imply setting consumption to be the same across time periods. Substituting back into the budget constraint implies

$$c_t = \frac{E_w\left(\sum_{t=0}^{T} \beta^t P_t W_t\right)}{\sum_{t=0}^{T} \beta^t}$$

for all t. Substituting for c_t in the utility function gives the result.

force participation.[13] Under these assumptions, (2) can be rewritten as

$$u_{wjk} = \alpha_w(\gamma_{wk1} + \gamma_{wk2}A + \gamma_{wk3}\overline{A}_j + \gamma_{wk4}X_w) + \alpha_w \log\left(E_w\left[\sum_{t=t'}^{T} \beta^{t-t'} P_{kt} \exp(g_{wkt} + \epsilon_{wt})\right]\right), \tag{3}$$

where the indirect utility of working can be written as a linear term plus a function of the trends in participation and earnings over time. Since I do not have good information on earnings growth rates across years and majors, I assume that the expected growth rates are common across individuals of the same gender and major.[14] Hence, this last term, which includes both the growth dates and future participation decisions, is captured by sex interacted with a major-specific constant.[15] Note further that no assumptions need to be made on the discount factor because it too is absorbed into the major-specific intercept.

2.2. *Stage 3: Choice of College and Major*

At Stage 3, individuals may choose a school from a set J_a that includes all the schools that accepted them. The colleges themselves are not important; it is only the characteristics of the colleges that are relevant to the model. That is, utility from attending Harvard can be captured by the characteristics of Harvard. Those who decide to attend college must also choose a major from the set K. The same set of majors exists at all colleges. When making the college and major decisions, individuals take into account the repercussions these decisions have on future earnings.

Define the flow utility, u_{cjk}, as the utility received while actually attending college j in major k. This flow utility includes the effort demanded in major k at school j as well as any compensating differentials that may take place (such as college quality being a consumption good). Each of the majors then varies in its demands on the students. Let v_{cjk} be the corresponding expected present discounted value of indirect utility:

$$v_{cjk} = u_{cjk} + u_{wjk}. \tag{4}$$

Individuals then choose the option that yields the highest present value of lifetime utility.

[13]The sample of blacks in the data set is small, so blacks are allowed to have different preferences for attending college, but not for particular majors. Differences in unemployment rates for blacks are embedded in this model as long as they are proportional to the employment rates for whites conditional on attending or not attending college.

[14]See Arcidiacono (2004) for a similar specification.

[15]Since the utility of attending college also has a major-specific constant, these two constant terms will not be separately identified.

Individuals also have the option not to attend college, with the utility given by

$$v_{co} = u_{wo}, \tag{5}$$

where the o subscript indicates that the individual chose the outside option of working immediately.

I now specify in more detail the components of u_{cjk}. Embedded in this flow utility is the effort required to accumulate human capital in college. I assume that each major requires a fixed amount of work that varies by the individual's ability, A, the ability of one's peers, $\overline{A}_j$, and the major chosen, k. Hence, individuals with identical characteristics, attending schools with peers of similar abilities, and in the same major will have identical effort levels. This cost of effort is given by c_{jk}. The flow utility for pursuing a particular college option is then

$$u_{cjk} = \alpha_{c1} X_{cjk} - c_{jk} + \epsilon_{cjk}, \tag{6}$$

where X_{cjk} is a vector of individual, school, and major variables that affect how attractive particular education paths are. These include such things as the monetary cost of the school net of financial aid,[16] college quality as a consumption good, and whether particular sexes have preferences for particular majors. The vector X_{cjk} also includes major-specific intercepts.[17] The individual's unobserved preference for particular schooling options is given by ϵ_{cjk}.

I assume the following functional form for the cost of effort:

$$c_{jk} = \alpha_{c2k}(A - \overline{A}_j) + \alpha_{c3}(A - \overline{A}_j)^2. \tag{7}$$

Note that the cost of effort function allows the costs of majoring in particular fields to vary by relative ability in the linear term, but not in the squared term. While I will be able to identify α_{c3}, I will not be able to separately identify α_{c2k}, because college quality can serve as a consumption good and high ability individuals may have preferences for particular majors independent of effort costs; college quality and abilities of the students may enter the utility of majoring in a particular field through X_{cjk} as well. This cost of effort may lead to optimal qualities that are on the interior: even if an individual was allowed to attend all colleges, the individual may not choose to attend the highest quality college because of the effort required. With different levels of effort required by different majors, optimal college qualities may vary by major. Individuals are then trading off the cost of obtaining the human capital with the future benefits.[18]

[16] Recall that the indirect utility of earning took the log form, but here is linear. The two can be reconciled by assuming that the indirect utility is linear for sufficiently small values of consumption and individuals cannot borrow against their future earnings while in college.

[17] When the mixture distribution is added, these major-specific intercepts will be allowed to vary by type.

[18] See Arcidiacono (2004) for a more detailed discussion of the identification of effort costs.

Note also that many of the variables that affect the utility of college also affect the utility of working. To identify α_w, the coefficient on the log of the expected value of lifetime earnings, an exclusion restriction will be necessary. This is discussed in more detail in the data and identification sections.

I assume that ϵ_{cjk}'s follow a generalized extreme value distribution. Special cases of the generalized extreme value distribution lead to multinomial logit and nested logit models. With nested logit models, errors in one nest cannot be correlated with errors in another nest. For example, if we nested the choice of school, then it would not be possible to nest the choice of major also. However, a paper in the industrial organization literature, Bresnahan, Stern, and Trajtenberg (1997) (henceforth BST), shows that, consistent with McFadden's (1978) framework,[19] another special case of the generalized extreme value distribution allows for errors to be correlated across multiple nests while still being consistent with random utility maximization. I use this specification to allow errors to have a component that is common across all majors at a particular school and a component that is common across all school choices that involve the same choice of major. In particular, it is possible to have ϵ_c's that are correlated across both schools and majors using the G function

$$G(e^{v'_c}) = a_{c1}\sum_j\left(\sum_k \exp\left(\frac{v'_{cjk}}{\rho_{c1}}\right)\right)^{\rho_{c1}} + a_{c2}\sum_k\left(\sum_j \exp\left(\frac{v'_{cjk}}{\rho_{c2}}\right)\right)^{\rho_{c2}} + \exp(v'_{co}), \tag{8}$$

where $\rho_{c1}, \rho_{c2} \in [0, 1]$, $a_{c1} + a_{c2} = 1$, and $v'_{cjk} = v_{cjk} - \epsilon_{cjk}$ (the indirect utility net of the unobservable preference). The a_{c1} and a_{c2} terms are defined as

$$a_{c1} = \frac{1-\rho_{c1}}{2-\rho_{c1}-\rho_{c2}}$$

[19] McFadden's (1978) framework is as follows. Let $r = 1, \ldots, R$ index all possible choices. Define a function $G(y_1, \ldots, y_R)$ on $y_r \geq 0$ for all r. If G is nonnegative, homogeneous of degree 1, approaches $+\infty$ as one of its arguments approaches $+\infty$, has nonnegative nth cross-partial derivatives for odd n, and nonpositive cross-partial derivatives for even n, then McFadden (1978) showed that

$$F(\epsilon_1, \ldots, \epsilon_R) = \exp\{-G(e^{-\epsilon_1}, \ldots, e^{-\epsilon_R})\}$$

is the cumulative distribution function for a multivariate extreme value distribution. Furthermore, the probability of choosing the rth alternative conditional on the observed characteristics of the individual is given by

$$P(r) = \frac{y_r G_r(y_1, \ldots, y_R)}{G(y_1, \ldots, y_R)},$$

where G_r is the partial derivative of G with respect to the rth argument.

and

$$a_{c2} = \frac{1-\rho_{c2}}{2-\rho_{c1}-\rho_{c2}}.$$

Note that as ρ_{c1} (ρ_{c2}) approaches 1, a_{c1} (a_{c2}) approaches 0, and the G function reverts to the one used to derive a nested logit with errors correlated only within majors (schools). While the ϵ_c's are known to the individual, they are not observed by the econometrician. Hence, from the econometrician's perspective, the probability of choosing the kth major at the jth school is then given by

$$(9) \qquad P(j,k) = \frac{a_{c1}\exp\left(\frac{v'_{cjk}}{\rho_{c1}}\right)\left(\sum_k \exp\left(\frac{v'_{cjk}}{\rho_{c1}}\right)\right)^{\rho_{c1}-1} + a_{c2}\exp\left(\frac{v'_{cjk}}{\rho_{c2}}\right)\left(\sum_j \exp\left(\frac{v'_{cjk}}{\rho_{c2}}\right)\right)^{\rho_{c2}-1}}{G(e^{v'_c})}.$$

2.3. *Stage 2: Admissions and Financial Aid*

Given a set of applicants, schools decide who is admitted and how much financial aid will be given to each student. Entering into the school's utility function is the average ability of their students, $\overline{A}$, the sum of tuition payments net of any scholarships, and a school's unobserved preference for a particular student. The school side does not fall directly out of a well specified optimization problem for the schools themselves. However, I am not interested in how schools respond to changes in the application pool, but rather in how changing the school admission and financial rules affects student behavior. If schools respond to the removal of affirmative action by attempting to obtain a diverse student body using other means, the effect on blacks of removing affirmative action will be somewhat mitigated.

I assume that the admission rules that result from the school's maximization problem yield logit probabilities. The probability of being admitted to school j is then given by

$$P(j \in J_a | j \in J) = \frac{\exp[\gamma_a X_{aj}]}{\exp[\gamma_a X_{aj}]+1},$$

where X_{aj} includes such things as the quality level of the school and the individual's own ability, and γ_a is a vector of coefficients to be estimated.

I assume that the stochastic part of these probabilities is independent across schools. Hence, the probability that an individual who applies to the set of schools J has the choice set J_a is given by

$$(10) \qquad P(J_a|J) = \prod_j^{\#J} \left(\frac{\exp[\gamma_a X_{aj}]}{\exp[\gamma_a X_{aj}]+1}\right)^{j\in J_a} \left(\frac{1}{\exp[\gamma_a X_{aj}]+1}\right)^{j\notin J_a},$$

the product of the logit probabilities of the individual outcomes.

I now turn toward the financial aid decision. Write the bill paid by the student as $s_j t_j$, where t_j is the actual cost of attending school j and s_j is the share of that actual cost. I assume that the optimal financial aid rule follows a tobit with s_j as the dependent variable. In particular, we have

$$(11) \qquad s_j^* = \gamma_f X_{fj} + \epsilon_{fj},$$
$$s_j = \begin{cases} 0, & \text{if } s_j^* \leq 0, \\ 1, & \text{if } s_j^* \geq 1, \\ s_j^*, & \text{if } 0 < s_j^* < 1, \end{cases}$$

where X_{fj} is a vector of individual and institutional characteristics that affect financial aid outcome. Hence, the share of the bill paid will have mass points at no aid and full aid, which is consistent with the data. The forecast error, ϵ_{sj}, is independent of X_{fj}, independent across schools, and is unknown to the student until after the application decision has been made.[20]

2.4. *Stage 1: Applying to College*

Let there be a set of $\mathbb{J}$ colleges to which an individual may submit an application. Since each school may accept or reject the student, the number of possible outcomes for applying to all the schools in $J \subset \mathbb{J}$ is $2^{\#J}$, where $\#J$ is the number of schools in subset J. Let J_a indicate the subset of schools at which the individual was accepted and let $P(J_a)$ be the corresponding probability that this outcome occurs. Individuals make their application decisions based on their expectations on the probability of acceptance, the expected financial aid conditional on acceptance, and an expectation of how well they will like particular college and major combinations in the future.[21] I assume that the expected utility of applying to the set J, v_{sJ}, is given by

$$v_{sJ} = \alpha_{s1} E_s(V_c|J) + u_{sJ} + \epsilon_{sJ},$$

where V_c is the value of the best alternative at the college and major choice stage, u_{sJ} is the flow utility (application cost) of applying to the set J, and ϵ_{sJ} is the unobserved preference for applying to the set J.[22] By assuming a particular

[20]Both this assumption and the assumption of independent errors in admissions is relaxed later in the paper as the intercept terms in both equations are allowed to vary by type. Hence, highly productive people may see higher earnings as well as higher probabilities of admittance and financial aid at all colleges.

[21]Note that at this stage the individual knows the gross costs of attending each school. However, he only has expectations with regard to how much of those gross costs he will actually have to pay, because financial aid is uncertain.

