Statement before the U.S. House of Representatives Committee on Education and the Workforce
Subcommittee on Higher Education and Workforce Development

Hearing Title: Empowering Students and Families to Make Informed Decisions on Higher Education

Data on the Federal Student Loan Program

Jason D. Delisle
Resident Fellow

Wednesday, May 24, 2017

Introduction

Good morning, Chairman Guthrie, Ranking Member Davis, and distinguished Members of the Subcommittee, and thank you for giving me the opportunity to testify about data on the federal student loan program.

My name is Jason Delisle and I am a resident fellow with the Center on Higher Education Reform at the American Enterprise Institute (AEI), a non-profit, non-partisan public policy research organization based here in Washington, DC. My comments today are my own and do not necessarily reflect the views of AEI.

The federal government's Direct Loan program dominates the student-loan market today, issuing 90 percent of all loans made across the country each year.[1] Students pursuing everything from short-term certificates to master's degrees qualify on a no-questions-asked basis for nearly $100 billion of these loans every year at terms more generous than most private lenders would offer.

The federal role in higher-education lending has grown ever since lawmakers enacted the first loan program under the National Defense Education Act of 1958. The Higher Education Act of 1965 expanded access to loans to more colleges and students through the Guaranteed Student Loan Program, but the interest rate subsidies it provided were restricted to students from low-income families. In 1980, Congress created a loan program for parents of undergraduates (Parent PLUS), and then in 1992, eliminated annual and lifetime borrowing limits for those loans. That year, lawmakers also authorized the Unsubsidized Stafford Loan program, which allows all undergraduate students to borrow federal loans regardless of their financial circumstances. In 2006, Congress created the Grad PLUS loan program, which removed limits on the amount graduate students could borrow.

This expansion, along with rising college costs and increases in student enrollments, have led to a rapid increase in the stock of outstanding debt in recent years. Now at $1.3 trillion, the student loan program rivals the Federal Housing Administration's largest mortgage program in size.[2]

Options to repay these loans have also exploded in number and in generosity. These include repayment plans with fixed or graduated monthly payments spread over 10 to 30 years, and a variety of plans with payments set according to borrowers' incomes (which I collectively refer to as Income-Based Repayment, or IBR). Payments in IBR are set at 10 percent of adjusted gross income after an exemption of 150 percent of the federal poverty guidelines ($18,090 for a single person). Unpaid balances are forgiven after 20 years, or 10 years for borrowers working in a nonprofit or government job.[3] While enrolled in any of these plans, borrowers can qualify for several types of deferments and forbearances that allow them to suspend payments for years.

Despite the ever-expanding benefits, loan types, and repayment options, delinquency and default rates suggest that the current system is not working. Over 8 million people are in default on their federal student loans today, a number that has continued to grow year after year, even though the country is now many years into an economic expansion with low rates of unemployment.[4] Estimates also suggest that over 40 percent of all borrowers whose loans have come due are in default, are delinquent, or are in forbearance or deferment.[5] Nearly one in four federal student

loans issued to undergraduates this year is eventually expected to enter default.[6]
Given the size, scope, and complexity of the student loan program, the data that the federal government makes available leaves much to be desired. While there have been some improvements in recent years, the data form only a patchwork rather than a complete picture. Many key questions about the federal student loan program cannot be answered with the data available to the public and researchers. Improving the quantity and quality of the data is imperative for ensuring that the program works well for all types of borrowers and does not waste taxpayer dollars. I'll provide two cases to illustrate this point.

The available information points to an ongoing student-loan default crisis, but without better data about borrowers after they leave school, it is nearly impossible to fully understand the program or even begin to develop solutions. For example, reports suggest that many of the borrowers who default never even make the first payment on their loans.[7] But it is impossible to analyze the data to better understand this issue. Some statistics also imply that a large share of defaulted loans are held by borrowers who left school over a decade ago, but many borrowers also leave default quickly and return to good standing.[8] The lack of data means we do not understand what explains those very different patterns, and how policymakers might tailor solutions to these two groups.

