

Strategic use of FAFSA list information by colleges

Stephen R. Porter
Department of Leadership, Policy, and Adult and Higher Education
North Carolina State University
Raleigh, NC 27695
srporter@ncsu.edu

Johnathan G. Conzelmann
RTI International
Research Triangle Park, NC 27709
jconzelm@rti.org

April 2015

Abstract

Critics are concerned that colleges are using financial aid data from the FAFSA to penalize students who unwittingly rank them high on the list of schools to which they wish their financial aid data sent. We use the most recent National Postsecondary Student Aid Study to test this hypothesis. We find that while the evidence suggests students are indeed ranking their most preferred institutions on the FAFSA, selective colleges and universities do not appear to be lowering their financial aid packages in response to student lists.

Strategic use of FAFSA list information by colleges

Competition within postsecondary education has increased dramatically during the past several decades. Worried about gaining admission to a quality institution, college applicants apply to more colleges with each passing year. Colleges face pressures to perform well in rankings such as *U.S. News & World Report*, while also grappling with declining revenues from research funding and state support. As a result, many colleges now use what is termed strategic enrollment management (SEM), a set of sophisticated techniques used to make admissions decisions (Hossler and Bontrager, 2014). One aspect of SEM that sets it apart from more traditional admissions decision-making is the strategic use of information about applicants and applicant behavior during the admissions process.

One common approach is to target admissions decisions and financial aid strategically, in order to boost yield rates while maximizing the impact of the financial aid budget. For example, an institution might deny admission to a student they deem very unlikely to enroll, in order to increase the proportion of admitted students who enroll, a key admissions metric used to judge college selectivity. Financial aid packages can be based, in part, on the estimated probability of matriculation. The aid package for students with high estimated probabilities of matriculation can be reduced (or shifted from grant aid to loans), without greatly reducing their probability of matriculation. The “saved” aid is then offered to students with a lower probability of matriculation, but who have characteristics desired by the institution (such as high SAT scores), in order to increase their probability of enrolling at the institution.

While some have criticized SEM in terms of its ethical treatment of students and its effect on institutional behavior, SEM is widely accepted within postsecondary education. However, the National Association for College Admission Counseling, comprised of high school counselors

and college admissions officers, is very critical of an alleged SEM practice that uses Free Application for Federal Student Aid (FAFSA) data. When students and their parents fill out the FAFSA, they can list up to ten colleges they wish to receive their financial information. The association argues that some colleges are using these lists as a revealed preference ranking for students, allowing them to identify which students view their college as their top choice. They then use this information when making admission and financial aid decisions. The association is so alarmed by the practice that they recently requested the federal government stop providing these FAFSA data elements to colleges, in order to prevent their use during the admissions process (Rivard, 2014).

However, it is not clear whether this practice is actually occurring. The purpose of this paper is to use the 2011-12 National Postsecondary Student Aid Study (NPSAS) to investigate whether some colleges are actually engaging in this strategic behavior. To date, whether colleges are using the list information to make admissions and financial aid decisions has been a matter of rumor amongst college admissions professionals, and has not been studied empirically.

We seek to answer two main questions with our analyses. First, do students use the FAFSA college lists to express their college preferences? While commentators assume this is the case, it is still an open question as to whether students are listing colleges in a preferred order; indeed, it is not even clear whether a significant number of students are listing multiple colleges on the FAFSA. Second, do some colleges appear to alter their aid packages to penalize students who list them in the first slot on the FAFSA form?

Data

The data used here are from the 2011-12 National Postsecondary Student Aid Study, a nationally representative sample of students enrolled in postsecondary education during the

2011-12 academic year. It is the most recent and comprehensive financial aid survey of college students in the country. The study data come from three main sources: 1) a student survey interview; 2) institutional student records collection; 3) and several administrative databases, including data from the FAFSA.

The full analytical sample consists of 111,057 students who attended a postsecondary institution in the 2011-12 academic year. We focus on a subset of respondents that meet several criteria. First, they must have a complete set of records. Students without a presence in the main undergraduate dataset were dropped, leaving 95,099 observations. Second, observations were also dropped if they did not match with a 2011-2012 FAFSA record, shrinking the sample further to 86,484 observations. We also limit to students enrolled in a bachelor's degree program at a four-year public or private nonprofit institution. In total, 21,526 observations remain in the sample at this point.