[22]Note that even though $E_s(V_c|J)$ is already denominated in utils, there is still a coefficient on the variable $\alpha_{s1} = \mu_c/\mu_s$, where μ_c and μ_s are the variance scale parameters for the choice of

functional form on the application costs and taking into account the probability of being admitted into each school in the set, the expression becomes

$$v_{sJ} = \alpha_{s1} \sum_{a=1}^{2^{\#J}-1} E_s(V_c|J_a)P(J_a|J) - \alpha_{s2}X_{sJ} + \epsilon_{sJ}, \tag{12}$$

where X_{sJ} represent the variables that affect the cost of applying to set J.[23]

I assume that unobservable tastes for particular schools and majors, the ϵ_c's, are independent from the ϵ_s's. All individuals then have the same expectations with regard to the realizations of the ϵ_c's. I need this assumption to make the expectations on future utility of applying to a particular set of schools J tractable.[24] Integrating out and discretizing the financial aid realizations into L categories, Rust (1987) showed that the conditional expectations have a closed form solution. With these assumptions, (12) can be rewritten as

$$\begin{aligned} v_{sJ} = \alpha_{s1} \sum_{a=1}^{2^{\#J}-1} \Bigg(\sum_{l=1}^{L} \ln \Bigg[& a_{c1} \sum_{j} \bigg(\sum_{k} \exp\bigg(\frac{v'_{cjk}}{\rho_{c1}}\bigg)\bigg)^{\rho_{c1}} \\ & + a_{c2} \sum_{k} \bigg(\sum_{j} \exp\bigg(\frac{v'_{cjk}}{\rho_{c2}}\bigg)\bigg)^{\rho_{c2}} \\ & + \exp(v'_{co}) \Bigg] \pi(s_{al}|X_{fJ})\, ds_a \Bigg) P(J_a|X_{aJ}) \\ - \alpha_{s2}X_{sJ} + \gamma + \epsilon_{sJ}, & \end{aligned}$$

where π is discretized probability density function (p.d.f.) of s_a, the financial aid decisions at each of the schools in the acceptance set, and γ is Euler's constant.

With the calculation of the expected value of lifetime utility in hand, I now specify the distribution of the taste parameters, the ϵ_s's. Similar to the college and major choice stage, it is reasonable to assume that unobserved preferences

school and major stage, and the application stage, respectively. Typically with multinomial logits these scale parameters are assumed to be 1 so as to identify the parameters of the utility function. Since we have, in a sense, two multinomial logits on two very different decisions (applying versus attending) that are connected by the expected utility term, we can only identify one variance term relative to the other. Any discounting across stages will not be separately identified, but will instead be incorporated into this parameter.

[23]The -1 in the summation takes the sum only over acceptance sets where the individual was accepted to at least one school.

[24]When unobserved heterogeneity is added through the use of mixture distributions, the expectations on the values of particular application sets will vary with type. That is, high unobserved ability individuals will have higher expectations with regard to the value of all application sets.

for application bundle where some of the schools overlap and should be correlated. I specify a distribution where each school has its own nest. Hence, the nest for the first school in $\mathbb{J}$ includes all J's that have as one of their elements the first school. An application of BST's framework once again applies. Namely, order the schools in $\mathbb{J}$ from $1, \dots, N$. Order the possible combinations of schools from $1, \dots, R$ and let J_r denote the set of schools in the rth combination. The G function I use is then given by

$$G(v'_s) = \sum_{n=1}^{N} \frac{1}{M} \left(\sum_{r=1}^{R} (n \in J_r) \exp\left(\frac{v'_{sr}}{\rho_s} \right) \right)^{\rho_s} + \sum_{r=1}^{R} \left(1 - \sum_{n=1}^{N} \frac{(n \in J_r)}{M} \right) \exp(v'_{sr}) + \exp(v'_{so}),$$

where $v'_{sr} = v_{sr} - \epsilon_{sr}$, M is the maximum number of schools to which one can apply, and $\rho_s \in [0, 1]$ is the nesting parameter. The last term allows the individual to apply to no schools, while the second to last term adjusts the G function for the fact that some combinations of schools have smaller numbers of schools than others. That the ρ_s is common across schooling nests restricts the correlation within each school's nest to be the same.

This G function then leads to the probability of choosing the application set J_r, which is given by

$$P(J_r) = \left(\sum_{n=1}^{N} \frac{1}{M} \left(\sum_{r'=1}^{R} (n \in J_r)(n \in J_{r'}) \exp\left(\frac{v'_{sr'}}{\rho_s} \right) \right)^{\rho_s - 1} \exp\left(\frac{v'_{sr}}{\rho_s} \right) + \left(1 - \sum_{n=1}^{N} \frac{(n \in J_r)}{M} \right) \exp(v'_{sr}) \right) \Big/ G(v'_s).$$

Note that if the number of schools in J_r equals M, then the second term disappears. As in the college and major choice stage, as ρ_s approaches 1, the model reverts to a multinomial logit.

2.5. *The Estimation Strategy*

With independent errors across the stages, the log-likelihood function can now be divided into five pieces:

$L_1(\gamma_w)$—the log-likelihood contribution of earnings;

$L_2(\gamma_a)$—the log-likelihood contribution of admissions decisions;

$L_3(\gamma_f)$—the log-likelihood contribution of financial aid;

$L_4(\alpha_c, \alpha_w, \gamma_w)$—the log-likelihood contribution of college and major decisions conditional on the acceptance set;

$L_5(\alpha_s, \alpha_c, \alpha_w, \gamma_a, \gamma_f, \gamma_w)$—the log-likelihood contribution of the application decision.

The total log-likelihood function is then $L = L_1 + L_2 + L_3 + L_4 + L_5$.

Note that consistent estimates of γ_w, γ_a, and γ_f can be found by maximizing L_1, L_2, and L_3 separately.[25] With the estimates of γ_w, consistent estimates of α_c and α_w can be obtained by maximizing L_4. All of these estimates can then be used in L_5 to find consistent estimates of α_s.[26]

The computational savings from employing this method are quite large. The expectation on the value of applying to any reasonable number of schools is very expensive to calculate, let alone calculate the derivative. This method minimizes the number of times this expectation needs to be calculated. The maximization then reduces to ordinary least squares for the earnings estimates, a logit at each school for the admissions estimates, a tobit for the financial aid estimates, and two multinomial logits for the college and major decision and the application decision.

2.6. *Serial Correlation of Preferences and Unobserved Ability*

One of the assumptions that seems particularly unreasonable is that the unobservable preference parameters are uncorrelated over time. That is, if one has a strong unobservable preference for engineering initially, he is just as likely as someone who has a strong unobservable preference for education initially to have an unobservable preference for education when it comes time to choose a college and a major. We would suspect that this is not the case. Furthermore, it is unreasonable to assume that there is no unobserved (to the econometrician) ability that is known to the individual.[27]

Following Heckman and Singer (1984), one method of dealing with this problem is to assume that there are R types of people with π_r being the proportion of the rth type in the population.[28] Types remain the same throughout all stages, individuals know their type, and preferences and abilities may vary across type.[29] The log likelihood for a particular individual then follows the

[25] See Rust and Phelan (1997) and Rothwell and Rust (1997).

[26] The standard errors, however, are not consistent. I take one Newton step on the full likelihood function to obtain consistent estimates of the standard errors.

[27] See Willis and Rosen (1979) for the importance of controlling for selection when estimating the returns to education.

[28] See Keane and Wolpin (1997), Eckstein and Wolpin (199), and Cameron and Heckman (1998, 2001) for other examples of using mixture distributions to control for unobserved heterogeneity in dynamic education models.

[29] An example would be if the parameters of the utility function do not vary across types except for the constant term. This would be the same as having a random effect that is common across everyone of a particular type.

mixture distribution

$$(13)\qquad L(\alpha_a, \alpha_c, \alpha_w, \gamma_a, \gamma_s, \gamma_w) = \ln\left(\sum_{r=1}^{R} \pi_r \mathcal{L}_{1r}\mathcal{L}_{2r}\mathcal{L}_{3r}\mathcal{L}_{4r}\mathcal{L}_{5r}\right).$$

Here, the α's and γ's can vary by type and $\mathcal{L}$ refers to the likelihood (as opposed to the log likelihood).

Now the parts of the log-likelihood function are no longer additively separable. If they were, a similar technique could be used as in the case of complete information: estimate the model in stages with the parameters of previous stages being taken as given when estimating the parameters of subsequent stages. Arcidiacono and Jones (2003) show that the expectation–maximization (EM) algorithm[30] restores the additive separability at the maximization step.

In particular, note that the conditional probability of being the rth type is given by

$$(14)\qquad P(r|\mathbf{X}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\pi}) = \frac{\pi_r \mathcal{L}_{1r}\mathcal{L}_{2r}\mathcal{L}_{3r}\mathcal{L}_{4r}\mathcal{L}_{5r}}{\sum_{r'=1}^{R} \pi_{r'} \mathcal{L}_{1r'}\mathcal{L}_{2r'}\mathcal{L}_{3r'}\mathcal{L}_{4r'}\mathcal{L}_{5r'}},$$

where $\mathbf{X}$ refers to the data on the decisions and the characteristics of the individual.

The EM algorithm has two steps: first, calculate the expected log-likelihood function given the conditional probabilities at the current parameter estimates; second, maximize the expected likelihood function, holding the conditional probabilities fixed. This process is repeated until convergence is obtained, but the expected log likelihood for a particular observation is now additively separable:

$$(15)\qquad \sum_{r=1}^{R} P(r|\mathbf{X}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\pi})\big(L_{1r}(\gamma_w) + L_{2r}(\gamma_a) + L_{3r}(\gamma_f) + L_{4r}(\alpha_c, \alpha_w, \gamma_w) + L_{5r}(\alpha_a, \alpha_c, \alpha_w, \gamma_a, \gamma_f, \gamma_w)\big).$$

Taking the conditional probabilities of being a particular type as given, I can obtain estimates of γ_w by maximizing the L_{1r}'s times the conditional probabilities of being a particular type. Similarly, estimates of γ_a and γ_f come from maximizing the conditional probabilities times the L_{2r}'s and L_{3r}'s, respectively. I then use only the L_{4r}'s and the conditional probabilities to find estimates of α_c and α_w—not needing L_{4r} to obtain estimates of γ_w. These estimates are

[30]See Dempster, Laird, and Rubin (1977) for the seminal paper on the EM algorithm.

taken as given and the L_{5r}'s are used only to find the α_s's. Note that all of the parts of the likelihood are still linked through the conditional probabilities, where the conditional probabilities are updated at each iteration of the EM algorithm. Arcidiacono and Jones (2003) show that this method produces consistent estimates of the parameters with large computational savings.

Note also that the population probabilities of being a particular type, the π_r's, can depend on characteristics of the individual. Here, type is allowed to depend on whether one comes from a high or low income family.[31] Allowing income to affect type probabilities in this way does not affect the sequential EM algorithm. Rather than updating the unconditional probability of being a particular type using the whole sample, the unconditional probabilities now vary by income and are updated individually using the sample of high and low income households respectively.

3. DATA

The National Longitudinal Study of the Class of 1972 (NLS72) is the primary data source for estimating the model. The NLS72 is a stratified random sample that tracks individuals who were seniors in high school in 1972. Individuals were interviewed in 1972, 1973, 1974, 1976, 1979, and 1986. Table I provides descriptive statistics for the whole sample, those who applied to college, and those who attended college. All statistics are disaggregated by race.[32]

The NLS72 has data on the top three schooling choices of each individual in 1972 and whether the individual was accepted to each of these schools. Then J_a is defined as the (up to three) schools that the individual listed as accepting him. Unfortunately, the NLS72 does not have data on whether an individual considered any other four-year institutions. Hence, I may only be partially observing J_a. However, this turns out to be not very restrictive, because the percentage of students who report applying to three schools is small: approximately 10% of those who apply to college submit three applications.

To keep the model computationally tractable, I need to restrict the number of schools where the individual can submit an application. It is not possible to

[31] Definitions of high and low income are given in Section 3.

[32] The sample was selected as follows. To be in the final data set, survey respondents must have participated in the first follow-up study as well as have valid test score information from taking either the SAT or the standardized test given to the NLS72 participants in the first year. Of these, 6,940 did not apply to a four-year college in their senior year of high school. Of the 7,877 who did apply, 4,645 had a valid schooling option. Of the 4,645, 3,670 had valid information on major choice and did not transfer schools. This last set of attrition resulted in a 24% reduction in the sample of those who submitted an application. I therefore reduced (randomly) the sample of those who did not apply by 24% as well, leaving the sample of nonapplicants at 5,244. This attrition had no effect on the ratio of men to women or blacks to whites. High school dropouts are ignored.

