

Navigating the Financial Aid Process: Borrowing Outcomes among First-Generation and Non-First Generation Students

Fernando Furquim (corresponding author)

University of Michigan
Center for the Study of Higher and Postsecondary Education
610 E. University Ave, Ann Arbor, MI 48109
312-268-4157 | ffurquim@umich.edu

Kristen M. Glasener

glasener@umich.edu

Meghan Oster

omeghan@umich.edu

Brian P. McCall

bpmccall@umich.edu

Stephen L. DesJardins

sdesj@umich.edu

Abstract

A growing number and proportion of students rely on student loans to assist with the costs of postsecondary education. Yet little is known about how first-generation students use federal loans to finance their education. In this article, we examine each of the decisions that culminate in student indebtedness: the decision to apply for aid, whether to borrow, and how much to borrow. We find significant differences by generational status at each step of the student borrowing process. First-generation students are more likely to apply for financial aid, borrow, and take out larger loans than their peers, after controlling for a rich set of covariates for costs and financial resources. We find that student characteristics cannot fully explain these observed differences in borrowing outcomes across generations.

Keywords: student loan; financial aid; first-generation students; hurdle model; decomposition

Citation: Furquim, F., Glasener, K. M., Oster, M., McCall, B. P., & DesJardins, S. L. (2017). Navigating the financial aid process: Borrowing outcomes among first-generation and non-first-generation students. *The ANNALS of the American Academy of Political and Social Science*, 671(1), 69-91.

NOTE:

We thank Julie Posselt, John Cheslock, Nick Hillman, and Laura Perna for comments on earlier drafts.

This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

Author Bios

Fernando Furquim is a doctoral student and research assistant at University of Michigan's Center for the Study of Higher and Postsecondary Education. His research interests include the regulatory and accountability systems in higher education, and adult education.

Kristen M. Glasener is a doctoral student and research assistant at University of Michigan's Center for the Study of Higher and Postsecondary Education. She studies higher education access and stratification, selective college admissions and enrollment management, and financial aid policy.

Meghan Oster is a doctoral student and research assistant at University of Michigan's Center for the Study of Higher and Postsecondary Education. She studies higher education finance and policy.

Brian McCall is a professor of higher education, economics, and public policy at the University of Michigan and is coeditor of Economics of Education Review. His research in higher education focuses on issues in college choice, persistence, and graduation.

Stephen DesJardins is a professor in the Center for the Study of Higher and Postsecondary Education, professor in the Ford School of Public Policy, and evaluation faculty lead in CEDER, an evaluation unit, all located at the University of Michigan. He has published widely in education and economics journals.

In recent decades, reliance on student loans to assist with the costs of postsecondary education has become the norm (Addo, Houle, and Simon 2016). In 2014, approximately two-thirds of graduating seniors took out at least one student loan, with the average graduate owing \$28,950 (Cochrane and Reed 2013). Students who entered repayment of undergraduate or graduate loans in 2014 (including those who did not earn a credential) carried a median debt obligation of \$19,647 (Looney and Yannelis 2015). The growth in prevalence and amount of borrowing by students stem in part from the rising cost of college, which has outpaced inflation rates and median family income growth for over a decade (College Board 2011). In addition, changes in federal financial aid policy have led to a shift away from a grants-based aid system to one based primarily on student loans (Hearn and Holdsworth 2004). In this context, students and families turn to a number of federal, state, and private sources of lending for education (Schmeiser, Stoddard, and Urban 2015); in so doing, they face an increasingly complex set of decisions regarding whether and how much to borrow for college.

Fundamental questions regarding the usage of student loans remain unanswered (Hillman 2015). One open question concerns the frequency and intensity of borrowing across student subpopulations (Avery and Turner 2012; Hillman 2015). In particular, we know relatively little about the borrowing behavior of first-generation (FG) students (those whose parents do not have a college degree), even though these students make up a growing proportion of the college population (Pappano 2015). Many scholars have explored first-generation students' college outcomes (e.g., Bui 2002; Chen and Carroll 2005; Ishitani 2003; Nunez, Cuccaro-Alamin, and Carroll 1998), but fewer researchers have examined differences in loan usage by students' parental education level. Because earnings are positively associated with educational attainment (Goldin and Katz 2007), children of highly educated parents are more likely to have access to the

financial resources required for a college education. First-generation students, on the other hand, often face numerous challenges in pursuit of postsecondary education (Bui 2002; Gofen 2009). These students tend to be at a disadvantage with respect to knowledge about higher education, level of family income and support, and academic preparation (Pascarella et al. 2004). In addition, FG students have lower enrollment rates and lower rates of retention and graduation compared to their non-first-generation peers (Ishitani 2003, 2006; Nunez, Cuccaro-Alamin, Carroll 1998). Differences between FG students and their peers regarding the expected value of their degrees, levels of family resources, and familiarity with the financial aid system may result in variations in student borrowing behavior (Avery and Turner 2012).

In this article we explore differences in borrowing behavior across students' generational status. To do so, we analyze differences at various stages in the process of financing college expenses using detailed information on more than 100,000 students in a large public university system. We begin by modeling financial aid application decisions, followed by the decision to borrow, and, conditional on borrowing, the decision of how much to borrow. This approach allows us to illuminate the sources of disparities in borrowing outcomes. Prior research demonstrates that FG students borrow more than their non-first-generation peers, net of controls for costs and financial resources (Addo, Houle, and Simon 2016; Hart and Mustafa 2008). Modeling each decision along the path to borrowing, we find that FG students are more likely than their peers to apply for financial aid, more likely to borrow, and more likely to take on larger loans. In addition to using models with indicator variables for FG students, we estimate separate models for the two populations and use an Oaxaca decomposition. We identify several differences across generational status. Unlike their peers, FG students borrow more as their expected family contribution rises. When receiving other forms of aid such as grants, FG

students do not reduce their borrowing as much as their peers. Finally, each additional dollar in tuition expenses is associated with a higher increase in borrowing for FG students than for their non-FG peers. These differences hold net of a robust set of controls for student characteristics, expenses, aid, and family resources. Student characteristics do not fully explain differences in borrowing outcomes. In other words, our findings suggest that FG students behave differently with respect to student loans. We provide estimates of how these differences in behavior affect student debt, and suggest a few possible explanations for them, but there is still much work to do to understand the mechanisms behind these differences.

Background

FG students differ from their peers on their path to and through college. Chen and Carroll (2005) found that parental education is positively associated with the probability of attending any college, and Toutkoushian, Stolberg, and Slaton (2015) concluded that FG students were significantly less likely than non-FG students to enroll in a postsecondary institution. When FG students do enroll in college, they are more likely to attend part time (Chen and Carroll 2005) and enroll at institutions with lower retention and graduation rates, such as two-year or for-profit institutions (Choy 2001; Nunez, Cuccaro-Alamin, Carroll 1998). Once in college, FG students are more likely to enroll in remedial classes, delay the declaration of a major, earn lower GPAs, and withdraw or repeat classes, resulting in slower accumulation of credits over time (Chen and Carroll 2005). FG students also persist and graduate from college at lower rates than their peers (Ishitani 2003, 2006; Nunez, Cuccaro-Alamin, Carroll 1998). Some researchers suggest that lower levels of perceived faculty support or student engagement (Longwell-Grice and Longwell-Grice 2008), lack of integration and socialization (Pascarella et al. 2004), and inadequate or insufficient advising (Swecker, Fifolt, and Searby, 2013) help to explain these differences in

educational outcomes for FG students. These differences in educational outcomes are important for the study of student borrowing, given the high default rates for noncompleters (Dynarski 2015) and the need for a degree to experience the benefits of higher education (Baum, Ma, and Payea 2013).

Federal student loans

Student loans are one of the financial aid options students and their families consider when paying for college. Students often rely on a combination of savings, parental contributions, employment income, work-study, scholarships, grants, and federal and private loans to help finance their education (Schmeiser, Stoddard, and Urban 2015). The Free Application for Federal Student Aid (FAFSA) governs students' access to many forms of need-based aid, serving as a gateway to billions of dollars in federal grants and loans. In addition to federal funds, states and institutions also commonly require completion of the FAFSA for students to qualify for financial aid programs. The FAFSA collects information about family income, assets, and household size, among other financial information, to estimate financial need and eligibility for means-tested aid programs (Dynarski and Scott-Clayton 2006). As such, the FAFSA is a critical first step in obtaining the financial aid that many students need to attend college (McKinney and Novak 2015).

Despite its important role, many eligible students do not complete the FAFSA, particularly low-income students, who are typically FG students (King 2004). By failing to complete the FAFSA, individuals face higher college costs by forgoing potentially significant amounts of financial aid (King 2004; Kofoed 2017), or they fail to enroll in college altogether (Bettinger et al. 2012). Researchers argue that the current financial aid system is needlessly convoluted, noting the complexity of the FAFSA (Dynarski and Wiederspan 2012) and lack of

clarity regarding aid eligibility (Avery and Kane 2004; Dynarski and Scott-Clayton 2006), both of which disadvantage lower-income and FG students. Recently, Kofoed (2017) found that nearly 20 percent of eligible students who attended college did not complete the FAFSA, most of whom came from families earning less than \$50,000 annually, suggesting that filing the FAFSA is a critical barrier to financial assistance for many disadvantaged students.

Eligibility for need-based grants and federal loans depends on completion of the FAFSA and calculations of a student's expected family contribution (EFC). The federal student loan system comprises several different programs with varying interest rates, repayment rules, and eligibility requirements, including direct subsidized and unsubsidized Stafford loans, and parent PLUS Loans (Dynarski and Scott-Clayton 2006).

Literature Review

Many researchers have explored the mechanisms leading to differences in students' borrowing behavior, pointing to unequal access to college-related information as one potential source of variation in loan take-up rates. Given the complexity of financial aid and nuances of the various student loan programs, knowledge of the financial aid process plays an important role in terms of students' ability to navigate the process (Goldrick-Rab et al. 2016). Students whose families are familiar with college-going have greater access to the resources that promote college enrollment, including knowledge about the financial aid process (McDonough 1997). Conversely, parents who lack personal experience with college tend to have lower levels of awareness and understanding about college prices and financial aid (Hossler, Schmidt, and Bouse 1991). As a result, students who are first in their families to attend college may behave differently than their peers with respect to borrowing and debt accumulation (Goldrick-Rab et al. 2016).

Few studies have focused on the usage of student loans by FG students, and studies that

do examine borrowing by generational status provide inconsistent findings. Gladieux and Perna (2005), for example, found that lower levels of parental education are associated with higher borrowing, but their study used aggregated enrollment data and bivariate analysis only.

Conversely, in interviews with high school students and their parents, Perna (2008) observed a pattern of uncertainty about student loans among parents who did not attend college. Taking into account the role of the high school context in students' decision making, she argued that students in low-resourced schools receive conflicting information about student loans from parents, counselors, and teachers, which may affect students' perceptions of the utility of loans.

