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

Peter Arcidiacono

Presented by Dan McLeod and Joel McMurry

December 4, 2017

Purpose of Paper

- Estimate structural model of college applications, admissions, attendance, and labor market outcomes
  - Include financial aid and major choice
- Estimate counterfactual black college choices and earnings if *affirmative action in admissions and financial aid were eliminated*

Purpose of Paper

- Estimate structural model of college applications, admissions, attendance, and labor market outcomes
  - Include financial aid and major choice
- Estimate counterfactual black college choices and earnings if *affirmative action in admissions and financial aid were eliminated*

Eliminating affirmative action has...

- small effects on the earnings of black males
- large effects on number of black students at elite schools

$\Longrightarrow$ Interpretation: marginal black college student enjoys low treatment effect from attending college

Model Overview

1. Potential students decide where to apply
2. Admissions and financial aid decisions are made
3. Given offers, potential students choose school and major (or enter labor market without schooling)
4. Students enter labor market

Model: Applications

Student applying to subset J of colleges gets expected utility:

$$v_{sJ} = \alpha_{s1} \sum_{a=1}^{2^{\#J}-1} \mathbb{E}_s(V_c|J_a)P(J_a|J) - \alpha_{s2}X_{sJ} + \epsilon_{sJ}$$

where

- $P(J_a|J)$ is probability of being accepted at subset $J_a \subset J$, and $\mathbb{E}_s(V_c|J_a)$ is expected value of optimal college and major choice
- $\alpha_{s2}X_{sJ}$ is parameterized utility cost of applying to set J
- ϵ_{sJ} is unobserved preference for applying to set J, distributed GEV as explained below

Model: Admissons and Aid Offers

School maximization is not modeled. Instead, assumed that admissions probabilities are logit

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

and financial aid (student pays $s_j t_j$) is tobit

$$s_j^* = \gamma_f X_{fj} + \epsilon_{fj}$$

$$s_j = \begin{cases} 0, & s_j^* \leq 0 \\ s_j^*, & 0 < s_j^* < 1 \\ 1, & s_j^* \geq 1 \end{cases}$$

Model: Attendance Decisions

Student accepted at set J_a who chooses college $j \in J_a$ and major $k \in K$ enjoys utility of attending, u_{cjk}, and PDV of future earnings. Utility of attendance given by

$$
\begin{aligned}
v_{cjk} &= u_{cjk} + u_{wjk} \\
&= \left[\alpha_{c1} X_{cjk} - \underbrace{\left(\alpha_{c2k}(A - \bar{A}_j) + \alpha_{c3}(A - \bar{A}_j)^2 \right)}_{\text{cost of effort}} + \epsilon_{cjk} \right] + u_{wjk}
\end{aligned}
$$

where

- A is ability and $\bar{A}_j$ is average ability at school j
- ϵ_{cjk} is unobserved preference for school j and major k, distributed GEV as explained below

Model: Labor Market

Working agents get log of expected PDV of earnings

$$u_{wjk} = \alpha_w \log\left(\mathbb{E}_w\left[\sum_{t=t'}^{T} \beta^{t-t'} P_{kt} W_{jkt}\right]\right)$$

where

- P_{kt} is probability of working in year t given majored in k
- W_{jkt} is earnings t years out of high school given college j and major k:

$$\log(W_{jkt}) = \gamma_{wk1} + \gamma_{wk2}A + \gamma_{wk3}\bar{A}_j + \gamma_{wk4}X_w + g_{wkt} + \epsilon_{wt}$$

  - A is individual ability and $\bar{A}_j$ is college quality (average ability)
  - $\epsilon_{wt} \sim N(0, \sigma_w^2)$

Unobserved Hetereogeneity: College and Major Choice

- McFadden (1978): $F(\epsilon) = \exp\{-G(e^{-\epsilon})\}$ is cdf for multivariate extreme value distribution
- Following Bresnahan, Stern, and Trajtenberg (1997) (BST), define

$$G(e^{-\epsilon}) = \left(\frac{1-\rho_{c1}}{2-\rho_{c1}-\rho_{c2}}\right)\sum_j\left(\sum_k\left(e^{-\epsilon_{cjk}}\right)^{\frac{1}{\rho_{c1}}}\right)^{\rho_{c1}} + \left(\frac{1-\rho_{c2}}{2-\rho_{c1}-\rho_{c2}}\right)\sum_k\left(\sum_j\left(e^{-\epsilon_{cjk}}\right)^{\frac{1}{\rho_{c2}}}\right)^{\rho_{c2}} + e^{-\epsilon_{co}}$$

- Notice as $\rho_{c1} \to 1$, we have

$$G(e^{-\epsilon}) \to \sum_k\left(\sum_j \exp\left(\frac{-\epsilon_{cjk}}{\rho_{c2}}\right)\right)^{\rho_{c2}} + e^{-\epsilon_{co}}$$

gives nested logit with majors as nests

Unobserved Hetereogeneity: Applications

Enumerate schools $1, ..., N$ and possible combinations $1, ..., L$, applying BST, define

$$G(e^{-\epsilon}) = \sum_{n=1}^{N} \frac{1}{M} \left(\sum_{\ell=1}^{L} (n \in J_\ell) \left(e^{-\epsilon_{s\ell}} \right)^{\frac{1}{\rho_s}} \right)^{\rho_s} + \sum_{\ell=1}^{L} \left(1 - \sum_{n=1}^{N} \frac{(n \in J_\ell)}{M} \right) e^{-\epsilon_{s\ell}} + e^{-\epsilon_{so}}$$

Nests defined by each school, and errors correlated among groups J that contain that particular school

Correlation Across Stages

- Assume preferences for colleges and majors persist between stages
- Assume ability is partially unobserved

Following Heckman and Singer (1984), assume R types, each occurring with probability π_r. Preferences and ability vary by type.