The third criterion involves the actual list of schools created by each student on their FAFSA applications. When filling out the FAFSA, either on paper or online, students can list up to ten schools using Federal School Codes, which are available through the Federal Student Aid website. To make use of these lists, we created a crosswalk of Federal School Codes and unit identifiers from the Integrated Postsecondary Education Data System (IPEDS). Of the 2,460 unique Title IV eligible institutions listed by students in our NPSAS sample, 98% were matched with IPEDS unit identifiers using a fuzzy string matching program and careful post-matching verification. The few schools without an IPEDS match created gaps in FAFSA school lists for a limited number of sample members. These 331 observations affected by the gaps were dropped from the sample, leaving 21,195 observations.

Fourth, the sample was subset further to include only full-time/full year, first-time beginning students, resulting in 5,257 observations. We focus on full-time students to ensure financial aid packages are comparable across students, and we use only first-time beginning students because we are interested in the first FAFSA that students filled out while they were making their college enrollment choice. Although FAFSA applications from other non-first-time beginners are included in the NPSAS database, these are not included in our analytical sample because they do not represent the first FAFSA filed by these students during the time they were deciding to enroll in college for the first time. Full-time/full year enrollment status is defined as students who were enrolled full-time for nine or more months during the 2011-12 school year, and those who only attended one institution in this time frame. This is to ensure that institutional aid awards are comparable, and unaffected by varying enrollment patterns. First-time beginning students were identified in the student interview portion of the NPSAS through a series of questions that determined eligibility for the Beginning Postsecondary Students (BPS) longitudinal follow-up studies. In order to be classified as a first-time beginner (FTB), students must not have enrolled in any other postsecondary institutions except for the institution on record for the NPSAS since completing their high school requirements.

Fifth, we limit the analysis to students who completed their FAFSA on or before May 1, 2011, using the FAFSA completion date in the NPSAS administrative data. Most schools require matriculation decisions by the beginning of May, and using FAFSAs submitted at a later date runs the risk of including students in the sample who were not considered for institutional admissions and financial aid in the same manner as students filing the FAFSA at an earlier date.

The sample for the descriptive analysis contains 5,239 full-time first-time beginning students seeking bachelor's degrees in the 2011-12 academic year. 656 institutions are

represented in this sample, for an average of 8 sample members who attended each institution. Because of estimation issues with schools with low numbers of students, the sample sizes for the analytic models vary.

FAFSA lists and student preferences

The National Association for College Admission Counseling believes that the list of colleges acts as a set of revealed college preferences: when students list colleges to which they wish their financial information sent, they start the list with their most-preferred college and end with their least-preferred college. But is this truly the case? Here, we assemble three pieces of evidence that collectively suggest this is indeed how students, on average, view the college list section of the FAFSA.

First, in order for students (and colleges) to use the list section as an implied preference ranking, most students should be listing multiple colleges when they submit the FAFSA. Figure 1 shows the distribution of the number of colleges listed by first-time, full-time students in Fall 2011. Just over one-third of students listed only a single institution, while over half listed at least three colleges. The median number of colleges listed on the FAFSA is three. If colleges adopt a simple view of the FAFSA list, and view these data as binary (e.g., “we are the first choice/we are not the first choice”), two-thirds of college students are listing enough colleges to imply a meaningful preference ranking.

Second, and somewhat surprisingly, the FAFSA itself instructs students to use the list as an implicit preference ranking. The instructions for this section tell students that, “For state aid, you may wish to list your preferred college first” (see Figure 2). At a minimum, the subset of students interested in state aid are likely placing their top choice in the top slot on the FAFSA.

Third, if students list their preferred college in the top slot, then higher proportions of students should be attending colleges near the top of their list compared to colleges at the bottom of their list. Table 1 shows the distribution of where students enrolled by FAFSA list position, separated out by the number of colleges they listed on the FAFSA. For example, the first row of the table shows that of the 1,920 students who only listed one college on the FAFSA (which was filed before May 1), 97% enrolled at the institution the following fall. For the 621 students who listed only two colleges on the FAFSA, 66% enrolled at the college listed in the first position, and 31% at the college listed in the second position.