Without better data, the government will continue to underestimate the cost of the loan program. Consider that when the Obama administration dramatically expanded the IBR program in 2010, Congress and the public were told this change might cost around $700 million a year.[9] We are now learning from the Government Accountability Office and other federal agencies that the costs are substantially larger, running in the billions.[10] And it is still unclear which types of borrowers (dropouts, graduate students, the unemployed, etc.) are benefiting from this program and its recent expansions. One can only wonder whether Congress and the Obama administration would have pursued different policies if they had known then what we know now.

The key problem is that the data are running far behind the policy, the exact opposite of how things should operate. Things are getting better: federal agencies have been working to make more data available to researchers and the public. But there are still dangerous blind spots in the information accessible to those outside the federal government.

Below I explain the type of questions that the available data can answer about the federal student loan program and, more importantly, which questions it cannot. Finally, I offer a few recommendations for how the government can improve the data it provides to researchers about the student loan program.

Office of Federal Student Aid

The Department of Education's Office of Federal Student Aid (FSA) offers two main categories of publicly-available data on the student loan program. Under one category, FSA provides information on the portfolio of all outstanding federal student loans—a measure of the "stock" of debt. The other category provides college and university-level data on quarterly loan disbursals, short-term cohort default rates, and typical monthly loan payments for graduates of some programs.

Loan Portfolio Summary Statistics

Since late 2013, FSA has provided summary statistics about the outstanding portfolio of federal student loans on its website. The statistics provide information on the number of borrowers and total loan balances in various states of repayment.[11] For example, the statistics show the total number of borrowers and outstanding debt enrolled in the IBR repayment plans and how those figures have changed since 2013.

But because the data are aggregate statistics and not student-level, they are only minimally useful. For example, we do not know how many borrowers enrolled in IBR are making $0 monthly payments; how many students' loan balances are growing because they are not covering interest on the debt (i.e., "negative amortization"); how many borrowers are using IBR to get out of defaulted; or how many borrowers in IBR have debt from graduate school.

The FSA summary statistics on the loan portfolio also include information about repayment status, such as the number of borrowers and amount of debt in active repayment, default, forbearance, hardship deferment etc. Therefore, the data reveal that some 40 percent of borrowers whose loans have come due are not making regular payments because they are in default, forbearance, or deferment. We also know that there are approximately 8 million borrowers in default on their federal loans currently, up from 6.5 million in 2013.[12] FSA has also recently begun to include information about the flow of borrowers into default status and about default resolutions.[13] Nevertheless, the information is still quite limited and tells us nothing about borrowers, how long they've been in default, what share of borrowers who used forbearance before defaulting, etc.

Title IV Volume Reports by School

The institution-level data that FSA provides on student loans offer a different perspective. These spreadsheets display how much and how many of each type of loan (Subsidized Stafford, Parent PLUS, etc.) FSA disburses to each college or university.[14] If one wants to know the total amount of Parent PLUS loans disbursed at a particular university, how many parents received them, and how those figures have changed over time, this data source provides that information. Yet because the data include only aggregate statistics, these reports are not very useful for answering questions other than the types I have listed here.

Cohort Default Rates by School

The other institution-level data on student loans on the FSA website provide information about cohort default rates.[15] These rates are part of an accountability regime for eligibility for federal student aid. Institutions whose students experience high rates of default may not participate in the aid programs.

These data are unique relative to what the federal government makes available regarding student loans because they are institution-level *and* longitudinal. That is, they provide information about what happens over time to borrowers and loans issued at a particular school. They cover only three years of loan repayment for each cohort, however. One cannot know how many students

default in years four, five, or later. (Department of Education documents show that the lifetime default rate on undergraduate loans exceeds 20 percent.[16]) Moreover, the data cannot answer even basic questions about students who default, such as the share of defaulters who completed their program versus the share who dropped out.

Gainful Employment Rule

FSA also provides data related to the "gainful employment" (GE) regulations issued under the Obama administration. Like the cohort default rate data, these data provide limited information about what happens to borrowers and loans after students leave school. The first set of GE data was released in late 2016 and early 2017, and mostly covers students who graduated in 2011 or 2012 as well as their most recent available earnings outcomes.[17] The Department of Education is expected to release data for subsequent cohorts on an ongoing basis.

