



Student Loan Data Deficits: Improving Borrower Data for Institutional Decision Makers

Matthew Soldner

JUNE 2016

Student Loan Data Deficits: Improving Borrower Data for Institutional Decision Makers

June 2016

Matthew Soldner



1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000

www.air.org

Copyright © 2016 American Institutes for Research. All rights reserved.

Contents

	Page
Introduction.....	1
The Case of Walden University.....	2
After Leaving Walden, How Much Debt Do Students Hold?	3
Estimating the Outstanding Debt of Borrower Cohorts.....	3
Institution-Based Estimation of Cumulative Outstanding Principal Balance.....	5
Improving Borrower Data for Institutions	7
Are Students Borrowing Too Much?	8
Using Loan Volume Reports to Understand Current Borrowing	9
Using NPSAS to Understand Current Borrowing	11
Improving Data for Institution Leaders to Understand Student Borrowing	17
Concluding Thoughts.....	18
Endnotes.....	19

Tables

Table 1. Current Cumulative Outstanding Principal Balance of Consolidation Loans, All Other Loans, and Total Loans by Fiscal Year Loans Entering Repayment: Through Fiscal Year 2015.....	6
Table 2. Typical FSA Data Center Loan Volume Report.....	9
Table 3. Average Disbursements to Walden University Students Based on FSA Data Center Statistics and Statistics Generated From Walden’s Student Information Systems, by Loan Program, Award Year, and Level of Study: 2011–12 to 2013–14	11
Table 4. Average Total Title IV Loans to Master’s Degree Students at Walden University, as well as Non-profit public, Non-profit private, and Proprietary Institutions, by Program of Study: 2011–12	14
Table 5. Average Total Title IV Loans to Research and Other Doctoral Degree Students at Walden University, as well as Non-profit public, Non-profit private, and Proprietary Institutions, by Program of Study: 2011–12	15
Table 6. Count of All Postsecondary Institutions Participating in Title IV Federal Student Aid Programs, Number Sampled for NPSAS, and Count of Students Sampled, by NPSAS Sector: 2011–12.....	16

Introduction

Authorized by the Higher Education Act of 1965 (HEA), the nation’s federal student loan program, known as Guaranteed Student Loans at its inception and most recently as the William D. Ford Federal Direct Loan Program, has supported millions of students in their goal of earning a college degree. For students and families who are unable to finance their postsecondary education out-of-pocket and find that even after a Pell Grant they have unmet financial need, access to federal student loans can be the difference between enrolling in college and watching a life-changing opportunity slip by. In addition to federal student aid programs’ undeniable private benefit, their contribution to the public good has been asserted from its earliest days. In signing HEA in to law, President Johnson remarked that the “nation can never make a wiser or more profitable investment anywhere.”ⁱ

Half a century later, the magnitude of that investment—and the tenor of the conversation about the student loan programs—has noticeably changed.

The size of the government’s federal student loan program has grown dramatically. This is particularly true over the last decade, as more students have entered college, the price of college has risen, private credit markets have tightened, and many families have seen their household income stagnate or decline. In the 2004–05 academic year, the federal student aid program made about 23 million separate loan disbursements totaling nearly \$46 billion. By the 2014–15 academic year, the last for which data are available, disbursements to students had grown by half, exceeding 33 million annually, and loan disbursements had reached nearly \$85 billion.

That growth has resulted in a substantial amount of student debt. The Federal Reserve Bank of New York’s most recent *Quarterly Report on Household Debt and Credit*, released in November 2015, estimated borrowers owed nearly \$1.2 trillion on their student loans, a \$13 billion increase over the prior quarter. Almost 12 percent of that debt, about \$144 billion, had been delinquent for more than 90 days or was in default.ⁱⁱ A decade ago, Maryland Senator Barbara Mikulski argued that “[c]ollege is part of the American Dream: it shouldn’t be part of the American financial nightmare.”ⁱⁱⁱ Times have not changed. It is nearly impossible to turn on the television, open a newspaper or magazine, or talk to a neighbor and not hear expressions of concern about the price of college, student debt, and the economic plight of college graduates.

Given the size of the nation’s investment in its federal student loan programs, the difficulty that some borrowers appear to be having in repaying their education debt, and the depth of the public’s concerns about both, one might expect stakeholders to be awash in actionable data about loan programs, borrowing, and debt. But this is not case.

Despite the recent increase in the availability of consumer-oriented information, most notably the U.S. Department of Education’s (ED’s) College Scorecard, the data available on student borrowing do not meet the urgency felt in most quarters to address the serious policy concerns related to student borrowing.

This report focuses specifically on two questions faced by institution leaders as they attempt to better understand borrowers and borrowing:

- After leaving the institution, how much debt do borrowers hold?
- Are current students borrowing too much to finance their education?

The Case of Walden University

This report is animated by American Institutes for Research's (AIR's) ongoing consulting engagement with Walden University. In fall 2015, the institution's senior leadership approached AIR after two recent analyses, one conducted by the Brookings Institution^{iv} and another by the Center for American Progress (CAP)^v, could have been interpreted as suggesting Walden students has high levels of borrowing..^{vi, vii} Walden believed the estimates from Brookings that the university's borrowers had a cumulative outstanding loan balance of \$9.8 billion to be inaccurate based on its own knowledge. And the CAP analysis, which focused on Walden borrowers' total loan volume in a single year without consideration of the number of students the university served, was a concerning portrayal.

As a result, Walden's leaders raised two simple questions about the administration of financial aid programs at their institution: Were the estimates presented in the analyses correct? And how did borrowing among Walden students compare with students at other institutions, both annually and in the aggregate?

To evaluate these questions, Walden provided AIR researchers fully deidentified borrowing data maintained in their student information systems related to their students' participation in Title IV loan programs. These included hundreds of thousands of loan-level records related to annual borrowing between award years 2010 and 2014, and several hundreds of thousands more loan-level records related to repayment between fiscal years (FYs) 1997 and 2015.

As we describe below, answering those questions, even with rich institutional data, was far from easy. Without institutional data—the situation most external policy analysts find themselves in—it would be virtually impossible.

Our work with Walden makes one thing clear: Institutional decision makers could benefit if more information about student borrowing was more readily available. To be sure, Walden is a unique university. It is a proprietary institution primarily offering graduate programs in business, education, and the health professions through distance education.

