American Economic Journal: Applied Economics 2023, 15(1): 1–25
https://doi.org/10.1257/app.20200362

The Impact of Income-Driven Repayment on Student Borrower Outcomes†

By Daniel Herbst*

In the United States, most student loans follow a fixed payment schedule that falls early in borrowers' careers. This structure provides no insurance against earnings risk and may increase student loan defaults. Income-driven repayment (IDR) plans are designed to help distressed student borrowers by lowering their monthly payments to a share of income. Using random variation in a loan servicer's automatic dialing system, I find that IDR reduces delinquencies by 22 percentage points and decreases outstanding balances within eight months of take-up. I find suggestive long-run impacts on borrower credit scores, mortgage-holding rates, and other measures of financial health. (*JEL* G23, G51, H52, I22)

In the United States, over one million student borrowers default each year, and millions more struggle with low homeownership (Mezza et al. 2020; Bleemer et al. 2017) and poor financial health (Gicheva and Thompson 2015). Many blame this crisis on inflexible student loan contracts, which require fixed, fully amortized payments that fall on borrowers early in their careers and provide no insurance against income shocks (Barr et al. 2017). The policy response has been income-driven repayment (IDR) programs, which set monthly minimum payments to a fixed portion of borrowers' income until debt is repaid or some forgiveness period has been reached. US enrollment in IDR has tripled since 2014 and more than $500 billion in debt is currently repaid through the program (US Department of Education 2020b).

Even as IDR enrollment continues to rise, its effects on social welfare are largely unknown. IDR can improve borrowers' liquidity by aligning their repayment burden with the wage returns to college. It may also provide insurance against lifetime

*Department of Economics, University of Arizona (email: dherbst@email.arizona.edu). David Deming was coeditor for this article. I am extremely thankful to my advisers Will Dobbie and Ilyana Kuziemko for their guidance on this project, as well as Henry Farber, Nathan Hendren, Alan Krueger, David Lee, Adam Looney, Alexandre Mas, Christopher Neilson, and Cecilia Rouse, who provided invaluable advice throughout its development. I also benefited from the helpful comments of David Arnold, Barbara Biasi, George Bulman, Felipe Goncalves, Steve Mello, David Price, Maria Micaela Sviatschi, and seminar participants at the APPAM Annual Research Conference, the CFPB Research Conference, the Federal Reserve Board, the IZA Economics of Education Workshop, the Jain Family Institute, Kansas State University, the MIT Golub Center, the National Academy of Education Research Conference, the National Tax Association Research Conference, Princeton University, the RAND Corporation, the University of Arizona, and Vanderbilt University. Financial support was provided by the Princeton Industrial Relations Section, the Jain Family Institute, the MIT Golub Center for Finance and Policy, and the National Academy of Education Spencer Dissertation Fellowship.

†Go to https://doi.org/10.1257/app.20200362 to visit the article page for additional materials and author disclosure statements or to comment in the online discussion forum.

income risk by forgiving the balances of persistently low earners. At the same time, IDR may distort labor-supply incentives, inducing borrowers into lower-paid careers or unemployment. These behavioral responses carry an efficiency cost that increases the long-run fiscal burden of the student loan program.

In this paper, I use administrative data from a large student loan servicing company to estimate IDR's causal effects on borrower outcomes. The data I use link monthly loan records from a large loan servicer (LLS) to credit bureau information from TransUnion, allowing me to investigate both short-term repayment behavior and long-term proxies for homeownership and financial health (LLS 2018). The data include monthly records of student loan balances, payments, delinquencies, and repayment plan enrollment; annual records of bankruptcies, credit scores, mortgages and credit cards; and borrower-level information on demographics, college attendance, and contact histories.

To estimate the causal effects of IDR on short-run outcomes such as delinquencies and balances, I use an instrumental-variables design exploiting the quasi-random assignment of loan-servicing calls. I also investigate potential impacts on long-run outcomes by comparing IDR enrollees to non-enrollees before and after receiving servicing calls. Trends in outcomes are very similar for these two groups prior to the call, and I find no impact of "placebo" calls prior to the true enrollment event.