The GE data are unique in that they provide information on student outcomes for individual *programs*, whereas most federal data sources summarize student outcomes only at the *institution* level. GE data include typical debt payments for program graduates, from which approximate debt balances at graduation can be inferred. The data also include mean and median annual earnings of graduates three to four years after completion, and the number of covered students in each program.

While FSA provides GE data for certificate programs at all types of institutions, it provides data on degree programs *only at for-profit colleges* (in other words, it excludes degree programs at public and private not-for-profit colleges). Debt, earnings, and cohort sizes are only reported for completers, meaning we know nothing about dropouts from GE programs. Finally, the debt statistics reported in GE include some private loans; it is impossible to separate out federal loans in the data. Thus the data do not present a clean picture of borrowing and repayment patterns with respect to federal student loan policy. The information reflects borrowing in the private market as well.

The College Scorecard

One signature data initiative of the Obama administration's Department of Education was the College Scorecard, intended as an informational tool for prospective students. Separate from the data that FSA provides, the Scorecard offers additional data on loan performance after students leave school, broken down by college and university, but not by programs within schools.[18] The Scorecard includes data on undergraduate student cohorts dating back to 1996; however, not all variables are available for all years.[19] The data are available publicly but housed separately from the FSA data. The Scorecard also excludes graduate and professional students; it covers only baccalaureate and sub-baccalaureate credentials.

The most important contribution of the College Scorecard data to our understanding of student loans is the repayment rate, defined as the percentage of borrowers from a particular university who have repaid at least one dollar of their original federal student loan balance.[20] Loan repayment rates for one, three, five, and seven years after borrowers leave school are available

via the College Scorecard. Repayment rates are also available for subcategories of students along various dimensions, including completion status, family income, dependency status, Pell Grant eligibility, gender, and first-generation status.

This is important because many borrowers reduce or postpone payments by switching repayment plans or using forbearance options. These borrowers are not captured in statistics on loan defaults, even though they too are not making progress on paying down their debts. The Scorecard data therefore provide a better measure of loan repayment than the default rates, indicating the institutions where students might be struggling to repay.

Aside from repayment rates, the Scorecard also reports students' median debt levels upon entering repayment, along with median debt by the subcategories mentioned above. Debt levels at the 10th, 25th, 75th, and 90th percentiles are available as well.[21] These are important data for understanding how much students borrow at particular institutions, but because it is aggregate data, it does not provide information about students who borrow the most and those who do not borrow at all.

National Center for Education Statistics Surveys

The National Center for Education Statistics (NCES), which is part of the Department of Education, collects and analyzes education data, including data on the federal student loan program. NCES administers several large, nationally-representative surveys of undergraduate and graduate students. These include the National Postsecondary Student Aid Study (NPSAS) and two "spin-off" datasets—the Beginning Postsecondary Study and the Baccalaureate and Beyond Study—that use a subsample from the NPSAS for longitudinal surveys.

National Postsecondary Student Aid Study

The NPSAS provides a comprehensive, student-level dataset about financial aid, demographics, and college prices.[22] NCES builds the dataset every four years (soon to be every two years) by sampling institutions and then students attending the selected institutions.[23] The data are collected from federal databases, institutions' administrative records, and student and family interviews.[24] Student loan data are collected primarily from the National Student Loan Data System (discussed later in this testimony) and is highly reliable; it is not self-reported by universities or borrowers and therefore less subject to error.

The NPSAS is a valuable source of data on federal student loans mainly because it provides student-level data, but also because of other information it includes about students. For example, it is possible to determine how much students from low-income families borrow in federal student loans compared with students from high-income families. Other permutations include borrowing patterns among dependent and independent students, or debt levels of students pursuing short-term certificate programs. It is also possible to examine how these data have changed over time by using earlier NPSAS datasets. And because the NPSAS data include nationally-representative samples for both undergraduate and graduate populations, it offers the most comprehensive data on graduate school debt available.