Some scholars contend that differences in families' financial resources, knowledge, attitudes, beliefs, and dispositions may lead to "loan aversion" among subpopulations. Loan averse individuals may avoid borrowing by reducing their enrollment intensity, working while in school, or attending lower-cost institutions (see Kelchen and Goldrick-Rab 2013, for a summary and critique of this literature). In interviews with students, Burdman (2005) found that FG students were more likely to work full time and avoid borrowing than their peers. Christou and Halisassos (2006) used a nationally representative survey and found no association between FG status and borrowing, but a higher propensity to work among FG students. Others argue that loan aversion may manifest itself via the decision not to enroll in college, attend lower-cost institutions, or attend college part time (Johnson 2013). Although the loan aversion literature has attracted considerable attention, additional work is necessary to understand its underlying causes. As Kelchen and Goldrick-Rab (2013) stated, "Loan aversion is often described as common, unfortunate, and not easily overcome, but many of these depictions are typically based on conjecture rather than evidence" (p. 11). Much of the evidence for loan aversion comes from other countries that do not reflect the U.S. context, and in many instances do not draw sharp

distinctions between loan aversion and differences in expected outcomes or risk preferences (Johnson 2013).

For college enrollees, the evidence suggests that FG students, most of whom come from lower-income families, borrow more frequently and at higher rates than their peers. Hart and Mustafa (2008) modeled student borrowing as a function of various demographic and financial characteristics, including parental educational attainment. They found that, compared to their peers, FG students borrow an additional \$500 in the first year of college. In a review of the evidence regarding student debt, Hillman (2015) found that students from the bottom half of the income distribution (where FG students are concentrated) borrow more frequently (more than 40 percent, compared to 25 to 30 percent for wealthier students), and, conditional on persistence, incur greater levels of debt. Javine (2013) found that FG college graduates are 8 percent more likely to owe at least \$10,000 in student loans, and 6 percent more likely to owe at least \$20,000 than their peers.

Persistent differences in borrowing by generational status are puzzling since most studies control for some form of family income. Addo, Houle, and Simon (2016) analyzed the role of parental wealth and race on student loan debt. Their analysis indicated that students whose parents have college degrees borrow 55 percent less than FG students. They found that disparities in parental net wealth and financial contributions to college expenses disadvantage students from underrepresented groups and FG students, partly explaining differences in borrowing outcomes. Parents who went to college are able to provide greater monetary support for their children (Grodsky and Jones 2007), and are more likely to incur debt on behalf of their children through parent PLUS loans or other financing vehicles (Cha, Weagley, and Reynolds 2005). Conditional on income, parents with a college education are also more likely to report

planning to help children pay for college, engage in some form of financial preparation to do so, and open college savings accounts (Hillman, Gast, and George-Jackson 2015).

Research on student loans for FG students has two important limitations. First, authors address nonborrowers in inconsistent ways. Hart and Mustafa (2008), for example, used Tobit regression as a way to deal with the bounded nature of student loan data (i.e., students cannot borrow less than \$0). However, Tobit regression implicitly assumes that a single mechanism governs the generation of zero and positive values of the dependent variable (Lin and Schmidt 1984). Tobit regression also assumes a linear relationship between all covariates and both the decision to borrow (to move from \$0 to a positive borrowing amount) and how much is borrowed (the amount above \$0; Lin and Schmidt 1984). Conceptually, students might not borrow for a number of reasons: they may fail to apply for aid, they might receive sufficient financial aid in other forms, or they might enroll in fewer courses to minimize costs. In each example, a student could have “zero” debt but with very different underlying mechanisms explaining why they did not borrow.

In addition to affecting the probability of borrowing, these same three factors (filing FAFSA, having sufficient aid, or taking fewer courses) may also affect the amount that students borrow. This distinction goes unobserved in Tobit regressions since all “zeros” would be censored from the analysis. Scholars have accounted for nonborrowers via exclusion (Addo, Houle, and Simon 2016) or by dichotomizing borrowing levels, such as comparing those with balances less than \$10,000 (including non-borrowers) to those with debt greater than \$10,000 (Javine 2013). Neither approach, however, provides insight into the series of decisions that students face when considering how to pay for college, such as whether to apply for financial aid, whether to take out loans, which loans to use, and how much to borrow. A final limitation of

prior research is that many authors examine student debt upon graduation (e.g., Addo, Houle, and Simon 2016; Javine 2013), without accounting for differences in college persistence rates among subpopulations. Given the evidence on higher attrition for FG students (Ishitani 2006), it is possible for these students to accrue less debt merely as a function of lower persistence rates, which is unobserved in studies focused on graduating students.

Theoretical Framework

We draw on basic tenets of demand theory to explain students' borrowing choices. Participating in higher education is costly. Typically, individuals forgo wages while enrolled in college and face significant direct costs in the form of tuition, living expenses, and other materials (Li 2013). Conditional on enrolling in college, students meet these expenses with some mixture of out-of-pocket funds, parental or family transfers, and sources of need- and merit-based financial aid (Schmeiser, Stoddard, and Urban 2015). Basic economic principles of demand theory suggest that students minimize costs, first by maximizing funds that need not be paid back (e.g., grants, scholarships), and then using the lowest-cost available bundle of funds from earnings, transfers, and student loans.

Under this framework, several factors may give rise to heterogeneity in borrowing outcomes. Means-tested grant programs such as Pell, need-based institutional grants, and merit aid create variation in the costs that students face, even under the same published tuition price (Kelchen and Goldrick-Rab 2013). Families may also be differentially able or willing to fund students' education. Addo, Houle, and Simon (2016) found that parental net worth, income, and level of education are all associated with larger parental contributions toward college expenses. Students may also differ in their access to various financing options. Researchers found evidence that more affluent families rely on housing equity to send their children to college and to higher

cost institutions (Lovenheim 2011; Lovenheim and Reynolds 2013). Because FG students are typically of lower socioeconomic status, it is plausible that their families have limited access to such forms of credit and, therefore, face greater constraints on their borrowing choices. This may result in greater utilization of federal student loan programs, which relax such constraints (Mankiw 1986). Students also have some control over costs through various mechanisms, such as choosing how many credits they take.

We express students' borrowing decisions as:

$$y_i = f(\theta_i C, \gamma_i S_i) \quad (1)$$

where y_i is the amount borrowed by the i th student as a function of costs associated with college enrollment and θ_i is a scalar measuring the proportion of direct costs (C) the student is unable to meet either out of pocket or with nonloan forms of financial aid (akin to the concept of “unmet need”). The product of these two measures captures the potential need for loans. The $\gamma_i S_i$ term is a student-specific scalar that reflects an individual's latent preferences for borrowing, risk, and so on. Using this framework, we hypothesize that higher levels of family income and greater amounts of nonloan forms of financial aid will reduce reliance on student loans, while costs will be positively associated with the probability of borrowing and amount borrowed.

Empirical Strategy

The dataset for this analysis is a repeated cross-section of the population of incoming new and transfer students at a large Midwestern public university system, for fall terms from 2002 to 2014.¹ Students enroll in one of four campuses, which comprise a flagship Research I institution; two regional campuses in distinct parts of the state; and a liberal arts college. Our primary focus is on the first year of college, so all variables used in the analysis are defined for that time period.

There are a total of 130,891 students in the initial sample; of those, 102,687 completed the FAFSA and 68,212 borrowed a positive amount in the first year.

Table 1 presents descriptive statistics of variables of interest for FG and non-FG students in our dataset (N=130,891). About 34 percent of the students are categorized as FG. FG status is defined based on parents' educational attainment; students whose parents do not hold a four-year degree are categorized as FG. We observe several differences between these groups with respect to demographic characteristics, financial resources, aid received, and borrowing. FG students are older on average and compose a larger proportion of underrepresented groups (e.g., black and Hispanic) compared to their peers. They are also more likely to have in-state residency and file taxes as independents. The expected family contribution of FG students is almost half that of students with college-educated parents (\$10,220 vs. \$19,550), reflecting the lower resources of FG students and their families. Nearly 35 percent of FG students are Pell grant recipients, compared to 15 percent of non-FG students. Unconditional mean borrowing among FG students is significantly higher than their peers (\$5,290 vs. \$4,043).

TABLE 1

Selected Descriptive Statistics for First-Year Students

[Insert table 1 here]

NOTE: Statistics reflect only students starting in fall terms between 2003 and 2014.

SOURCE: Authors' calculations based on administrative student records.

Analysis

We model the decision to apply for financial aid, the decision to borrow, and the amount borrowed as three separate steps to reflect each of the decisions that students make regarding borrowing for college.

This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

FAFSA filing. We model the dichotomous decision of submitting the FAFSA (1=completed FAFSA; 0=otherwise) as a function of student background characteristics (e.g., gender, ethnicity, age, FG status) and items collected in surveys administered during standardized college entrance exams. The survey measures capture student expectations for postsecondary attainment; plans for seeking out financial aid; where to live during college and whether to work; and parental income. To minimize unexplained variance and isolate the variable of interest, we include fixed effects at the campus, year, and campus-by-year levels. Respectively, these account for time-invariant differences among campuses (e.g., differences in mission, financial aid practices); time-varying changes affecting all campuses equally (e.g., changes to state appropriations); and unobserved characteristics of each campus that vary over time (e.g., the cost of living in particular locations). The main variable of interest is the coefficient associated with FG status. We estimate this model using logistic regression, where the probability of student i applying is:

$$\Pr(\text{Apply}_i = 1 | X'_i \beta_1) = \frac{\exp(X'_i \beta_1 + \varepsilon_i)}{1 + \exp(X'_i \beta_1 + \varepsilon_i)} \quad (2)$$

Borrowing outcomes. Students who submit the FAFSA must subsequently decide whether and how much to borrow. As a result of these separate decisions, students' loan balances are censored at zero. Most prior research has accounted for this truncation by using Tobit regression (e.g., Hart and Mustafa 2008), or by excluding nonborrowers (e.g., Addo, Houle, and Simon 2016). The former approach, Tobit regression, imposes assumptions about the relationship between dependent and independent variables. Specifically, the Tobit assumes that any variable that increases the probability of (in this case) borrowing also increases the average amount borrowed, and the increases are linear for both outcomes (Lin and Schmidt 1984). The

latter approach, exclusion of nonborrowers, ignores the active borrowing decision that individuals face.

To overcome these limitations, we use the double-hurdle model introduced by Cragg (1971). The double-hurdle model allows us to estimate separate parameters for the decision to borrow and the (conditional) amount borrowed. This approach accounts for the simultaneous participation and consumption decisions that students make when considering student loans, and treats zeroes as genuine outcomes reflecting a decision not to participate (Jones 1992). The double-hurdle technique is a less restrictive approach than the Tobit, allowing for different underlying mechanisms to affect the decision to borrow and the amount borrowed, and permitting the computation of the probability of borrowing and the conditional and unconditional amounts borrowed. This method has been employed widely in studies of consumption (Jones 1992; Lin and Schmidt 1984), loan default (Moffatt 2005), and medical outcomes (Hu, Pavlicova, and Nunes 2011), though its application in education research remains relatively uncommon. Cha and Weagley (2002) and Cha, Weagley, and Reynolds (2005) used the double-hurdle model to investigate the use of loans to finance postsecondary education; however, their analyses focused on college completers only, did not account for parental education, and did not consider the decision to apply for financial aid that precedes the double-hurdle. The double-hurdle we employ is estimated using a probit specification for the decision to borrow and a linear truncated regression for the amount borrowed.

