Dynamic discrete choice models: application 1

Olivier De Groote

TSE

MRes: EEE, 2021-2022

Introduction I

- Paper to be discussed: Arcidiacono, P. 2005. "Affirmative action in higher education: How do admission and financial aid rules affect future earnings?" *Econometrica* 73 (5): 1477–1524.
- Uses dynamic discrete choice model to study how race-based advantages for blacks help them to have better educational outcomes and higher wages
- Models decision of students to apply for which colleges and to go to which college and major
- Models decision of schools to accept students and provide financial aid
- Estimate wage equation

Introduction II

- Data from the 70s but still relevant discussion, see also Arcidiacono, P., Lovenheim, M., and Zhu, M. 2015. "Affirmative Action in Undergraduate Education". *Annual Review of Economics* 7 (1): 487–518.
- Empirical problem 1: changing probability of acceptance/aid will change probability students will apply
  - Need dynamic model
- Empirical problem 2: self-selection and heterogeneous treatment effects
  - Need persistent unobserved heterogeneity
  - Conditional independence assumption Rust (1987) too restrictive but can allow for unobserved heterogeneity in the form of unobserved types (Heckman & Singer (1984))
  - This was also used in influential paper of Keane and Wolpin (1997)

Introduction III

- This paper applies recent advances (Arcidiacono & Jones (2003)) to decrease computational burden (later we also cover Arcidiacono & Miller (2011) for further improvements).

Data I

- NLS72: high school graduates of 1972

- Don't observe, nor model, every year but rather different "stages"

- Data on applications and decisions of schools (1972), attendance of students (1974), and wages (1986)

Data II

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 of applications	0.5924 (0.8312)	0.5809 (0.8066)	1.4397 (0.6777)	1.4006 (0.6503)	1.5772 (0.7443)	1.5491 (0.7246)
Math school quality			535.6 (55.2)	460.3 (100.2)	538.2 (51.2)	466.5 (104.5)
Verbal school quality			508.2 (52.8)	438.8 (97.5)	509.6 (48.9)	446.1 (101.8)
School cost[a]			11,505 (4,192)	10,596 (3,843)	11,403 (4,003)	10,632 (4,124)
Financial aid			1,250 (2,736)	2,180 (3,796)	1,456 (2,930)	3,195 (4,594)
State college premium[b]	0.2684 (0.0621)	0.2949 (0.0566)	0.2705 (0.0623)	0.2959 (0.0586)	0.2722 (0.0621)	0.2999 (0.0600)
SAT math	442.1 (104.2)	334.3 (70.0)	500.3 (105.9)	360.9 (82.1)	529.8 (101.2)	378.9 (85.6)
SAT verbal	410.3 (100.7)	305.5 (67.5)	465.4 (102.8)	332.4 (79.5)	489.8 (99.5)	356.9 (87.5)
HS class rank	0.5589 (0.2780)	0.4677 (0.2749)	0.6983 (0.2351)	0.5710 (0.2629)	0.7633 (0.2032)	0.6212 (0.2546)
Unknown HS class rank	0.1143	0.2322	0.1373	0.2634	0.1249	0.2486
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.

Model: overview

- 4 stages

1. Students choose where to submit applications.
2. Colleges make admission and financial aid decisions
3. Students choose school (or labor market)
4. Students enter labor market

- The model is solved (and discussed) backwards
- We first discuss a model that satisfies Rust's CI assumption
- We then generalize to allow for unobserved ability and how this can be identified

Model: stage 4 (labor market) I

- Log earnings t years after high school:

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

- j: school, k: major, A: (observed) ability, $\overline{A}_j$: college quality, X_w: other individual characteristics
- g_{wkt} trend and ϵ_{wt} a normal shock