TABLE I

SAMPLE MEANS

	Full Sample		Applied		Attended	
	White	Black	White	Black	White	Black
Prob. of applying	0.4115	0.4133				
Prob. of attending	0.2114	0.1667	0.5137	0.4033		
Prob. of admission			0.9121	0.8609		
Number	0.5924	0.5809	1.4397	1.4006	1.5772	1.5491
of applications	(0.8312)	(0.8066)	(0.6777)	(0.6503)	(0.7443)	(0.7246)
Math school			535.6	460.3	538.2	466.5
quality			(55.2)	(100.2)	(51.2)	(104.5)
Verbal school			508.2	438.8	509.6	446.1
quality			(52.8)	(97.5)	(48.9)	(101.8)
School cost[a]			11,505	10,596	11,403	10,632
			(4,192)	(3,843)	(4,003)	(4,124)
Financial aid			1,250	2,180	1,456	3,195
			(2,736)	(3,796)	(2,930)	(4,594)
State college	0.2684	0.2949	0.2705	0.2959	0.2722	0.2999
premium[b]	(0.0621)	(0.0566)	(0.0623)	(0.0586)	(0.0621)	(0.0600)
SAT math	442.1	334.3	500.3	360.9	529.8	378.9
	(104.2)	(70.0)	(105.9)	(82.1)	(101.2)	(85.6)
SAT verbal	410.3	305.5	465.4	332.4	489.8	356.9
	(100.7)	(67.5)	(102.8)	(79.5)	(99.5)	(87.5)
HS class rank	0.5589	0.4677	0.6983	0.5710	0.7633	0.6212
	(0.2780)	(0.2749)	(0.2351)	(0.2629)	(0.2032)	(0.2546)
Unknown HS	0.1143	0.2322	0.1373	0.2634	0.1249	0.2486
class rank						
Low income[c]	0.4410	0.7139	0.3508	0.6737	0.3169	0.6416
Female	0.4950	0.5732	0.4736	0.6107	0.4757	0.6358
Natural science					0.2246	0.1329
Business					0.1628	0.1734
Social science					0.4450	0.4855
Education					0.1676	0.2081
Observations	7,876	1,038	3,241	4,29	1,665	173

[a] Costs and aid are in 1999 dollars. Cost is defined as tuition + books + room and board. Financial aid is scholarships only. Both costs and financial aid are for all schools applied to in the second column and only the school attended in the third column.

[b] Defined as family's before tax income being less than $36,000 (1999 dollars).

[c] Taken from the 1973–1975 March Current Population Surveys.

estimate a model where an individual may apply to all combinations of schools in the United States. I restrict the set of schools to eight, with individuals being able to apply to any combination of up to three schools from these eight. Schools were assigned randomly, with 70% of the draws coming from schools in the same state as the student. All colleges where an individual actually submitted an application were included in the choice set. This leaves 92 possible application sets.

The first four rows in Table I then show how the probability of applying, being admitted, and attending vary by race. Blacks and whites were equally likely to submit any applications and the average number of applications conditional on applying was 1.4 for both groups. The unconditional probability of being admitted is higher for whites, with blacks being over 50% more likely to be rejected.

I use data on decisions made in 1974 as to whether to attend college. This should roughly correspond to the junior year of college. The data for those who chose a schooling option are then restricted to students who were attending a college in their original choice set: transferring schools is not modeled. Individuals who attend college but drop out before the junior year are treated as not having attended college. With this measure, Table I shows that the college attendance rates conditional on applying are 25% higher for whites.[33]

The next set of rows shows the characteristics of the schools applied to and attended as well as how much aid was offered. The schools' math and verbal SAT scores, as reported by the schools themselves, are used as my measures of school quality.[34] James et al. (1989) find that the only measure of college quality that significantly affects earnings is student quality. This does not necessarily imply peer effects because student quality may be driven by the quality of the instruction, which is not easily measured. Costs are calculated as tuition plus books plus room and board. Both the college quality and the cost data are taken from the schools themselves. Individuals list their general scholarships as well as school-specific scholarships. The only measure of financial aid I use is this scholarship data. Scholarships are constrained to be less than the total cost of the school. There is much censoring, because over 60% of individuals receive no financial aid from scholarships.

The table shows that whites apply to and attend colleges with much higher SAT math and verbal scores than their black counterparts. Cost may have something to do with this because the blacks applied to college that were on average less expensive and also received more aid than whites. Although the out of pocket expenses were on average lower for blacks who chose to attend, this was not the case for whites. A possible explanation is that there are more high income whites and this group may have received poor financial aid packages yet still planned to attend college.

With blacks applying to worse schools and being admitted at lower rates, the role of affirmative action may seem limited. However, Brewer, Eide, and

[33]The percentage of individuals attending college is slightly lower than their counterparts in the Current Population Survey. This has two sources. First, to be counted here, no breaks are allowed in college attendance. Second, our sample does not use information on community colleges or vocational schools.

[34]These data, as well as the tuition and various other costs of attending the school, are found in The Basic Institutional Source File. This file has information from the 1973–74 Higher Education Directory, the 1973–74 Tripariate Application Data file, the 1972–73 HEGIS Finance Survey, and the 1972 ACE Institutional Characteristics File, all of which are surveys of colleges.

Goldhaber (1999) document that affirmative action in college during this time period (early seventies) was actually much stronger than in later years. These facts may be reconciled by examining the third set of rows, which points toward large differences in backgrounds of the white and black populations.

The difference in SAT scores is particularly striking. Examining columns 1 and 6 shows that blacks who attended college performed over 50 points worse on both the math and verbal sections of the SAT[35] than whites in the population at large. These gaps increase to over 150 points and 130 points for the math and verbal scores when comparing white and black attendees. For both groups, SAT scores are higher for those who attended than for those who just applied and higher for those who applied than for those who did not.

Large differences also exist in both high school class rank[36] and in the percentage of families that are low income, where low income is defined as being below the median family income in the data for those who report a family income.[37] In both cases and in all groupings, blacks have lower high school class ranks and are more likely to come from low income families than whites. Consistent with the results on SAT math score, both higher incomes and class ranks are found as we move from the full sample to the sample that attended college. Black females were more likely to attend college than their male counterparts, but this is not the case for whites.[38]

To identify the coefficient on future earnings, I need a variable that affects choice of school and major only through earnings; X_w must have an element not contained in X_c. I use state differences in the log college premium from 1973–1975 for workers aged 22–35 as a variable that affects choice of school and major only through earnings. This variable is calculated from the March Current Population Survey (CPS). Some small states are aggregated in the CPS, leading to differences in the college premiums across 22 regions. The descriptive statistics indicate this may be a good choice because for both blacks and whites the highest college premiums are found for those who attended col-

[35]Many individuals, particularly those who did not apply to college, did not take the SAT. However, participants in the NLS72 took both a math and a verbal test in the base year of the survey. For those who did take the SAT, I regressed SAT math and SAT verbal on the score of the math and the verbal tests, respectively. These results were then used to forecast SAT math and verbal scores for those who did not take the SAT. Those who took neither the SAT nor the NLS72 test were removed from the sample. Some individuals did not take this standardized test. These individuals were slightly more likely to be black (18% compared to the final sample of 12%), but were equally likely to be female. Individuals in this group were much more likely not to have responded to the other survey questions.

[36]This variable is taken from the high schools themselves and is on a percentage basis where a 1 indicates the student was at the top of the class.

[37]About one-third of the sample did not report a family income. Based on their behavior, these individuals look like they come from high income families and are coded as such.

[38]Note that the percentage of females is much higher for blacks in the full sample than for whites. This is true in the survey itself and is not a result of the selection rules used to obtain the sample.

lege, with higher values for those who applied than in the population at large. It is interesting to note that blacks were more likely to live in states where the college premium was high.

Motivated by the work in Arcidiacono (2004), majors are aggregated into four groups: engineering, physical sciences, and biological sciences in group 1, business and economics in group 2, social sciences, humanities, and other in group 3, and education in group 4. The maximum number of choices available at the college decision stage is then thirteen: four majors for each of three schools and a work option. Differences in major choice exist across the races, with blacks being substantially less likely to choose natural science majors.

The final row gives the number of observations in each category. Note that the number of blacks is quite small, with only 173 attending college. This limits what race effects can be identified in the model. At the earnings stage, I allow earnings by blacks to differ from those of whites and allow this difference to vary based on whether the individual attended college. At the college stage, I allow blacks to have a preference for attending (or not attending) college. In both the earnings and college choice stages, the limited observations make it impossible to estimate major-specific wage premiums or preferences for particular majors that differ by race. At both the admissions and the financial aid stage, race is taken into account in the intercept term, in an interaction with college quality, and in an interaction with low income. The interaction with college quality is implemented because past research (see Bowen and Bok (1998), Brewer, Eide, and Goldhaber (1999), and Kane (1998)) has found that racial preferences are only present at top-tier colleges. The black effect on admissions and aid is also constrained to be positive. Hence, blacks may face the white rules up to a particular college quality level and receive advantages after that point. This issue is discussed in more detail in the results section.

4. IDENTIFICATION

As discussed in the model section, whereas the schooling decisions are choices, it is important to control for selection at all stages of the model. All characteristics of the individuals themselves are taken as exogenous, including such things as test scores and parental income. With no correlations across the various stages of the model, selection into schools and majors is only controlled for by these exogenous characteristics. With the mixture distribution, however, errors are allowed to be correlated across the various stages, accounting for selection into schools and majors on unobservables. That is, each individual is of a particular type, where type is unobserved and integrated out of the likelihood function. With the data described in the previous section, I now discuss what features of the data are used to identify types. As is standard in the dynamic discrete choice literature, selection is controlled for via a mixture over unobserved types. In particular, the types are identified by the dynamics of the model and exclusion restrictions.[39]

The dynamics of the model help to identify types through realizations of the admittance and financial aid decisions as well as through educational decisions. For example, someone who has a strong preference to attend college but is weak on unobservable ability may apply to many schools, be rejected by many schools, and have low earnings. Similarly, a person with high unobserved ability may apply to one school and get an outstanding financial aid package. These realizations on financial aid and admissions do not directly enter into the log wage regression, yet those who receive higher than expected (by the econometrician) realizations on these two variables may have higher unobserved ability. Similarly, not being admitted to a particular school does not enter the college and major choice decision. Consider two individuals who are identical in their observable characteristics and apply to the same school. Suppose one of these individuals also applies to a different school but is rejected. Without unobserved heterogeneity, these two individuals would be predicted to make the same decisions on the choice of college and major. However, with unobserved heterogeneity we learn something about both the preferences of each individual for attending college and their unobserved ability through the rejection. It is this kind of variation that helps to identify the unobserved types.

Exclusion restrictions from the characteristics of the individuals themselves also are used to identify the unobserved types. Namely, family income, discretized into high and low, appears in the admissions and financial aid decisions, and also as an interaction with net cost in the college and major decisions, but does not appear in the log wage regression. Parental income is directly related to the probability of financial aid and may make paying for college less attractive due to liquidity constraints. Parental income, however, may also be related to an individual's unobserved ability. This correlation then helps to identify type. I also make assumptions on the role of math and verbal college quality and math and verbal ability. Arcidiacono (2004) found that math ability and college quality were much more important than their verbal counterparts in major selection and future earnings. I assume that the effects of ability and college quality operate only through the math channel for major selection and future earnings. However, for admissions and aid, the total SAT score is used. Hence, SAT verbal scores will dictate whether some individuals go to college and others do not due to the opportunity to get into better colleges and receive better aid packages. At the margin, high unobserved ability and high verbal ability may serve as substitutes. Finally, the premium college graduates receive in particular states is assumed to affect earnings, but affects the choice of college and major only through earnings. In states where the college premium is higher, those with lower unobserved ability may be induced

[39]Functional form invariably helps to identify types as well, although we would like to identify the types by more than functional form.

to choose a college option. These individuals know they will have to worker harder in college than their high unobserved ability counterparts, but are willing to do so for the higher wages.

5. RESULTS

In this section, I present the estimation results. I begin with the school side—admissions and financial aid—before proceeding with the student side. Throughout, two models are presented. One does not place any controls for unobserved heterogeneity; there is only one type of person. This model has correlations in the errors within the application, and college and major choice stages, but has no correlation across the stages of the model. The other allows for two types, where one's type affects all aspects of the problem from the application decision to expected earnings through a dummy term.[40] These types work as random effects and link all stages of the model. The population probability of being a particular type is allowed to vary with income level.[41]

5.1. *Admissions*

Table II presents estimates of the admissions logit with and without controls for unobserved heterogeneity. Comparing the two models shows that, while type 2's have a higher probability of being admitted,[42] the coefficient estimates of the rest of the parameters are similar across the two specifications. Increasing one's own SAT score as well as one's high school class rank both increase the probability of being accepted. However, increasing both one's own SAT score and the average SAT score of the school where the individual is applying by the same amount results in decreasing the probability of being accepted. College admissions are not particularly competitive until high levels of college quality are reached.[43] Neither the individual's gender nor whether the school was private had a significant effect on the probability of being admitted.