Decision to borrow:

$$Borrow_i^* = X'_{bi}\beta_1 + \gamma_c + \delta_t + \varphi_{ct} + u_i \quad (3)$$

Level of borrowing:

$$(Amount)_i^* = X'_{ai}\beta_2 + \gamma_c + \delta_t + \varphi_{ct} + r_i \quad (4)$$

This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

$$Amount_i = \begin{cases} Amount_i^* & \text{if } Borrow_i^* > 0 \text{ and } Amount_i^* > 0 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where X'_{bi} and X'_{ai} are vectors of explanatory variables for the i th student used for the decision to borrow and amount borrowed, respectively; γ_c , δ_t , and φ_{ct} are campus, year, and campus-by-year fixed effects; β_1 and β_2 are vectors of coefficients; and u_i and r_i are the student-level error terms, clustered at the high school level to account for possible unobserved shared characteristics of students within schools. The standard double-hurdle approach does not account for the possible correlation of u_i and r_i . It is plausible that among those individuals for whom we underestimate the probability of borrowing we would also underestimate the amount borrowed. For example, some students may complete the FAFSA solely to gain access to student loans, borrowing with probability of 1. We control for the correlation of errors across the two equations by incorporating the inverse Mills ratio into the second equation as suggested by Jones (1992).

The explanatory variables included in the model are the same for both hurdles, and derive from our theoretical framework and from rules governing financial aid practices. We control for background characteristics, including age, gender, race/ethnicity, and transfer status. Net tuition charges and students' residency status (in-state, out-of-state, in tuition reciprocity agreement states, and states participating in the Midwest Student Exchange program) reflect the costs that students face. All else being equal, students facing greater costs are expected to borrow at higher rates. To control for financial resources available to students, we capture tax filing status (dependent or independent) and expected family contribution (EFC), which is mainly determined by adjusted gross income and household size. The EFC reflects both the resources students may have access to, and the funds that the institution assumes the student will be able to provide to cover costs. Filing status and EFC are also used to define eligibility for various types of aid, as well as the loan limits individuals qualify for. Finally, we control for whether and how much

financial aid students receive, including grants, scholarships, and work-study awards. These components of a financial aid package reflect need and/or merit, and (unlike loans) need not be paid back. We expect that students receiving higher amounts of nonloan aid would borrow less, all else equal. The main coefficient of interest is the indicator variable for first-generation status.

Decomposition analysis

The models specified above use an indicator variable for FG status to identify (conditional) average differences in FAFSA filing, the probability of borrowing, and the amount borrowed. To more fully understand the drivers of these differences, we estimate separate models for the FG and non-FG student populations. This approach allows each individual parameter in the logistic and double-hurdle models to vary by generational status. We then perform an Oaxaca decomposition of the expected value for each outcome (probability of filing FAFSA and expected total amount borrowed). The decomposition apportions differences in the outcome of interest into two sources: an “explained” portion associated with differences in population characteristics and a “residual” component associated with differences in model parameters. A common application of the Oaxaca decomposition analysis is the study of wage disparities by gender or race (e.g., Oaxaca and Ransom 1994; Shatnawi, Oaxaca, and Ransom 2014). In higher education research, the Oaxaca decomposition has been used to understand bachelor’s degree attainment differences between students who begin their postsecondary education at four-year institutions and students who begin at two-year institutions (Sandy, Gonzalez, and Hilmer 2006), the underrepresentation of women in elite higher education institutions (Bielby et al. 2014), and racial bias in standardized tests (Duncan and Sandy 2013).

There are several approaches to decompositions. We use a nonlinear pooled decomposition for the model of FAFSA filing. For the double-hurdle, we decompose the

unconditional expected amount borrowed, which is the product of the probability of borrowing and the conditional amount borrowed.

Limitations

This study has a number of limitations. First, the analysis is correlational, not causal, and meant only to explore baseline differences across generational status. We do not formally test possible mechanisms that could account for the differences between FG and non-FG students' observed borrowing outcomes. In addition, the use of data from a single university system raises concerns about external validity and may limit the scope of inferences from this analysis. For example, private universities that are typically higher cost may also provide greater discounts and institutional aid, reducing differences by generational status. For-profit providers, which serve a high number of FG students, typically provide little institutional aid, resulting in greater borrowing for students. In those settings, students may make different choices regarding FAFSA filing or borrowing. Finally, this study considers a single state context; findings may not be generalizable to other states because of differences in higher education governance structure, tuition and financial aid policies, and student composition.

There is also an issue of endogeneity between enrollment and financial aid. As found in DesJardins, Ahlburg, and McCall (2006), students behave strategically with respect to financial aid. Enrollment choices are informed by expectations about financial aid offers, and the relative generosity of financial aid packages across institutions (Kim, DesJardins, and McCall 2009). Because our sample includes only enrolled students, we cannot observe FG student applicants who choose to enroll elsewhere or not to attend college at all, or the financial aid offers (and potential loans) that may have influenced such decisions. A related issue is that students'

borrowing decisions are constrained by the financial aid offer letter, which places limits on the loans an individual student can take.

Finally, the functional forms in our models carry several assumptions. The double-hurdle assumes normality of errors, for example, and that the errors are uncorrelated across the two equations. We correct for the latter, and cluster the errors, but alternative functional forms could impose less restrictive assumptions on errors. We also impose linear relationships for many of the explanatory variables and borrowing. We tested alternative specifications that included quadratic terms for age, EFC, and aid, but it may be valuable to explore more flexible parametrizations of student loan outcomes in future research.

Results

FAFSA filing

Our initial analysis focuses on students' decision to file the FAFSA. Table 2 reports the results (as odds ratio and average marginal effects) for FG status across three model specifications. Model 1, which includes only controls for generational status, indicates that FG students have 2.7 times the odds of filing a FAFSA as their peers, translating to a probability that is 15 percentage points higher than non-FG students. Controlling for other demographic characteristics, measured academic performance in high school, educational expectations, and a number of measures of students' financial plans and resources, we find that FG students have 1.6 times greater odds of filing the FAFSA (an AME of 4 percentage points) than their peers (Model 3).

TABLE 2

Logistic Regression Results and Average Marginal Effects for FAFSA Filing (n=75,784)

[Insert table 2 here]

NOTE: Standard errors in parentheses clustered at the high school level. Model 1 includes indicator variable only.

This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

Model 2 adds gender, race/ethnicity, age, and measures of academic performance in high school. Model 3 adds measures of postsecondary expectations, plans for living arrangements, work, and applying for aid, and family income. All models include institution, year, and institution x year fixed effects. Full model results are found in Appendix Table A1.

*** $p < 0.001$.

Two parameters that appear to account for much of the unexplained variance are students' reported FAFSA filing plans and family income in 11th grade. As depicted in Figure 1, FG students are more likely than non-FG students to file a FAFSA regardless of reported intent to apply for financial aid. Figure 2 shows that, for all family income brackets above \$36,000, FG students are also more likely to file for financial aid (all other variables held at observed values).

FIGURE 1

Predicted Probability of FAFSA Filing by First-Generation Status and Reported Intention of Filing

FIGURE 2

Predicted Probability of FAFSA Filing by First-Generation Status and Family Income

[Insert figures 1 and 2 here]

Using a nonlinear pooled decomposition, we find that student characteristics account for 68 percent of the differences in FAFSA filing observed across generational status, with 32 percent of the difference owing to the effects (parameters) of the determinants of FAFSA filing. As completion of the FAFSA is a necessary precondition for taking out federal student loans, a greater proportion of FG students are at risk of borrowing. Selection into the FAFSA accounts for some of the observed unconditional differences in borrowing outcomes by generation status that are found in Table 1.

There are also important analytical implications for differences in FAFSA completion. Research about the effects of different forms of financial aid is conditioned on certain pieces of

This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

information, such as the EFC, that are calculated from FAFSA. The nonrandom selection of students into FAFSA filing thus has implications for what data researchers observe. Students and families have discretion over where to submit their FAFSA, selecting into eligibility for student loans differently at different institutions. For example, it is plausible that at a high-cost private college, more relatively well-off students would have submitted the FAFSA.

Borrowing outcomes

Decision to borrow. Students who complete the FAFSA may receive financial aid offers that include a number of loan types (e.g., subsidized loans, unsubsidized loan, parent plus loans). These students face a decision of whether to borrow, and thus constitute the sample for the double-hurdle models that we estimate. Table 3 reports the coefficients for FG status, both as the probit score and an average marginal effect, for the loan take-up decision (the probability of borrowing a positive amount). Results indicate that, compared to their peers, the average marginal effect for FG students is around 5 percentage points ($p < 0.001$), controlling for expected family contribution, tuition costs, and all other forms of aid.

TABLE 3

Double-hurdle Regression Results for Decision to Borrow (n=101,157)

[Insert table 3 here]

NOTE: Standard errors in parentheses clustered at the high school level. Model 1 includes indicator variable only. Model 2 adds gender, race/ethnicity, and age. Model 3 adds expected family contribution and tax filing status. Model 4 includes tuition expenses and grant, scholarship, and work-study aid. All models include institution, year, and institution x year fixed effects. Full model results are found in Appendix Table A2.

*** $p < 0.001$.

To better understand differences in aggregated borrowing, we investigated differences in student debt by type of loan (results available by request). FG students are less likely to rely on parent PLUS loans, with probabilities about one-half of a percentage point lower than their non-FG peers ($p < 0.1$). Parent PLUS loans require that borrowers have favorable credit histories or

This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

that they provide an endorser who qualifies for the loan (Federal Student Aid n.d.). These loans represent a repayment obligation for parents or guardians, not students. Additional research into differential access to parent PLUS loans and its implications for other forms of borrowing seems warranted. In particular, it is important to investigate how differences in parental knowledge of the PLUS program, willingness of parents to take on such loans, and PLUS loan eligibility rules affect loan take-up and students' ability to minimize direct borrowing through use of the parent PLUS loans.

For other types of loans (subsidized, unsubsidized, and state), we find small and typically insignificant differences in loan take-up and amounts by generational status. The higher prevalence of borrowing and higher balances among FG students is in large part a result of a greater likelihood of taking up at least one, and typically multiple, loan programs. This pattern of loan take-up has implications for the portfolio of loans that different students carry, and the repayment terms and rules that they face after enrollment.

As part of the decomposition analysis, we estimated separate models for FG and non-FG students (reported in the appendix). We identify significant differences by generational status for several parameters. For example, though EFC is negatively associated with the probability of borrowing for non-FG students, the relationship is essentially flat for FG students. As a result, we observe that FG students from higher-resourced families borrow more frequently, all else held constant. We discuss this finding in greater detail in the next section.