Model: stage 4 (labor market) II

- Expected utility of working

$$
\begin{aligned}
u_{wjk} &= \alpha_w log\left(E_w\left[\sum_{t=t'}^{T}\beta^{t-t'}P_{kt}W_{jkt}\right]\right) \\
&= \alpha_w(\gamma_{wk1}+\gamma_{wk2}A+\gamma_{wk3}\overline{A}_j+\gamma_{wk4}X_w) \\
&+ \underbrace{\alpha_w log\left(E_w\left[\sum_{t=t'}^{T}\beta^{t-t'}P_{kt}exp(g_{wkt}+\epsilon_{wt}\right]\right)}_{gender-major-time\ FE}
\end{aligned}
$$

- How to identify? γ simply comes from the regression of log wages on covariates, α_w-> see next stage

Model: stage 3 (choice of college+major) I

- Flow utility attending college j and major k

$$u_{cjk} = \alpha_{c1} X_{cjk} - c_{jk} + \epsilon_{cjk}$$
$$with\ c_{jk} = \alpha_{c2k}(A - \overline{A}_j) + \alpha_{c3}(A - \overline{A}_j)^2$$

- ϵ_{cjk} follows an extension of nested logit that allows for correlation both within major and within college (BST)

- Students who have the option to go to college choose the option with the highest available

$$v_{cjk} = u_{cjk} + u_{wjk}$$

- Others have to go for $v_{c0} = u_{w0}$

Model: stage 3 (choice of college+major) II

- This implies that labor market is assumed to be an absorbing state
- Note that X_w needs to contain a variable not included in X_{cjk} (i.e. an exclusion restriction) to identify α_w
  - Exclusion restriction: state college premium
  - Choices differ between students because of differences in college premia -> must go through α_w

Model: stage 2 (admissions and financial aid)

- Admissions: logit probabilities, independent across schools

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

- Financial aid (share of the bill paid s): tobit between 0 (no aid) and 1 (fully cover tuition cost)
  - tobit to have mass points at 0 and 1
  - $s_j^* = \gamma_f X_{fj} + \epsilon_{fj}$
  - $s_j = 0$ if $s_j^* \leq 0$
  - $s_j = 1$ if $s_j^* > 1$
  - $s_j = s_j^*$ if $0 < s_j^* < 1$

Model: stage 1 (applying) I

- Let the flow utility of application to a set of options J be

$$u_{sJ} = -\alpha_{s2} X_{sJ}$$

- Then students choose the option with the highest expected lifetime utility

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

- Make it tractable using Rust (1987)
  - Conditional independence: $\epsilon_s^{\prime}s$ are independent from the ϵ_c's
  - Discretize the financial aid realizations
  - -> closed form solution for this expectation (see page 1489)

Model: stage 1 (applying) II

- Apply BST framework again, now each school is its own nest so different application sets affect different nests, this still leads to a closed form for the probability (page 1490)

Estimation I

- We can construct a loglikelihood function of the entire model (I index part of the likelihood function wrt stages which is different from paper):

$$
\begin{aligned}
lnL(\alpha_s, \alpha_c, \alpha_w, \gamma_a, \gamma_f, \gamma_w) &= ln \prod_i L_{i,4}(\gamma_w) \times L_{i,3}(\alpha_c, \alpha_w, \gamma_w) \times L_{i,2,admission}(\gamma_a) \\
&\times L_{i,2,aid}(\gamma_f) \times L_{i,1}(\alpha_s, \alpha_c, \alpha_w, \gamma_a, \gamma_f, \gamma_w) \\
&= \sum_i lnL_{i,4}(\gamma_w) + \sum_i lnL_{i,3}(\alpha_c, \alpha_w, \gamma_w) \\
&+ \sum_i lnL_{i,2,admission}(\gamma_a) + \sum_i lnL_{i,2,aid}(\gamma_f) \\
&+ \sum_i lnL_{i,1}(\alpha_s, \alpha_c, \alpha_w, \gamma_a, \gamma_f, \gamma_w)
\end{aligned}
$$