Two trends are apparent in the table. First, the proportion of students enrolling in the first position college is generally double the proportion enrolling in the second position college. Second, the proportion enrolling at a listed institution shows a clear decline from the first position to the bottom position, with the proportions tending to stabilize after the third position. If we assume that students enroll in their most preferred college when extended an offer of admission, these numbers suggest that some students are listing their most preferred school first, followed by their second and third preferred choices.

In sum, the FAFSA provides students the ability to send their financial aid information to up to ten colleges. Many students use this to send information to multiple colleges. The FAFSA form itself encourages students to rank order colleges by their preferred choices, and a comparison of enrollment behavior and list position reveals that students are far more likely to enroll in their first listed college than any other college listed on the FAFSA. Together, these data suggest that many students are indeed using the FAFSA list as a college preference ranking. The next question is whether colleges are using this information in the financial aid process.

FAFSA lists and aid allocations

In an ideal world, we would obtain applicant data from schools suspected of using FAFSA lists when allocating financial aid. We could then estimate models predicting financial aid, to test whether aid allocations were significantly lower for students who indicated the school was their preferred choice. Clearly, schools engaging in this behavior will be unwilling to share data for this purpose, and financial aid and applicant data are typically difficult to obtain from most institutions for any research purpose, given the sensitivity of these data. Thus, we turn to the NPSAS, a national dataset of college students, for our analysis.

We face three main issues in determining the appropriate methods for estimating models of financial aid allocation. First, the NPSAS contains data for students from hundreds of different colleges, each of which will be allocating aid in a different fashion. Some schools may place a greater weight on GPA when making aid decisions, while others place a greater weight on need. Some schools have a lot of institutional aid at their disposal to award to prospective students; other schools award very little institutional aid. We use school-level fixed effects to take into account differences between schools.

Second, given that selectivity varies widely across institutions, it seems unlikely that less selective institutions have the ability or motivation to alter financial aid packages based on students' revealed preferences. Less selective schools may not have the resources available to allocate aid strategically, nor would they necessarily have the desire to do so. If only the most selective schools are engaging in this behavior, the effect will only be detected by the interaction between preference and school selectivity. This, in turn, suggests the use of an interaction term between the student-level preference variable and the school-level selectivity variable. Such interactions with group-level variables are possible with fixed effects models, as long as the group-level variable is interacted with an individual-level variable.

Third, the dependent variable, institutional grant aid, is characterized by a distribution with many zeroes (reflecting no institutional grant aid allocated) and some very high values (see Figure 3). Such distributions cannot be analyzed using traditional OLS, and a variety of models have been proposed for these types of data (Buntin and Zaslavsky, 2004; Cameron and Trivedi, 2005; Min and Agresti, 2002). The main distinction between the different methodological approaches is whether the data generating process can be considered a one-step or two-step process. In the current application, the question is whether financial aid offices examine applicants and simply allocate aid based on applicant characteristics, or whether applicants are first sorted into aid/no-aid groups, with the aid group then allocated an amount of aid based on their characteristics. One-step models assume that a single set of covariates has the same effect on both the sorting decision and allocation decision; two-step models can allow the covariates and their coefficients to differ across the two processes. We are agnostic as to whether financial aid offices allocate aid in one or two steps; thus, we estimate models that assume both types of data generating processes.

While Poisson regression have been suggested for these types of expenditure data (Buntin and Zaslavsky, 2004), our single step model is a negative binomial regression, because the variance in institutional grant aid far exceeds the mean. For two-step models, hurdle models are the most popular choice (Buntin and Zaslavsky, 2004; Cameron and Trivedi, 2005). These use two separate models to estimate the probability of a non-zero outcome, and then for units with non-zero outcomes, model the now continuous outcome. Here, we use a linear probability model for the first step and OLS with a logged dependent variable for the second step.

Our dependent variable is the amount of grant aid awarded by the college (in dollars). This aid is under the direct control of the institution, and if colleges are reducing aid for students with a strong preference for their institution, we should see it here.