Amount borrowed. Students who decide to borrow must also choose the total amount of loans they take up. The coefficients in Table 4 reflect the association between FG status and the amount borrowed in the first year among borrowers. For comparison, we include results from a Tobit specification of the same model. The Tobit regression consistently estimates larger

differences than the double-hurdle, suggesting the importance of taking into account differences in the first stage (decision to borrow). The double-hurdle model makes it evident that there is a higher probability of borrowing (by about 5 percentage points, Table 3), and (among borrowers), \$574 larger loans for FG students than for their peers. As was the case for the probability of borrowing, we find that these differences hold only for aggregated borrowing. By type of loan, generational status is weakly or essentially uncorrelated with borrowing amounts.

TABLE 4

Tobit and Double-hurdle Regression Results for Amount Borrowed (n=101,157)

[Insert Table 4 here]

Notes: Standard errors in parentheses clustered at the high school level. Model 1 includes indicator variable only. Model 2 adds gender, race/ethnicity, and age. Model 3 adds expected family contribution and tax filing status. Model 4 includes tuition expenses and grant, scholarship, and work-study aid. All models include institution, year, and institution x year fixed effects. Full model results are found in Appendix Table A2.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Turning to the separate models estimated for each of the two populations, we find that a few parameters seem to drive much of the borrowing differences across student groups. The coefficients for EFC, tuition charges, and nonloan aid all have different relationships to borrowing by generational status. For EFC, we find that, for values up to \$20,000, FG students borrow more as EFC increases. For non-FG students, on the other hand, EFC is consistently negative in its association with loan amounts. This difference is visualized in Figure 3. Similarly, we find that tuition is positively associated with borrowing for all students, but the relationship is stronger for FG students, who borrow \$50 more than their peers for each \$1,000 increase in tuition (Figure 4). Finally, we find differences in the relative decrease in borrowing associated

with increases in other forms of aid. For non-FG students, a \$1,000 increase in aid is correlated with a \$575 decrease in borrowing, compared to \$480 for FG students (Figure 5).

FIGURE 3

Predicted Conditional Borrowing Amount by First-Generation Status and Expected Family Contribution

FIGURE 4

Predicted Conditional Borrowing Amount by First-Generation Status and Tuition Expenses

FIGURE 5

Predicted Conditional Borrowing Amount by First-Generation Status and Non-Loan Aid

[Insert figures 3 to 5 here]

The decomposition analysis shows that differences in model parameters account for more than 100 percent of the observed variation in expected borrowing (158 percent), implying that, if FG students “behaved,” or had the same coefficients, as non-FG students, they would actually borrow more than they do.

Conclusion

As tuition rises and more students rely on an increasingly complex financial aid system to support their postsecondary education, scholars, administrators, and policy-makers must understand the financial choices that students face. Student decisions to borrow are now an essential part of students’ choices about college. Borrowing decisions depend on a number of factors, including financial aid knowledge and family resources. As suggested by this study, we

This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

find that students' borrowing behaviors significantly differ across subpopulations. FG students borrow more frequently and in greater amounts than their peers, even when we control for a rich set of covariates including ability to pay, costs, and all other aid. While this finding is consistent with prior research (e.g., Hart and Mustafa 2008; Javine 2013), we gain new insights by examining the multiple decisions required to take out student loans through investigating FAFSA filing patterns, using a double-hurdle approach, and conducting decomposition analyses.

Higher borrowing among FG students is largely driven by an increased likelihood of: 1) applying for aid, 2) borrowing, and 3) taking on multiple types of loans. However, there are relatively small differences in borrowing amounts when disaggregated by loan type. In addition, by conducting separate analyses by generational status, we identify a stronger positive relationship between tuition and borrowing for FG students and a stronger negative association between aid and borrowing for non-FG students. We also find higher levels of borrowing for FG students with relatively high EFC compared to their peers. Prior research has found that parents without a college degree are less likely to plan to pay for college, prepare for the financial costs of college, or save for college (Hillman, Gast, and George-Jackson 2015). Our results suggest that one possible consequence of this pattern is that FG students are less able to cover their EFC, leading to higher levels of student debt compared to their non-FG peers. Researchers and financial aid practitioners should consider to what extent families are aware of, willing, and able to meet their expected financial contribution, and the consequences to students of over-estimated EFCs that are unmet by parental transfers.

Future research could benefit from the application of causal methods that uncover the mechanisms explaining differences in borrowing amounts by generational status. One potential mechanism is differential knowledge and access to other forms of credit that could help to offset

the amount borrowed by students. For example, Lovenheim (2011) demonstrates that many families use housing equity to cover college costs, and that housing equity is associated with enrollment in pricier, more selective institutions and an increased likelihood of graduation. Differences in access to various forms of credit may result in greater reliance on federal and state loans for FG students or could constrain families' ability to meet their EFC.

Students with college-educated parents may have a distinct advantage over FG students in accessing knowledge about the financial aid process and financing college because they have an immediate family member with experiential knowledge of college. Additionally, FG students are at the nexus of many of the subgroups most likely to need financial support, because of the association of FG status with lower socioeconomic standing and membership in marginalized groups. Research is necessary to understand the relationships between student loans, the intersection of these various characteristics, and how information interventions could play a role in improving financial aid knowledge of FG students.

Our findings also have practical implications. First, financial aid practitioners need better tools to improve forecasts of the burden that the EFC represents to different kinds of families, inform families lacking experience with the postsecondary sector on the EFC, and understand the contingencies that students face should family financial support be lacking. It may also be possible to structure financial aid packages in ways that minimize the need for loans for FG students. For example, we find differences in the association between types of aid and student loan amounts. Work-study awards have the weakest association with student borrowing (each dollar in work-study is associated with \$0.18 reduction in borrowing); each dollar in grants is associated with a \$0.62 reduction in borrowing, suggesting that work-study awards may not be the best mechanism for reducing students' loan burdens. Scholarship funds, which may combine

both need-based and merit criteria, are associated with a \$0.81 reduction in amount borrowed. Financial aid package composition could present an opportunity to reduce the debt burden of students by substituting away from particular types of aid that do not seem to reduce borrowing (such as work-study) in favor of increasing grants. Any increases in state allocations for financial aid could also be targeted at decreasing student debt. It is also important for policy-makers and practitioners to consider that any such changes could have consequences for each of the intertwined decisions outlined here, affecting students' choices regarding application for aid, whether to borrow, and how much to borrow.

Finally, it is important to note that our analysis is concerned exclusively with the first year of college. The differences we identify here may understate the debt burden of FG students who persist in college, as evidence suggests that some forms of financial aid are associated with increased likelihood of persistence and graduation (e.g., DesJardins, Ahlburg, and McCall 2006; Dynarski 2003). The packaging of financial aid, then, has implications both for student indebtedness and student outcomes, and presents an opportunity for longitudinal research into packaging strategies that simultaneously minimize indebtedness and maximize student success, especially in an era of rising costs and shrinking financial support for higher education.

References

- Addo, Fenaba R., Jason N. Houle, and Daniel Simon. 2016. Young, black, and (still) in the red: Parental wealth, race, and student loan debt. *Race and Social Problems* 8 (1): 64–76.
- Avery, Christopher, and Thomas J. Kane. 2004. Student perceptions of college opportunities. The Boston COACH program. In *College choices: The economics of where to go, when to go, and how to pay for it*, ed. Caroline Hoxby, 355–94. Chicago, IL: University of Chicago Press.
- Avery, Christopher, and Sarah Turner. 2012. Student loans: Do college students borrow too much—or not enough? *The Journal of Economic Perspectives* 26 (1): 165–92.
- Baum, Sandy, Jennifer Ma, and Kathleen Payea. 2013. Education pays 2013. Washington, DC: The College Board.
- Bettinger, Eric P., Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu. 2012. The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment. *The Quarterly Journal of Economics* 127 (3): 1205–1242.
- Bielby, Rob, Julie Renee Posselt, Ozan Jaquette, and Michael N. Bastedo. 2014. Why are women underrepresented in elite colleges and universities? A non-linear decomposition analysis. *Research in Higher Education* 55 (8): 735–60.
- Bui, Khanh Van T. 2002. First-generation college students at a four-year university: Background characteristics, reasons for pursuing higher education, and first-year experiences. *College Student Journal* 36 (1): 3–12.
- Burdman, Pamela. 2005. *The student debt dilemma: Debt aversion as a barrier to college access*. Berkeley, CA: Center for Studies in Higher Education.
- Cha, Kyung-Wook, and Robert O. Weagley. 2002. Higher education borrowing. *Journal of*

Financial Counseling and Planning 13 (1): 61-74.

Cha, Kyung-Wook, Robert O. Weagley, and Laura Reynolds. 2005. Parental borrowing for dependent children's higher education. *Journal of Family and Economics* 26 (3): 299–321.

Chen, Xianglei, and C. Dennis Carroll. 2005. *First-generation students in postsecondary education: A look at their college transcripts*. Postsecondary Education Descriptive Analysis Report. NCES 2005-171. Washington, DC: National Center for Education Statistics.

Choy, Susan. 2001. *Students whose parents did not go to college: Postsecondary access, persistence, and attainment*. Findings from the Condition of Education, 2001. Washington, DC: National Center for Education Statistics.

Christou, Costas, and Michael Haliassos. 2006. How do students finance human capital accumulation? The choice between borrowing and work. *Journal of Policy Modeling* 28 (1): 39–51.

Cochrane, Debbie and Michael Reed. 2014. *Student debt and the class of 2014*. Washington, DC: The Institute for College Access and Success: Project on Student Debt.

College Board. 2011. *Trends in college pricing 2011*. Washington, DC: The College Board.

Cragg, John G. 1971. Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica: Journal of the Econometric Society* 39 (5): 829–44.

DesJardins, Stephen L., Dennis A. Ahlburg, and Brian P. McCall. 2006. An integrated model of application, admission, enrollment, and financial aid. *Journal of Higher Education* 77 (3): 381–429.

- Duncan, Kevin, and Jonathan Sandy. 2013. Using the Blinder-Oaxaca decomposition method to measure racial bias in achievement tests. *The Review of Black Political Economy* 40 (2): 185–206.
- Dynarski, Susan M. 2003. Does aid matter? Measuring the effect of student aid on college attendance and completion. *American Economic Review* 93 (1): 279–88.
- Dynarski, Susan. 31 August 2015. Why students with smallest debts have the larger problem. *New York Times*.
- Dynarski, Susan M., and Judith E. Scott-Clayton. 2006. The cost of complexity in federal student aid: Lessons from optimal tax theory and behavioral economics. National Bureau of Economic Research, No. w12227. Cambridge, MA.
- Dynarski, Susan, and Mark Wiederspan. 2012. Student aid simplification: Looking back and looking ahead. National Bureau of Economic Research, No. w17834. Cambridge, MA.
- Federal Student Aid. n.d. PLUS loans are federal loans that graduate or professional students and parents of dependent undergraduate students can use to help pay for college or career school. Available from <https://studentaid.ed.gov> (accessed 10 August 2016).
- Gladieux, Lawrence, and Laura Perna. 2005. Borrowers who drop out: A neglected aspect of the college student loan trend. National Center Report# 05-2. Washington, DC: National Center for Public Policy and Higher Education.
- Gofen, Anat. 2009. Family capital: How first-generation higher education students break the intergenerational cycle. *Family Relations* 58 (1): 104–20.
- Goldin, Claudia, and Lawrence F. Katz. 2007. The race between education and technology: The evolution of U.S. educational wage differentials, 1890 to 2005. National Bureau of Economic Research. No. w12984, Cambridge, MA.