I find evidence of large and persistent financial benefits to borrowers. IDR reduces delinquency rates (i.e., late payments) by 22 percentage points within eight months of take-up, relative to a pre-call mean of 66 percent. This increased repayment likelihood leads IDR borrowers to pay down \$36 more debt each month than standard borrowers, despite facing \$172-lower monthly minimums. For comparison, average balances *grew* by \$4 in the month prior to the servicing call. Long-term analysis provides suggestive evidence of lasting effects on financial outcomes and homeownership. IDR enrollees have credit scores that are 5.6 points higher and are 9 percent more likely to hold a mortgage compared to non-enrollees three years after the servicing call. Relative to non-enrollees, IDR enrollees are 2 percentage points more likely to move to a higher-income zip code.

Despite its persistent effects on financial outcomes, the actual increase in cash-on-hand through IDR is remarkably short-lived. Likely due to the burdensome income-recertification process, most borrowers fail to re-enroll in IDR after one year and quickly return to their pre-call repayment patterns. This pattern suggests IDR's financial benefits operate through a short-term liquidity channel. IDR was designed to provide comprehensive, "equity-like" restructuring of student-debt contracts, but in practice it serves only as a temporary cash infusion to distressed borrowers. This increased cash-on-hand has large and lasting benefits, but if policymakers want IDR to provide more persistent income smoothing or insurance against lifetime-earnings risk, they must reform recertification requirements and loan forgiveness rules.

This paper complements a small but growing literature on student loan contracts and IDR. As early as Friedman (1955), many researchers have documented the benefits of income-contingent student debt (Chapman 2006; Barr et al. 2017). A related stream of literature documents the revenue implications of various loan contracts by simulating repayment paths and loan forgiveness-incidence across different populations (Lucas and Moore 2010; Johnston and Barr 2013; Britton, van der Erve, and

Higgins 2019). Chapman and Leigh (2009) and Britton and Gruber (2019) both use bunching designs to estimate labor-supply responses to marginal changes in the income-share rates charged by Australian and UK student loan systems, respectively. Both find small or null effects of increased rates on earned income. Several studies look at ex ante selection into IDR and take-up effects of alternative enrollment procedures (Abraham et al. 2018a; Field 2009; Cox, Kreisman, and Dynarski 2018; Mueller and Yannelis 2019). Finally, Herbst and Hendren (2021) argue that adverse selection prevents the private market from offering contracts with IDR-like repayment terms. They use survey data on individuals' expected income to demonstrate how these markets have unraveled, providing a rationale for government provision of income-contingent contracts.

This paper makes three contributions to the existing literature. First, I provide the first causal estimates of IDR's treatment effects on loan repayment and balances, as well as suggestive evidence of effects on homeownership and financial health. Second, I document high attrition rates arising from re-enrollment frictions in IDR, which carry important fiscal implications for IDR. Third, my findings provide evidence of liquidity constraints among student borrowers, which suggests incomplete credit markets may hamper efficient investments in human capital.

The remainder of this paper is organized as follows. Section I provides a brief overview of federal student loans, IDR, and student loan servicing in the United States. Section II describes the administrative data and sample-selection criteria used in my analysis. Section III describes my empirical strategy. Section IV presents results, and Section V concludes.

I. Background

A. *Federal Student Loans and Repayment Plans*

Over 90 percent of student loans in the United States are federally subsidized and guaranteed.[1] The government holds the liability on student loans, and interest rates are set by Congress.[2] Student loans are not secured by collateral or subject to any credit check. While the amount one can borrow from federal sources is capped by semester, virtually anyone attending an accredited institution is eligible to borrow at the same subsidized rate.[3]

[1] A small private student loans market constitutes around 10 percent of total student debt, mostly for creditworthy graduate students or borrowers who have exhausted their federal loan limits. In most cases, however, private lenders cannot compete with the subsidized rates offered by the government under the Federal Family Education Loan (FFEL) and Federal Direct Loan programs. Unless stated otherwise, I will use "student loans" to refer to loans originating from these federal programs.