The main variable of interest is a dummy variable indicating whether the enrolled school was listed as the student's first choice on the FAFSA form. Covariates included as controls are the student's gender, minority status, SAT score (with ACT scores converted using a concordance table), high school GPA, and the student's unmet need (in dollars).

Many colleges in this country are not selective enough to be able to engage in SEM, so it is possible that the negative effects of FAFSA preference may only be occurring at the most selective institutions. To test for this possibility, we include a measure of selectivity included in the NPSAS dataset, and interact this variable with the first choice dummy variable. The variable is measured on a four-point scale, with 1 for very selective institutions, 2 for moderately selective, 3 for minimally selective, and 4 for institutions with open admissions.

Results

Table 2 presents the results of the fixed effects models estimating the effect of preference on institutional grant aid. The top panel presents results for the one-step and two-step models, keeping the effect of listing a school as first choice constant across schools. The effect of listing a school as first choice is not significant in the negative binomial model and the linear model predicting receipt of aid; both coefficients are also zero. The effect on logged grant aid (conditional on aid receipt) is marginally significant ($p < .07$), and suggests a 5% reduction in aid for listing a school in the top slot on the FAFSA listing.

The bottom panel of Table 2 presents models with the school selectivity interaction term. Both terms are statistically significant in the negative binomial regression model, but the

substantive results are somewhat nonsensical. Combining the terms and estimating 95% confidence intervals for each level of selectivity indicates that preference indication has no effect on aid for students at the most selective schools (selectivity variable equal to 1 or 2); it has a positive effect for minimally selective and a larger, positive effect for schools with open admissions. Thus, the estimated effect is the opposite of what critics have suggested is happening.

Neither term is significant in the linear probability model, and calculations indicate that the confidence intervals bracket zero for the coefficient for every value of the selectivity indicator. Neither term is significant in the logged regression model conditional on aid receipt, while the calculations indicate a growing negative effect as schools become less selective. The estimated coefficient for listing a school as the first choice is significantly different from zero only for schools that are moderately selective, and indicates 5.5% penalty in aid receipt.

Discussion

The evidence presented here suggests that at least some students are following the instructions on the FAFSA and are listing schools in order of their preference. Students are much more likely to be enrolled at the school they listed in the first position on the FAFSA form than schools at any other position. Yet our models as whole suggest that students do not seem to be paying a price by providing an implicit preference ranking to schools when they fill out the FAFSA, as some critics have suggested. At most, these students receive institutional grants that are 5% lower than students who do not list an institution as their first choice on the form; this result is contingent on model specification and holds, at most, for only moderately selective schools. To put this percentage in practical terms, if a student would typically receive \$10,000 in institutional grant aid, the results presented here suggest the penalty they pay is around \$500.

Given the intense pressures on admissions at many schools, and limited financial aid budgets, such limited results may seem surprising.

The most likely explanation lies in how we view the school listing on the FAFSA. Critics have implicitly assumed that these rankings are absolutely correct for every student who fills out the FAFSA: colleges can easily find out a student's preference by looking at their FAFSA list. Yet the instructions on the FAFSA instructing students to list their preferred college first are written in tiny print; given the voluminous instructions for filling out the FAFSA, it is possible that many students and their families overlook them. The instruction about state aid is somewhat ambiguous; students in states with few aid programs may assume this instruction does not apply to them. The implication is that these rankings are somewhat noisy, and that colleges cannot assume that they are correct for any particular student.

Instead, colleges that allocate financial aid strategically are likely combining FAFSA list data with other information about students that reveals their preferences. For example, colleges can track student attendance at college fairs, email and telephone contacts with the admission, campus visits, as well as other information, such as application for early admission. These data can collectively demonstrate that a particular student is very interested in the college. For predictive models that forecast matriculation probabilities based on different types of financial aid packages, inclusion of these variables along with the FAFSA preference ranking can only increase the predictive ability of the model.

If the FAFSA preference ranking is only one piece of information that colleges use to reduce aid for "committed" students, then comparing aid packages for students who rank a college first to those that ranked it below first may not yield significant differences. Some of the students who rank a college as below first place because they did not read the FAFSA

instructions carefully may visit a campus more than once, while some of those ranking a college in first place may not engage in any other behaviors that signal interest in the college, besides the application itself.