The NPSAS is not, however, without shortcomings. For example, NPSAS is nationally representative, but its sample size is too small to be used for state-level or institution-level analyses.[25] And when the data are filtered across multiple categories, the small sample sizes provides less reliable information about the larger population. For example, the dataset includes rich information about borrowing among graduate students. But using it to examine how much graduate students who pursue *education degrees* borrow, broken out by *race and type of institution*, results in a small subsample of students. The data about their debt will thus be of limited use.

The NPSAS provides only a snapshot in time of the undergraduate and graduate populations, resulting in another limitation. Data about how students repay and manage their federal student loans cannot be included because the survey only covers students who are *currently enrolled*. For example, researchers can use the NPSAS to figure how much debt different categories of students have when they complete their degree (a variable included in the dataset) but not how much debt *students have when they drop out* (there is no way to determine if a first-year student eventually drops out).

The NPSAS dataset is available only to researchers who obtain a restricted-use license from NCES. To obtain the license, researchers must agree to follow certain protocols to protect any potentially personally identifiable information in the dataset. However, NCES maintains a data retrieval tool that is available to the public, called DataLab, which allows users to perform some, but not all, of the analyses possible with a restricted-use license.

Beginning Postsecondary Study and Baccalaureate and Beyond Study

NCES uses the NPSAS to obtain baseline data for longitudinal studies that follow a subset of students who participated in the survey. These include the Beginning Postsecondary Students (BPS) longitudinal study[26] and the Baccalaureate and Beyond (B&B) longitudinal study.[27] Because they are derived using the NPSAS—with an even smaller subsample—these studies are subject to the same limitations as the NPSAS with respect to data on student loans, except that they add a longitudinal element. The studies are also conducted less frequently than the NPSAS.

The BPS study follows a sample of first-time students at the end of their first, third, and fifth academic years.[28] The data provide researchers a way to examine how much students borrow during their academic careers and compare it with a number of other student characteristics and enrollment patterns (graduation, transfers, part-time status). But the timeframe of the study is too short to allow for data on student loan repayment patterns.

The B&B study uses a sample of students included in the NPSAS who complete bachelor's degrees. It offers insight into post-graduation income and loan repayment, as well as graduate school enrollment and debt.[29]

The B&B provides a unique source of data for tracking students' borrowing and repayment patterns. But a series of issues severely limits its usefulness. The timeframe it covers is short relative to the timeframe over which students are likely to repay loans. Students who dropped out or completed two-year degrees and certificates—a group that makes up the majority of

students—are excluded from the study. Finally, sample sizes are small, making it difficult to examine student debt data by subgroup.

Non-Public Federal Agency Data on Student Loans

Thus far, my testimony has covered data sources for federal student loans available to the public or researchers via a restricted-use license. There are two other major data sources that federal agencies use to study the loan program but are not currently available outside the government. These include a sample file extracted from the Department of Education's National Student Loan Database System (NSLDS), a record-keeping system that tracks the status of individual loans and borrowers, and a dataset developed by the U.S. Treasury Department that links NSLDS data to Internal Revenue Service (IRS) tax records for a sample of borrowers.

The Department of Education, the Congressional Budget Office, and the Treasury Department have periodically issued statistics and reports using these data, providing some of the most insightful analysis on the federal student loan program to date.[30] Treasury and the Department of Education also use the data to inform policy proposals and formulate budget estimates. These datasets overcome many of the limitations noted for other sources, but they are not without their flaws. The analyses are infrequent, cannot be scrutinized by third parties, and often serve to answer a limited set of questions motivated by the policy agenda of whatever administration happens to control the Executive Branch. Below I briefly discuss the two datasets, keeping in mind that because the data are not available outside the government, information about them must be gleaned from reports.

National Student Loan Data System and Sample File

NSLDS is the Department of Education's central database for tracking federal student aid programs.[31] Institutions of higher education, lenders holding federally-backed loans, loan servicers, and the Department of Education all submit information to the database. Students and borrowers can also access it to view their loan balances, loan statuses, and disbursements.