Data and Restrictions

Use NLSY72, which tracks students who were HS seniors in 1972 from 1972-1986. Following restrictions...

1. Each student can apply to three schools out of eight randomly assigned $\implies$ 92 combinations
2. Majors grouped: (engineering, physical sciences, biological sciences), (business and economics), (social sciences and humanities), (education)

Identification

Primary issue is *selection*

- With no serial correlation in preferences, selection is on observables
- With serial correlation, selection is on observables and unobservable type

How to identify types?

Identification

Primary issue is *selection*

- With no serial correlation in preferences, selection is on observables
- With serial correlation, selection is on observables and unobservable type

How to identify types?

- Dynamics: observing combinations of applications, acceptances/aid, and labor market outcomes
- Exclusion Restrictions: total SAT affects admissions, but only math SAT affects earnings

Estimation: No Serial Correlation, Observed Ability

- ϵ_{cjk} correlated *within majors, across schools* and *across majors, within schools*
- ϵ_{sr} correlated *across groups of schools with common element*
- Stages not connected

Then, log-likelihood is

$$L = \underbrace{L_1(\gamma_w)}_{\text{earnings}} + \underbrace{L_2(\gamma_a)}_{\text{admissions}} + \underbrace{L_3(\gamma_f)}_{\text{aid}} + \underbrace{L_4(\alpha_c, \alpha_w, \gamma_w)}_{\text{offer acceptance}} + \underbrace{L_5(\alpha_s, \alpha_c, \alpha_w, \gamma_a, \gamma_f, \gamma_w)}_{\text{applications}}$$

1. Maximize L_1, L_2, L_3 to get $\hat{\gamma}_w, \hat{\gamma}_a, \hat{\gamma}_f$
2. Maximize L_4 given $\hat{\gamma}_w$ to get $\hat{\alpha}_c, \hat{\alpha}_w$
3. Maximize L_5 given $\hat{\gamma}_w, \hat{\gamma}_a, \hat{\gamma}_f, \hat{\alpha}_c, \hat{\alpha}_w$ to get $\hat{\alpha}_s$

Estimation: Serial Correlation, Unobserved Ability

$$L = \log\left(\sum_{r=1}^{R} \pi_r \mathcal{L}_{1r}\mathcal{L}_{2r}\mathcal{L}_{3r}\mathcal{L}_{4r}\mathcal{L}_{5r}\right)$$

Estimate via expectation-maximization (EM) algorithm

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

EM algorithm:

1. Given $\hat{\alpha}_m, \hat{\gamma}_m, \hat{\pi}_m$, calculate $P_m(r|X, \hat{\alpha}_m, \hat{\gamma}_m, \hat{\pi}_m)$ and update $\hat{\pi}_{m+1}$
2. Given $P_m(r|X, \hat{\alpha}_m, \hat{\gamma}_m, \hat{\pi}_m)$, solve

$$\begin{aligned}(\hat{\alpha}_{m+1}, \hat{\gamma}_{m+1}) = \text{argmax}_{\alpha,\gamma} \sum_{i=1}^{I}\sum_{r=1}^{R} P(r|X, \hat{\alpha}_m, \hat{\gamma}_m, \hat{\pi}_m)(L_{1r}(\gamma_w) \\ + L_{2r}(\gamma_a) + L_{3r}(\gamma_f) + L_{4r}(\alpha,\gamma) + L_{5r}(\alpha,\gamma))\end{aligned}$$

where additive separability is again exploited. Iterate.

Results: Treatment Effect of College

% earnings premiums for attending college calculated using average characteristics by race and group

Group	Nat. Sci.	Bus.	Soc. Sci.	Edu.
Whites (One-Type/Two-Type)				
Did not apply	21.8/14.0	19.6/16.6	13.1/4.2	4.0/-9.5
Did not attend	23.4/14.4	22.0/18.6	12.7/3.5	3.8/-9.1
Attended college	24.8/17.0	24.3/20.5	12.3/3.2	3.6/-10.6
Blacks (One-Type/Two-Type)				
Did not apply	25.9/26.9	24.9/28.3	20.9/18.7	14.1/6.6
Did not attend	26.9/27.1	26.3/29.6	20.8/18.3	14.1/7.1
Attended college	27.5/29.1	27.4/30.4	20.5/18.2	14.0/5.7

Results: Expected Annual Earnings Losses for Black Males (One Type)

Experiment: Black males face white admissions and financial aid policies

Quantile	Application Not Fixed	Application Fixed
25th	\$28 (20)	\$15 (10)
50th	\$70 (35)	\$29 (15)
75th	\$146 (68)	\$59 (27)
90th	\$410 (143)	\$145 (57)
95th	\$606 (203)	\$213 (89)
99th	\$1320 (393)	\$610 (170)

Results: Expected Annual Earnings Losses for Black Males (One Type)

Experiment: Black males face white admissions and financial aid policies

Quantile	Application Not Fixed	Application Fixed
25th	\$28 (20)	\$15 (10)
50th	\$70 (35)	\$29 (15)
75th	\$146 (68)	\$59 (27)
90th	\$410 (143)	\$145 (57)
95th	\$606 (203)	\$213 (89)
99th	\$1320 (393)	\$610 (170)

	Baseline	Experiment
College	12.03% (1.15%)	10.68% (1.06%)
School avg. SAT $\geq$ 1100	1.93% (0.19%)	1.26% (0.11%)
School avg. SAT $\geq$ 1200	0.67% (0.13%)	0.32% (0.03%)

Questions?