If true, then this raises the question of why it is so important to prevent the school lists from the FAFSA to the schools themselves. Colleges and universities that wish to discriminate amongst students based on the probability of matriculation already have rich datasets to use for this practice; the FAFSA list data may not add much information to voluminous data they already collect about applicants. More importantly, changing aid allocations based on campus visits and other signals that occur during the application process is relatively uncontroversial within higher education. If this practice is widely accepted, it is not entirely clear why using FAFSA list information in a similar manner is such an undesirable practice.

References

- Buntin, M.B., and Zaslavsky, A.M. (2004). Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures. *Journal of Health Economics*, 23: 525-542.
- Cameron, A.C. and Trivedi, P.K. (2005). *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.
- Hossler, D. and Bontrager, B. (2014). *Handbook of Strategic Enrollment Management*. San Francisco: Jossey-Bass.
- McKinney, L. and Novak, H. (2015). FAFSA filing among first-year college students: Who files on time, who doesn't, and why does it matter? *Research in Higher Education*, 56: 1-28.
- Min, Y. and Agresti, A. (2002). Modeling nonnegative data with clumping at zero: A survey. *Journal of the Iranian Statistical Society*, 1(1-2): 7:33.
- Rivard, R. (2014). FAFSA transparency. *Inside Higher Ed*. Downloaded from <https://www.insidehighered.com/news/2014/10/06/group-asks-federal-government-stop-giving-colleges-information-students-choices>.

Table 1. Enrollment Choice by FAFSA List Position

Number of schools listed on FAFSA	List position of enrolled school, Fall 2011										Did not enroll in any listed schools	N students	
	1	2	3	4	5	6	7	8	9	10			
1	0.97											0.03	1,920
2	0.66	0.31										0.02	621
3	0.56	0.24	0.19									0.02	568
4	0.47	0.25	0.12	0.14								0.02	462
5	0.45	0.19	0.14	0.11	0.11							0.01	367
6	0.41	0.17	0.10	0.10	0.10	0.11						0.02	280
7	0.30	0.16	0.11	0.13	0.08	0.06	0.15					0.01	259
8	0.24	0.15	0.13	0.13	0.06	0.07	0.09	0.10				0.02	230
9	0.20	0.13	0.10	0.10	0.09	0.10	0.10	0.05	0.11			0.01	175
10	0.24	0.12	0.10	0.10	0.05	0.08	0.08	0.07	0.06	0.08		0.03	375
All students	0.64	0.13	0.07	0.05	0.03	0.02	0.02	0.01	0.01	0.01		0.02	5,257

Table 2. Effect of FAFSA Preference on Institutional Grant Aid

	Grant aid		Awarded aid Y/N		Logged grant aid >0	
	B	<i>p</i>	B	<i>p</i>	B	<i>p</i>
1st choice	-0.0023	0.95	0.0064	0.55	-0.0494	0.07
Female	0.0355	0.33	-0.0046	0.65	0.0173	0.53
Minority	-0.2420	0.00	0.0256	0.12	0.0398	0.38
SAT score	0.0001	0.37	0.0002	0.00	0.0007	0.00
HA GPA	0.2467	0.00	0.0839	0.00	0.1950	0.00
Unmet need	0.0000	0.00	0.0000	0.00	0.0000	0.00
Intercept	-3.4467	0.00	-0.0164	0.82	7.1980	0.00
N students		4,734		5,239		3,296
N schools		524		656		557
1st choice	-0.2554	0.00	-0.0067	0.81	-0.0077	0.91
1st choice*selectivity	0.1440	0.00	0.0074	0.60	-0.0238	0.51
Female	0.0351	0.34	-0.0048	0.64	0.0177	0.52
Minority	-0.2390	0.00	0.0257	0.12	0.0396	0.38
SAT score	0.0002	0.09	0.0002	0.00	0.0007	0.00
HA GPA	0.2523	0.00	0.0839	0.00	0.1949	0.00
Unmet need	0.0000	0.00	0.0000	0.00	0.0000	0.00
Intercept	-3.5777	0.00	-0.0161	0.83	7.1983	0.00
N students		4,734		5,239		3,296
N schools		524		656		557

Note: all models include school fixed effects.