Aside from being a tool for borrowers, NSLDS provides the most comprehensive source of data on the federal student loan program. It includes student- and borrower-level data that covers the entire life of a borrower's loans. It includes records and dates for each loan's status changes such as when the loan is disbursed; when it is in the in-school period; when it is paid in full; or if it enters repayment, default, deferment, or forbearance. It therefore provides information on patterns of repayment over long periods of time. NSLDS also includes information on the repayment plan for borrowers under the Direct Loan program.

One major limitation of the data is that NSLDS does not track cash flow.[32] It reports a borrower's loan status, but not his monthly payments over time. Such information must be inferred from annual changes in the borrower's loan balance. Finally, the NSLDS only includes information on a borrower's loans, other federal student aid, and the school he attended. It does not include other information about the borrower during repayment, such as income, employment status, etc. Those are key variables needed to more fully understand repayment

patterns.

The Congressional Budget Office, the Department of Education, and the Treasury Department use an annual subsample of loan and borrower records from NSLDS to inform cost estimates, develop policy, and conduct other analyses. Thus, a sample file from the database is generated each year that could allow outside entities to conduct similar analyses. Until recently, no effort was made on the part of the Department of Education to release a version of this file to outside entities. However, earlier this year the Obama administration announced that it had been working to "create a privacy-protected, public-use microdata file from the National Student Loan Data System (NSLDS) that can facilitate valuable research and other studies of higher education."[33] It is not clear what the status of this effort is under the Trump administration.

U.S. Treasury Department NSLDS-IRS Match

In recent years, the Treasury Department has developed a dataset that adds to the data housed in the NSLDS. The Department matched a random sample of NSLDS records to de-identified tax data from the IRS Compliance Data Warehouse.[34] This dataset combines individual-level data on student loans and repayment with information contained in each borrower's tax records, such as income, marital status, family size, and use of tax credits and deductions. Thus, it represents the most comprehensive and detailed source of information on the federal student loan program and borrowers. Income data from tax records matched to a borrower's student loan history is especially useful for analyzing the impact of the IBR plan. It also allows researchers to analyze the burden of student loan payments relative to a borrower's income over a long time period (1999 to 2014).

Researchers have produced two important studies using this data. One revealed how the demographics of borrowers have changed over time, skewing more towards "non-traditional students" who are older, are independent of their parents, have lower incomes, and are less likely to enroll full-time.[35] These borrowers are also more likely to attend non-selective institutions, particularly public two-year colleges and for-profit institutions. The study found that loan performance among this demographic of borrowers has been extremely poor in recent years, with high rates of default, negative amortization, and reliance on benefits that allow borrowers to postpone repayment for long periods of time without defaulting. The other study that used these data provided some of the first analyses of borrowers who use IBR, suggesting that the program helps reduce defaults.[36]

Conclusion and Recommendations

To conclude, I will reiterate that the existing, publicly available data on federal student loans are limited in two main ways. First, they are often not broken down at the student level and therefore provide only high-level summary statistics. Second, the data generally reflect snapshots in time and are not longitudinal, meaning information about what happens to loans and borrowers after the money is disbursed is not observable. The best available data sources—those that are student-level and track borrowers over time—are derived from infrequent surveys with small sample sizes and short time horizons.

Fortunately, a ready-made solution exists to these problems. The Department of Education, in cooperation with Treasury, could make the two datasets that the federal government already compiles—the NSLDS sample file and the NSLDS-IRS matched sample file—available in the same manner as other restricted-use datasets, like the NPSAS. Researchers and organizations that agree to follow the National Center for Education Statistics privacy-protection rules could obtain the same de-identified data the government uses to analyze the federal student loan program and formulate policy. While that falls short of full availability, it is likely necessary to address concerns over privacy.

I should also note that this solution does not fill in all of the gaps in the data. For example, it will not provide institution-level loan performance information. But it need not do so to vastly improve what we know about the federal loan program and help us discern what policies could strengthen it for borrowers and taxpayers alike.