But we do not believe its story and its data problems are unfamiliar. We contend that lessons learned during our analysis of its data are broadly generalizable and should be shared with the field because they can inform a set of recommendations for data policy that will benefit the larger higher education community and the students they serve.

After Leaving Walden, How Much Debt Do Students Hold?

As noted earlier, the estimated \$1.2 trillion in outstanding education debt has raised concerns among policymakers, educators, and the general public about college affordability and the capacity of borrowers to meet their student loan obligations after leaving the institution. In response to those concerns, a variety of research efforts have emerged, with the goals of (1) better understanding changes in student borrowing behaviors and subsequent student loan default, and (2) identifying student- and institution-level factors associated with changes in those behaviors and outcomes over time.

Work done by Adam Looney (U.S. Department of the Treasury) and Constantine Yannelis (New York University) originally released in fall, 2015, and revised in May, 2016, is a recent example. It is notable not only for its key findings but also for a byproduct of the research: estimates of how much loans associated with a selected list of institutions contribute to the national education debt. Table 5 in the authors' report ranked institutions by the "accumulated federal loans of their students," focusing on the top 25 contributors in both 2000 and 2014.^{viii} And although the authors place neither emphasis nor judgment on these rankings, it should come as no surprise that they caught the attention of the media, higher education experts, and institution leaders.

According to Looney and Yannelis's original analysis, Walden University borrowers held \$9.8 billion in debt as of 2014, second only to the University of Phoenix and in company with other well-known non-profit private and non-profit public institutions such as New York University (#8 on the list), the University of Southern California (#13 on the list), and Pennsylvania State University (#14 on the list).

To some, Walden's relative position on Looney and Yannelis's original list may not be particularly surprising: It is a large institution, enrolling more than 50,000 students annually, and more than 80 percent of its students are pursuing graduate study, which is typically more costly than undergraduate education. But to Walden leadership, the estimate was unexpectedly high. Their intuition that something might be amiss with the original \$9.8 billion figure begged a simple question: After leaving the institution, how much debt *do* Walden students hold?

As we will see, they had good reason to wonder.

Estimating the Outstanding Debt of Borrower Cohorts

To generate their estimates, Looney and Yannelis rely on a data set maintained by ED's Budget Service, known to its users as CEAD-STAB (that is, the *Cost Estimation and Analysis Division's Statistical Abstract*). As the authors note on page 42 of their original report:

The analysis in this paper is based on a 4 percent random sample of student loan borrowers ... The educational records are sampled as of the end of fiscal years from 1970 to 2013 from transactions [sic] records from FSA's operational database. The sample is intended to reflect loan balances and status as of the close of the fiscal year ... [and] institutions corresponding to the years in which loans were disbursed ... The sample includes federal direct and federally guaranteed students [sic] loans, including both the Federal Family Education Loan Programs and the Direct Loan Program. The sample does not include Perkins loans. Parent PLUS loans are included in the analysis but the outcomes of parent borrowers are excluded from our analysis ...

Armed with this ED resource, estimating the outstanding principal balance (OPB) associated with successive cohorts of borrowers from a given institution would seem to be a relatively straightforward activity. With access to information about each loan held by borrowers, including the school that originated the loan and its current OPB, Looney and Yannelis needed only to select years on which to focus and start adding. And it would be just that simple, except for decisions that must be made in handling two groups of students: those who attended multiple institutions and those who consolidated their student loans.

In their original analysis, Looney and Yannelis attributed all of a student's loans to the last institution he or she attended.^{ix} Because the authors were not focused on specific institutions, this does not necessarily reflect a flaw in their larger analysis. But, as the authors acknowledge, it has serious implications for how student debt is—or is not—allocated to the institution where that debt was actually incurred.^x

For institutions such as Walden, which are largely graduate-serving, the authors' analytic choice shifts potentially large amounts of undergraduate debt to a student's postbaccalaureate institution, opening the door to mischaracterizations. But any institution that admits students with prior coursework, from small regional campuses accepting community college transfers to the most well-known graduate programs at elite universities, might be affected in a similar way. The potential scope of this problem is substantial. Research by the National Student Clearinghouse has suggested, for example, that 37 percent of first-time undergraduate students transfer within six years of entry to college.^{xi}

In May, 2016, the authors' revised their analytic approach, presenting new findings in an on-line data appendix. Rather than simply assigning loans to the last institution students' attended, the authors leveraged ED resources to more accurately associate loan balances with the institutions at which those loans were originally taken. The revised results—that Walden students' total indebtedness was actually \$6.1 billion—suggest the original \$9.8 billion estimate was a substantial overestimate.

Unfortunately, if Walden (or any other institution) wanted to verify or extend Looney and Yannelis's work, determining exactly how much outstanding debt its former students had accumulated specifically for their Walden education, it would have difficulty doing so. It does not have access to CEAD STAB, and neither does any other institution that would want to verify—or calculate anew—its students' outstanding principal balance.

Institution-Based Estimation of Cumulative Outstanding Principal Balance

Institutions wanting to create their own estimate of their borrowers' OPB can do so using a resource provided by ED's Federal Student Aid (FSA) office. Authorized users at each institution can request their institution's *School Portfolio Report (SPR)* using a Web-based portal to FSA's National Student Loan Data System. By stitching together a series of *SPRs* that include successive cohorts of loans entering repayment, it is possible to create a data set that details the current status of all loans associated with a particular institution, including those loans' OPB.^{xii} Then, as was the case for Looney and Yannelis, creating an estimate of outstanding student debt would seem to be a matter of simple addition.