- Goldrick-Rab, Sara, Robert Kelchen, Douglas N. Harris, and James Benson. 2016. Reducing income inequality in educational attainment: Experimental evidence on the impact of financial aid on college completion. *American Journal of Sociology* 121 (6): 1762–1817.
- Grodsky, Eric, and Melanie T. Jones. 2007. Real and imagined barriers to college entry: Perceptions of cost. *Social Science Research* 36 (2): 745–66.
- Hart, Natala K., and Shoumi Mustafa. 2008. What determines the amount students borrow? Revisiting the crisis–convenience debate. *Journal of Student Financial Aid* 38 (1): 17-39.
- Hearn, James C., and Janet M. Holdsworth. 2004. Federal student aid: The shift from grants to loans. In *Public funding of higher education: Changing contexts and new rationales*, eds. Edward P. St. John and Michael D. Parsons, 40-59. Baltimore, MD: Johns Hopkins University Press.
- Hillman, Nicholas W. 2015. Borrowing and repaying student loans. *Journal of Student Financial Aid* 45 (3): 35-48.
- Hillman, Nicholas W., Melanie J. Gast, and Casey George-Jackson. 2015. When to begin? Socioeconomic and racial/ethnic differences in financial planning, preparing, and saving for college. *Teachers College Record* 117:1–28.
- Hossler, Don, Jack Schmidt, and Gary Bouse. 1991. Family knowledge of postsecondary costs and financial aid. *Journal of Student Financial Aid* 21 (1): 4-17.
- Hu, Mei-Chen, Martina Pavlicova, and Edward V. Nunes. 2011. Zero-inflated and hurdle models of count data with extra zeros: Examples from an HIV-risk reduction intervention trial. *The American Journal of Drug and Alcohol Abuse* 37 (5): 367–75.
- Ishitani, Terry T. 2003. A longitudinal approach to assessing attrition behavior among first-generation students: Time-varying effects of pre-college characteristics. *Research in*

- Higher Education* 44 (4): 433–49.
- Ishitani, Terry T. 2006. Studying attrition and degree completion behavior among first-generation college students in the United States. *Journal of Higher Education* 77 (5): 861–85.
- Javine, Victoria. 2013. Financial knowledge and student loan usage in college students. *Financial Services Review* 22 (4): 367-87.
- Johnson, Matthew T. 2013. Borrowing constraints, college enrollment, and delayed entry. *Journal of Labor Economics* 31 (4): 669–725.
- Jones, Andrew M. 1992. A note on computation of the double-hurdle model with dependence with an application to tobacco expenditure. *Bulletin of Economic Research* 44 (1): 67–74.
- Kelchen, Robert, and Sara Goldrick-Rab. 2013. *Accelerating college knowledge: Examining the feasibility of a targeted early commitment Pell grant program*. IRP Discussion Paper No. 1405-13. Madison, WI: Institute for Research on Poverty.
- Kim, Jiyun, Stephen L. DesJardins, and Brian P. McCall. 2009. Exploring the effects of student expectations about financial aid on postsecondary choice: A focus on income and racial/ethnic differences. *Research in Higher Education* 50 (8): 741–74.
- King, Jacqueline E. 2004. *Missed opportunities: Students who do not apply for financial aid*. Washington, DC: American Council on Education.
- Kofoed, Michael S. 2017. To apply or not to apply: FAFSA completion and financial aid gaps. *Research in Higher Education* 58 (1): 1–39.
- Li, Wenli. 2013. The economics of student loan borrowing and repayment. *Business Review* 3:1–10.
- Lin, Tsai-Fen, and Peter Schmidt. 1984. A test of the Tobit specification against an alternative

- suggested by Cragg. *The Review of Economics and Statistics* 66 (1): 174–77.
- Longwell-Grice, Rob, and Hope Longwell-Grice. 2008. Testing Tinto: How do retention theories work for first-generation, working-class students? *Journal of College Student Retention: Research, Theory & Practice* 9 (4): 407–20.
- Looney, Adam, and Constantine Yannelis. 2015. A Crisis in student loans? How changes in the characteristics of borrowers and in the institutions they attended contributed to rising loan defaults. *Brookings Papers on Economic Activity* 2015 (2): 1–89.
- Lovenheim, Michael F. 2011. The effect of liquid housing wealth on college enrollment. *Journal of Labor Economics* 29 (4): 741–71.
- Lovenheim, Michael F., and C. Lockwood Reynolds. 2013. The effect of housing wealth on college choice: Evidence from the housing boom. *Journal of Human Resources* 48 (1): 1–35.
- Mankiw, N. Gregory. 1986. The allocation of credit and financial collapse. *Quarterly Journal of Economics* 101 (3): 455–70.
- McDonough, Patricia M. 1997. *Choosing colleges: How social class and schools structure opportunity*. Albany, NY: SUNY Press.
- McKinney, Lyle, and Heather Novak. 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): 1–28.
- Moffatt, Peter G. 2005. Hurdle models of loan default. *Journal of the Operational Research Society* 56 (9): 1063–1071.
- Nunez, Anne-Marie, Stephanie Cuccaro-Alamin, and D. Dennis Carroll. 1998. First-generation students: Undergraduates whose parents never enrolled in postsecondary education.

Statistical Analysis Report. Postsecondary Education Descriptive Analysis Reports.

Washington, DC: National Center for Education Statistics.

Oaxaca, Ronald L., and Michael R. Ransom. 1994. On discrimination and the decomposition of wage differentials. *Journal of Econometrics* 61 (1): 5–21.

Pappano, L. 8 April 2015. First generation students unite. *New York Times*.

Pascarella, Ernest T., Christopher T. Pierson, Gregory C. Wolniak, and Patrick T. Terenzini.

2004. First-generation college students: Additional evidence on college experiences and outcomes. *Journal of Higher Education* 75 (3): 249–84.

Perna, Laura W. 2008. Understanding high school students' willingness to borrow to pay college prices. *Research in Higher Education* 49 (7): 589–606.

Sandy, Jonathan, Arturo Gonzalez, and Michael J. Hilmer. 2006. Alternative paths to college completion: Effect of attending a 2-year school on the probability of completing a 4-year degree. *Economics of Education Review* 25 (5): 463–71.

Schmeiser, Maximilian D., Christiana Stoddard, and Carly Urban. 2015. Does salient financial information affect academic performance and borrowing behavior among college students? Federal Reserve Board Finance and Economic Discussion Series, Working Paper No. 2015-075, Washington, DC.

Shatnawi, Dina, Ronald Oaxaca, and Michael Ransom. 2014. Movin' on up: Hierarchical occupational segmentation and gender wage gaps. *The Journal of Economic Inequality* 12 (3): 315–38.

Swecker, Hadyn K., Matthew Fifolt, and Linda Searby. 2013. Academic advising and first-generation college students: A quantitative study on student retention. *NACADA Journal* 33 (1): 46–53.

Toutkoushian, Robert K., Robert S. Stolberg, and Kelly A. Slaton. 2015. Talking 'bout my generation: Defining "first-generation students" in higher education research. Paper presented at the Association for the Study of Higher Education, November. Denver, CO.

Table 1:
Selected Descriptive Statistics for First-Year Students

Variable	First-generation students			Non-first-generation students		
	N	Mean	Median	N	Mean	Median
White	34,753	0.739	–	89,179	0.873	
FAFSA filed	36,484	0.872		94,407	0.751	
EFC (\$)	31,791	\$10,220	\$5,639	70,858	\$19,550	\$13,121
Received aid (%)	36,484	0.748		94,407	0.623	
Aid \$	36,484	\$5,483	\$3,101	94,407	\$3,275	\$994
Borrowed (%)	36,484	0.625		94,407	0.481	
Borrowed \$	36,484	\$5,290	\$3,900	94,407	\$4,043	\$0

Notes: Statistics reflect only students starting in fall terms between 2003 and 2014.

Source: Authors' calculations based on administrative student records.

Table 2:
Logistic Regression Results and Average Marginal Effects for FAFSA Filing (n=75,784)

Variable	Model 1 (first-gen only)	Model 2 (+ demographics & academic)	Model 3 (+ plans & income)
First-generation status	2.702 ^{***} (0.113)	2.595 ^{***} (0.100)	1.578 ^{***} (0.062)
AME (at observed values)	0.146	0.136	0.043

Notes: Standard errors in parentheses clustered at the high school level.

Model 1 includes indicator variable only.

Model 2 adds gender, race/ethnicity, age, and measures of academic performance in high school.

Model 3 adds measures of postsecondary expectations, plans for living arrangements, work, and applying for aid, and family income.

All models include institution, year, and institution*year fixed effects.

Full model results are found in Appendix Table A.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3:
Double-hurdle Regression Results for Decision to Borrow (n=101,157)

Variables	Model 1 (first-gen only)	Model 2 (+ demographics)	Model 3 (+ resources)	Model 4 (+ tuition & financial aid)
<i>Decision to borrow (Probit)</i>				
First-generation status	0.259*** (0.015)	0.246*** (0.022)	0.123*** (0.009)	0.151*** (0.010)
AME (at observed values)	0.093	0.088	0.042	0.049

Notes: Standard errors in parentheses clustered at the high school level.

Model 1 includes indicator variable only.

Model 2 adds gender, race/ethnicity, and age.

Model 3 adds expected family contribution and tax filing status.

Model 4 includes tuition expenses and grant, scholarship, and work-study aid.

All models include institution, year, and institution*year fixed effects.

Full model results are found in Appendix Table B.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 4:
Tobit and Double-hurdle Regression Results for Amount Borrowed (n=101,157)

Variables	Model 1 (first-gen only)	Model 2 (+ demographics)	Model 3 (+ resources)	Model 4 (+ tuition & financial aid)
<i>Tobit</i>				
First-generation status	1031.728*** (101.806)	1223.939*** (106.473)	1173.424*** (169.869)	899.413*** (101.645)
<i>Double-hurdle (amount borrowed)</i>				
First-generation status	-616.335*** (125.662)	-472.221*** (68.981)	356.161** (122.618)	574.153*** (113.881)

Notes: Standard errors in parentheses clustered at the high school level.

Model 1 includes indicator variable only.

Model 2 adds gender, race/ethnicity, and age.

Model 3 adds expected family contribution and tax filing status.