[2] Congress has set rates on student loans since 1965, though it automated the process in 2013 with the Bipartisan Student Loan Certainty Act, which sets interest rates equal to the 10-year Treasury bond rate plus 205 basis points (360 bps for graduate students). Interest rates are fixed throughout the life of a loan and accrue as simple daily interest on principal only.

[3] A small portion of borrowers who exceed their borrowing caps supplement their federal student loans with private loans, parent-cosigned PLUS loans, or risk-rated Grad PLUS loans for graduate schools. All of these "top-up" loan types are excluded from my analysis, though some borrowers in my sample might hold them in addition to their federal loans. While all borrowers are subject to the same federal borrowing caps, short-term borrowing costs can vary by financial need, as the Subsidized Stafford Loan program forgives interest accrued while the borrower is still in school, up to a means-tested limit.

The Department of Education sets repayment terms for student loans through repayment plans. Repayment plans specify the monthly minimum payments borrowers must make, though borrowers can pay more than the minimum without penalty if they wish to pay down their debt early. The default repayment plan into which all borrowers are automatically enrolled is known as "standard repayment." Under standard repayment, minimum monthly payments follow a flat repayment schedule over ten years. Until 2010, the vast majority of borrowers in repayment were enrolled in standard repayment plans, with only a small fraction of borrowers choosing alternative financing options.

Income-driven repayment (IDR) plans were first offered in 1994 as an alternative to standard repayment. Since then, several versions of IDR have become available, including Income-Based Repayment (IBR), Pay-As-You-Earn (PAYE), and Revised-Pay-As-You-Earn (REPAYE). Eligibility criteria and repayment terms can vary across these plans, though they share the same general structure.[4] Minimum payments under IDR are pegged to 15 percent of borrowers' discretionary income, defined as the difference between adjusted gross income (AGI) and 150 percent of the federal poverty line (FPL).[5] Specifically,

$$\text{(1)} \qquad \text{Monthly IDR Payment} = 15\% \times \left(\frac{\text{AGI} - 1.5 \times \text{FPL}}{12}\right).$$

Payments for a married borrower who files jointly are prorated to their share of combined household student debt. Minimum monthly payments are capped at the standard minimum payment amount, and payments continue until the borrower's balance reaches zero. If a borrower successfully makes 300 payments under IDR, any remaining balance is forgiven, though any forgiven debt is treated as taxable income.

Borrowers can switch to IDR at any point in the repayment process. Opting in requires completing an online form through the Department of Education, which verifies income and family size using information from a borrower's most recent federal tax return. Borrowers must recertify their income on a yearly basis, though they can adjust their payments more frequently with proof of income. If a borrower on IDR goes more than one year without recertifying income and family size, their payments automatically return to the standard payment amount, though their repayment plan is still classified as IDR (US Department of Education 2020c).

Borrowers who fail to meet their monthly payments (i.e., "fall delinquent") under any repayment plan face penalties that increase in severity with the number of days past due. Between 1 and 10 days past due, borrowers receive delinquency notices by email and post. Between 10 and 90 days past due, borrowers are charged late fees and contacted by phone at increasing frequency to encourage repayment and discuss repayment options. At 91, 181, and 271 days past due, borrowers are reported to credit bureaus, damaging their credit scores. Loans more the 270 days past due are

[4] For the purposes of this study, I focus on the largest IDR plan, Income-Based Repayment (IBR), as borrowers in my sample are ineligible for newer IDR plans, though the discussion generalizes to the broader concept of IDR.

[5] Online Appendix Figure A1 provides a graphical comparison of IDR versus standard repayment plans under alternative income scenarios.

considered eligible for default. Once in default, all remaining balance on student debt becomes due, and the Department of Education can garnish up to 15 percent of borrowers' wages or withhold their tax returns to collect on defaulted debt. In 20 states, the federal government can block the renewal of professional licenses for defaulted borrowers working in health care, education, and/or other licensed fields. Unlike other forms of consumer debt, student loans cannot be discharged in bankruptcy, except in rare circumstances. Defaulted borrowers are ineligible for any future federal student aid (US Department of Education 2020a).