However, loan consolidations, which allow borrowers to bundle multiple student loans into one, complicate things. At the point of consolidation, all data on the outstanding balance of the loans being consolidated disappear from the *SPR*. This makes it impossible for institutions to determine what share of a consolidation loan is due to borrowing specific to their institution.^{xiii} Were consolidations rare, this would not be an issue. But they are not: FSA data show that, as of the first quarter of 2016, these loans comprise 22 percent of the value of ED's entire student loan portfolio.^{xiv}

As a result of how consolidation loans are treated in *SPR*, two problems arise:

- Loans originated at one institution that are later consolidated elsewhere appear to the original institution as being "paid-in-full by consolidation," but no information is available about the OPB of the consolidation loan itself. This makes it impossible for an institution to know whether consolidated borrowers are successfully meeting their loan obligations.
- *SPR* data on consolidation loans attributed to a given institution show details for only the underlying, consolidated loans that were originated there. Consolidated loans originated elsewhere, if any exist, are not detailed. As a result, the analyst must compare the original loan amounts for the loans that *are* detailed with the amount originated for the consolidation loan, and then decide whether the consolidation loan's entire OPB is likely attributable to the borrower's institution. (A better approach would be to compare the OPB of the consolidated loans with the origination amount of the consolidation loan; however, bear in mind that those underlying OPBs disappear upon consolidation.)

For institutions seeking to tease out what fraction of a consolidation loan is due to a student attending their institution, the best-case scenario is that the origination amounts of the underlying loans approximate the origination amount of the consolidation loan, making it tenable that no external loans have been consolidated. In the event that the total origination amounts of the underlying loans are *significantly less* than the consolidation loan, one might conclude that at least some external loans have been included in the consolidation loan. If they are only *slightly less*, then the consolidation loan might have been originated at the higher amount due to capitalized interest or fees. If, on the other hand, the underlying loans appear to be *greater* than the consolidation amount, then the analyst can only conclude that the borrower had made at least some progress toward repaying some portion of his or her loan obligation.

Institution-based analysts are left to their best professional judgment when trying to understand how much of a consolidation loan's OPB is genuinely attributable to their institution. One can only hope that the amount of OPB mistakenly attributed to an institution is cancelled out by the portion of that institution's OPB that is misattributed elsewhere. But there is really no way to tell.

Using *SPR* data for loans entering repayment as early as 1997, AIR independently calculated Walden borrowers' OPB. Based on those data, and after attributing all consolidation loans to Walden, we estimated that the university's borrowers held approximately \$5.1 billion in education debt at the end of FY 2014. We estimate that figure would be \$6.1 billion if the OPB associated with students entering repayment in FY 2015 were included (see Table 1).

This \$5.1 billion in outstanding Walden student borrower debt as of FY 2014 is about half the figure reported by Looney and Yannelis. Yet, because it attributes all consolidated debt to Walden borrowers, it is arguably an overestimate. AIR attempted to disentangle the contribution of Walden-originated loans to borrowers' cumulative OPB. To do so, we scaled each consolidation loan's OPB by the ratio of its origination amount to the total originated amount of all Walden loans consolidated within it. That approach yielded a cumulative OPB of \$3.8 billion, less than 40 percent of Looney and Yannelis's original estimate.

Table 1. Current Cumulative Outstanding Principal Balance of Consolidation Loans, All Other Loans, and Total Loans by Fiscal Year Loans Entering Repayment: Through Fiscal Year 2015

Fiscal Year Loans Entering Repayment	Consolidation Loans	All Other Loans	Total
1997	\$4,065,397	\$324,475	\$4,389,872
1998	7,638,660	912,968	8,551,628
1999	15,082,044	1,548,095	16,630,139
2000	21,883,346	2,809,084	24,692,430
2001	31,136,037	3,926,998	35,063,035
2002	41,262,452	6,238,020	47,500,472
2003	51,984,156	7,115,947	59,100,103
2004	73,843,112	9,705,828	83,548,940
2005	127,583,411	14,397,558	141,980,969
2006	272,446,794	27,973,048	300,419,842
2007	397,436,727	53,508,094	450,944,821
2008	537,711,968	143,699,830	681,411,798
2009	767,195,827	311,323,363	1,078,519,190
2010	1,052,984,952	584,129,343	1,637,114,295
2011	1,380,447,200	949,647,562	2,330,094,762
2012	1,672,372,877	1,413,852,664	3,086,225,541
2013	2,006,505,073	2,030,383,136	4,036,888,209
2014	2,328,316,351	2,749,632,644	5,077,948,995
2015	2,539,635,321	3,571,695,064	6,111,330,385

SOURCE: Authors' calculation from Walden University *School Portfolio Reports*.

These are large differences. Our best estimate is that, as of the end of FY 2014, Walden borrowers had an outstanding principal balance of between \$3.8 and \$5.1 billion, not the original \$9.8 billion the authors estimated (or even their revised \$6.1 billion). Although this disparity is notable, the deeper issue is that arriving at the answer to such a fundamental question was far from easy and, due to the uncertainties regarding consolidation loans, still highly imprecise.

Institution executives face difficult questions that pose substantive challenges every day. In many instances, the hard data needed to answer those questions do not exist. But they do—or at least they *could*—here. Leaders need, and students deserve, better.

Improving Borrower Data for Institutions

Motivated by our attempt to verify estimates of Walden borrowers' OPB, we offer two suggestions for improved data policy and resources.

Provide additional information to institutions about consolidation loans.

Nearly half of Walden borrowers' OPB is wrapped up in consolidation loans. For institutions such as Walden, this ups the ante for making sure institutions know what share of that OPB is actually attributable to them and what share results from students having incurred debt elsewhere. Institutions need detailed information about all loans that are associated with them, regardless of whether the institution originated that loan or whether the loan has been inherited through consolidation.

A potentially promising option would be to establish an allocation factor for consolidated loans upon their consolidation, reflecting the share of the new loan's origination amount contributed by each consolidated loan. That factor could persist with loan detail records for the life of the consolidation, along with the properly prorated share of the consolidation loan's current OPB. This would have two benefits. First, it keeps consolidated loans "on an institution's books," even though they are ostensibly paid in full upon consolidation, so that the institution can continue to monitor students' repayment outcomes. Second, it allows institutions to easily see what share of a consolidation loan associated with their institution is really comprised of debt accrued elsewhere, providing a better answer to the question of how much debt is held by an institution's borrowers.

A related solution would involve FSA maintaining information about a consolidated loan's balance at the time of consolidation. Particularly if paired with information about the resulting consolidation loan, this would allow the institution to estimate whether a loan was, and continues to be, in negative amortization.

Improve the usefulness of NSLDS-based data products.