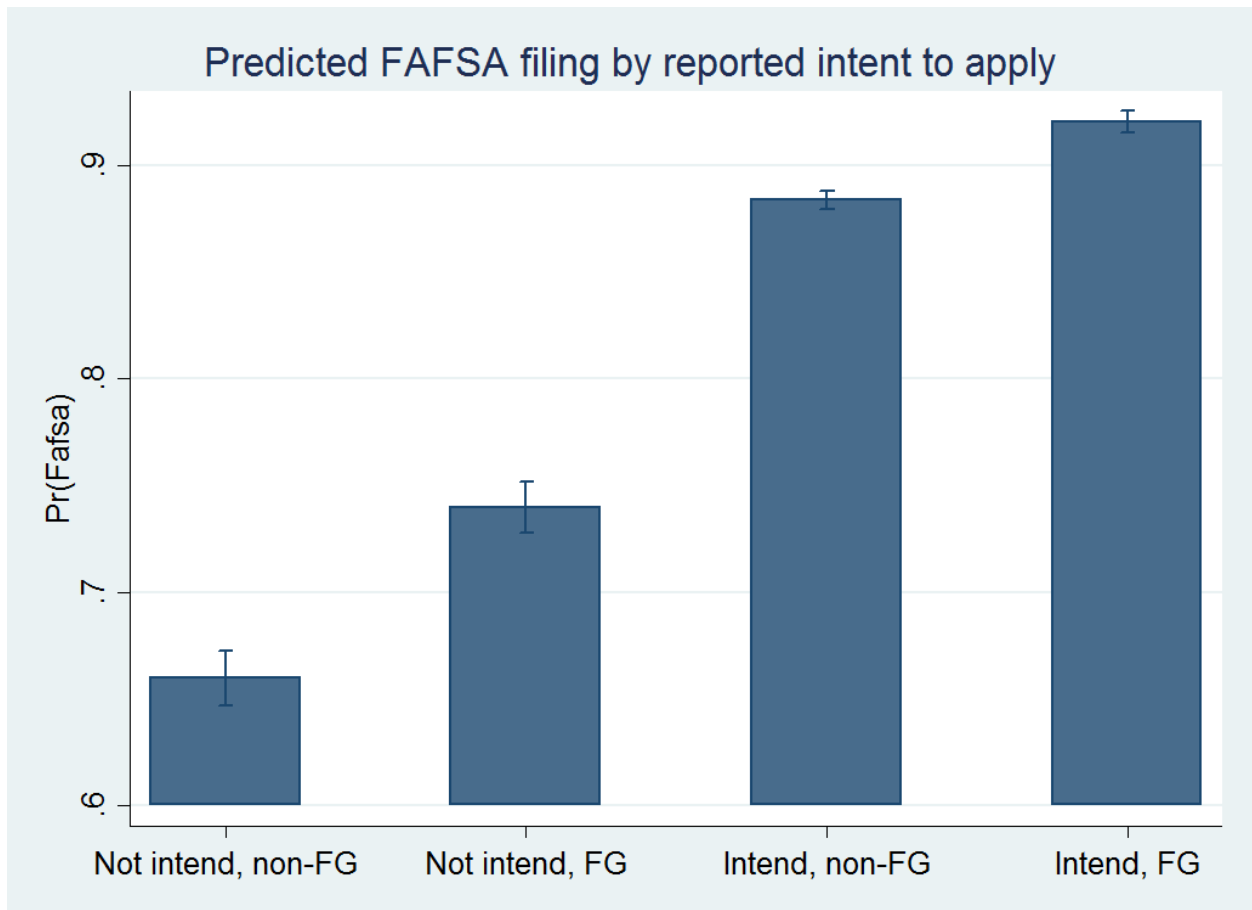
Model 4 includes tuition expenses and grant, scholarship, and work-study aid.

All models include institution, year, and institution*year fixed effects.

Full model results are found in Appendix Table B.

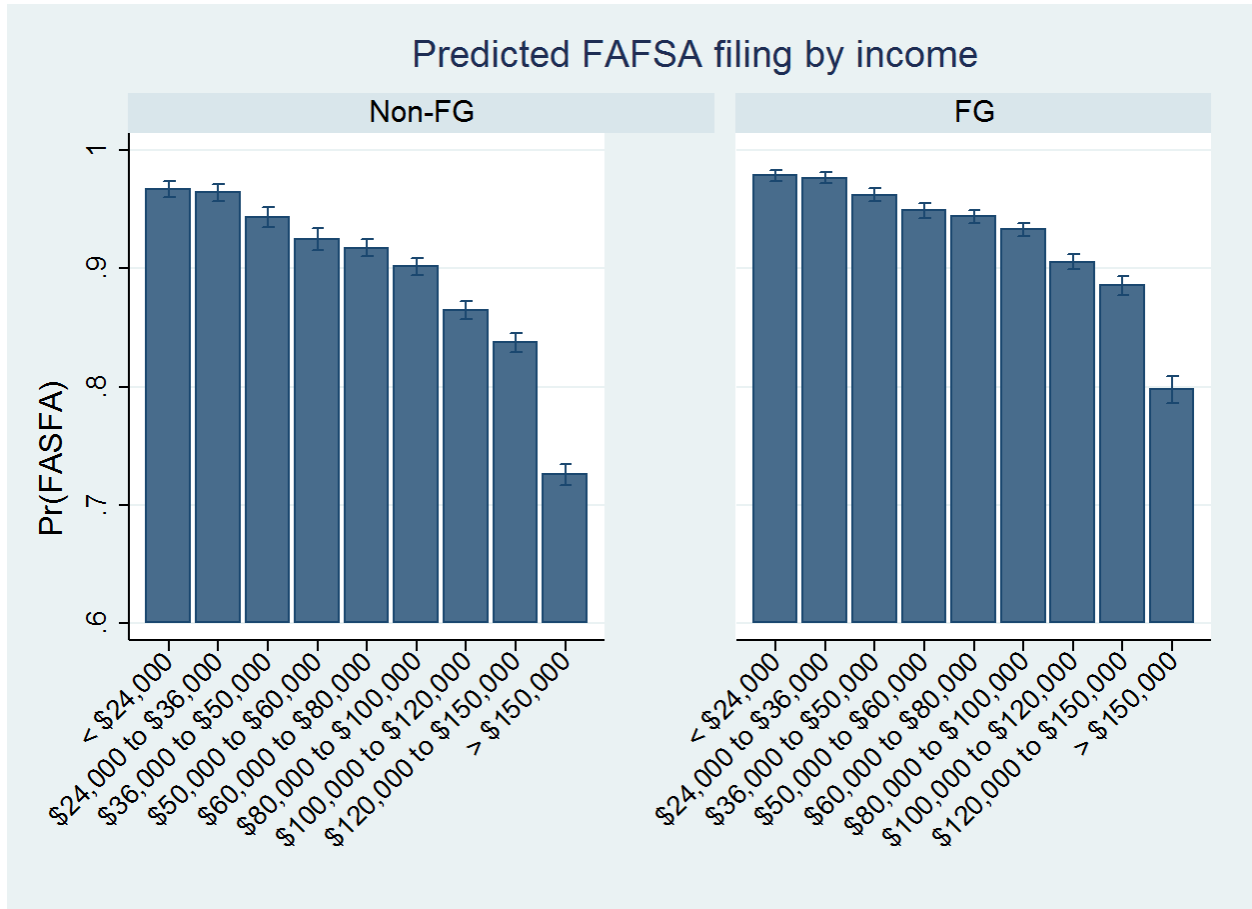
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure 1:
Predicted Probability of FAFSA Filing by First-Generation Status and Reported Intention of Filing



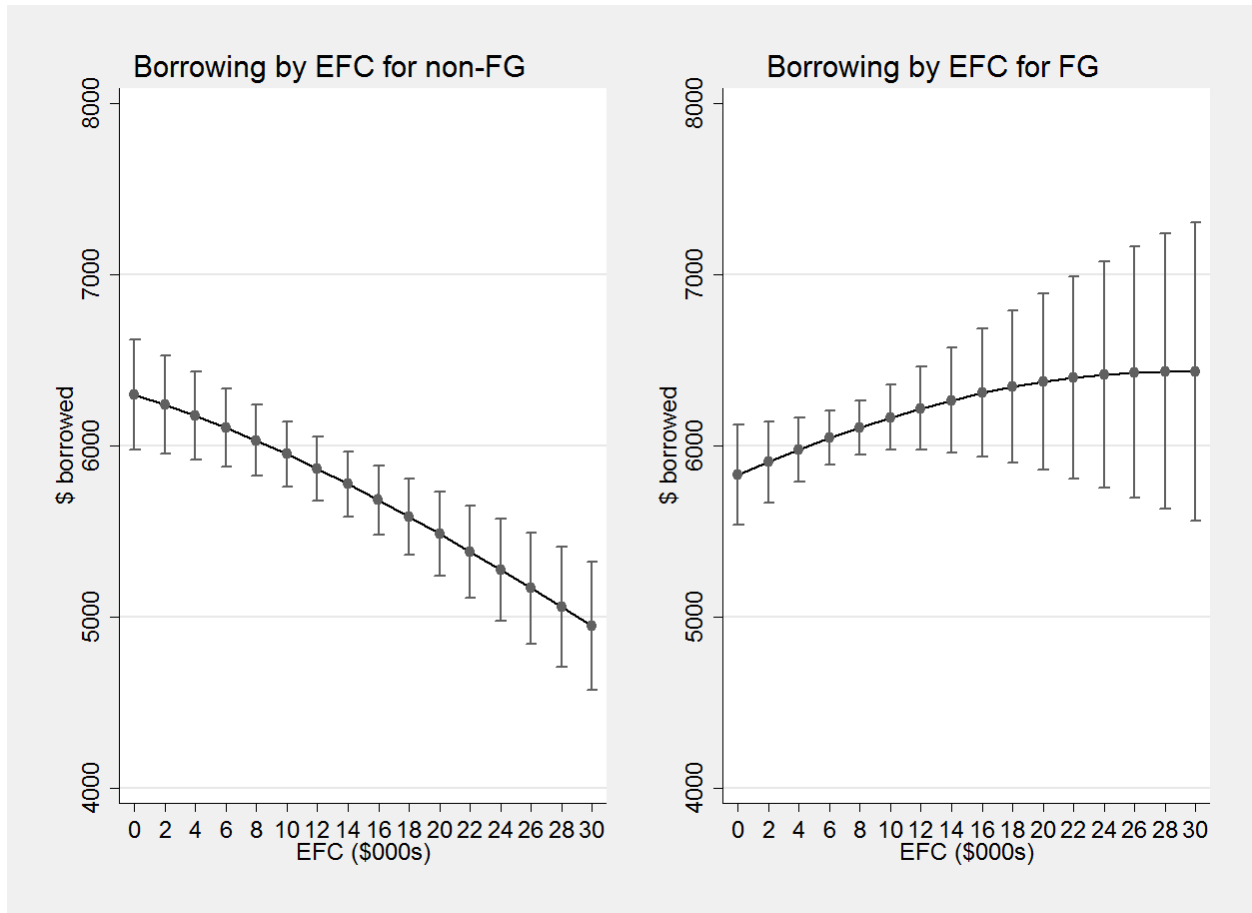
This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

Figure 2:
Predicted Probability of FAFSA Filing by First-Generation Status and Family Income



