

Summer Melt and Free Application for Federal Student Aid Verification



Brian Holzman, Ph.D., and Vansa Shewakramani Hanson, M.A.

June 2020

Appendix A. Variables Used in the Analyses

Variable	Description
Outcomes	
FAFSA Verification	Binary: Whether a student was flagged for FAFSA verification.
Summer Melt	Binary: Whether a student experienced summer melt by November 1 of the fall semester after completing high school.
Delayed College Enrollment	Binary: Whether a student who experienced summer melt enrolled in college during a later term (e.g., spring, summer, or sometime thereafter) and within two years of high school graduation.
Student and School Characteristics	
Age	Continuous.
Female	Binary.
Race/Ethnicity	Categorical: White (ref.), Black, Hispanic, Asian/Pacific Islander, and Other. Other includes Native American and multiracial students.
Immigrant	Binary: Foreign-born.
Non-English Home Language	Binary: Primary language spoken at home was a foreign language.
Special Education	Binary.
Economically Disadvantaged	Binary: Was eligible for the free and reduced-price lunch program or other federal poverty programs, or living below the federal poverty line.
Median Household Income of Block Group	Continuous (reported in 10,000s). Based on geocoded home address data linked to Census Block Group data from the 2015 American Community Survey.
Household Structure	Categorical: Two Parents (ref.), One Parent, and Unknown.
SAT Score	Continuous (reported in 100s).
Grade Point Average	Continuous.
Number of College-Level Courses Taken	Categorical: None (ref.), One-Two Courses, Three-10 Courses, and 11 or More Courses. Number of Advanced Placement (AP), International Baccalaureate (IB), and academic dual enrollment courses taken during high school. Academic dual enrollment courses were dual enrollment courses that were not also Career & Technical Education courses.
Targeted for College Advising	Binary: Whether a student was identified by College Advisors as a target student. The criteria for being a target student varied by school and cohort.
Number of Advising Sessions	Categorical: None (ref.), One Session, Two Sessions, and Three or More Sessions. Number of advising sessions a student participated in, according to the district's tracking system.
School Percentage Economically Disadvantaged	Continuous: Percentage of economically disadvantaged students (reported in 10s). Calculated by aggregating student data to the school level.
School Average SAT Score	Continuous: Average SAT score (reported in 100s). Calculated by aggregating student data to the school level.

Appendix B. Methodology

Data

The study used data from Houston ISD, made available through the Houston Education Research Consortium (HERC), from 2013-2018. The analyses focused on two cohorts of high school seniors who graduated during the 2015-2016 and 2016-2017 school years. All students were tracked through the fall semester after graduating from high school: fall 2016 for the first cohort and fall 2017 for the second cohort. Students who experienced summer melt were tracked for additional time in order to examine delayed college enrollment: summer 2018 for the first cohort and summer 2019 for the second cohort. (Data dating back to the 2012-2013 school year were pulled in order to generate a measure of college-level course-taking during high school.)

Sample

HISD's leaver file was used to determine high school graduation. According to the leaver file, there were 20,836 high school graduates across the two cohorts.¹ Since summer melt research has focused on college-intending students — students who have completed the steps to college enrollment and appear ready to attend — the sample was limited to students who 1) applied to a postsecondary institution and 2) filed the FAFSA. Completing these two tasks might indicate a student was college-intending and planning to go to college. Limiting the sample to college-intending students reduced the analytic sample to 9,715.

Additional students were excluded from the sample because they could not be matched from the leaver file to the Public Education Information Management System (PEIMS) for background characteristics (N = 28) or were missing information on immigration status (from enrollment card data, N = 29), household structure (from enrollment card data, N = 15), grade point average (from grade point average data, N < 5), or college-level course-taking (from course grades data, N < 5). The final analytic sample consisted of 9,682 students. For Research Question 3, which examined delayed college enrollment among the students who experienced summer melt, the sample was further restricted to the 2,496 students who experienced summer melt.