An institution's *SPR* is one of its most important tools in understanding the behavior of individual borrowers and its loan portfolio in the aggregate. But working with the *SPR* is not easy. First, similar to the bulk of the data maintained by FSA, the *SPR* is focused on the current status of a loan, not its history over time. As a result, creating a longitudinal portrait of an institution's loan portfolio requires the institution to schedule and archive the *SPR* at regular

intervals. Second, once downloaded, the institution must manipulate the *SPR* data itself. This can include the need to link students and loans across multiple *SPRs*, parse nearly 50 different loan status codes, and, as we have noted, wrestle with consolidation loans.

Reports such as the *SPR* make a huge amount of data available to institutions. But data and information are not synonymous. We believe that FSA could improve the usefulness of *SPR* data by creating complementary reports that can help users understand, at a glance, the status of their loan portfolio. Reports that reflected change over time, either of the portfolio as a whole or of specific repayment cohorts, would add additional value. A preformatted report that automatically produced numbers such as those found in Looney and Yannelis’s analysis, for example, could be informative to institutional decision makers. There is no doubt that FSA subject-matter experts, with the assistance of institution leaders, financial aid professionals, and higher education advocates, could identify more, and perhaps more useful, preformatted reports.

To determine what kinds of data would be most useful to institution leaders, FSA should rely on a strategy routinely used by other offices within ED: convening technical working groups (TWGs). A TWG that includes institution executives and institutional research professionals, advocates for students and families, and career professionals from FSA and elsewhere within ED could identify what types of borrower characteristic-based statistics would be most useful to members of each community and offer up a prioritized action plan for FSA. Subsequent meetings of the TWG could provide opportunities for monitoring FSA’s progress, reprioritization of existing activities, or the identification of new ones.

Are Students Borrowing Too Much?

Understanding aggregate outstanding loan volume is important for an institution. Even more important on a routine basis is understanding how an institution’s per-student borrowing trends compare with those at peer institutions. Although institutions have relatively complete information on how current enrollees are financing the cost of their education, they have little information about how that information compares with students at similar institutions, in similar programs of study, and with similar financial circumstances. As a result, institution leaders—not to mention policymakers or consumers—have little ability to answer a fundamental question: Are students borrowing too much, particularly when compared with their peers at other colleges and universities?

There are two national sources of data on student borrowing that, because they are drawn directly from data systems maintained by ED, can be considered definitive. The first is a collection of quarterly campus-level loan volume reports available on the FSA Data Center website.^{xv} The second is the National Center for Education Statistics’ (NCES’) quadrennial National Postsecondary Student Aid Study (NPSAS), which combines data held by FSA, individual institutions, and student survey responses to create a research data set that can be used to answer a variety of questions about how students and families finance their postsecondary education.^{xvi}

Unfortunately, neither data source meets the needs of institution leaders who need to understand whether their students’ borrowing behaviors are comparable to trends at the national or state level, or among their institutional peers. Using data provided by Walden University, we demonstrate that the value of the information gleaned from the FSA Data Center is undercut by a

reporting system that focuses on loans and loan programs, not the students they are meant to support. Thankfully, students are the unit of analysis in NPSAS. But because that study is only executed every four years and is representative at the level of institutional sectors, not individual institutions, its usefulness as a basis for benchmarking or comparison is sorely limited. We discuss both data systems in more detail below.

Using Loan Volume Reports to Understand Current Borrowing

The information made available by FSA through its FSA Data Center is a peek into the federal student loan program's checkbook register. Each row of that register corresponds to an institution whose students are receiving funds. Also, each set of columns tells us just how big that check is, broken out by specific federal loan program. Additional details include how many students received that type of loan and how many separate loan disbursements those students received. An example of some of the information available through the FSA Data Center is depicted in Table 2. Other columns not shown would present similar information for Unsubsidized Direct Loans, Unsubsidized Direct Loans to Graduate Students, and loans made under the GRAD PLUS Direct Loan program.

Table 2. Typical FSA Data Center Loan Volume Report

OPEID	Institution	School Type	Direct Loans: Subsidized				
			Recipients	# of Loans Originated	\$ of Loans Originated	# of Loans Disbursed	\$ of Loans Disbursed
12345600	State Public University	Non-profit public	18	20	\$10,000	44	\$9,642
98765400	Private University	Non-profit private	14	14	7,000	31	6,749
34567800	National University	Proprietary	16	18	9,000	39	8,678

NOTE: OPEID=Office of Postsecondary Education Identification code

At first glance, the FSA Data Center would seem to be an ideal source of data for the institution leader who wants to understand his or her students' borrowing behavior in comparison with an institution's peers. But the barrier to using loan volume reports for this purpose is simple: The unit of analysis is the loan program as a whole at a particular institution, not the student. As a result, it is not possible to:

- Create an accurate estimate of the total annual amount borrowed by a typical student because students can receive varying levels of financial support from multiple programs within a given year.
- Identify differences in borrowing behaviors across students pursuing different fields of study, even though differential tuition and fees can affect the prices students must pay.
- Distinguish borrowing behaviors of students at different levels of study (e.g., less-than two-year certificates versus bachelor's degrees for programs serving undergraduates, or master's students versus doctoral students for programs serving graduate students) within an institution, even though the price of those programs may vary widely.

Because loan volume reports cannot make distinctions between the characteristics of borrowers, the composition of their financial aid packages, or the programs of study in which they have enrolled, most types of analyses that rely upon them are unavoidably flawed. In Table 3 below, we highlight the problems that can arise from this type of analysis by contrasting publicly available data about the borrowing of Walden University students, based on loan volume reports, with data extracted from Walden’s student information systems.

As Table 3 shows, using loan volume reports to create borrower-level estimates can be a risky proposition, depending on an analyst’s population of interest. Based on our analysis of data from the FSA Data Center and Walden University’s financial aid systems, the level of bias introduced by using loan volume reports varies substantially by program. At least for the three award years we examined, borrowing statistics for undergraduate programs generated using data from the FSA Data Center compare favorably with those created using institutionally held data. Per-borrower averages are virtually identical: The largest discrepancy, a \$50 gap between estimates of average undergraduate subsidized loans in 2011–12—\$4,051 according to FSA and \$4,102 according to Walden—is trivial.

When one examines statistics related to graduate borrowing, however, it quickly becomes apparent just how misleading aggregate estimates based on FSA Data Center loan volume reports can be. For an institution such as Walden, where upwards of 80 percent of learners are in graduate programs, this matters.