Estimation II

- As in Rust (1987), the likelihood function is additively separable and the model can be estimated stage-by-stage
  - Wages and admission and aid probabilities can be estimated separately because they only depend on the parameters introduced in their own stage
  - Once we have these, can estimate stage 3 (choice of college+major)
  - Finally, we can use all we have and estimate stage 1 (applying)
- To correct standard errors he uses one Newton step of the full likelihood function

Adding unobserved ability I

- It remains important to consider ability bias
- If we do not control for ability (or more generally heterogeneity), we cannot make causal statements, in particular we might be overestimating returns to college
- He controls for a measure of ability (SAT scores) but this might not capture everything
- Unobserved ability then enters the ϵ's which violates Rust's CI assumption

Adding unobserved ability II

- Solution: keep the ϵ's and CI the way it is but add an unobserved state variable to the model
  - Unobserved state is here a type (Heckman and Singer (1984))
  - Each student belongs to 1 of the R types
  - Type enters the model as if it was an observed student characteristic
  - The econometrician specifies the number of types
  - The model estimates the distribution over the population (type probabilities π_r) and how each type differs in each stage of the model

Estimation with unobserved ability I

- We lose additive separability

$$lnL(\alpha_s, \alpha_c, \alpha_w, \gamma_a, \gamma_f, \gamma_w)$$
$$= ln \sum_{r=1}^{R} \pi_r \prod_i L_{i,4,r} \times L_{i,3,r} \times L_{i,2,admission,r} \times L_{i,2,aid,r} \times L_{i,1,r}$$

Estimation with unobserved ability II

- Arcidiacono & Jones (2003) show that this can be restored using the EM algorithm
  - Start from arbitrary parameter values $(\alpha^0, \gamma^0, \pi^0)$
  - Step 1: Calculate the probability to belong to each type, conditional on the data and parameters

$$Pr(r|\mathbf{X_i}, \alpha^0, \gamma^0, \pi^0) = \frac{\pi_r L_{i,4,r} L_{i,3,r} L_{i,2,admission,r} L_{i,2,aid,r} L_{i,1,r}}{\sum_{r'=1}^{R} \pi_{r'} L_{i,4,r'} L_{i,3,r'} L_{i,2,admission,r'} L_{i,2,aid,r'} L_{i,1,r'}}$$

Estimation with unobserved ability III

- Step 2: Find new parameters using the expected log-likelihood function, holding the conditional probabilities fixed

$$\sum_i \sum_{r=1}^{R} Pr(r|\mathbf{X_i}, \alpha^0, \gamma^0, \pi^0)$$
$$[lnL_{i,4,r}(\gamma_w^1) + lnL_{i,3,r}(\alpha_c^1, \alpha_w^1, \gamma_w^1)$$
$$+lnL_{i,2,admission,r}(\gamma_a^1) + lnL_{i,2,aid,r}(\gamma_f^1)$$
$$+lnL_{i,1,r}(\alpha_s^1, \alpha_c^1, \alpha_w^1, \gamma_a^1, \gamma_f^1, \gamma_w^1)]$$

- Repeat until convergence
- Additive separability restored because in step 2, $Pr(r|\mathbf{X_i}, \alpha^0, \gamma^0, \pi^0)$ does not depend on parameters to be estimated, we just need to use it as a weight in the estimation of the different parts of the model

Estimation with unobserved ability IV

- (Note: can converge to local maximum so repeat for different starting values)

- Note that $Pr(r|\mathbf{X}_i, \alpha, \gamma, \pi)$ will be different for everyone but the population probability is assumed to be the same here π_r
  - This can be generalized to be conditional on an observed characteristic, we then have to calculate step 1 separately for each realization of that characteristic to obtain the weight
  - Usually this is not done for interpretation issues, if all observables enter everywhere, type is capturing what the observables are not
  - However, when observables do not enter everywhere, one might need to condition on them here to avoid an initial conditions problem (see e.g. Keane and Wolpin (1997))
  - In this paper, types are conditioned on income while income is not included in the wage regression to help identification

Estimation with unobserved ability V

- How is it identified?
  - Dynamics (similar to a fixed effect in panel data)
    - We model many choices of which the $\epsilon' s$ are independent, however in the data we observe correlations, the model allows for this through the unobserved types
    - 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."
  - Exclusion restrictions (see above)

Estimation results: stage 4

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.