About four percent of students were missing information on the median household income of their Census Block Group. These students' home addresses might have been missing information from the enrollment card data, or their addresses could not be geocoded and matched to the 2015 American Community Survey data. About four percent of students were also missing SAT score information. These students either did not take the test or took the test outside HISD. Instead of excluding students missing median household income or SAT score data, dummy variables that indicated whether a student was missing these data were created. Students missing these data were then assigned the mean value for the sample: students missing median household income were assigned a value of \$49,707 and students missing SAT scores were assigned a value of 971. The statistical models controlled for the recoded

¹ Some observations (N < 5) were dropped because it appeared the students were recorded as graduates of both cohorts.

median household income and SAT score variables (`medhhincimp` and `satimp`), as well as the dummy variables that indicated whether a student was missing data (`nomedhhinc` and `nosat`).

Additional details on sample selection and missing data are available from the authors upon request.

Analytic Strategy

Research Question 1: What types of students were more likely to be flagged for FAFSA verification?

To address this research question, logistic regression models with standard errors clustered at the school-by-cohort level were estimated. The first model controlled for student and school characteristics. The second model was a conditional fixed-effects model that accounted for observed and unobserved school characteristics that also changed over time. This was the preferred specification since schools might have implemented different policies and practices in college advising, for instance, in 2015-2016 and 2016-2017.² The variables included in the statistical models are described in Appendix A. Graphs were produced using Stata's `margins` command.

Research Question 2: Did FAFSA verification predict summer melt?

For this research question, logistic regression models with standard errors clustered at the school-by-cohort level were estimated. The first model controlled for FAFSA verification only, the second model added in student and school characteristics, and the third model was a conditional fixed-effects model that accounted for observed and unobserved school characteristics that also changed over time. The conditional fixed-effects approach was the preferred model. The variables included in the statistical models are described in Appendix A. Graphs were produced using Stata's `margins` command.

Research Question 3: Did students who experienced summer melt attend college at a later date? If so, was FAFSA verification related to those decisions?

Logistic regression models with standard errors clustered at the school-by-cohort level were estimated. The first model controlled for FAFSA verification only, while the second model added in student and school characteristics. The third model was preferred specification: a conditional fixed-effects model that accounted for observed and unobserved school characteristics that also changed over time. The variables included in the statistical models are described in Appendix A. Graphs were produced using Stata's `margins` command.

² During the 2015-2016 school year, HISD recorded how many advising sessions each student attended, regardless of the service provider (e.g., College Success Advisor, Advise Texas Advisor). In 2016-2017, HISD recorded advising sessions with College Success Advisors only. It was impossible to accurately distinguish the service provider in the 2015-2016 data and limit the advising sessions measure to sessions with College Success Advisors only; details are available from the authors upon request. However, the school-by-cohort effects accounted for this change in the district's advising tracking over time.

Appendix C. Summary Statistics

Variable	Mean	SD
FAFSA Verification	0.33	(0.47)
Summer Melt	0.26	(0.44)
Delayed College Enrollment	0.77	(0.42)
Age	0.15	(0.46)
Female	0.57	(0.50)
Race/Ethnicity		
White	0.10	(0.30)
Black	0.30	(0.46)
Hispanic	0.53	(0.50)
Asian/PI	0.06	(0.24)
Other	0.01	(0.10)
Immigrant	0.08	(0.28)
Non-English Home Language	0.41	(0.49)
Special Education	0.04	(0.19)
Economically Disadvantaged	0.69	(0.46)
Median Household Income of Block Group (in 1,000s)	4.97	(3.41)
Household Structure		
Two Parents	0.55	(0.50)
One Parent	0.34	(0.47)
Unknown	0.11	(0.31)
SAT Score (in 100s)	9.71	(2.15)
Grade Point Average	3.19	(0.75)
Number of College-Level Courses Taken		
None	0.21	(0.41)
One-Two Courses	0.11	(0.31)
Three-10 Courses	0.33	(0.47)
11 or More Courses	0.35	(0.48)
Targeted for College Advising	0.35	(0.48)
Number of Advising Sessions		
None	0.33	(0.47)
One Session	0.27	(0.44)
Two Sessions	0.17	(0.37)
Three or More Sessions	0.23	(0.42)
School Percentage Economically Disadvantaged (in 10s)	6.96	(1.95)
School Average SAT Score (in 100s)	9.19	(1.54)
Number of Students	9,682	