What is not apparent in the FSA Data Center-based estimates, but evident when institution-level data are analyzed, is that there can be substantial variation in borrowing based on a student’s field and level of study (see Table 3). Higher cost programs, particularly higher cost, high-enrollment programs, can artificially inflate per-borrower averages that would be naively created using the FSA Data Center. And although institution practitioners know these differences exist and can create more accurate and nuanced portraits of their students’ borrowing, there is no way for federal or state policymakers, advocates and analysts, or families to do the same.

Opening a family’s (or the government’s) checkbook register can tell us where and how much money is being spent. But it tells us little about what that money is buying. The same is true for data housed in the FSA Data Center. For some purposes, it may well be useful to know how much money was disbursed to enrollees at particular colleges and universities, under particular loan programs. But for institution leaders seeking to understand the borrowing behaviors of their students in context, data about how many dollars are flowing to other institutions are not enough.

As our analyses show, the best that can be done using data from the FSA Data Center is the creation of an institutional average for each loan program, a blunt measure that belies significant variation in borrowing driven by both program and level of study. If their goal is to answer the question “Are our students borrowing too much?,” institution leaders need data that are more granular than those FSA Data Center currently provides.

Table 3. Average Disbursements to Walden University Students Based on FSA Data Center Statistics and Statistics Generated From Walden’s Student Information Systems, by Loan Program, Award Year, and Level of Study: 2011–12 to 2013–14

	FSA Data Center	Walden Student Information Systems				
		All Four Fields	Averages, by Program of Study			
Loan Program, Award Year, and Level of Study	Avg.	Avg.	Educ.	Health Sciences	Business	Human Services
<i>Subsidized Direct Loans to Undergraduates</i>						
2011–12 Award Year	\$4,051	\$4,102	\$4,116	\$4,242	\$4,001	\$4,114
2012–13 Award Year	4,024	4,051	4,010	4,205	3,940	4,058
2013–14 Award Year	3,919	3,909	3,878	4,196	3,753	3,798
<i>Unsubsidized Direct Loans to Graduate Students</i>						
2012–13 Award Year	17,023	15,638				
Master’s Degrees	—	—	13,701	15,922	14,451	16,487
Doctoral Degrees	—	—	14,287	17,493	15,679	16,586
2013–14 Award Year	17,198	16,802				
Master’s Degrees	—	—	13,933	18,297	14,926	17,433
Doctoral Degrees	—	—	15,690	18,768	16,067	17,581
<i>PLUS Loans for Graduate Students</i>						
2011–12 Award Year	25,689	25,228				
Master’s Degrees	—	—	15,881	11,782	12,785	10,017
Doctoral Degrees	—	—	28,437	28,933	30,776	30,062
2012–13 Award Year	26,827	26,984				
Master’s Degrees	—	—	16,791	12,069	14,464	12,882
Doctoral Degrees	—	—	30,576	30,014	31,578	31,347
2013–14 Award Year	26,620	25,081				
Master’s Degrees	—	—	18,902	11,263	14,263	13,967
Doctoral Degrees	—	—	28,857	28,065	29,581	30,248

— Not available; FSA Data Center does not disaggregate by level of study.

NOTES: Total includes four largest programs of study, detailed to the right. Unsubsidized Direct Loans to Graduate Students are not included for 2011–12 because changes in loan program rules hamper comparability

SOURCE: Authors’ calculations from U.S. Department of Education, Federal Student Aid, Federal Student Aid Data Center and Walden University Student Information Systems.

Using NPSAS to Understand Current Borrowing

Unlike the FSA Data Center, ED’s NPSAS combines data from a variety of sources to paint a picture of how individual students finance their postsecondary education. This includes data from students’ Free Application for Federal Student Aid (FAFSA); loan and grant detail records maintained in a variety of FSA data systems; demographic, academic, and financial aid information housed in individual institutions’ student information systems; and a survey of

students themselves that gathers information about work for pay, nonfederal grants and loans, and educational goals and aspirations. Once stripped of personally identifiable information, the resulting data files can be licensed by qualified researchers approved by NCES. An easy-to-use Web application, PowerStats, allows the general public to securely create tabular reports about student borrowing without access to the underlying data. We use it here.

The additional information about borrowers that NPSAS makes available to analysts solves a significant problem associated with the program-level data found in the FSA Data Center. Above, we used institutionally held data provided by Walden University to demonstrate that loan program-level borrower averages can be misleading at best. At their worst, these averages can wholly misrepresent the debt an overwhelming majority of an institution's borrowers should expect to incur. Because of the robust data that it includes and because its unit of analysis is an individual student, not a loan or loan program, NPSAS allows analysts to:

- Create an accurate estimate of the total annual amount borrowed by a typical student, even if that student receives financial support from multiple loan programs.
- Disaggregate students by important academic characteristics, identifying relationships between student borrowing behaviors and students' field and level of study.
- Understand how student borrowing may be related to other student characteristics, such as income, employment status, and family size and composition.

Most importantly, NPSAS provides a potential benchmark against which institution leaders can place their students' borrowing into context. In Table 4, we explore the borrowing behavior of Walden University's single largest group of students—those seeking master's degrees.^{xvii} Table 5 presents analogous results for doctoral programs, which also enroll significant numbers of students.

To illustrate the usefulness of NPSAS data for an institution seeking to benchmark its student borrowing, we used NPSAS data from 2011-12 (the latest available) to compare with Walden student borrowing for master's and doctoral degree programs in education; business, management, and marketing; health services; and human services. The results, shown in Tables 4 and 5, suggest that institutions may use NPSAS data to obtain rough measures of comparability with other institutions.

Point estimates shown in Table 4 suggest that Walden federal student loan borrowers in master's degree programs overall borrow less than their peers at non-profit private institutions (\$17,520 versus \$18,930), with variation among programs of study. Walden students in business and human services fields borrow more than their counterparts at non-profit public institutions, but less than those at non-profit private institutions. The point estimates further suggest that while Walden federal student loan borrowers in master's degree education programs borrow more than their counterparts at either non-profit public or non-profit private institutions, those in health fields borrow less.