Recall that the black effect is constrained to be positive. While a more flexible functional form would perhaps yield this result on its own, the small num-

[40]Models with more types were also estimated. The standard errors on these models blew up and yielded nonsensical results. The reason for the difficulty is because we observe only one college and major decision. Identifying additional preference parameters for that particular stage requires additional structure. A three-type model, where the effect of type in the college and major choice stage was constrained to be proportional to the effect of type on earnings, yielded similar results to the two-type model.

[41]Keane and Wolpin (1997, 2000, 2001) also allow the population type probabilities to depend on income level through a logit specification. In this paper, income level affects the population type probabilities by allowing the means to differ across high and low income individuals.

[42]A white male with a 1,200 SAT score, at the 90th percentile for high school class rank, and applying to a college with a 1,200 average SAT score would be admitted 87% of the time if he was type 1, but 91% of the time if type 2.

[43]Venti and Wise (1982), Bowen and Bok (1998), and Kane (1998) found similar results.

TABLE II

LOGIT ADMISSION PROBABILITIES[a]

	One Type		Two Type	
	Coefficient	Standard Error	Coefficient	Standard Error
Female	−0.0925	0.0847	−0.1048	0.0849
Black	−4.2959	1.2933	−4.1743	3.0992
SAT (000's)	2.6531	0.2484	2.6930	0.2479
HS class rank	1.5728	0.2166	1.4943	0.2163
Do not know rank	0.8740	0.1749	0.8164	0.1745
Low income	−0.0134	0.0908	0.0178	0.0919
Black × Low income	0.2782	0.5265	0.2700	0.5579
School quality (000's)	−8.2513	0.2398	−8.3171	0.2478
Black × School quality	3.8633	1.0452	3.7505	2.5449
Private	0.0579	0.0910	0.0507	0.0910
Type 1	7.3661	0.2104	7.2475	0.2154
Type 2			7.6856	0.2281

[a]In this stage, 5,269 observations are used from 3,670 individuals.

ber of black applicants in the data prohibits a less restrictive functional form. Hence, if particular levels of school quality would imply a negative effect of being black, the white admissions rule is used. The black variable and the corresponding interactions with low income and school quality show large differences from the coefficients for whites. Low income black students are given a small advantage over other black students. The overall effect of being black on the probability of admission depends on the quality of the school. The estimates for both models show advantages for high (low) income blacks beginning at college quality levels above 1,100 (1,040).[44]

Figure 1 displays the admissions probabilities as a function of college quality for five representative agents: two white high income males, with SAT scores of 800 and 1,200, respectively, two high income black males, again with SAT scores of 800 and 1,200, respectively, and one low income black male with an SAT score of 800. From the graph, it is clear that the advantages blacks have in admissions occur only at very high levels of college quality. Furthermore, these advantages have little to do with the income level of the black student: an increase in SAT score of 400 points yields a much higher increase in the probability of admission than being low income.

[44]Note that for the black variables, although not affecting the magnitudes of the coefficients, adding controls for unobserved heterogeneity substantially increased the standard errors for these variables. While the individual variables are no longer significant, the black variables as a whole are jointly significant. Even with these large standard errors, the standard errors on the policy simulations are small.

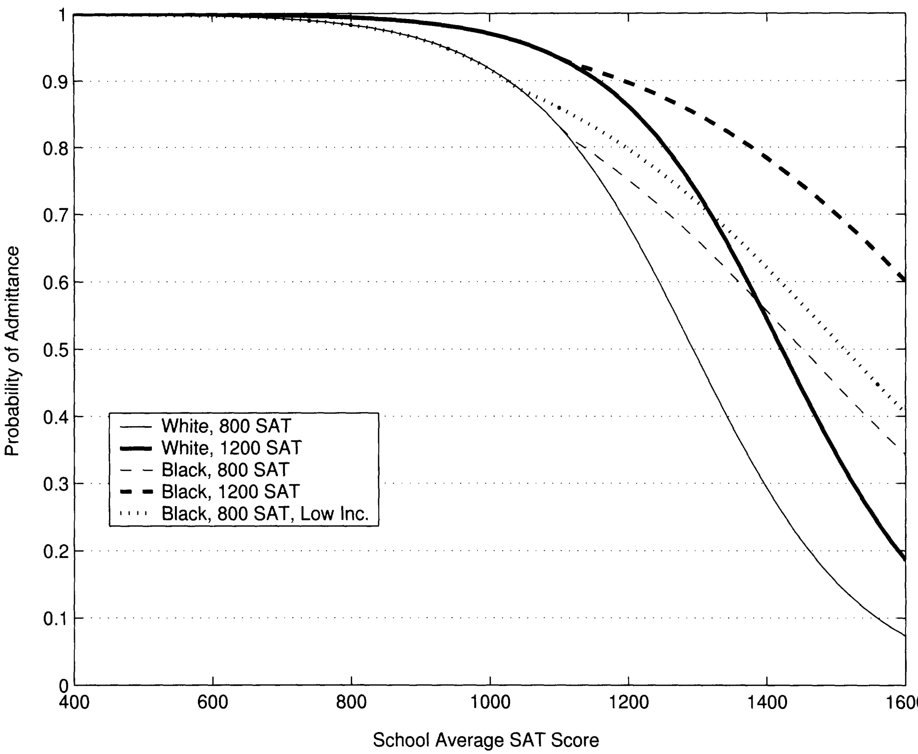

FIGURE 1.—Admissions as a function of race, school quality, and ability. High school class rank is held at the 75th percentile. The graph is for estimates without controls for unobserved heterogeneity.

5.2. *Financial Aid*

Table III gives estimates of the financial aid tobit. Similar to the admissions results, while type 2's receive more aid, adding controls for unobserved heterogeneity did not significantly affect the other parameter estimates. Those who have high SAT math scores and class ranks have higher probabilities of receiving good aid packages. College quality reverses here as high quality colleges appear to be more generous in offering to pay for a percentage of the total costs. Private schools also offer larger financial aid packages. As expected, low income students receive better packages than those who are not low income. Gender was again insignificant.

In addition to admissions, black students also face different financial aid rules. The black coefficient is positive and significant. However, the interaction between black and low income is negative. The interaction of black and college quality is positive but insignificant.

To get a sense of the magnitude of the black advantage in financial aid, I calculate the probability of receiving aid for a male with an 800 SAT score and

TABLE III

TOBIT ESTIMATES OF THE SHARE OF COSTS PAID BY THE SCHOOL[a]

	One Type		Two Type	
	Coefficient	Standard Error	Coefficient	Standard Error
Female	0.0115	0.0140	0.0073	0.0142
Black	0.3218	0.0876	0.3063	0.0920
SAT (000's)	0.3236	0.0428	0.3365	0.0437
HS class rank	0.4066	0.0377	0.3869	0.0384
Do not know rank	0.2950	0.0326	0.2794	0.0331
Low income	0.3491	0.0158	0.3566	0.0162
Black × Low income	−0.2413	0.0372	−0.2441	0.0374
School quality (000's)	0.3737	0.1060	0.3459	0.1131
Black × School quality	0.1046	0.1682	0.1235	0.1777
Private	0.1789	0.0169	0.1766	0.0171
Type 1	−1.4676	0.0604	−1.4915	0.0634
Type 2			−1.3856	0.0632
Variance	0.5150	0.0103	0.5128	0.0104

[a]In this stage, 4,710 observations are used from 3,459 individuals.

a 0.8 high school class rank, who applies to a public school with school quality equal to 800 for the model without controls for unobserved heterogeneity.[45] The probability of receiving aid given the preceding characteristics is 47.2% for a low income black student, 42.9% for a high income black student, and 40.7% for a low income white student. Conditional on receiving some aid, but not full aid, a school with an 800 average SAT score pays an additional 16.4% and 5.6% of the total bill for low income and high income black students than a similar low income white student.

5.3. *Earnings*

Estimates of the earnings parameters are given in Table IV. The 1986 earnings are used as the base year; the coefficients on the year dummies and the year dummies interacted with sex are omitted. Log mean state earnings conditional on education,[46] which is our one variable that affects schooling choices only through earnings, is positive and significant in both specifications. Math ability is positive for all majors and for those who do not attend college regardless of controls for unobserved heterogeneity.

The constraint that math college quality has a positive effect on earnings binds solely for education regardless of whether controls for unobserved het-

[45]Similar results hold when the controls for heterogeneity are implemented.

[46]Recall that this variable is calculated from the March Current Population Survey and is conditional on whether an individual attended college for more than two years.

TABLE IV

LOG EARNINGS ESTIMATES [a]

	One Type		Two Type	
	Coefficient	Standard Error	Coefficient	Standard Error
Log state earnings	0.4313	0.0077	0.3015	0.0171
Black	−0.0588	0.0026	−0.0644	0.0058
Black × College	0.0852	0.0081	0.1458	0.0176
SAT math interactions (000's)				
Natural science	0.5414	0.0427	0.5066	0.0994
Business	0.6656	0.0535	0.6606	0.1255
Soc/Hum	0.2570	0.0298	0.2794	0.0609
Education	0.2942	0.0591	0.3608	0.1315
No college	0.3361	0.0086	0.3808	0.0188
Math school quality interactions (000's)				
Natural science	0.5848	0.0723	0.1022	0.1600
Business	0.2153	0.0836	0.1672	0.2028
Soc/Hum	0.4271	0.0577	0.2907	0.1283
Education	0.0000	—	0.0000	—
Female interactions				
Natural science	−0.2873	0.0150	−0.2787	0.0277
Business	−0.2057	0.0155	−0.1851	0.0281
Soc/Hum	−0.2255	0.0126	−0.2331	0.0190
Education	−0.2147	0.0184	−0.1954	0.0343
No college	−0.3575	0.0077	−0.3382	0.0108
Constant				
Natural science	5.2667	0.0781	6.4005	0.1720
Business	5.3943	0.0802	6.3625	0.1785
Soc/Hum	5.3838	0.0745	6.3234	0.1644
Education	5.5092	0.0749	6.3452	0.1662
No college	5.5670	0.0673	6.4834	0.1487
Type 2 interactions				
Natural science		0.5362	0.0157	
Business		0.4557	0.0159	
Soc/Hum		0.4694	0.0106	
Education		0.3885	0.0214	
No college		0.4564	0.0029	
Variance	0.1421		0.0917	

[a] Year effects and sex × year effects are also included. All year and sex × year effects are interacted with college. The base year is 1986. In this stage, 31,616 observations are used from 7,859 individuals.

erogeneity are implemented. Controlling for unobserved heterogeneity substantially reduces the return to math college quality for all majors.

The black coefficient is negative and significant: conditional on the controls, blacks without a college degree earn less than whites without a college degree. This is not true for blacks who obtain a college degree because the coefficient on black interacted with college is positive and larger than the black intercept.

Affirmative action in the workplace may account for the result. With employers valuing diversity and the number of blacks graduating from college being small, this leads to higher earnings for blacks conditional on having a degree. It also suggests that blacks may be liquidity constrained and therefore unable to take advantage of the higher premiums from attending college.

The coefficient on black interacted with college almost doubles when controls for unobserved heterogeneity are implemented. All else equal, college-educated blacks earn over 7% more than their white counterparts. Affirmative action in college may contribute to the coefficient increasing once we account for selection. By introducing advantages for blacks in the admissions and aid processes, schools may attract black students whose academic backgrounds, both observed and unobserved, are weaker at the margin.

In addition to advantages in admissions and aid, type 2's have higher earnings in all majors as well as in the no-college sector. This is true particularly for natural science majors. The two-type model improves the fit of the earnings model substantially, reducing the variance on the residual component of earnings by a third.

