

## Informing Policy Options through Improved Student Loan Data

### Summary

*There has been widespread discussion both within and outside government about the shortcomings of the current Federal Direct Student Loan program and potential changes to the system. If policymakers are serious about reworking the student loan system, their positions should be informed by sound non-partisan research. Unfortunately, data limitations currently inhibit most potential researchers outside of government from conducting analysis and contributing insights in this arena. While data sufficient for making informed policy decisions about student loans do exist, they need to be better shared to realize their potential.*

With the arrival of a new Congress and a new administration, change may be in store for current and former students who have borrowed to pay for higher education expenses. On the campaign trail, Donald Trump indicated a desire to ease the debt burden carried by former students, and Republicans at times have sought to rein in the amount of funds being loaned by the government. Add to this environment two high-profile reports from the Government Accountability Office – [one](#) that pegged the taxpayer cost of the popular Income-Based Repayment program at a shocking \$108 billion and [another](#) that found that in the last year 114,000 people had their federal benefits garnished to pay their student loan debt – and the system seems poised for a change of some sort. So while student loan reform is hardly a sure thing, an educated guess suggests we could soon be seeing legislative and/or administrative action on this front.

Of course, if and when policymakers take up this tremendously important issue, they should be equipped with the best tools befitting a topic of such consequence. Reliable, accurate data and the analytics around it will be key to crafting any new approaches to student lending. Unfortunately, data limitations threaten to undermine such efforts for dispassionate, fact-based analysis. That is not to say that the data does not exist. On the contrary, there is a treasure trove of recorded data on lending, servicing, and repayment that would facilitate a hard look at student lending. The problem is that it does not all lay under one big X on the map.

*This policy brief was prepared by Kyle Shohfi, a research associate at the Golub Center for Finance and Policy. He can be reached at [shohfi@mit.edu](mailto:shohfi@mit.edu).*

MIT Golub Center for Finance and Policy  
One Broadway  
Suite 1375  
Cambridge, MA 02142  
(617) 324-7367  
[gcfp@mit.edu](mailto:gcfp@mit.edu)  
[gcfp.mit.edu](http://gcfp.mit.edu)  
@MITCFP

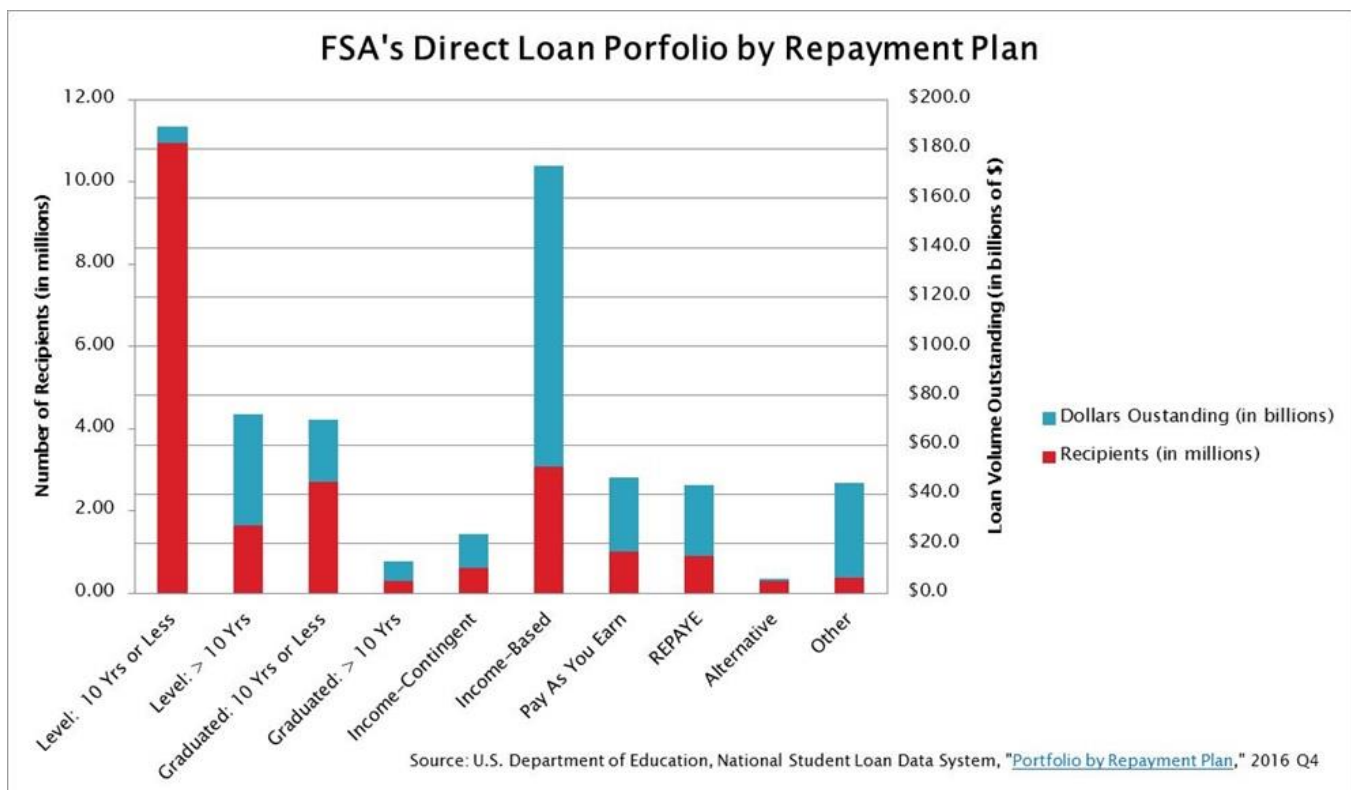
The student loan process is unlike any other, with many participants filling the various lender roles in the extension of credit. For instance, students submit their financial aid applications to the Department of Education, which lends them the money. But the funds are actually disbursed via the students' financial aid offices at their schools. When borrowers enter the repayment period, they interact with yet another entity when they make payments (or not) to their loan servicer(s), a third-party firm. With each of these distinct players come proprietary data sets about borrowers and their loans that, in isolation from each other, do little to inform a broad understanding of the federal student loan program overall.