The NSLDS data also has shortcomings that stem from its primary purpose as a database for reporting and tracking the status of students and their loans. It was not designed to collect data for research, analysis, and policy development. That is why, for example, it does not contain information about borrowers' monthly payments or other cash flow information. These issues suggest a further recommendation. In addition to releasing the NSLDS sample file, the Department of Education could overhaul its data collection systems so that they capture information that researchers and policymakers need to better understand the program—and how it affects borrowers.

Moreover, to avoid unnecessary controversy, I want to emphasize that these efforts are distinct from any broader agenda to create a student unit-record or link federal aid to performance metrics beyond existing policies. Congress need not repeal the ban on a student-unit record system for federal agencies to release the NSLDS sample file and the NSLDS-IRS matched file to researchers. The federal government can make the existing datasets available regardless of the merits of those broader goals and lawmakers' views on them.

Finally, releasing and improving these existing datasets will require that the Department of Education and the Treasury Department commit scarce time and resources to the goal. Private loan servicing companies, with whom the Department of Education has contracted to operate much of the loan program, will also need to commit additional time and resources and submit to the government a significant amount of new information about borrowers to improve the data. This is where Congress could be helpful, both in making its interest known in such a project and in ensuring that sufficient resources are provided to the agencies. Far too much is at stake for lawmakers to be satisfied with the existing data. Taxpayers and students deserve better than policies developed through anecdotes and assumptions, and these recommendations offer a straightforward path to get there.

Notes

[1] For more information about the history and expansion of the federal student loan program, see Jason D. Delisle, *Private in Name Only: Lessons from the Defunct Student Loan Program,* American Enterprise Institute, February 2017, www.aei.org/wp-content/uploads/2017/02/Private-in-Name-Only.pdf.

[2] Board of Governors of the Federal Reserve System, "Mortgage Debt Outstanding," March 2017, www.federalreserve.gov/econresdata/releases/mortoutstand/current.htm.

[3] For more information about federal income-driven repayment plans, see U.S. Department of Education, Federal Student Aid, "Income-Driven Plans," https://studentaid.ed.gov/sa/repay-loans/understand/plans/income-driven.

[4] Author's calculations using U.S. Department of Education, "Federal Student Loan Portfolio," 2016, https://studentaid.ed.gov/sa/about/data-center/student/portfolio.

[5] Rohit Chopra, "A Closer Look at the Trillion," Consumer Financial Protection Bureau, August 5, 2013, www.consumerfinance.gov/about-us/blog/a-closer-look-at-the-trillion/.

[6] U.S. Department of Education, *Student Loans Overview: Fiscal Year 2017 Budget Proposal,* https://www2.ed.gov/about/overview/budget/budget17/justifications/q-sloverview.pdf.

[7] Josh Mitchell, "More Than 40% of Student Borrowers Aren't Making Payments," *Wall Street Journal,* April 7, 2016, www.wsj.com/articles/more-than-40-of-student-borrowers-arent-making-payments-1459971348.

[8] Author's calculations using U.S. Department of Treasury, "Report on Initial Observations from the Fiscal-Federal Student Aid Pilot for Servicing Defaulted Student Loan Debt," 2016, https://www.treasury.gov/connect/blog/Documents/student-loan-pilot-report-july-2016.pdf.

[9] White House Office of Management and Budget, "The Budget for Fiscal Year 2011: Summary Tables," www.gpo.gov/fdsys/pkg/BUDGET-2011-BUD/pdf/BUDGET-2011-BUD-28.pdf.

[10] Government Accountability Office, *Education Needs to Improve Its Income-Driven Repayment Plan Budget Estimates,* November 2016, www.gao.gov/assets/690/681064.pdf.

[11] U.S. Department of Education, "Federal Student Loan Portfolio."

[12] Ibid.

[13] U.S. Department of Education, "Default Rates," https://studentaid.ed.gov/sa/about/data-center/student/default.

[14] U.S. Department of Education, "Title IV Program Volume Reports," https://studentaid.ed.gov/sa/about/data-center/student/title-iv.