The point estimates for doctoral-level programs in Table 5 display virtually the same pattern, with Walden doctoral students borrowing less overall than their non-profit private peers (\$22,000 versus \$22,810), with variation among programs of study. Walden doctoral students in business

programs borrow about the same as their counterparts at non-profit private institutions; however, comparison with doctoral students at non-profit public institutions is not possible because of data suppression (see below). As with master's degree programs, Walden doctoral students in health fields borrow less than their counterparts at either public or private non-profit institutions, while Walden doctoral students in education borrow more. Walden doctoral students in human services borrow more than their counterparts at non-profit public institutions, but less than those at non-profit private institutions.

These estimates – though an improvement over those obtained using data from the *FSA Data Center* – are still lacking in several respects, however. Namely, they are imprecise (due to survey sampling error), infrequent, and only representative at the sector level. These shortcomings are described next.

NPSAS estimates include survey error.

As can be seen in Tables 4 and 5, NPSAS estimates suggest that, taken together, students enrolled in Walden's four graduate fields of study borrow more than their peers at non-profit public and other proprietary institutions, but often less than students who are studying at non-profit private colleges and universities. The fact that Walden University students borrow more than their peers at non-profit public institutions is not surprising, if for no other reason than the price of non-profit public universities is offset by contributions to institutional revenue from state taxpayers. But should these analyses give Walden University leaders, already sensitive to questions of affordability for their students, reason to worry?

Unfortunately, the answer is not entirely clear. Because NPSAS is based on a survey—rather than a complete census of all borrowers—estimates are surrounded by a “margin of error.” These margins become particularly large at the level of individual fields: up to 13,000 in one instance (i.e., doctoral-level human service programs at non-profit private institutions). In some instances, they become so large that NCES will not release an estimate to the public, warning the user that doing so would not meet NCES's reporting standards (e.g., doctoral programs in business at proprietary institutions).

Due to the margins of error, we cannot assert with a strong degree of confidence that Walden student borrowing is less than their peers at non-profit private institutions in specific programs of study. However, with the exception of education programs, they certainly do not borrow more.

NPSAS is only conducted every four years.

To an institution leader trying to make sense of his or her students' borrowing behavior in the 2015–16 academic year, the information in Table 4, which summarizes data from 2011–12, may seem less than compelling. Unfortunately, NCES is only required, and is only provided resources from Congress, to conduct NPSAS every four years.^{xviii} A lot can happen in four years, and during the four years between each NPSAS cycle, institution leaders are left to wait and wonder what the next set of benchmarks for their institution will look like. Even when NPSAS data become available, they have already begun to show their age: Data from NPSAS 2011–12 were not released until December 2013.

Table 4. Average Total Title IV Loans to Master's Degree Students at Walden University, as well as Non-profit public, Non-profit private, and Proprietary Institutions, by Program of Study: 2011–12

Program of Study and Institutional Sector	Total Title IV Loans		
	Average	Lower Bound	Upper Bound
<i>All Four Fields</i>			
Walden University	\$17,520	N/A	N/A
Non-profit public and non-profit private universities	17,450	\$16,780	\$18,120
Non-profit public universities	15,930	15,150	16,710
Non-profit private universities	18,930	17,860	20,000
Proprietary universities	14,390	13,810	14,970
<i>Education</i>			
Walden University	15,670	N/A	N/A
Non-profit public and non-profit private universities	13,470	12,400	14,540
Non-profit public universities	13,570	12,010	15,140
Non-profit private universities	13,360	11,980	14,750
Proprietary universities	14,770	‡	‡
<i>Business, Management, and Marketing</i>			
Walden University	17,230	N/A	N/A
Non-profit public and non-profit private universities	17,650	16,100	19,190
Non-profit public universities	15,820	13,560	18,090
Non-profit private universities	19,080	16,910	21,260
Proprietary universities	14,150	12,140	16,160
<i>Health Professions</i>			
Walden University	17,980	N/A	N/A
Non-profit public and non-profit private universities	21,050	19,380	22,710
Non-profit public universities	18,340	16,190	20,490
Non-profit private universities	23,660	21,180	26,140
Proprietary universities	15,040	13,220	16,860
<i>Human Services</i>			
Walden University	19,190	N/A	N/A
Non-profit public and non-profit private universities	19,650	17,510	21,790
Non-profit public universities	17,270	15,670	18,870
Non-profit private universities	22,380	18,630	26,130
Proprietary universities	13,540	11,230	15,850

‡ NCES reporting standards not met.

SOURCES: Authors' calculations from data provided by Walden University and NCES PowerStats, replicable via http://nces.ed.gov/datalab/index.aspx?ps_x=cacbgm91 and http://nces.ed.gov/datalab/index.aspx?ps_x=bedbge8a

Table 5. Average Total Title IV Loans to Research and Other Doctoral Degree Students at Walden University, as well as Non-profit public, Non-profit private, and Proprietary Institutions, by Program of Study: 2011–12

Program of Study and Institutional Sector	Total Title IV Loans		
	Average	Lower Bound	Upper Bound
<i>All Four Fields</i>			
Walden University	\$22,000	N/A	N/A
Non-profit public and non-profit private universities	20,390	\$18,770	\$22,010
Non-profit public universities	18,180	16,340	20,010
Non-profit private universities	22,810	20,120	25,500
Proprietary universities	21,390	19,850	22,930
<i>Education</i>			
Walden University	20,200	N/A	N/A
Non-profit public and non-profit private universities	16,320	14,890	17,760
Non-profit public universities	15,990	14,370	17,600
Non-profit private universities	16,700	13,970	19,430
Proprietary universities	19,350	14,780	23,920
<i>Business, Management, and Marketing</i>			
Walden University	22,620	N/A	N/A
Non-profit public and non-profit private universities	19,620	14,720	24,520
Non-profit public universities	‡	‡	‡
Non-profit private universities	22,740	15,450	30,020
Proprietary universities	17,470	13,920	21,020
<i>Health Professions</i>			
Walden University	22,730	N/A	N/A
Non-profit public and non-profit private universities	25,650	21,370	29,920
Non-profit public universities	23,740	18,420	29,070
Non-profit private universities	28,870	22,150	35,600
Proprietary universities	23,720	19,390	28,050
<i>Human Services</i>			
Walden University	23,330	N/A	N/A
Non-profit public and non-profit private universities	22,550	19,280	25,820
Non-profit public universities	16,870	13,720	20,010
Non-profit private universities	27,520	22,920	32,130
Proprietary universities	26,750	23,330	30,170

‡ NCES reporting standards not met.