Figure 1. Number of Colleges Listed on the 2011-2012 FAFSA

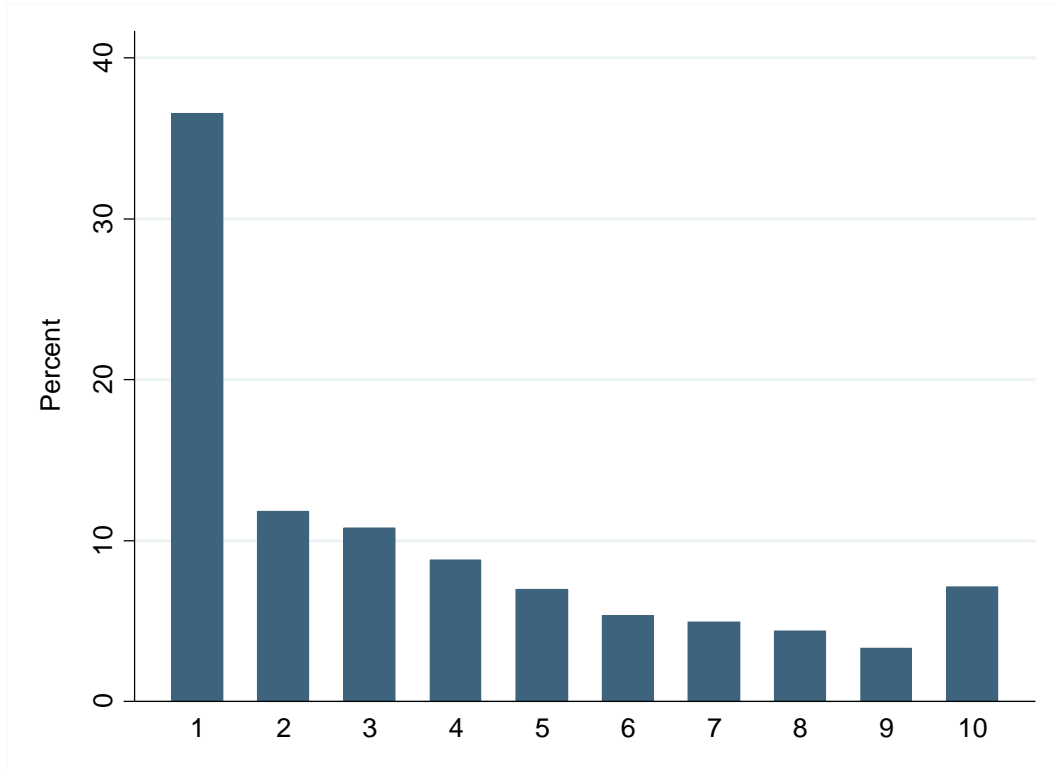


Figure 2. College List Section of the 2011-2012 FAFSA

Step Six (Student): Indicate which colleges you want to receive your FAFSA information.

Enter the six-digit federal school code and your housing plans. You can find the school codes at www.fafsa.gov or you can call 1-800-4-FED-AID. If you cannot get the code, write in the complete name, address, city and state of the college. For state aid, you may wish to list your preferred college first. To have more colleges receive your FAFSA information, read *What is the FAFSA?* on page 10.

	1st FEDERAL SCHOOL CODE	OR	NAME OF COLLEGE ADDRESS AND CITY	STATE	HOUSING PLANS
101.a	<input type="text"/>		<input type="text"/>	<input type="text"/>	101.b on campus <input type="radio"/> 1 with parent <input type="radio"/> 2 off campus <input type="radio"/> 3
101.c	<input type="text"/>		<input type="text"/>	<input type="text"/>	101.d on campus <input type="radio"/> 1 with parent <input type="radio"/> 2 off campus <input type="radio"/> 3
101.e	<input type="text"/>		<input type="text"/>	<input type="text"/>	101.f on campus <input type="radio"/> 1 with parent <input type="radio"/> 2 off campus <input type="radio"/> 3
101.g	<input type="text"/>		<input type="text"/>	<input type="text"/>	101.h on campus <input type="radio"/> 1 with parent <input type="radio"/> 2 off campus <input type="radio"/> 3

Figure 3. Distribution of Institutional Grant Aid