Source: HERC multi-year data.

Appendices

Appendix D. Odds Ratios from Logistic Regression Models Predicting FAFSA Verification

Variable	Model 1		Model 2	
	O.R.	Sig.	O.R.	Sig.
Age	1.01		1.04	
Female	0.96		0.97	
Race/Ethnicity (ref. = White)				
Black	1.99	***	2.01	***
Hispanic	1.92	***	1.92	***
Asian/PI	1.51	***	1.41	**
Other	2.15	***	2.27	***
Immigrant	0.91		0.90	
Non-English Home Language	1.03		1.07	
Special Education	1.04		1.03	
Economically Disadvantaged	1.16	*	1.12	
Median Household Income of Block Group (in 1,000s)	1.00		1.00	
Household Structure (ref. = Two Parents)				
One Parent	1.11		1.12	
Unknown	1.08		1.17	+
SAT Score (in 100s)	0.90	***	0.94	**
Grade Point Average	1.02		0.97	
Number of College-Level Courses Taken (ref. = None)				
One-Two Courses	1.24	*	1.23	*
Three-10 Courses	1.31	***	1.16	+
11 or More Courses	1.40	***	1.15	
Targeted for College Advising	0.96		0.98	
Number of Advising Sessions (ref. = None)				
One Session	0.96		0.95	
Two Sessions	0.98		1.04	
Three or More Sessions	0.96		1.17	*
School Percentage Economically Disadvantaged (in 10s)	0.91	**		
School Average SAT Score (in 100s)	0.86	***		
School-by-Cohort Fixed-Effects				X
Pseudo R ²	0.03		0.01	
Log Likelihood	-5,954		-5,614	
Number of Students	9,682		9,668	

Source: HERC multi-year data.

Note: Model 1 controlled for school characteristics and Model 2 controlled for school-by-cohort effects. Both models included missing dummies for a small proportion of students with incomplete data on median household income and SAT scores. Standard errors were clustered at the school-by-cohort level.

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001 (two-tailed tests)

Appendix E. Odds Ratios from Logistic Regression Models Predicting Summer Melt

Variable	Model 1		Model 2		Model 3	
	O.R.	Sig.	O.R.	Sig.	O.R.	Sig.
FAFSA Verification	1.58	***	1.43	***	1.52	***
Age			1.29	***	1.27	***
Female			1.03		1.01	
Race/Ethnicity (ref. = White)						
Black			0.73	+	0.70	*
Hispanic			0.73	*	0.82	
Asian/PI			0.61	*	0.65	*
Other			0.84		0.83	
Immigrant			1.33	**	1.38	**
Non-English Home Language			0.88	+	0.84	*
Special Education			0.92		0.88	
Economically Disadvantaged			1.17	*	1.22	**
Median Household Income of Block Group (in 10,000s)			0.97	**	0.98	*
Household Structure (ref. = Two Parents)						
One Parent			1.13	*	1.11	+
Unknown			1.20	*	1.04	
SAT Score (in 100s)			0.93	**	0.91	***
Grade Point Average			0.46	***	0.50	***
Number of College-Level Courses Taken (ref. = None)						
One-Two Courses			0.95		0.94	
Three-10 Courses			0.79	***	0.80	**
11 or More Courses			0.65	***	0.66	***
Targeted for College Advising			1.18	**	1.17	**
Number of Advising Sessions (ref. = None)						
One Session			0.84	*	0.83	*
Two Sessions			0.89		0.81	*
Three or More Sessions			0.85	+	0.69	***
School Percentage Economically Disadvantaged (in 10s)			1.14	***		
School Average SAT Score (in 100s)			0.92	*		
School-by-Cohort Fixed-Effects					X	
Pseudo R ²	0.01		0.15		0.09	
Log Likelihood	-5,482		-4,675		-4,368	
Number of Students	9,682		9,682		9,669	

Source: HERC multi-year data.