NOTES: Estimates do not include doctors of professional practice; Walden does not offer M.D., J.D., D.D.S., or other professional doctoral degree programs.

SOURCES: Authors' calculations from data provided by Walden University and NCES PowerStats, replicable via http://nces.ed.gov/datalab/index.aspx?ps_x=bfdbgh65 and http://nces.ed.gov/datalab/index.aspx?ps_x=bfdbgh65.

NPSAS data are only representative at the sector level.

Benchmarking is only useful to an institution leader when the comparison group against which it is being compared is meaningful. Unfortunately, for those who seek to use NPSAS to understand how their institution stacks up to its peers, the same resource constraints that limit the study to an every-four-year cycle also limit the number of institutions it can include in the data. As a result, NPSAS estimates are only statistically representative at the sector level (see a list of institution sectors in Table 6). State-level estimates, although called for in law, are not available.^{xix} Also, it is certainly not possible to generate reliable estimates for individual institutions. As can be seen in Table 6, the average number of students sampled per institution varies. At four-year proprietary institutions, the sector to which Walden University belongs, about 100 students per institution are sampled.

Table 6. Count of All Postsecondary Institutions Participating in Title IV Federal Student Aid Programs, Number Sampled for NPSAS, and Count of Students Sampled, by NPSAS Sector: 2011–12

Institution Type	Institutions		Students Sampled
	Total	Sampled	
Total	7,050	1,690	128,120
<i>Non-profit public</i>			
Less-than-2-year	270	20	790
2-year	1,110	380	37,000
4-year nondoctorate-granting	360	130	8,180
4-year doctorate-granting	310	230	20,530
<i>Non-profit private</i>			
Less-than-4-year	260	20	1,090
4-year nondoctorate-granting	1,030	260	8,520
4-year doctorate-granting	560	220	10,070
<i>Proprietary</i>			
Less-than-2-year	1,510	60	5,270
2-year	1,030	120	10,280
4-year	620	260	26,390

NOTE: Detail may not sum to totals because of rounding.

SOURCE: Table 2 of Wine, J., Bryan, M., & Siegel, P. (2014). *2011–12 National Postsecondary Student Aid Study (NPSAS:12) data file documentation* (NCES 2014-182). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

In a statistical sense, each NPSAS sector is representative of the institutions that it is meant to exemplify. Just how meaningful institutional comparisons against sector-level estimates truly are, though, are likely in the eye of the beholder. For an institution leader who feels his or her institution is generally similar to others in the same sector—perhaps four-year, doctorate-granting non-profit public universities—NPSAS may serve as a strong foundation for

benchmarking. But for institutions whose uniqueness is not bound up in their tax status or the highest level of degree offered, NPSAS may offer poor options for comparison.

Improving Data for Institution Leaders to Understand Student Borrowing

Institution leaders have good reason to want to understand how much students are borrowing to finance their education. To be sure, at least some of that motivation is provided by accountability efforts launched by the federal government, such as regulations governing maximum loan default rates (for all programs) and debt-to-income ratios (for programs subject to “Gainful Employment” provisions). We have every reason to believe that, at the overwhelming majority of institutions—and certainly in our experience with Walden—the need for better data about borrowing is also motivated by leaders’ genuine commitment to balancing students’ access needs with the long-term consequences of student debt.

To be good stewards of federal student loan programs, leaders need access to information about their students’ borrowing behavior relative to their peers in similar programs, at similar levels of study, and at similar institutions. Below, we identify two steps ED might consider to meet that need.

Make borrower characteristic-based statistics available on the FSA Data Center.

The primary shortcoming of the student borrowing data made available on the FSA Data Center is that they use loans—not types of students—as the unit of analysis. As a result, it is impossible to disaggregate those data into segments that are informative within, or comparable across, institutions. We do not presume that the segments we have identified in our analysis, field and level of study, are the only possibilities or even among the most important. Borrowing behaviors likely vary along several important dimensions, including student demographics and the characteristics of the institutions and academic programs in which they are enrolled.

Data on segments of student borrowers must begin to figure more prominently into the information made available on the FSA Data Center. As we have suggested earlier, FSA should convene TWGs to determine what kinds of data would be most useful to institution leaders.

Supplement NPSAS with biennial updates based on administrative data.

Although it is a critical component of the nation’s postsecondary data infrastructure, NPSAS is fielded too infrequently, and is not statistically representative at the right level, to be useful for institutional decision making. (Indeed, these same limitations undermine NPSAS’s usefulness for federal and state policy analysis.) Minimally, NPSAS should be augmented to comply with its authorizing legislation, which requires its data to be representative on a state-by-state basis.^{xx} Ideally for institution leaders, NPSAS would be made representative at the level of more narrowly tailored institutional sectors, or at the level of institutions themselves. Reducing NPSAS’s periodicity from once every four years to once every two years also would be a boon.

A more frequently occurring NPSAS that is representative at finer levels of detail is possible, but requires additional resources for NCES and a willingness to rethink critical elements of the study. One promising option is to maintain NPSAS in its current form, but field an “administrative NPSAS” two years in to each quadrennial cycle. Student-level data on grant and loan program use could be generated directly from the ED’s data systems, and sampled to create estimates that are representative at the state level, as required by law, or at even finer levels of detail, such as sectors within states or even individual institutions. Depending on the availability of resources, NCES could seek to augment these administrative data with additional information collected from institutional data systems.

Concluding Thoughts

The nation’s federal student loan programs have allowed tens of millions of learners to realize their educational goals and dreams. Despite the undeniable private and public good these loans represent, the seemingly inexorable uptick in Americans’ total education indebtedness and unsure employment and wage prospects for recent graduates have led many to worry that too many students are borrowing too much. Institution leaders are among those concerned, and are rightly asking questions about their own students’ borrowing behaviors and the total debt held by borrowers from their institution. Unfortunately, the information that institution executives and financial aid professionals have at their disposal is insufficient to answer the former question in a meaningful way, and the latter with any degree of accuracy. That does not have to be the case.