With the returns to abilities and type varying across majors, heterogeneous treatment effects exist. Table V calculates the earnings premiums for males by race and major. To see the difference in treatment effects, these premiums are

TABLE V

MALE EARNINGS PREMIUMS BY RACE AND MAJOR[a]

	One Type				Two Type[b]			
	Natural Science	Business	Soc Sci/ Humanities	Education	Natural Science	Business	Soc Sci/ Humanities	Education
Whites								
Did not apply	21.8%	19.6%	13.1%	4.0%	14.0%	16.6%	4.2%	−9.5%
	(1.1%)	(1.2%)	(1.0%)	(1.7%)	(2.4%)	(1.4%)	(1.4%)	(3.3%)
Apply, did not attend	23.4%	22.0%	12.7%	3.8%	14.4%	18.6%	3.5%	−9.1%
	(1.0%)	(1.2%)	(1.0%)	(1.7%)	(1.6%)	(1.2%)	(1.3%)	(2.9%)
Attend college	24.8%	24.3%	12.3%	3.6%	17.0%	20.5%	3.2%	−10.6%
	(1.0%)	(1.2%)	(0.9%)	(1.9%)	(1.6%)	(1.2%)	(1.3%)	(2.9%)
Blacks								
Did not apply	25.9%	24.9%	20.9%	14.1%	26.9%	28.3%	18.7%	6.6%
	(1.4%)	(1.6%)	(1.4%)	(1.9%)	(3.3%)	(2.1%)	(1.9%)	(3.8%)
Apply, did not attend	26.9%	26.3%	20.8%	14.1%	27.1%	29.6%	18.3%	7.1%
	(1.3%)	(1.5%)	(1.4%)	(1.9%)	(3.1%)	(2.0%)	(1.9%)	(3.7%)
Attend college	27.5%	27.4%	20.5%	14.0%	29.1%	30.4%	18.2%	5.7%
	(1.3%)	(1.4%)	(1.3%)	(1.9%)	(2.7%)	(1.9%)	(1.8%)	(3.6%)

[a]Premiums are relative to no college and are calculated 14 years after high school graduation using the average characteristics by race conditional on not applying, applying but not attending, and attending. Calculations are made using the race-specific mean of college quality for those who attended.

[b]For the two-type model, the average probability of being a particular type conditional on the cell (e.g., white male, did not apply) is used.

calculated using the average characteristics by race for three groups: those who did not apply to college, those who applied but did not attend, and those who attended college.

Regardless of the groupings, premiums are always highest for natural science and business, and lowest for education. The differences in premiums for the three groups are small for both blacks and whites. Adding unobserved heterogeneity leads to lower premiums in all majors for whites, with the drops particularly large in the social sciences and in education. The decrease in the premiums for blacks is not as large, again suggesting that, perhaps due to affirmative action, college works as a better sorting device for whites than for blacks. Since premiums still differ dramatically across majors, other factors, such as compensating differentials or the effort required in the major, must be leading individuals to choose majors with lower premiums.

To compare these premiums to those found in the literature, we can multiply the major-specific premia by the probabilities of choosing particular fields to obtain an average college premium. For white males, the average college premium for those who attended college is around 10% in the two-type model. This number is substantially smaller than those found in the previous literature. One of the reasons for the lower premiums is how returns to experience are reported. Typically estimates of this type are presented with experience held constant. However, the premiums calculated here are at a particular point in time: 14 years after graduation. Hence, there is a trade-off between the college premium and the lost years of experience.

Even accounting for this experience effect, assuming individuals work 2,000 hours a year in all years after completing their education, the corresponding college premium from Keane and Wolpin (2001), for example, is 15.5%—still much larger than the estimates reported here. The difference between the two estimates can be explained by the rough cut used here on years of schooling that was made necessary because of the detailed modeling of the college application and the choice of college processes. Here, I take whether an individual is enrolled in a four-year college three years after graduating from high school. For those who chose a college option, the average years of postsecondary schooling is four because those who drop out cancel out with those who obtain an advanced degree.[47] However, those who chose the no-college option may have dropped out of a four-year college or enrolled in a community college. The average years of schooling for whites who chose the no-college option was 1.3 years of schooling and correspondingly 12.7 years of work experience. Hence, the estimated premium here is actually for an additional 2.7 years of schooling. The corresponding premium calculated from Keane and Wolpin is virtually identical to the estimate reported here.

[47]The average years of schooling was remarkably similar across fields as well. The highest average years of schooling was in the natural sciences at 4.10 and the lowest was in education at 3.84.

5.4. *College and Major Choice*

I now use the estimates of the earnings regression in the calculation of the parameters of the utility for attending a particular college in a particular major. These estimates are reported in Table VI.

The first set of rows shows the coefficients that are common across majors. Both with and without unobserved heterogeneity, the monetary cost of attending college is significantly negative and more negative for those who come from

TABLE VI

UTILITY ESTIMATES[a]

	One Type		Two Type	
	Coefficient	Standard Error	Coefficient	Standard Error
Black × College	0.2457	0.0625	−0.2034	0.0613
Net cost	−1.5127	0.1735	−1.4399	0.1742
Coefficients, common across majors				
Low income × Net cost	−1.5561	0.2352	−1.4434	0.2301
Private school	0.2701	0.0253	0.2473	0.0253
School in State	0.0976	0.0215	0.1016	0.0219
(SAT math quality)^2	−8.9823	1.1471	−8.9094	1.1285
Expected log earnings	2.3429	0.4790	4.4129	0.7018
SAT math interactions (000's)				
Natural science	7.6766	0.6984	7.7854	0.6893
Business	3.0954	0.4935	2.6853	0.5009
Soc/Hum	3.4154	0.3322	3.7370	0.3390
Education	1.6333	0.5042	1.6775	0.5072
Math school quality interactions (000's)				
Natural science	5.1813	0.6758	5.8723	0.6856
Business	2.6140	0.7815	2.2538	0.7852
Soc/Hum	3.9808	0.5026	3.4662	0.5016
Education	0.9296	0.7504	0.4926	0.7400
Type 1 interactions				
Natural science	−8.6664	0.6086	−9.9144	0.6380
Business	−4.5538	0.4438	−5.3511	0.4937
Soc/Hum	−4.9208	0.3113	−6.2336	0.3606
Education	−3.1011	0.3632	−4.5652	0.4520
Type 2 interactions				
Natural science			−9.0711	0.6012
Business			−4.8048	0.4474
Soc/Hum			−5.3187	0.3140
Education			−3.6386	0.3750
Nesting parameters				
ρ_{c1} (school)	0.5040	0.1280	0.5105	0.1293
ρ_{c2} (major)	0.6676	0.0837	0.6274	0.0856

[a] Also includes sex indicator variables interacted with major choice. In this stage, 3,670 observations are used.

low income families. Private schools and schools in the same state both make choosing a schooling option more attractive, all else equal.

Important to the specification is allowing for the relative math ability squared to affect the decision as to which college to attend and whether to attend college at all. The negative coefficient on this variable implies that, even if the very best schools were free and allowed everyone to attend, some students would find it optimal to attend schools that were better suited to their own abilities.[48]

Unobserved heterogeneity does, however, substantially affect both the black coefficient and the coefficient on the log of expected present value of lifetime earnings. Without controls for unobserved heterogeneity, blacks have a preference for attending college. However, with controls for unobserved heterogeneity, the sign reverses: blacks actually prefer not to attend college, all else equal. This is driven by blacks having higher monetary returns for attending college when controls for unobserved heterogeneity are included. These increased returns are also magnified by the increase in the effect of future earnings on the probability of choosing a particular schooling combination. This coefficient almost doubles in the model with unobserved heterogeneity. Blacks are more likely to choose college conditional on their characteristics, but do so because of the high college premium, not due to preferences.

The next two sets of rows show how the individual's SAT math score and the average SAT math score of the school affect the choice of school and major. Higher SAT math scores make college in general and the natural sciences, in particular, attractive. The same pattern is also observed for the effect of the average SAT math score for the school.

Differences in preferences for particular school–major combinations also exist across types. In addition to enjoying a comparative advantage in earnings from going to college, type 2's also have a preference for college relative to type 1's. This preference differential is similar across fields, with the exception of business, where the effect is half that of the other majors.

The last two rows show the estimates of the nesting parameters. Both estimates of the nesting parameters are significantly less than 1 whether or not controls for unobserved heterogeneity are implemented, suggesting that unobservable preferences have a component that is correlated across school and another component that is correlated across major.

Recall that the treatment effect of attending college on earnings depended on the choice of major, with higher premiums occuring in the natural science and business. Crucial to evaluating the treatment effect of attending college for those who did not attend is knowing what majors these individuals would have chosen. The probability of a particular individual choosing major k conditional

[48]See Arcidiacono (2004) for a similar result.

on attending college is given by

$$(16)\qquad P(k|\text{attend}) = \frac{\sum_{r=1}^{92} P(J_r)\big(\sum_{a=1}^{2^{\#J_r}-1} P(J_a|J_r)\big(\sum_{j=1}^{\#J_a} P(j,k|J_a)\big)\big)}{\sum_{r=1}^{92} P(J_r)\big(\sum_{a=1}^{2^{\#J_r}-1} P(J_a|J_r)\big(\sum_{j=1}^{\#J_a}\sum_{k=1}^{K} P(j,k|J_a)\big)\big)}.$$

The denominator gives the probability of attending college. To make this calculation, it is necessary to calculate the probabilities of applying to each application set,[49] $P(J_r)$, the probabilities of each acceptance set given the application set, $P(J_a|J_r)$, and the probabilities of each possible schooling bundle (a school j and a major k) conditional on the acceptance set, $P(j,k|J_a)$. The numerator is identical except that we now do not sum over all the majors, but rather look only at the probability of choosing a particular major. Averaging over the individuals in a particular group, for example, white males who chose not to attend college, gives the average probability of a group member choosing a particular major conditional on attending college. These average probabilities are reported for different groups in Table VII.[50]

The first set of rows shows the probabilities of choosing each of the majors conditional on attending college for whites. Those who did not apply to college would be most likely to choose a major with a lower premium had they attended, while those who attended were also most likely to choose one of the more lucrative majors conditional on attending. Regardless of controls for unobserved heterogeneity, moving down the rows shows decreases in the probabilities of choosing majors exclusive of the natural sciences and correspondingly large increases in choosing one of the natural sciences.

Although blacks are predicted to be much less likely to choose a natural science major, the trends are similar to those of whites. Moving down the rows again shows increases in the expected probabilities of choosing majors in natural science and decreases in the probabilities of choosing majors in business. Similar decreases are found for social science majors and education majors when there are no controls for unobserved heterogeneity. However, once unobserved heterogeneity is added, no clear patterns emerge for these two majors.

5.5. *Applications*

Using estimates from all the previous stages, I now estimate the parameters of the utility function for applying to college. Table VIII presents these estimates. The coefficient on the present value of future utility is both strongly

[49]The 92 application sets result from calculating all of the possible combinations of up to three schools in a set given that there are eight schools to choose from. Recall also that if an individual applies to $\#J_r$ colleges, then there are $2^{\#J_r}$ possible acceptance sets. However, one of these is the acceptance set that includes being rejected by all schools.

[50]Note that calculating these probabilities relies on estimates from the next section to form the probabilities of applying to various combinations of schools.

TABLE VII

EXPECTED MAJOR CHOICES CONDITIONAL ON ATTENDING COLLEGE[a]

	One Type				Two Type			
	Natural Science	Business	Soc Sci/ Humanities	Education	Natural Science	Business	Soc Sci/ Humanities	Education
Whites								
Did not apply	14.2%	28.4%	44.3%	13.2%	14.2%	30.3%	43.3%	12.2%
	(1.1%)	(1.9%)	(2.2%)	(1.6%)	(0.8%)	(2.2%)	(2.5%)	(2.1%)
Applied, did not attend	26.8%	25.7%	38.7%	8.8%	26.8%	25.5%	38.8%	8.9%
	(1.1%)	(1.1%)	(1.6%)	(1.1%)	(1.2%)	(1.3%)	(1.5%)	(1.2%)
Attend college	31.2%	24.6%	36.6%	7.6%	31.2%	22.8%	37.6%	8.4%
	(1.2%)	(1.0%)	(1.5%)	1.0%)	(1.5%)	(1.3%)	(1.5%)	(1.0%)
Blacks								
Did not apply	7.2%	28.4%	46.0%	18.5%	7.3%	31.5%	44.8%	16.4%
	(1.0%)	(2.9%)	(3.2%)	(2.3%)	(0.7%)	(3.5%)	(4.1%)	(3.9%)
Applied, did not attend	11.5%	28.1%	44.3%	16.1%	11.5%	29.1%	43.8%	15.6%
	(0.9%)	(2.3%)	(2.6%)	(2.0%)	(0.7%)	(2.7%)	(2.9%)	(2.9%)
Attend college	13.0%	28.1%	44.0%	15.0%	12.9%	25.9%	44.7%	16.5%
	(1.0%)	(2.1%)	(2.4%)	(1.8%)	(0.7%)	(2.4%)	(2.3%)	(2.4%)

[a]Males only. Estimates are calculated by simulating the probabilities of choosing particular educational paths divided by the probability of attending college. See the text for details.

positive and significant. Since we believe that individual's discount rates are less than 1 and the coefficient is much greater than 1, this suggests that the

TABLE VIII

APPLICATION ESTIMATES[a]

	One Type		Two Type	
	Coefficient	Standard Error	Coefficient	Standard Error
PV of future utility	4.1636	0.2316	4.2749	0.2380
Application ≥ 1	−4.7757	0.1387	−4.3408	0.1275
Application ≥ 2	−3.1387	0.1827	−3.3736	0.2484
Application $= 3$	−1.5650	0.2299	−1.9501	0.3133
Low income $\times$ (Application ≥ 1)	0.0574	0.0890	0.0232	0.0883
Low income $\times$ (Application ≥ 2)	0.0852	0.0800	0.0912	0.0851
Low income $\times$ (Application $= 3$)	−0.0956	0.1076	−0.1494	0.1258
Type 2 $\times$ (Application ≥ 1)			−1.1783	0.1508
ρ_s (nesting parameter)	0.6671	0.0744	0.8283	0.1068
Prob. type 1 | Low income			0.6204	0.0099
Prob. type 1 | High income			0.5288	0.0101
Log likelihood for full model	−44,978		−37,722	

[a]In this stage, 8,914 observations are used. Each individual has 92 application sets from which to choose.

variance of the unobservable preferences at the application stage is smaller than at the college and major choice stage. Increasing the number of applications submitted is costly with a falling marginal cost, but these application costs are no higher for those who come from low income families.