Note: Model 1 controlled for FAFSA verification only, Model 2 controlled for school characteristics, and Model 3 controlled for school-by-cohort effects. Models 2 and 3 included missing dummies for a small proportion of students with incomplete data on median household income and SAT scores.

Standard errors were clustered at the school-by-cohort level.

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001 (two-tailed tests)

Appendices

Appendix F. Odds Ratios from Logistic Regression Models Predicting Delayed Enrollment among Students Who Experienced Summer Melt

Variable	Model 1		Model 2		Model 3	
	O.R.	Sig.	O.R.	Sig.	O.R.	Sig.
FAFSA Verification	1.28	**	1.28	**	1.21	*
Age			0.98		0.97	
Female			0.87		0.86	
Race/Ethnicity (ref. = White)						
Black			0.82		0.83	
Hispanic			0.99		1.01	
Asian/PI			0.80		0.91	
Other			1.44		1.16	
Immigrant			0.98		0.92	
Non-English Home Language			1.25	+	1.35	*
Special Education			0.76		0.75	
Economically Disadvantaged			1.09		1.02	
Median Household Income of Block Group (in 1,000s)			1.01		1.00	
Household Structure (ref. = Two Parents)						
One Parent			0.87		0.86	
Unknown			0.89		0.97	
SAT Score (in 100s)			0.95		0.98	
Grade Point Average			1.25	*	1.14	
Number of College-Level Courses Taken (ref. = None)						
One-Two Courses			1.24		1.25	
Three-10 Courses			1.36	*	1.35	+
11 or More Courses			1.50	*	1.43	+
Targeted for College Advising			1.15		1.16	
Number of Advising Sessions (ref. = None)						
One Session			0.80		0.85	
Two Sessions			0.93		0.94	
Three or More Sessions			0.95		1.15	
School Percentage Economically Disadvantaged (in 10s)			0.93			
School Average SAT Score (in 100s)			0.98			
School-by-Cohort Fixed-Effects					X	
Pseudo R ²	0.00		0.02		0.02	
Log Likelihood	-1,433		-1,402		-1,211	
Number of Students	2,496		2,496		2,456	

Source: HERC multi-year data.

Note: The sample was limited to students who experienced summer melt. Model 1 controlled for FAFSA verification only, Model 2 controlled for school characteristics, and Model 3 controlled for school-by-cohort effects. Models 2 and 3 included missing dummies for a small proportion of students with incomplete data on median household income and SAT scores. Standard errors were clustered at the school-by-cohort level.

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001 (two-tailed tests)

Note on the authors: Vansa Shewakramani Hanson, M.A., is currently a doctoral student at Texas A&M University. Brian Holzman developed the research questions, analyzed the data, and wrote the brief. Vansa Shewakramani Hanson developed the research questions.

About HERC. Focusing on the most pressing challenges facing the region, the Houston Education Research Consortium (HERC) is a research-practice partnership between Rice University and 11 Houston-area school districts. HERC research is developed directly alongside district leaders with findings shared with decision makers – culminating in long-term, equity-minded solutions, opportunities and growth for Houston and beyond.



Houston Education Research Consortium
a program of the Kinder Institute for Urban Research
MS-258 Rice University | Houston, Texas 77005
713-348-2532 | herc@rice.edu
Find us online:
kinder.rice.edu/houston-education-research-consortium