Policymakers have rightly demanded that institutions have a high level of accountability for student borrowing and indebtedness. At least \$89 billion in new Direct Loans to students is expected to be disbursed in the upcoming fiscal year.^{xxi} Now is the time for institutional leaders to demand that policymakers provide access to accurate and comprehensive data that allow them to manage their responsibility for those billions effectively.

Endnotes

- ⁱ Remarks on Signing the Higher Education Act of 1965, retrieved from <http://www.txstate.edu/commonexperience/pastsitearchives/2008-2009/lbjresources/higheredact.html>
- ⁱⁱ Research and Statistics Group, Microeconomic Studies. (2015). *Quarterly report on household debt and credit, November 2015*. New York, NY: Federal Reserve Bank of New York.
- ⁱⁱⁱ 151 Cong. Rec 5224 (2005) (statement of Sen. Barbara Mikulski)
- ^{iv} Looney, A., & Yannelis, C. (2015). *A crisis in student loans? How changes in the characteristics of borrowers and in the institutions they attended contributed to rising loan defaults*. Washington, DC: The Brookings Institution. Retrieved from <http://www.brookings.edu/about/projects/bpea/papers/2015/looney-yannelis-student-loan-defaults>
- ^v Baylor, E. (2015, July 8). As graduate-student debt booms, just a few colleges are largely responsible. *The Chronicle of Higher Education*. Retrieved from <http://chronicle.com/article/As-Graduate-Student-Debt/231415/>
- ^{vi} Douglas-Gabriel, D. (2015, July 9). These 20 schools are responsible for a fifth of all graduate school debt. *The Washington Post*. Retrieved from <https://www.washingtonpost.com/news/get-there/wp/2015/07/09/these-20-schools-are-responsible-for-a-fifth-of-all-graduate-school-debt/>.
- ^{vii} Lorin, J. (2015, November 4). Laureate's U.S. students struggling to repay loans as IPO looms. *Bloomberg News*. Retrieved from <http://www.bloomberg.com/politics/articles/2015-11-04/laureate-loan-repayment-rate-shows-u-s-debt-burden-as-ipo-looms>
- ^{viii} See Table 5, *Institutions Ranked by Accumulated Federal Loans of their Students 2000 and 2014*, in Looney & Yannelis (2015).
- ^{ix} On Page 16, the authors write: “The table aggregates the total federal loan liabilities of all borrowers based on the last institution they had borrowed to attend (including undergraduate, graduate, and Parent PLUS loans) in each of those two years.”
- ^x Although not discussed in the original version of the authors’ report, a forthcoming April, 2016, revision provides a more detailed rationale for their analytic decision, and acknowledges the concern we raise here.
- ^{xi} Shapiro, D., Dundar, A., Wakhungu, P.K, Yuan, X., & Harrell, A. (2015). *Transfer and mobility: A national view of student movement in postsecondary institutions, Fall 2008 cohort* (Signature Report No. 9). Herndon, VA: National Student Clearinghouse Research Center.
- ^{xii} As of December, 2015, the ad-hoc *School Portfolio Report* can report on the current status of loans entering repayment during any three-year period of time. For more information, consult technical documentation related to the *School Portfolio Report* available at <https://ifap.ed.gov/nsldsmaterials/attachments/NSLDSSchoolPortfolioFileLayoutSCHPR1FW.pdf>.
- ^{xiii} Looney and Yannelis sidestepped this issue by attributing all debt to the last institution attended. It is unknown whether the data they used may have permitted them to connect balances on consolidation loans to the institutions where the underlying loans were originated.
- ^{xiv} The most recent data on ED’s loan portfolio can be found at <https://studentaid.ed.gov/sa/sites/default/files/fsawg/datacenter/library/PortfolioByLoanType.xls>.
- ^{xv} More information about the FSA Data Center can be found at <https://studentaid.ed.gov/sa/data-center>.
- ^{xvi} More information about the National Postsecondary Student Aid Study can be found at <http://nces.ed.gov/surveys/npsas>.
- ^{xvii} Walden University students are included in this estimate as excluding a specific institution is not possible using the PowerStats tool.
- ^{xviii} NCES’ statutory authority to conduct NPSAS is derived from successive reauthorizations of the Higher Education Act, which call for a “Student Aid Recipient Survey,” most recently the Higher Education Act of 2008, as codified at 20 U.S.C. § 1015(a)(k).
- ^{xix} NPSAS’ authorizing legislation requires state-level estimates: “The Secretary, acting through the Commissioner for Education Statistics, shall conduct, on a State-by-State basis, a survey of recipients of Federal student aid ...” (see 20 U.S.C. § 1015(a)(k)). Some years of NPSAS have been representative of a subset of larger states. NPSAS 2011–12 has no state representative samples.

^{xx} As noted above, 20 U.S.C. § 1015(a)(k) requires NPSAS to be state-representative.

^{xxi} ED’s fiscal year 2017 budget justification for student loan programs, located at <http://www2.ed.gov/about/overview/budget/budget17/justifications/q-sloverview.pdf>, includes \$23.5 billion for Direct Stafford Loans, \$25.9 billion for Unsubsidized Stafford Loans for Undergraduates, \$29.6 billion for Unsubsidized Stafford Loans for Graduate Students, and \$9.8 billion for Direct PLUS Loans for Graduate Students.

ABOUT AMERICAN INSTITUTES FOR RESEARCH

Established in 1946, with headquarters in Washington, D.C., American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance both domestically and internationally. As one of the largest behavioral and social science research organizations in the world, AIR is committed to empowering communities and institutions with innovative solutions to the most critical challenges in education, health, workforce, and international development.



AMERICAN INSTITUTES FOR RESEARCH®

1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000

www.air.org

Making Research Relevant

LOCATIONS

Domestic

Washington, D.C.
Atlanta, GA
Austin, TX
Baltimore, MD
Cayce, SC
Chapel Hill, NC
Chicago, IL
Columbus, OH
Frederick, MD
Honolulu, HI
Indianapolis, IN
Metairie, LA
Naperville, IL
New York, NY
Rockville, MD
Sacramento, CA
San Mateo, CA
Waltham, MA

International

Egypt
Honduras
Ivory Coast
Kyrgyzstan
Liberia
Tajikistan
Zambia