The government's most important database on student loans is the National Student Loan Data System (NSLDS) operated by the office of Federal Student Aid. The NSLDS contains data on all individuals who receive aid from FSA, such as their school(s) and grant and loan amounts. Once a borrower begins repayment, information on loan balance and status is updated periodically by servicers. From this data, FSA publishes [reports](#) on aid volume by institution, portfolio by loan status, portfolio by repayment plan, and other aggregate statistics. Unfortunately, the data driving these reports has traditionally been challenging to access for researchers

outside of government, limiting the potential to conduct independent program and policy analysis. The data also has certain limitations such as not revealing how loans tend to move between various repayment options through their lifetimes or how each option might influence propensity to default.

Within government, the Department of Education uses a four percent sample of loan and borrower records linked to NSLDS data to inform budgetary and subsidy cost estimates. A similarly sized dataset was accessed and analyzed for delinquency and default incidence in 2015 by Adam Looney of the U.S. Department of the Treasury and Stanford's Constantine Yannelis. Their report, "[A Crisis in Student Loans? How Changes in the Characteristics of Borrowers and the Institutions they Attend Contributed to Rising Loan Defaults](#)," benefited from linking data on student borrowing to earnings data from tax records. Ideally, data of this sort will become increasingly available as the Department of Education continues to improve its systems.

Fortunately, there are signs that such improvements could be on the way. The Department of Education is in the early stages of procuring a [new federal loan servicing system](#), mainly intended to improve the borrower experience. Such a portal will allow borrowers to access



information, make payments, and apply for benefits across all of their loans and servicers in a single place. Moreover, the new system could also potentially facilitate the transaction-level detail of loan performance that has thus far proven elusive. While details pertaining to the new system are scant, there is hope that the new system will provide the government with the broad view it needs to inform sound policy decisions.

Furthermore, a better flow of data between servicers and the government will have benefits that reach even farther than just Capitol Hill. Of course, the executive branch would be able to craft legislative proposals to pitch to the Congress, but the implications of enhanced data extend to more than just policy changes. On an operational level, FSA would stand to benefit from more precise, up-to-the-minute data on its loan portfolio. Just imagine how cash flow forecasting and related budgetary projections could be improved – with continuously updated, longitudinal data, FSA could model future repayment by looking at a combination of a borrower’s repayment history and macroeconomic factors.

Symmetrically, enhanced data flow between the government and loan servicers would also help servicers do their jobs better. Many studies have found that certain personal characteristics – such as age, marital status, family income, family education levels, gender, and household size – are useful predictors of repayment behavior (and all are reported on a student’s FAFSA). By combining continuously updated repayment data with personal data about the borrowers, dynamic models could trigger alerts when a borrower appears to be in danger of default. FSA could then notify the appropriate servicer to reach out to the borrower and help them manage their debt to avoid default.

Most importantly, all of these advantages would serve the interests of the primary stakeholders: current and former students. By improving the flow of data and maximizing its analytical potential, the government can

craft better policies, support smoother operations, and encourage better customer service from loan servicers and debt collectors. Regardless of what specific changes might be ahead, unlocking the potential of relevant data will be key.

## References

- Government Accountability Office. (2016). Education Needs to Improve Its Income-Driven Repayment Plan Budget Estimates. (GAO Publication No. 17-22). Washington, D.C.: U.S. Government Printing Office. Available at <http://www.gao.gov/products/GAO-17-22>
- Government Accountability Office. (2016). Improvements to Program Design Could Better Assist Older Student Loan Borrowers with Obtaining Permitted Relief. (GAO Publication No. 17-45). Washington, D.C.: U.S. Government Printing Office. Available at <http://www.gao.gov/products/GAO-17-45>
- 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. *Brookings Papers on Economic Activity*, 2015(2), 1-89. Available at <https://www.brookings.edu/wp-content/uploads/2015/09/LooneyTextFall15BPEA.pdf>

You can view other GCFP reports at:  
<http://gcfp.mit.edu>