[15] U.S. Department of Education, "Official Cohort Default Rates for Schools," September 28, 2016, https://www2.ed.gov/offices/OSFAP/defaultmanagement/cdr.html.

[16] U.S. Department of Education, *Student Loans Overview.*

[17] U.S. Department of Education, "Gainful Employment Information," https://studentaid.ed.gov/sa/about/data-center/school/ge.

[18] U.S. Department of Education, "College Scorecard Data," https://collegescorecard.ed.gov/data/.

[19] U.S. Department of Education, "Data Documentation for College Scorecard," January 2017, https://collegescorecard.ed.gov/assets/FullDataDocumentation.pdf.

[20] Ibid.

[21] In January 2017, the Department of Education revealed that a coding error had significantly inflated repayment rates reported in the Scorecard. The data have since been fixed. See U.S. Department of Education, "Updated Data for College Scorecard and Financial Aid Shopping Sheet," January 13, 2017, https://ifap.ed.gov/eannouncements/011317UpdatedDataForCollegeScorecardFinaidShopSheet.html.

[22] U.S. Department of Education, "National Postsecondary Student Aid Study," https://nces.ed.gov/surveys/npsas/.

[23] NCES conducted the NPSAS every three years between the 1986-87 and 1995-96, and every four years since. The 2015-16 survey data will be available later this year.

[24] NCES has announced plans to add a more limited NPSAS two years after it releases the full NPSAS using only administrative data.

[25] In some years, however, NCES oversampled several states and plans to oversample states again starting with the 2017-18 analysis. Notwithstanding those exceptions, the existing set of NPSAS data are only suited for national-level analyses.

[26] National Center for Education Statistics, "Beginning Postsecondary Students," https://nces.ed.gov/surveys/bps/.

[27] National Center for Education Statistics, "Baccalaureate & Beyond*,"* https://nces.ed.gov/surveys/b&b/.

[28] The current BPS (12/17) surveyed students in 2012, 2014, and will survey them for the final time in 2017. Past cohorts are BPS 90/94 (8,000 students), BPS 96/01 (12,000 students) and BPS 04/09 (16,700 students).

[29] The first cohort was drawn from the 1993 NPSAS, with follow-ups in 1994, 1997, and 2003 (11,000 students).

Another cohort was drawn from the 2000 NPSAS, with a follow up in 2001 (10,000 students). The most recent cohort was drawn from 2008 NPSAS, with follow ups in 2009 and 2012, and one planned for 2018 (19,000 students).

[30] For example, see Congressional Budget Office, *January 2017 Baseline Projections for the Student Loan Program,* January 2017, www.cbo.gov/sites/default/files/recurringdata/51310-2017-01-studentloan.pdf

[31] U.S. Department of Education, "National Student Loan Data System*,*" www.nslds.ed.gov/nslds/nslds_SA/.

[32] Congressional Budget Office Working Paper, *Guaranteed Versus Direct Lending: The Case of Student Loans,* June 2007, www.cbo.gov/sites/default/files/110th-congress-2007-2008/workingpaper/2007_09_studentloans_0.pdf.

[33] Lynn Mahaffie, "More Transparency in Higher Education Will Help Improve Student Outcomes," U.S. Department of Education, January 12, 2017, https://blog.ed.gov/2017/01/transparency-higher-education-will-help-improve-student-outcomes/.

[34] For a full description of this process, see Adam Looney and Constantine Yannelis, *A Crisis in Student Loans? How Changes in the Characteristics of Borrowers and the Institutions They Attended Contributed to Rising Loan Defaults,* The Brookings Institution, September 2015, www.brookings.edu/wp-content/uploads/2016/07/PDFLooneyTextFallBPEA.pdf.

[35] Looney and Yannelis, *A Crisis in Student Loans?*.

[36] Holger Mueller and Constantine Yannelis, "Students in Distress: Labor Market Shocks, Student Loan Default, and Federal Insurance Programs" (working paper, National Bureau for Economic Research, Cambridge, MA, March 2017), www.nber.org/papers/w23284.