Estimation results: stage 3

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.

Estimation results: stage 2

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.

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.

Estimation results: stage 1(+)

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.

Counterfactual choices

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.

Counterfactual average income loss

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.

Conclusion

- Results of paper
  - Affirmative action affects where students go to college, not whether they go
  - Returns to college quality are low (major more important)
  - Result: no large effects on wages, mainly shifting between colleges
- Empirical problem 1: changing probability of acceptance/aid will change probability students will apply
  - Need for a dynamic model nicely illustrated by comparing results with and without adjustment in application decision
- Empirical problem 2: self-selection and heterogeneous treatment effects
  - Nice application of Arcidiacono & Jones (2003) to combine Rust's (1987) benefits of CI, but relaxing this assumption using Heckman and Singer (1984) types and illustrate its importance by comparing results with 1 and 2 types

References I

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

Peter Arcidiacono and John Bailey Jones.
Finite mixture distributions, sequential likelihood and the EM algorithm.
Econometrica, 71(3):933–946, 2003.

Peter Arcidiacono, Michael Lovenheim, and Maria Zhu.
Affirmative Action in Undergraduate Education.
Annual Review of Economics, 7(1):487–518, August 2015.

References II

Peter Arcidiacono and Robert A. Miller.
Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity.
Econometrica, 79(6):1823–1867, 2011.

Christian Belzil and Jörgen Hansen.
A structural analysis of the correlated random coefficient wage regression model.
Journal of Econometrics, 140(2):827–848, October 2007.

J. Heckman and B. Singer.
A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data.
Econometrica, 52(2):271, March 1984.

References III

Michael P. Keane and Kenneth I. Wolpin.
The Career Decisions of Young Men.
Journal of Political Economy, 105(3):473–522, June 1997.

John Rust.
Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher.
Econometrica, 55(5):999–1033, September 1987.

Kenneth Train.
Discrete choice methods with simulation.
Cambridge University Press, Cambridge; New York, 2009.

Derivation utility labor I

- $u_{wjk} = \alpha_w log\left(E_w\left[\sum_{t=t'}^{T} \beta^{t-t'} P_{kt} W_{jkt}\right]\right)$
- with $ln(W_{jkt}) = \gamma_{wk1} + \gamma_{wk2}A + \gamma_{wk3}\overline{A}_j + \gamma_{wk4}X_w + g_{wkt} + \epsilon_{wt}$
- Note that
  $W_{jkt} = exp(\gamma_{wk1} + \gamma_{wk2}A + \gamma_{wk3}\overline{A}_j + \gamma_{wk4}X_w)exp(g_{wkt} + \epsilon_{wt})$
  and $ln(exp(a)) = a$

Derivation utility labor II

- Therefore

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

Derivation utility labor III

- Second term captured by fixed effects so no need to model employment
- Alternative: $u = ln(e * w) = ln(e) + ln(w)$ and model each, see Belzil and Hansen (2007)

Scaling

- In the application stage, we obtain an estimate of 4.2749 for the PV of future utility
- Arcidiacono claims this cannot be just the discount factor (because <1) but differences in variance
- Note that with standard exteme value type 1 errors with variance $(\sigma_t)^2 \frac{\pi^2}{6}$ in a 2 period model we have

$$v_{ijt} = \frac{u_j(x_{it})}{\sigma_1} + \frac{\sigma_2}{\sigma_1} ln \sum_{j'} exp\left(\frac{u_{j'}(x_{it+1})}{\sigma_2}\right)$$

- usually we normalize $\sigma_1 = \sigma_2 = 1$ (or we say we are actually estimating parameters divided by some scale, see also section 3.2 in Train (2009)) but note that if we have something shifting utility in period 2, not 1, we can actually identify $\frac{\sigma_2}{\sigma_1}$.