Adding unobserved heterogeneity shows that type 2's have higher application costs than their type 1 counterparts. Recall that type 2's had higher admissions probabilities, were more likely to obtain lucrative financial aid packages, and also had a comparative advantage in the college sectors. Individuals then apply to college either because their expected benefits are high (type 2's) or because the costs of applying are low (type 1's).[51]

Also shown in Table VIII are the log likelihoods for the two models as well as the population probabilities of being a particular type conditional on high and low income. Those who come from high income families are more likely to be type 2. Adding the mixture distribution to control for unobserved heterogeneity substantially increases the likelihood.

6. MODEL FIT

Given the parameter estimates, it is possible to see how the model matches the key features of the data from Table I. Table IX displays the actual data and the predictions of the model both for whites and blacks and with and without unobserved heterogeneity. Conditional on race, the predictions are calculated for three groups: the full sample, those who applied to college, and those who chose a college option.[52] Throughout, the general trends in the data are matched quite well: the model with unobserved heterogeneity matches particularly well for blacks.

[51]Without unobserved heterogeneity in application costs, the model forces too much selection into college, yielding college premia that are unrealistic and inconsistent with the previous literature. Furthermore, adding unobserved heterogeneity here substantially improves the log likelihood of the model.

[52]Each of these involve calculating the probability of applying to each of the 92 possible schooling combinations and often involve calculating the admittance, financial aid, and choice of college and major probabilities as well. For example, order the application sets such that the first eight involve applying to one school, options nine through twenty-eight involve applying to two schools, and options greater than twenty-eight involve applying to three schools. The average number of schools applied to by the full sample, the sample of those who applied to college, and the sample of those who attended college is given by

$$E(\text{applications}) = \frac{\sum_{n=1}^{N}\sum_{r=1}^{92} P(J_{rn})(1 + (r > 8) + (r > 28))}{N},$$

$$E(\text{applications}|\text{applied}) = \frac{\sum_{n=1}^{N}\sum_{r=1}^{92} P(J_{rn})(1 + (r > 8) + (r > 28))}{\sum_{n=1}^{N}\sum_{r=1}^{92} P(J_{rn})},$$

$$E(\text{applications}|\text{attended}) = \frac{\sum_{n=1}^{N}\sum_{r=1}^{92} P(J_{rn})(1 + (r > 8) + (r > 28))P^*(r, n)}{\sum_{n=1}^{N}\sum_{r=1}^{92} P(J_{rn})P^*(r, n)},$$

TABLE IX

COMPARING MODEL PREDICTIONS OF INDIVIDUAL CHOICES WITH THE DATA[a]

Group Variable	Whites			Blacks		
	Actual	One Type	Two Type	Actual	One Type	Two Type
Full sample						
Prob. of applying	0.4115	0.4194	0.4222	0.4133	0.3532	0.3597
Prob. of attending	0.2114	0.2085	0.2124	0.1667	0.1536	0.1611
Number of applications	0.5924	0.6039	0.6112	0.5809	0.4912	0.5076
Applied to college						
Prob. of attending	0.5137	0.4972	0.5032	0.4033	0.4350	0.4480
Number of applications	1.4397	1.4401	1.4478	1.4006	1.3908	1.4112
Low income	0.3508	0.3593	0.3584	0.6737	0.6551	0.6540
Female	0.4736	0.5045	0.5016	0.6107	0.6366	0.6371
SAT math	500.3	496.1	496.4	360.9	367.3	367.0
Attended college						
Number of applications	1.5772	1.6631	1.5653	1.5491	1.5231	1.5512
Low income	0.3169	0.3107	0.3058	0.6416	0.6643	0.6068
Female	0.4757	0.5001	0.4974	0.6358	0.6643	0.6763
SAT math	529.8	527.7	528.4	378.9	393.1	391.8
Math school quality	538.2	524.2	523.8	466.5	501.3	489.4

[a]See text for details of the calculations.

The first set of rows shows how the model matches the application and attendance decisions of the full sample. For whites, both models match the trends in the data. For blacks, both models underpredict the probability of applying and the number of applications, although the model with unobserved heterogeneity is closer to the actual data.

The second set of rows is for those individuals who applied to college, where the models' predictions are not for the subsample who actually applied to college, but for those who the models predict will apply to college. As with the full sample, the predicted probabilities of attending and the number of applications for whites is very similar to what is observed in the data. Both models overpredict the number of females who choose to apply, although both also

where

$$P^*(r,n) = \left(\sum_{a=1}^{2^{\#J_{rn}}-1} P(J_{an}|J_{rn}) \left(\sum_{1}^{\#J_{an}} P(j|J_{an}) \right) \right),$$

J_{rn} refers to the application set for the nth individual, J_{an} is the subset of J_{rn} where the individual was accepted, and j indexes a school in the set J_{an}. The first part of the numerator in each of these fractions gives the probability of applying to each of the 92 applications sets. The denominator in the second equation is the probability of applying at all. The extra term in both the numerator and the denominator in the third equation takes into account that the person must choose an option that involves attending college to be counted.

closely match the mean black and white SAT math scores for those who applied. For blacks, the probability of attending conditional on applying is too high. This must be the case, because the model with unobserved heterogeneity underpredicted the probability of applying and the number of applications submitted, yet matched the mean probability of attending for the full sample.

The final set of rows compares the data on those who attended college with those who were predicted to attend. Both models underpredict school quality for whites and overpredict for blacks. This overprediction for blacks is stronger for the model without unobserved heterogeneity. That the school qualities are overpredicted for blacks means that we may be overstating the losses due to removing affirmative action policies because advantages are much greater for blacks at higher quality schools.

Table X examines how well the two models predict the admissions and financial aid decisions of the schools. For admissions, I examine the overall admit rate by race as well as looking at those individuals with SAT scores above the mean. For financial aid, Table X focuses on the share of total cost paid by the school conditional on being admitted, with separate results by race and by whether the individual came from a low income family.[53] The two models yield very similar predictions for financial aid and admissions, and generally match

TABLE X

COMPARING MODEL PREDICTIONS OF ADMISSIONS AND FINANCIAL AID WITH THE DATA[a]

	Data	One Type	Two Type
Whites			
Prob. of admit	0.8978	0.8955	0.8961
Prob. of admit | SAT > mean	0.9110	0.9177	0.9199
Share of costs[b]	0.1036	0.1063	0.1078
Share of costs | Low income	0.1653	0.1830	0.1850
Blacks			
Prob. of admit	0.8640	0.8612	0.8622
Prob. of admit | SAT > mean	0.8830	0.9045	0.9085
Share of costs	0.1906	0.1977	0.2007
Share of costs | Low income	0.1948	0.2128	0.2156

[a] See text for details on the calculations.
[b] Share of costs covered is given by scholarships/total cost.

[53] The probability of being admitted and the share of costs of paid are calculated using:

$$P(\text{admit}|\text{applied}) = \frac{\sum_{n=1}^{N}\sum_{r=1}^{92} P(J_{rn})P^{*}(a,r,n)}{\sum_{n=1}^{N}\sum_{r=1}^{92} P(J_{rn})(1+(\#J_{rn}>1)+(\#J_{rn}>2))},$$

$$E(\text{cost share}|\text{accepted}) = \frac{\sum_{n=1}^{N}\sum_{r=1}^{92} P(J_{rn})\big(\sum_{a=1}^{2^{\#J_{rn}}-1} P(J_{an}|J_{rn})\big(\sum_{j=1}^{\#J_{an}} s_{jn}\big)\big)}{\sum_{n=1}^{N}\sum_{r=1}^{92} P(J_{rn})P^{*}(a,r,n)},$$

the trends in the data. For both blacks and whites, however, financial aid is overestimated, most likely as a result of overestimating college quality.

Recall that significantly lower premiums were found for attending college once controls for unobserved heterogeneity were implemented. In particular, returns to education were estimated to be significantly negative. Table XI compares earnings data with data predicted by each of the models for males. The predictions again are based on forecasting each stage of the model and associating particular earnings streams with the various educational paths.[54]

TABLE XI

COMPARING MODEL PREDICTIONS OF EARNINGS WITH THE DATA[a]

Earings	Data	One Type	Two Type
1986			
Natural sciences	\$57,552	\$61,028	\$60,495
Business	\$56,409	\$57,387	\$56,564
Soc sci/Humanities	\$51,879	\$51,077	\$50,840
Education	\$37,118	\$46,612	\$44,514
No college	\$42,009	\$43,358	\$42,345
1979			
Natural sciences	\$39,317	\$41,839	\$41,532
Business	\$36,309	\$39,343	\$38,883
Soc sci/Humanities	\$33,238	\$35,017	\$34,904
Education	\$29,356	\$31,956	\$30,560
No college	\$34,370	\$36,452	\$35,790

[a]See text for details of the calculations. Earnings are annual and in 1999 dollars.

where

$$P^*(a,r,n) = \left(\sum_{a=1}^{2^{\#Jrn}-1} P(J_{an}|J_{rn})\big(1 + (\#J_{an} > 1) + (\#J_{an} > 2)\big) \right).$$

For the acceptance probabilities, I take the total number of acceptances in the population divided by the total number of applications submitted. For the share of costs paid conditional on being accepted, I sum the shares in the population and divide by the total number of acceptances. The -1 in the summations takes the sum only over acceptance sets where the individual was accepted to at least one school.

[54]Let W_{jkt} be the expected earnings in year t conditional on choosing school j and major k, where the expectation is with respect to the transitory earnings shocks. The predicted earnings for major k are calculated using

$$E(W_t|k) = \frac{\sum_{n=1}^{N} \sum_{r=1}^{92} P(J_{rn}) \big(\sum_{a=1}^{2^{\#Jrn}-1} P(J_{an}|J_{rn}) \big(\sum_{j=1}^{\#J_{an}} P(j,k|J_{an}) W_{jkt}\big)\big)}{\sum_{n=1}^{N} \sum_{r=1}^{92} P(J_{rn}) \big(\sum_{a=1}^{2^{\#Jrn}-1} P(J_{an}|J_{rn}) \big(\sum_{1}^{\#J_{an}} P(j,k|J_{an})\big)\big)}.$$

Here, the sums in the numerator are taken over the probabilities of applying and attending all the various schools and choosing major k multiplied by the associated expected earnings with that particular school in major k. Dividing through by the probability of choosing major k gives the

The first set of rows shows the data and predictions for 1986 earnings, while the second set shows data and predictions for 1979. The data show that even ten years after college, education majors lag significantly behind their noncollege counterparts. Coupled with positive selection into college, it is not surprising that the estimated premia for choosing an education major is negative for males. Even social science/humanities majors have lower earnings than those who do not attend college three years into their labor market years, although they rapidly catch up and pass those who do not choose a college option. Both models perform well in predicting the overall trends in 1979 and in 1986. The one exception is that both models overpredict earnings for male education majors, although this effect is somewhat mitigated once controls for unobserved heterogeneity are implemented.

7. POLICY SIMULATIONS

With the model predicting the trends in the data reasonably well, I now proceed with the policy simulations. In particular, I use the estimates of the earnings process, financial aid and admission rules, and the parameters of the utility function to simulate how changes in the financial aid and admission rules affect college decision-making and, in turn, future earnings. I primarily focus on earnings ten years after college, although the Appendix provides estimates of the effects of affirmative action on the present value of both lifetime earnings and utility. I perform three policy simulations all related to the removal of affirmative action: giving blacks the same financial aid rules as whites, giving blacks the same admissions rules as whites, and giving blacks both the same admissions rules and the same financial aid rules. All policy simulations are under a partial equilibrium setting. Hence, they should be interpreted as what would happen if we changed the rules for a random person as opposed to changing the rules for the population or a large portion of the population.[55] However, as we will see shortly, affirmative action affects such a small percentage of the population that any general equilibrium effects are expected to be very small.