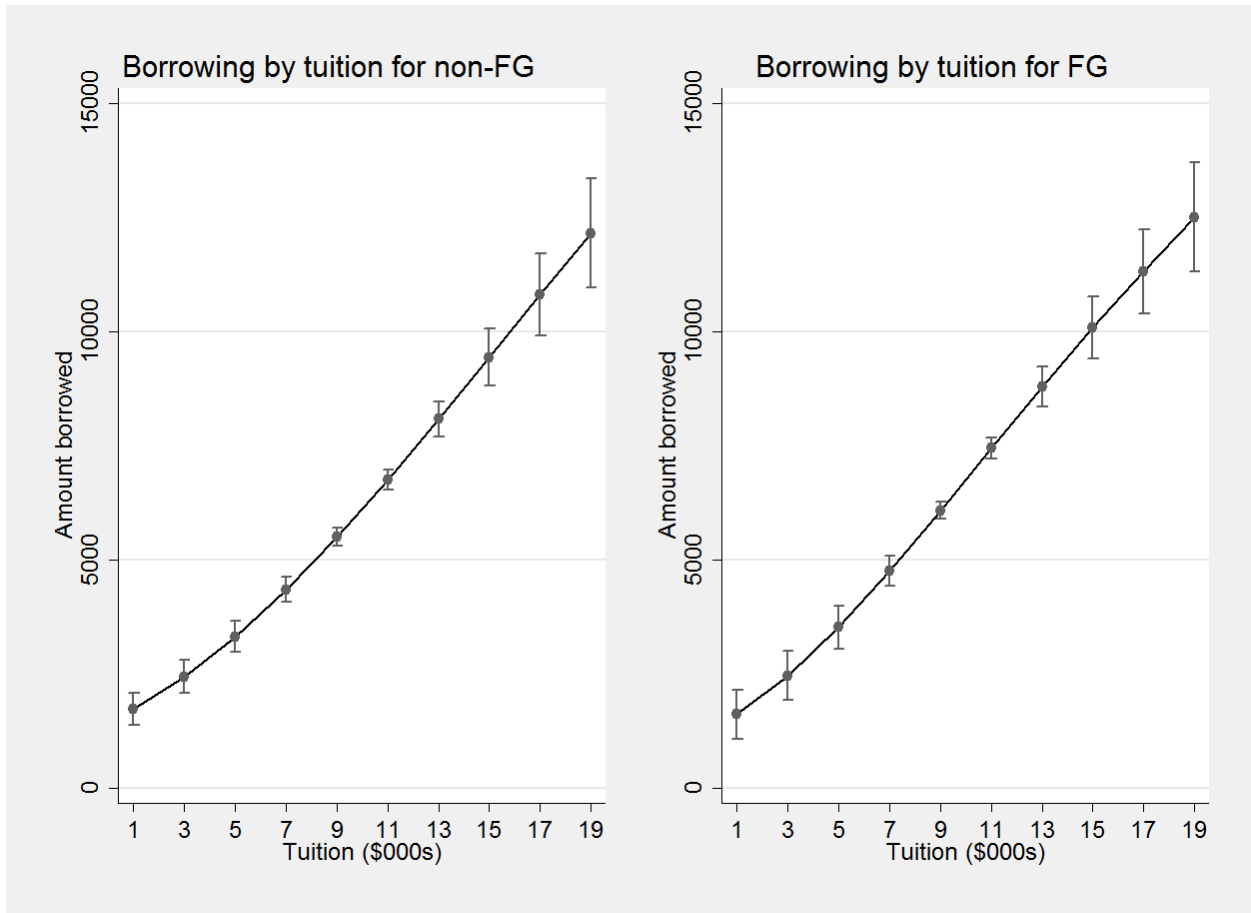
This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

**Figure 3:
Predicted Conditional Borrowing Amount by First-Generation Status and Expected Family Contribution.**



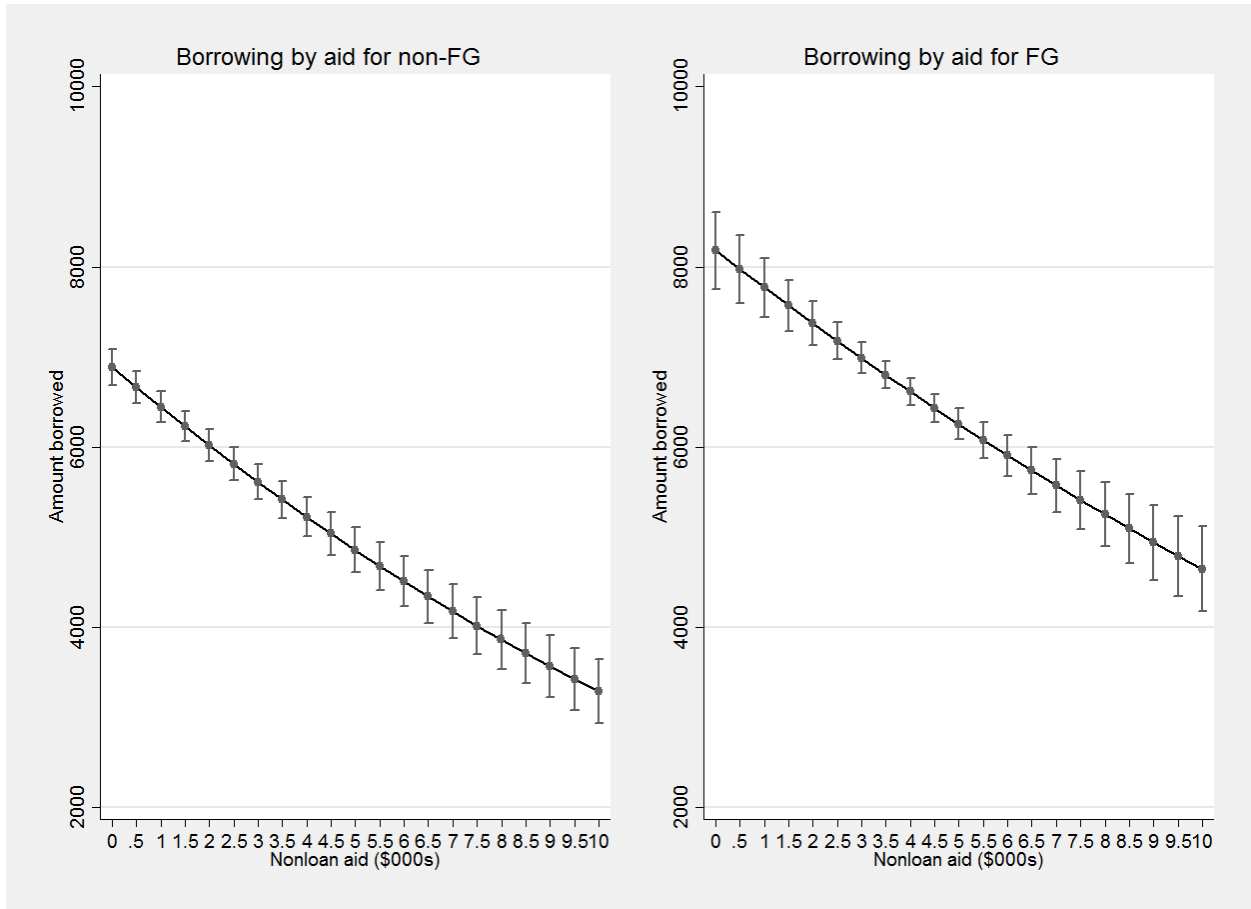
This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

Figure 4:
Predicted Conditional Borrowing Amount by First-Generation Status and Tuition Expenses



This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

Figure 5:
Predicted Conditional Borrowing Amount by First-Generation Status and Non-Loan Aid



This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

Appendix

Table A
Logistic Regression results for FAFSA filing by generational status

Variables	First-Generation	Non-First-Generation
Transfer	0.578*** (0.066)	0.564*** (0.040)
Gender (ref. male)		
Female	1.141* (0.070)	1.116*** (0.032)
Race (ref. White)		
American Indian	0.880 (0.218)	1.347* (0.179)
Asian	0.864 (0.113)	0.846* (0.068)
Black	3.661*** (1.176)	1.712*** (0.247)
Hispanic	0.771 (0.147)	0.981 (0.094)
Age	0.829 (0.287)	0.339*** (0.103)
Age ²	1.004 (0.008)	1.026*** (0.007)
High School GPA (ref. GPA>3.5)		
HS GPA <2.0	0.716 (0.338)	0.833 (0.223)
HS GPA 2.0-3.0	0.705** (0.091)	0.823** (0.053)
HS GPA 3.0-3.5	0.781*** (0.054)	0.823*** (0.025)
ACT Composite Percentile	0.997 (0.002)	0.999 (0.001)
Expect to apply for financial aid	4.751*** (0.369)	4.871*** (0.153)
Expect to work	1.634*** (0.106)	1.532*** (0.043)
Family income (ref. less than \$24,000)		
about \$24,000 to \$36,000	0.546* (0.155)	1.254 (0.242)
about \$36,000 to \$50,000	0.415*** (0.107)	0.643** (0.100)
about \$50,000 to \$60,000	0.301*** (0.073)	0.459*** (0.072)
about \$60,000 to \$80,000	0.250*** (0.061)	0.428*** (0.060)
about \$80,000 to \$100,000	0.246*** (0.056)	0.324*** (0.045)
about \$100,000 to \$120,000	0.150***	0.222***

This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

	(0.033)	(0.031)
about \$120,000 to \$150,000	0.125***	0.173***
	(0.028)	(0.023)
more than \$150,000	0.049***	0.082***
	(0.011)	(0.011)
Expected living arrangement in college (ref. college dorm)		
Own home	1.164	1.010
	(0.113)	(0.051)
Parents' or relatives' home	1.374**	0.952
	(0.150)	(0.051)
Married student housing	1.393	1.171
	(0.652)	(0.375)
Fraternity or sorority	1.861*	0.918
	(0.559)	(0.106)
Expected level of education (ref. Associate's or below)		
Bachelor's degree	1.023	0.965
	(0.331)	(0.203)
1-2 years of graduate school	1.016	0.908
	(0.321)	(0.193)
Doctoral or professional	1.115	0.925
	(0.366)	(0.195)
Other	0.890	1.041
	(0.352)	(0.249)
Term FE	YES	YES
Institution FE	YES	YES
Term X Institution FE	YES	YES
N	20,295	55,414

Note: Exponentiated coefficients; Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B
Regression results for FAFSA filing, expressed as odds ratios (n=75,784)

