

Replication Notes for “The Out-of-State Tuition Distortion”

Brian Knight and Nathan Schiff

July 2, 2018

The code and necessary files are in three separate directories: 1) “create_zipcode_dists” 2) HERI_analysis and 3) “ELS_NPSAS_analysis”. We have set up these directories so that each directory stands alone and does not need to be run in any particular sequence (however, see discussion of “create_zipcode_dists” directory).

0.1 Create Zipcode Dists

This directory creates files with distances between zipcodes and state borders: “zipDists_1990.dta”, “zipDists_2000.dta”, and “zipDists_2010.dta”. The Stata DO file “run_zipcode_dists.DO” in the root of the directory calls all programs in the correct order. To do so it uses code run in R (“calc_zip_dists.R”) along with a number of packages for spatial calculations. However, we have also included the files created by this directory in the input subfolder of the “HERI_analysis” directory, and thus it is not necessary for researchers to run any programs in “create_zipcode_dists” in order to replicate our analysis. There is additional documentation on the calculation in the documentation subfolder of “create_zipcode_dists”.

Input files: i) “gaz_zips_YEAR.txt” (1990, 2000, 2010-2014) ii) “state_fips.dta”
iii) “states_LAND_ONLY_2010” shapefile folder iv) “states_with_WATER_2010” shapefile folder

Code files: i) “run_zipcode_dists.do” ii) “parse_gazetteer_files.do” iii) “calc_zip_dists.R” iv) format_zipdists.DO
v) label_zipdists.DO vi) gaz_zip_2000.dct

0.2 HERI Analysis

This directory cleans the data from the “Higher Education Research Institute” (HERI) and then runs the analysis, creating main tables 3-7, Appendix tables B3 and B4, and figures 3-6. The Stata DO file “run_HERI_analysis.DO” in the root of the directory calls all programs in the correct order. All necessary files are included in the “input” subfolder *except* the data from HERI, which we do not have permission to include (this data file is referred to as “TFS_KNFINAL2.dta” in our code). At the end of this document we have included our HERI data request listing all variables.

Input files: i) “state_fips.dta” ii) “tuition_exchanges.dta” iii) “zip_gaz_2000.dta” iv) “zipDists_YEAR.dta” (1990, 2000, 2010)

Code files: i) “run_HERI_analysis.do” ii) “create_distBands.do” iii) “return_distBand.do”
iv) “create_analysis_file.do” v) “return_SAT.do” vi) “analyze_HERI.do”

0.3 ELS NPSAS Analysis

This directory contains the code necessary to reproduce our analysis using the Educational Longitudinal Study and National Postsecondary Student Aid Study datasets. We used these datasets to create tables 2, 8, and 9 in the main document; the code is labelled accordingly.

Input files: i) “aej_table2.do” ii) “aej_table8.do” iii) “aej_table9.do”