Sampling from the GEV distributions associated with the errors in the application, and college and major choices is not computationally feasible.[56] However, the advantages of the GEV distribution are the closed form expressions

---

result. For the case with unobserved heterogeneity, the calculation is then

$$E(W_t|k) = \frac{\sum_{n=1}^{N}\sum_{l=1}^{L} P(l_n)\sum_{r=1}^{92} P(J_{rn})\left(\sum_{a=1}^{2^{\#J_{rn}}-1} P(J_{an}|J_{rn})\left(\sum_{j=1}^{\#J_{an}} P(j,k|J_{an})W_{jkt}\right)\right)}{\sum_{n=1}^{N}\sum_{l=1}^{L} P(l_n)\sum_{r=1}^{92} P(J_{rn})\left(\sum_{a=1}^{2^{\#J_{rn}}-1} P(J_{an}|J_{rn})\left(\sum_{1}^{\#J_{an}} P(j,k|J_{an})\right)\right)},$$

where $P(l_n)$ is the conditional probability that the nth individual is type l and all other probabilities are now type specific.

[55] See Heckman, Lochner, and Taber (1998) for an analysis of the general equilibrium effects of a tuition subsidy program.

[56] Cardell (1997) shows how to take advantage of the independence assumptions to draw from a nested logit specification. No similar results are available for the BST specification. Metropolis–

for the probabilities of choosing particular actions. The simulations are conducted by calculating the probabilities of choosing particular educational paths rather than sampling over the ϵ's and then finding the optimal educational path for those specific ϵ's.[57] Calculating the probabilities will produce the same simulation results as if one took an infinite number of draws for each individual on the ϵ's. That is, what the probabilities give us is the fraction of individuals who would make particular educational choices. The disadvantage of this method is that it is not possible to calculate ex post utility losses; I cannot say how much worse off blacks who chose a college option under affirmative action would be if affirmative action was removed.

Table XII gives the ex ante losses in expected earnings fourteen years after high school graduation for black males from switching to the white admissions and/or financial aid rules at various quantiles. Ex ante expected earnings under a particular set of admissions and financial aid rules for a particular individual—and in the case of unobserved heterogeneity a particular individual of a particular type—are given by

$$E(W_t) = \sum_{r=1}^{92} P(J_r)\left(\sum_{a=1}^{2^{\#J_r}-1} P(J_a|J_r)\left(\sum_{j=1}^{\#J_a}\sum_{k=1}^{K} P(j,k|J_a)E(W_{jkt})\right)\right) + (1-P(C))E(W_{ot}),$$

where the expectations on the right-hand side are taken over the transitory portion of earnings. The first part of the equation gives the probabilities of applying to particular sets of schools,[58] $P(J_r)$, the probabilities of the set of accepted schools conditional on the application set, $P(J_a|J_r)$, and the probabilities of choosing particular school and major combinations conditional on the acceptance set, $P(j,k|J_a)$, times the associated expected earnings for each

---

Hastings is also not computationally possible because the number of terms in the joint p.d.f. increases exponentially with the number of alternatives. To see this, consider the cumulative distribution function (c.d.f.) of a nested logit with N alternatives in the nest. The c.d.f. in this case is

$$F(\epsilon_0,\dots,\epsilon_N) = \exp\left(\left(\sum_{n=1}^{N}\exp\left(\frac{\epsilon_n}{\rho}\right)\right)^{\rho} + \epsilon_0\right).$$

To find the p.d.f., we need to differentiate with respect to each of the arguments. The numbers of terms will be 2^N. The situation is worse when the BST framework is used, because now alternatives can belong to multiple nests.

[57] In the case of unobserved heterogeneity, I calculate the conditional probability of each individual being a particular type. The simulations are then conducted assuming that these conditional probabilities give the type distribution for individuals with those particular characteristics.

[58] Recall that an individual may apply to one of 92 sets of schools, where there is a maximum of three schools per set. Conditional on applying to the set J_r, there are $2^{\#J_r}$ possible acceptance sets, one of which involves being rejected at all schools.

TABLE XII

EX ANTE EXPECTED EARNINGS LOSSES FOURTEEN YEARS AFTER HIGH SCHOOL FOR BLACK MALES FROM SWITCHING TO WHITE ADMISSION AND FINANCIAL AID RULES[a]

			Adjustment in Application Decision			No Adjustment in Application Decision		
		Admission Rules:	Black	White	White	Black	White	White
	Quantile	Aid Rules:	White	Black	White	White	Black	White
One type:	25th		\$23	\$1	\$28	\$11	\$1	\$15
			(10)	(20)	(20)	(5)	(10)	(10)
	50th		\$60	\$9	\$70	\$24	\$6	\$29
			(26)	(34)	(35)	(10)	(16)	(15)
	75th		\$126	\$27	\$146	\$46	\$16	\$59
			(55)	(60)	(68)	(20)	(27)	(27)
	90th		\$330	\$86	\$410	\$101	\$46	\$145
			(140)	(117)	(143)	(43)	(53)	(57)
	95th		\$507	\$195	\$606	\$161	\$107	\$213
			(217)	(183)	(203)	(72)	(90)	(89)
	99th		\$827	\$506	\$1,320	\$281	\$330	\$610
			(362)	(317)	(393)	(120)	(184)	(170)
Two type:	25th		\$19	\$1	\$22	\$9	\$0	\$12
			(10)	(11)	(15)	(5)	(6)	(8)
	50th		\$44	\$8	\$52	\$18	\$5	\$23
			(22)	(20)	(27)	(9)	(10)	(13)
	75th		\$127	\$21	\$142	\$40	\$12	\$48
			(53)	(37)	(57)	(19)	(17)	(23)
	90th		\$324	\$67	\$360	\$93	\$31	\$124
			(116)	(70)	(111)	(41)	(32)	(42)
	95th		\$449	\$155	\$580	\$133	\$80	\$199
			(146)	(104)	(130)	(55)	(53)	(57)
	99th		\$747	\$373	\$1,157	\$249	\$276	\$526
			(259)	(231)	(243)	(100)	(156)	(137)

[a]Calculations are made before making any college decisions and are relative to expected earnings given current affirmative action policies. Earnings are in 1999 dollars and are annual. Standard errors are given in parentheses.

educational outcome, $E(W_{jkt})$. The last term gives the probability of choosing the no-college option times the expected earnings from doing so. Then $P(C)$ is the probability of choosing college and is given by

$$P(C)=\sum_{r=1}^{92}P(J_r)\left(\sum_{a=1}^{2^{\#J_r}-1}P(J_a|J_r)\left(\sum_{j=1}^{\#J_a}\sum_{k=1}^{K}P(j,k|J_a)\right)\right).$$

Changing the admissions rules then affects the probability of applying to a particular set of schools, $P(J_r)$, and the probabilities of particular acceptance sets conditional on the application decision, $P(J_a|J_r)$. Changing the financial aid rules also affects the probability of applying to a particular set of schools as well as affecting the probability of a school in a particular major, $P(j,k|J_a)$.

The first set of rows does not control for unobserved heterogeneity, whereas the second set of rows does. To assess the effect the application decision has on the gains and losses associated with affirmative action, the first set of columns allows the individuals to change their application decisions based on the change in admissions or aid rules. The second set of columns, however, restricts the application decision to be the same as before the policy change. Adding the application decisions leads to losses that are roughly one and a half to three times as large as the losses when individuals cannot adjust their application decision. Hence, rather than applying to more high quality colleges when affirmative action is removed, and thereby undoing some of the negative effects of the policy change on earnings, the effect of the policy is reenforced by individuals who choose either to apply to lower quality colleges or not to apply at all. This occurs because individuals are not maximizing earnings, but rather utility where earnings is one component. In the Appendix I report losses in expected utility from policy changes and, after converting utils to earnings, find similarly small effects.

Adding controls for unobserved heterogeneity generally reduced the expected losses by around 10% with larger drops associated with switching to the white admissions rules at the higher quantiles. However, regardless of controls for selection or adjustments in the application decision, the expected losses in earnings are quite small. At the 90th percentile of losses, for example, removing both affirmative action in financial aid and in admissions reduces yearly earnings ten years after college by $410 without unobserved heterogeneity and $360 with unobserved heterogeneity. To put these numbers in perspective, the gap between black male expected earnings for the 50th percentile versus the 60th percentile is $1,320. It is only at the 99th percentile that significant earnings changes occur; the largest losses are at $1,157 without unobserved heterogeneity and $1,163 with unobserved heterogeneity.

Similar small effects of financial aid have been found by Keane and Wolpin (1997, 2000, 2001). Their studies found very little effect from tuition subsidy programs. These results are confirmed here using an entirely different methodology—allowing individuals to attend different types of schools and choosing different majors at the expense of modeling decisions far out into the life cycle. What is noticeable about the results in Table XII are the very small effects of changes in admissions rules on future earnings: the effect of financial aid is always larger except sometimes at the 99th percentile.

It should be noted that the results in Table XII are expectations. How these decisions play out will lead some students to have made the exact same decisions as they made without affirmative action. The students who are at the margin and therefore do change their decisions based on the rules may have much higher (or lower) future earnings than those predicted here. Table XIII looks at the expected distribution of earnings for black males after individuals have already made their college decisions. This is accomplished by calculating expected earnings for each individual from choosing all the different major

TABLE XIII

EX POST EXPECTED EARNINGS LOSSES FOURTEEN YEARS AFTER HIGH SCHOOL FOR BLACK MALES FROM SWITCHING TO WHITE ADMISSION AND FINANCIAL AID RULES[a]

			Base	Adjustment in Application Decision			No Adjustment in Application Decision		
		Admission Rules:	Black	Black	White	White	Black	White	White
	Quantile	Aid Rules:	Black	White	Black	White	White	Black	White
One type:	25th		$35,227	$35,227	$35,227	$35,227	$35,227	$35,227	$35,227
			(223)	(224)	(223)	(225)	(222)	(223)	(222)
	50th		$36,440	$36,409	$36,440	$36,409	$36,440	$36,440	$36,409
			(235)	(234)	(239)	(236)	(238)	(235)	(234)
	75th		$38,352	$38,289	$38,333	$38,264	$38,333	$38,352	$38,330
			(263)	(272)	(264)	(266)	(267)	(264)	(268)
	90th		$43,359	$41,917	$43,020	$41,546	$42,856	$43,146	$42,614
			(1,278)	(1,509)	(1,225)	(1,281)	(1,383)	(1,251)	(1,327)
	95th		$48,480	$47,889	$48,235	$47,644	$48,323	$48,339	$48,122
			(825)	(948)	(755)	(852)	(854)	(789)	(812)
	99th		$56,333	$56,091	$56,003	$55,459	$56,311	$56,214	$55,956
			(706)	(802)	(686)	(798)	(724)	(688)	(708)
Two type:	25th		$28,687	$28,687	$28,687	$28,687	$28,687	$28,687	$28,687
			(206)	(207)	(208)	(210)	(206)	(207)	(208)
	50th		$30,253	$30,182	$30,239	$30,182	$30,195	$30,253	$30,195
			(213)	(214)	(210)	(217)	(213)	(212)	(212)
	75th		$45,010	$45,010	$45,010	$45,010	$45,010	$45,010	$45,010
			(343)	(346)	(342)	(348)	(342)	(342)	(343)
	90th		$47,812	$47,722	$47,804	$47,605	$47,804	$47,804	$47,799
			(362)	(350)	(356)	(336)	(360)	(360)	(362)
	95th		$54,288	$51,531	$53,744	$51,278	$53,588	$54,055	$53,356
			(1,831)	(1,726)	(1,747)	(1,420)	(1,837)	(1,781)	(1,782)
	99th		$69,187	$68,726	$68,980	$68,569	$69,060	$68,980	$68,895
			(387)	(430)	(360)	(350)	(400)	(364)	(378)

[a]Ex post distribution of earnings are after making all college decisions. Earnings are in 1999 dollars and are annual. Standard errors are given in parentheses.

and school combinations. The probabilities associated with these outcomes—and in the case of unobserved heterogeneity, the conditional probabilities of being a particular type— then characterizes the ex post expected earnings distribution.[59] Not surprisingly, the model with unobserved heterogeneity yields a much larger spread on earnings across the quantiles.