Variables	Model 1 (first-gen only)	Model 2 (+ demographics & academic)	Model 3 (+ expectations & plans)
First Gen	2.702*** (0.113)	2.595*** (0.100)	1.578*** (0.062)
Transfer		0.590*** (0.031)	0.567*** (0.035)
Gender (ref. male)			
Female		1.219*** (0.029)	1.120*** (0.029)
Race (ref. White)			
American Indian		1.546*** (0.164)	1.234 (0.143)
Asian		1.363*** (0.118)	0.862* (0.063)
Black		3.710*** (0.481)	1.955*** (0.262)
Hispanic		1.289** (0.101)	0.929 (0.076)
Age		0.631* (0.140)	0.463** (0.115)
Age ²		1.011* (0.005)	1.018** (0.006)
High School GPA (ref. GPA>3.5)			
HS GPA <2.0		0.716 (0.152)	0.804 (0.199)
HS GPA 2.0-3.0		0.714*** (0.044)	0.802*** (0.048)
HS GPA 3.0-3.5		0.787*** (0.022)	0.815*** (0.023)
ACT Composite Percentile		0.995*** (0.001)	0.999 (0.001)
Expect to apply for financial aid			4.858*** (0.141)
Expect to work			1.545*** (0.040)
Family income (ref. less than \$24,000)			
about \$24,000 to \$36,000			0.919 (0.138)
about \$36,000 to \$50,000			0.546*** (0.072)
about \$50,000 to \$60,000			0.391*** (0.050)
about \$60,000 to \$80,000			0.352*** (0.040)
about \$80,000 to \$100,000			0.284*** (0.033)
about \$100,000 to \$120,000			0.189*** (0.022)
about \$120,000 to \$150,000			0.149*** (0.017)

This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

more than \$150,000			0.069*** (0.008)
Expected living arrangement in college (ref. college dorm)			
Own home			1.038 (0.044)
Parents' or relatives' home			1.038 (0.046)
Married student housing			1.205 (0.307)
Fraternity or sorority			1.027 (0.108)
Expected level of education (ref. Associate's or below)			
Bachelor's degree			0.974 (0.165)
1-2 years of graduate school			0.925 (0.159)
Doctoral or professional			0.955 (0.161)
Other			0.994 (0.200)
<hr/>			
Term FE	YES	YES	YES
Institution FE	YES	YES	YES
Term X Institution FE	YES	YES	YES

Note: Coefficients exponentiated as odds ratios

Standard errors in parentheses clustered at the high school level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C
Regression results for Tobit regression of total amount borrowed in the first year
(n=101,157)

Variables	Model 1 (first-gen only)	Model 2 (+ demographics)	Model 3 (+ resources)	Model 4 (+ tuition & financial aid)
First Gen	1223.939*** (106.473)	1173.424*** (169.869)	899.413*** (101.645)	1031.728*** (101.806)
Transfer		1546.785*** (102.585)	1322.489*** (80.722)	299.042*** (71.671)
Gender (ref. male)				
Female		259.661*** (50.678)	259.106*** (50.881)	261.964*** (47.854)
Race (ref. White)				
American Indian		-1884.269*** (222.058)	-1962.648*** (195.806)	1591.600*** (364.071)
Asian		-2792.964*** (266.391)	-2849.710*** (291.137)	-2171.289*** (168.398)
Black		-1355.242*** (179.917)	-1417.268*** (178.558)	-454.728* (190.075)
Hispanic		73.136 (191.401)	18.687 (199.326)	474.898** (145.186)
Age		486.402*** (63.844)	660.073*** (51.018)	602.787*** (56.264)
Age ²		-7.520*** (0.927)	-9.594*** (0.797)	-8.188*** (0.859)
EFC			52.024*** (9.762)	-101.934*** (29.003)
EFC ²			-1.531*** (0.078)	-0.088 (0.230)
Independent student			-866.271*** (245.158)	-481.710* (188.716)
Residency (ref. in-state)				
Out-of-state				-3237.611*** (424.952)
Reciprocity state				847.552*** (174.637)
Midwest state				-584.076* (225.229)
Net tuition				1095.115*** (62.445)
Grant recipient				218.538 (139.256)
Scholarship recipient				338.255*** (62.151)
Work study recipient				1398.792*** (97.778)
Grant amount				-0.453*** (0.044)
Scholarship amount				-0.073 (0.050)
Work study amount				-0.747*** (0.027)

This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

Term FE	YES	YES	YES	YES
Institution FE	YES	YES	YES	YES
Term X Institution FE	YES	YES	YES	YES

Note: Standard errors in parentheses clustered at the high school level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table D
Double-hurdle regression results for probability of borrowing (Probit regression)
(n=101,157)

Variables	Model 1 (first-gen only)	Model 2 (+ demographics)	Model 3 (+ resources)	Model 4 (+ tuition & financial aid)
First Gen	0.259*** (0.015)	0.246*** (0.022)	0.123*** (0.009)	0.151*** (0.010)
Transfer		0.300*** (0.018)	0.246*** (0.019)	0.147*** (0.017)
Gender (ref. male)				
Female		0.047*** (0.008)	0.042*** (0.008)	0.041*** (0.009)
Race (ref. White)				
American Indian		-0.302*** (0.041)	-0.380*** (0.035)	0.189** (0.072)
Asian		-0.271*** (0.026)	-0.385*** (0.036)	-0.374*** (0.031)
Black		-0.061* (0.029)	-0.231*** (0.026)	-0.178*** (0.026)
Hispanic		0.020 (0.023)	-0.042 (0.025)	0.018 (0.024)
Age		0.058*** (0.008)	0.088*** (0.009)	0.108*** (0.010)
Age ²		-0.001*** (0.000)	-0.001*** (0.000)	-0.002*** (0.000)
EFC			-0.014*** (0.002)	-0.022*** (0.004)
EFC ²			0.000 (0.000)	0.000*** (0.000)
Independent student			-0.305*** (0.042)	-0.288*** (0.043)
Residency (ref. in-state)				-0.617*** (0.059)
Out-of-state				0.197*** (0.022)
Reciprocity state				-0.138*** (0.023)
Midwest state				0.168*** (0.010)
Net tuition				0.054* (0.025)
Grant recipient				0.048*** (0.010)
Scholarship recipient				0.367*** (0.023)
Work study recipient				-0.000*** (0.000)
Grant amount				0.000 (0.000)
Scholarship amount				0.000 (0.000)
Work study amount				-0.000*** (0.000)

This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

Term FE	YES	YES	YES	YES
Institution FE	YES	YES	YES	YES
Term X Institution FE	YES	YES	YES	YES

Notes: Standard errors in parentheses clustered at the high school level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table E
Double-hurdle regression results for conditional amount borrowed (Truncated regression)
(n=101,157)

Variables	Model 1 (first-gen only)	Model 2 (+ demographics)	Model 3 (+ resources)	Model 4 (+ tuition & financial aid)
First Gen	-616.335*** (125.662)	-472.221*** (68.981)	356.161** (122.618)	574.153*** (113.881)
Transfer		70.012 (97.590)	75.509 (138.405)	-844.450*** (130.313)
Gender (ref. male)				
Female		16.542 (63.713)	76.712 (60.106)	149.589** (52.913)
Race (ref. White)				
American Indian		-1451.403*** (249.846)	-857.266*** (251.383)	1123.426*** (319.642)
Asian		-3574.761*** (558.149)	-2276.976*** (379.071)	-1486.050*** (183.186)
Black		-2833.573*** (346.077)	-1203.870*** (205.481)	55.249 (226.891)
Hispanic		-117.081 (326.192)	422.669 (277.832)	616.457*** (154.973)
Age		423.391*** (59.024)	485.694*** (47.601)	339.977*** (55.073)
Age ²		-5.778*** (1.003)	-7.065*** (0.687)	-4.679*** (0.787)
EFC			292.499*** (17.775)	29.060** (9.460)
EFC ²			-2.974*** (0.209)	-0.418*** (0.091)
Independent student			1598.342*** (133.403)	1812.848*** (191.755)
Residency (ref. In-state)				
Out-of-state				-1306.137*** (394.620)
Reciprocity state				483.911*** (145.550)
Midwest state				611.895* (295.821)
Net tuition				841.944*** (39.684)
Grant recipient				116.853 (81.407)
Scholarship recipient				82.936 (74.629)
Work study recipient				-78.352 (118.853)
Grant amount				-0.622*** (0.062)
Scholarship amount				-0.188** (0.069)
Work study amount				-0.811*** (0.045)

This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

Term FE	YES	YES	YES	YES
Institution FE	YES	YES	YES	YES
Term X Institution FE	YES	YES	YES	YES

Notes: Standard errors in parentheses clustered at the high school level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table F
Double-hurdle regression results for probability of borrowing and amount borrowed

Variables	Decision to Borrow (Probit)		Amount Borrowed (Truncated)	
	FG	Non-FG	FG	Non-FG
Transfer	0.106*** (0.023)	0.156*** (0.020)	-1240.642*** (124.828)	-953.958*** (138.430)
Gender (ref. male)				
Female	0.051* (0.023)	0.037*** (0.011)	65.278 (82.890)	112.131 (65.598)
Race (ref. White)				
American Indian	0.099 (0.077)	0.219** (0.076)	314.870 (249.350)	1353.364** (463.741)
Asian	-0.451*** (0.044)	-0.286*** (0.025)	-1233.184*** (366.005)	-554.725* (250.666)
Black	-0.337*** (0.047)	-0.001 (0.042)	-560.312*** (157.568)	1172.659*** (286.231)
Hispanic	-0.038 (0.045)	0.054 (0.033)	389.095* (176.414)	641.294** (222.639)
Age	0.101*** (0.020)	0.120*** (0.011)	119.804* (58.821)	304.844*** (77.924)
Age ²	-0.001*** (0.000)	-0.002*** (0.000)	-1.508 (0.822)	-4.391*** (1.085)
EFC	-0.003 (0.007)	-0.027*** (0.003)	95.698*** (16.776)	43.685** (16.527)
EFC ²	-0.000 (0.000)	0.000*** (0.000)	-0.693*** (0.135)	-0.375*** (0.113)
Independent student	-0.123*** (0.033)	-0.389*** (0.069)	2128.064*** (153.938)	2335.942*** (259.321)
Residency (ref. in-state)				
Out-of-state	-0.723*** (0.105)	-0.556*** (0.046)	-184.886 (663.220)	-713.706 (525.750)
Reciprocity state	0.342*** (0.035)	0.151*** (0.021)	254.139 (130.800)	296.330 (156.254)
Midwest state	0.017 (0.066)	-0.169*** (0.025)	1424.742*** (357.144)	585.552 (352.961)
Net tuition	0.192*** (0.017)	0.156*** (0.009)	613.824*** (91.179)	686.485*** (82.229)
Grant recipient	0.105* (0.048)	0.043* (0.021)	-90.716 (121.673)	16.559 (118.733)
Scholarship recipient	0.027 (0.022)	0.056*** (0.013)	46.516 (114.764)	-81.623 (127.191)
Work study recipient	0.430*** (0.042)	0.323*** (0.030)	-406.706 (263.410)	-650.177** (234.301)
Grant amount	-0.000** (0.000)	-0.000*** (0.000)	-0.495*** (0.063)	-0.609*** (0.072)

This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

Scholarship amount	-0.000*** (0.000)	-0.000*** (0.000)	-0.282*** (0.083)	-0.128 (0.100)
Work study amount	-0.000 (0.000)	0.000* (0.000)	-0.551*** (0.081)	-0.758*** (0.075)
Term FE	YES	YES	YES	YES
Institution FE	YES	YES	YES	YES
Term X Institution FE	YES	YES	YES	YES
N	31,316	69,841	31,316	69,841

Notes

ⁱ We include fall starts only for consistency with the financial aid year.