1. Name of principal investigator:			
2. Which survey and which year(s)		Yes / No	Year (s)
The Freshman Survey (TFS)		Yes	1982-2011
Your First College Year (YFCY)		No	
College Senior Survey (CSS)		No	
The Faculty Survey (FAC)		No	
3. Please describe the population/subset in your study: (i.e., first-time full-time freshman only, women only, graduating seniors, STEM majors only, responded to freshman survey and senior survey, etc)		all freshman respondents	
		Please use variable names based on your type of request. If requesting data for one year only, use the variable names from the annual codebook. If requesting data for two or more years, use the variable names from the trends file codebook. If requesting longitudinal data (TFS/YFCY or TFS/CSS), use only variables from the YFCY or CSS codebooks. YFCY and CSS codebooks contain matching TFS responses and variables (e.g., ACT01_TFS) found at the end of the codebooks. HERI codebooks: http://www.heri.ucla.edu/researchers/Tools/Codebooks.php	
4. Please list all variables included in your study:		Dependent Variable? (Yes/No)	Variable (i.e., COLACT01, SATIS15)
1	YES	STATE	Institution state
2	NO	HOMESTATE	Student's home state
3	NO	CHOOSE22	Choose to Attend: The athletic department recruited me
4	NO	CHOOSE05	Choose to Attend: Could not afford first choice
5	NO	CHOOSE08	Choose to Attend: I wanted to live at home
6	NO	CHOOSE09	Choose to Attend: I wanted to live near home
7	NO	CHOOSE12	Choose to Attend: I was offered financial assistance
8	NO	CHOOSE18	Choose to Attend: Not accepted anywhere else
9	NO	CHOOSE19	Choose to Attend: Not offered aid by first choice
10	NO	CHOOSE22	Choose to Attend: The cost of attending this college
11	NO	CHOOSE25	Choose to Attend: This college has low tuition
12	NO	CHOICE	Choice: Is this college your:
13	NO	ACCP1ST	Were you accepted by your first choice college?
14	NO	DISTHOME	How many miles is this college from your permanent home?
15	NO	NUMAPPLY	To how many colleges other than this one did you apply for admission this year?
16	NO	NUMACPT	How many other acceptances did you receive this year?
17	NO	NUMACPT01_N	From how many did you receive acceptance?
18	NO	NUMACPT02_N	How many other acceptances did you receive this year?
19	NO	DORM	DORM
20	NO	DOB	DOB Year
21	NO	HOMZIP	Student's home zip
22	NO	SEX	Your sex
23	NO	RACEGROUP	Race/Ethnicity Group
24	NO	RRACE	Responded to race
25	NO	CITIZEN	Citizenship status
26	NO	MARITAL	Are you presently married?
27	NO	INCOME	What is your best estimate of your parents' total income last year?
28	NO	FATHEDEC	Father's education
29	NO	MOTHEDEC	Mother's education
30	NO	NUMDEP	How many persons are currently dependent on you and attending college?
31	NO	NUMDEP01	How many of these persons are currently dependent on you and attending college?
32	NO	FULLSTAT	Are you enrolled (or enrolling) as a:
33	NO	SIBLIVE1_T2	Number of older brothers
34	NO	SIBLIVE2_T2	Number of older sisters
35	NO	SIBLIVE3_T2	Number of younger brothers
36	NO	SIBLIVE4_T2	Number of younger sisters
37	NO	LEGACY_T2	Did either of your parents or legal guardians attend the institution that you are now attending?
38	NO	AID1	Aid: Family resources (parents, relatives, spouse, etc.)
39	NO	AID2	Aid: My own resources (savings from work, work-study, other income)
40	NO	AID3	Aid: Aid which need not be repaid (grants, scholarships, military funding, etc.)
41	NO	AID4	Aid: Aid which must be repaid (loans, etc.)
42	NO	AID5	Aid: Other than above
43	NO	AID01	Aid: Parents, other relatives or friends
44	NO	AID02	Aid: Spouse
45	NO	AID03	Aid: Savings from summer work
46	NO	AID04	Aid: Other savings
47	NO	AID05	Aid: Part-time job on campus
48	NO	AID06	Aid: Part-time job off campus
49	NO	AID07	Aid: Full-time job while in college
50	NO	AID08	Aid: Pell Grant
51	NO	AID09	Aid: Supplemental Educational Opportunity Grant
52	NO	AID10	Aid: State scholarship or grant
53	NO	AID11	Aid: College work-study grant
54	NO	AID12	Aid: College grants/scholarship (other than above)
55	NO	AID13	Aid: Vocational Rehabilitation funds
56	NO	AID14	Aid: Other private grant
57	NO	AID15	Aid: Other government aid (ROTC, BIA, Gilmilitary benefits, etc.)
58	NO	AID16	Aid: Stafford loan (GSL)
59	NO	AID17	Aid: Perkins loan
60	NO	AID18	Aid: Other college loan
61	NO	AID19	Aid: Other loan
62	NO	AID20	Aid: Other than above
63	NO	YRGRADHS	In what year did you graduate from high school?
64	NO	HSGPA	What was your average grade in high school?
65	NO	SATV	SAT Critical Reading/Verbal
66	NO	SATM	SAT Mathematics
67	NO	SATW	SAT Writing
68	NO	ACTCOMP	ACT Composite
69	NO	HSGPA01_N	What was your average grade in high school?
70	NO	PLANIVE	Where do you plan to live during the fall term?
71	NO	FUTACT21	Future Act: Play intercollegiate athletics (e.g., NCAA or NAIA-sponsored)
72	NO	FUTACT28	Future Act: Play varsity/intercollegiate athletics
73	NO	DEPOA1_T2	Last year - Living with your parents for more than five consecutive weeks
74	NO	DEPOA2_T2	Last year - Listed as a dependent on your parents' Federal Income Tax Return
75	NO	DEPOA3_T2	Last year - Receiving assistance worth \$600 or more from your parents
76	NO	DEPOB1_T2	This year - Living with your parents for more than five consecutive weeks
77	NO	DEPOB2_T2	This year - Listed as a dependent on your parents' Federal Income Tax Return
78	NO	DEPOB3_T2	This year - Receiving assistance worth \$600 or more from your parents
79	NO	YEAR	Survey year
80	NO	ACT	College I.D.
81	NO	SUBID	Survey I.D.
82	NO	STRAT	CRP Stratification Cell
83	NO	SELECTIVITY	Institutional Selectivity
84	NO	STUDSTAT	Student status
85	NO	NORMSTAT	Norms status
86	NO	STUDWGT	Student weight