One interesting feature of Table XIII is which part of the black distribution of earnings is affected by changing the affirmative action policies. In particular, both the largest changes and the largest standard errors are found at the 90th percentile without unobserved heterogeneity and at the 95th percentile with unobserved heterogeneity. Individuals above these percentiles are relatively unaffected by affirmative action policies because they are likely to attend college with or without affirmative action. Individuals significantly below these percentiles are unlikely to attend college regardless of what policies are in place. Still, the differences in the ex post earnings across even these quantiles are quite small, with the largest difference being at around 3,000 when advantage for blacks in both financial aid and admissions are removed.

Affirmative action may have effects in the labor market even if the channel is not through earnings. For example, attending college may lead to lower unemployment rates. Analyzing the present value of lifetime earnings requires making assumptions about growth rates on earnings, the discount factor, and unemployment rates both with and without a college education. Because of the assumptions entailed, I discuss this analysis in the Appendix. I show there that only under the most extreme assumptions is there any economically significant effect of affirmative action on the present value of lifetime earnings.

If affirmative action in higher education has little effect on the labor market, what does it affect? Table XIV addresses this question by showing how the different admissions and financial aid rules affect the educational decisions of blacks. Without unobserved heterogeneity, removing the black advantage in financial aid results in a 9.5% drop in the number of blacks attending college. This drop is smaller (8.6%) when unobserved heterogeneity is added. That the enrollment effects are larger than the effects on earnings points to the largest decreases in the probabilities of college attendance coming from those individuals whose expected earnings would increase the least from attending college.

Removing the black advantage in admissions has a much smaller effect on the number of blacks attending college, with drops of 2.3% and 1.9% for the models without and with unobserved heterogeneity, respectively. However, removing affirmative action in admissions would have a dramatic effect on the number of blacks attending the very best colleges. In particular, both models predict an over 45% drop in the number of black males attending colleges with

[59]Note that the individual at the 90th percentile with affirmative action may not be at the 90th percentile after affirmative action is removed. A disadvantage of using the GEV distribution is not being able to identify the differences in ex post earnings at the individual level.

average SAT scores greater than or equal to 1,200. While this drop is large, it is counteracted somewhat by the increase in the number of blacks attending lower tiered colleges, where the average SAT score is below 1,100.

8. CONCLUSION

Affirmative action in higher education is a very controversial topic, yet little is known about how these programs affect the earnings of their intended beneficiaries. The reason for this lack is that the path by which earnings are affected is complicated: affirmative action affects admissions and financial aid rules, not earnings directly. Individuals can undo or reenforce the effect of changes in admissions and financial aid rules on earnings through their application behavior. This paper provides a first step to understanding how both admissions and financial aid rules affect expected future earnings.

On the school side, I model the admissions and financial aid decisions. On the student side, I model the choice as to where to submit applications, where to attend, and what major to choose conditional on the acceptance set. I also model the relationship between these choices and earnings. With the estimates of all the parts of the model, I simulate how constraining the admissions and financial rules blacks face to be the same as the rules for whites affects future earnings and college decisions of blacks.

Simulating the effects of removing black advantages in admission and in financial aid showed surprisingly little effect on black male earnings, despite blacks enjoying much larger premiums to attending college than their white counterparts. The small effects on expected earnings from removing black advantages in financial aid occur because those individuals who are at the margin of attending are also the ones who have the lowest treatment effect; their abilities are relatively more rewarded in the noncollege market and they are ones who are most likely to chose majors with low premiums. On the admissions side, black advantages occur only at high quality schools. Removing black advantages in admissions has little effect on earnings, because the return to college quality is small and the blacks affected by the policy are most likely to attend college regardless of whether affirmative action is in place.

There are two extensions of the model that would be interesting to pursue. The first is gains in diversity. That is, if blacks would prefer to attend schools with other blacks, an affirmative action program may have a reenforcing effect, where letting in one black student encourages another black student to attend. This is currently not taken into account in the policy simulations and significantly adds to the complexity of the model. Now, not only do we have to keep track of each individual's education decisions, but also how those decisions aggregate into distributions of minorities at each school.

The second extension deals with the issue of "fit." One criticism of affirmative action in higher education is that it leads minorities into environments where they cannot succeed. The only way that this can be consistent with rational expectations is if individuals receive information in the admissions and

TABLE XIV

BLACK MALE CHOICES UNDER DIFFERENT ADMISSIONS AND AID RULES[a]

	One Type				Two Type			
Admission Rules:	Black	Black	White	White	Black	Black	White	White
Aid Rules:	Black	White	Black	White	Black	White	Black	White
Natural science	1.94%	1.77%	1.86%	1.69%	1.92%	1.77%	1.85%	1.70%
	(0.18%)	(0.18%)	(0.16%)	(0.16%)	(0.14%)	(0.14%)	(0.13%)	(0.12%)
Business	3.31%	3.01%	3.27%	2.97%	3.33%	3.04%	3.28%	3.01%
	(0.46%)	(0.43%)	(0.45%)	(0.41%)	(0.43%)	(0.41%)	(0.41%)	(0.38%)
Soc/Hum	5.16%	4.66%	5.03%	4.55%	5.07%	4.60%	4.94%	4.51%
	(0.57%)	(0.61%)	(0.56%)	(0.57%)	(0.57%)	(0.52%)	(0.53%)	(0.46%)
Education	1.62%	1.48%	1.61%	1.47%	1.58%	1.44%	1.57%	1.43%
	(0.36%)	(0.34%)	(0.34%)	(0.32%)	(0.34%)	(0.29%)	(0.34%)	(0.30%)
College	12.03%	10.91%	11.76%	10.68%	11.89%	10.86%	11.66%	10.66%
	(1.15%)	(1.14%)	(1.12%)	(1.06%)	(1.16%)	(1.06%)	(1.10%)	(0.97%)
School avg. SAT score ≥ 1,100	1.93%	1.62%	1.49%	1.26%	1.88%	1.60%	1.48%	1.27%
	(0.19%)	(0.22%)	(0.17%)	(0.11%)	(0.20%)	(0.20%)	(0.16%)	(0.11%)
School avg. SAT score ≥ 1,200	0.67%	0.54%	0.38%	0.32%	0.66%	0.55%	0.39%	0.34%
	(0.13%)	(0.14%)	(0.06%)	(0.03%)	(0.11%)	(0.10%)	(0.05%)	(0.03%)

[a]Standard errors are given in parentheses.

financial aid decisions of the schools. Individuals who are considering attending top colleges are used to succeeding. They may, however, have incomplete information as to how well their abilities match up with those attending top colleges. Individuals then use information from college admissions and financial aid to update their expectations on their own abilities. Affirmative action programs then provide a trade-off between larger choice sets and less information. While the first extension would most likely lead to increases in the gains of affirmative action, this latter extension would not.

Dept. of Economics, Duke University, 201A Social Sciences Building, Durham, NC 27708-0097, U.S.A.; psarcidi@econ.duke.edu; http://www.econ.duke.edu/ ˜psarcidi.

Manuscript received November, 2002; final revision received January, 2005.

REFERENCES

ALTONJI, J. (1993): "The Demand for and Return to Education when Education Outcomes Are Uncertain," *Journal of Labor Economics*, 11, 48–83.

ARCIDIACONO, P. (2004): "Ability Sorting and the Returns to College Major," *Journal of Econometrics*, 121, 343–375.

ARCIDIACONO, P., AND J. B. JONES (2003): "Finite Mixture Distributions, Sequential Likelihood, and the EM Algorithm," *Econometrica*, 71, 933–946.

BERGER, M. (1988): "Predicted Future Earnings and Choice of College Major," *Industrial and Labor Relations Review*, 41, 418–429.

BOWEN, W., AND D. BOK (1998): *The Shape of the River: Long-Term Consequences of Considering Race in College and University Admissions.* Princeton, NJ: Princeton University Press.

BRESNAHAN, T., S. STERN, AND M. TRAJTENBERG (1997): "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s," *RAND Journal of Economics*, 28, S17–S44.

BREWER, D., E. EIDE, AND D. GOLDHABER (1999): "An Examination of the Role of Student Race and Ethnicity in Higher Education Since 1972," Unpublished Manuscript, Public Policy Institute of California.

BREWER, D., E. EIDE, AND R. EHRENBERG (1999): "Does It Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings," *Journal of Human Resources*, 34, 104–123.

CAMERON, S., AND J. HECKMAN (1998): "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy*, 106, 262–311.

——— (2001): "The Dynamics of Educational Attainment for Black, Hispanic, and White Males," *Journal of Political Economy*, 109, 455–499.

CARD, D. (1999): "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics*, Vol. 3A, ed. by O. Ashenfelter and D. Card. New York: Elsevier, 1801–1864.

——— (2001): "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, 69, 1127–1160.

CARDELL, N. S. (1997): "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity," *Econometric Theory*, 13, 185–213.

DALE, S. B., AND A. B. KRUEGER (2002): "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables," *Quarterly Journal of Economics*, 117, 1491–1527.

DANIEL, K., D. BLACK, AND J. SMITH (1997): "College Quality and the Wages of Young Men," Unpublished Manuscript, University of Western Ontario.

DEMPSTER, A. P., M. LAIRD, AND D. B. RUBIN (1977): "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society*, Ser. B, 39, 1–38.

ECKSTEIN, Z., AND K. I. WOLPIN (1999): "Why Youths Drop Out of High School: The Impact of Preferences Opportunities and Abilities," *Econometrica*, 67, 1295–1339.

FULLER, D., C. MANSKI, AND D. WISE (1982): "New Evidence on the Economic Determinants of Postsecondary Schooling Choices," *Journal of Human Resources*, 17, 477–498.

GROGGER, J., AND E. EIDE (1995): "Changes in College Skills and the Rise in the College Wage Premium," *Journal of Human Resources*, 30, 280–310.

HECKMAN, J., L. LOCHNER, AND C. TABER (1998): "General-Equilibrium Treatment Effects: A Study of Tuition Policy," *American Economic Review*, 88, 381–386.

HECKMAN, J., AND B. SINGER (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52, 271–320.

HECKMAN, J., AND E. VYTLACIL (1998): "Instrumental Variables Methods for the Correlated Random Coefficient Model," *Journal of Human Resources*, 33, 974–987.

HECKMAN, J., J. TOBIAS, AND E. VYTLACIL (2000): "Simple Estimators for Treatment Parameters in a Latent Variable Framework with an Application to Estimating the Returns to Schooling," Working Paper 7950, NBER.

JAMES, E., A. NABEEL, J. CONATY, AND D. TO (1989): "College Quality and Future Earnings: Where Should You Send Your Child to College?" *American Economic Review: Papers and Proceedings*, 79, 247–252.

KANE, T. J. (1998): "Racial and Ethnic Preferences in College Admission," in *The Black–White Test Score Gap*, ed. by C. Jencks and M. Phillips. Washington, DC: Brookings Institution, 431–456.

KANE, J., AND L. SPIZMAN (1994): "Race, Financial Aid, and College Attendance: Parents and Geography Matter," *American Journal of Economics and Sociology*, 53, 85–97.

KEANE, M. P., AND K. I. WOLPIN (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105, 473–522.

——— (2000): "Eliminating Race Differences in School Attainment and Labor Market Success," *Journal of Labor Economics*, 18, 614–652.

——— (2001): "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment," *International Economic Review*, 42, 1051–1103.

LEE, L. F. (1983): "Generalized Econometric Models with Selectivity," *Econometrica*, 51, 507–512.

LIGHT, A., AND W. STRAYER (2002): "From *Bakke to Hopwood*: Does Race Affect College Attendance and Completion?" *Review of Economics and Statistics*, 85, 34–44.

LOURY, L. D., AND D. GARMAN (1995): "College Selectivity and Earnings," *Journal of Labor Economics*, 13, 289–308.

MANSKI, C., AND D. WISE (1983): *College Choice in America*. Cambridge, MA: Harvard University Press.

MCFADDEN, D. L. (1978): "Modelling the Choice of Residential Location," in *Spatial Interaction Theory and Planning Models*, ed. by A. Karlqvist, L. Lundqvist, F. Snikcars, and J. Weibull. New York: North-Holland, 75–96.

ROTHWELL, G., AND J. RUST (1997): "On the Optimal Lifetime of Nuclear Power Plants," *Journal of Business and Economic Statistics*, 15, 195–208.

RUST, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999–1033.

——— (1994): "Structural Estimation of Markov Decision Processes," in *Handbook of Econometrics*, Vol. 4, ed. by R. F. Engle and D. L. McFadden. Amsterdam: North–Holland, 3081–3143.

RUST, J., AND C. PHELAN (1997): "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets," *Econometrica*, 65, 781–831.

VENTI, S., AND D. WISE (1982): "Test Scores, Educational Opportunities, and Individual Choice," *Journal of Public Economics*, 18, 35–63.

WILLIS, R., AND S. ROSEN (1979): "Education and Self-Selection," *Journal of Political Economy*, 87, S7–36.