B. *Study Setting: Student Loan Servicers and LLS*

As one of ten federal student loan servicing companies, LLS manages disbursal, billing, and processing of over $300 billion in federal student loans on behalf of the Department of Education. As a part of its servicing operations, LLS makes frequent contact with delinquent borrowers to encourage repayment. When borrowers become 15 or more days past due on their payments, their phone numbers are placed in a dialing queue. An automatic dialer then places calls to queued numbers in rapid succession. If a call is unanswered, the dialer places it back at the bottom of the queue. Each answered call is immediately connected to a debt-servicing agent randomly selected from the pool of available agents not already on a call. If no agents are available, the dialer places the borrower on hold until one becomes available. Such instances are rare, however, as the dialer places calls at a rate to match agent availability, which is highly predictable over large numbers of agents.

LLS employs over 300 servicing agents across 4 call centers. Agents are tasked with informing borrowers of their delinquent status, inquiring about their ability to repay, and informing them of repayment options. During a call session, the questions and responses of the agent are guided by a decision tree. The agent first asks if a borrower can make payments under their current plan. If not, the agent asks if the borrower is unemployed or a full-time student, as such borrowers can typically qualify for interest-free unemployment deferments. Finally, the agent "models-out" IDR payments for the borrower, eliciting information on annual income, marital status, and family size. Borrower responses are entered into the agent's computer, which provides an estimate of monthly IDR payments according to equation (1). If these payments are lower than what the borrower is paying under the standard plan, the agent provides the borrower with their "modeled-out" IDR payment estimate as well as instructions for online IDR enrollment with the Department of Education. Agents are incentivized to bring delinquent accounts current, but face penalties if they fail to present borrowers with their best available options. Supervisors periodically monitor agents' calls to ensure they meet federal compliance standards. If an agent does not offer IDR to a borrower deemed suitable for the option during a monitored call, the agent's pay is reduced that month.

II. Student Loan Servicing Data

The data I use in this paper link administrative student loan repayment and contact data to credit bureau records for over one million borrowers. Data are drawn from LLS's FFEL loan portfolio, which includes over $90 billion in loans. The LLS loan data contain detailed repayment records for each borrower, including principal borrowing amounts, loan balances, minimum payments due, and dates of delinquency at a monthly frequency. They also include indicators for type of loan (e.g., Subsidized Stafford, PLUS), current repayment plan, and current loan status (e.g., deferment, grace period, default). In addition to loan information, the LLS data contain borrower characteristics, including year of birth, 9-digit zip code, OPE ID for attended institutions, college attendance dates, and graduation status. Gender is inferred using first names.[6]

I merge demographic and loan information with LLS contact histories from 2011 onward. Contact history data provide a single observation for each point of contact and include all incoming and outgoing calls in which the line was connected to a borrower in the sample. For each call in the data, I observe the date, time of day, incoming/outgoing status, and servicing agent identifier associated with the call. Agent identifiers are linked to a small set of agent characteristics, including work site location and work group (e.g., "claims aversion," "skip tracing," etc.).

Finally, borrowers in the LLS data are linked to yearly TransUnion credit bureau records from 2010 through 2018. The TransUnion data provide yearly balances, credit limits, delinquencies, and number of accounts for several categories of consumer debt, including mortgages, credit cards, and auto loans. They also include broader measures of financial health, like credit scores and bankruptcies.[7] TransUnion data are merged to borrowers in the LLS data by last name and last four digits of SSN. Ninety-two percent of borrowers are successfully matched to TransUnion records.

A. *Sample Selection*

The analysis sample used in this study consists of 133,630 individuals selected to best represent the general population of borrowers eligible for reduced payments under IDR. To construct this sample, I begin with the universe of LLS's FFEL borrowers with positive balances as of December 2011, excluding those who hold any private or Direct loans.[8] From this population of 5.8 million borrowers, I remove anyone whose loans were canceled, discharged, or paid-in-full by December 2013, leaving 3.8 million borrowers. I then select those borrowers who answered a delinquency call between 2014 and 2018, limiting the sample to 631,273 borrowers. I then remove borrowers who cannot be matched by zip code or first name to inferred measures of gender or income, or whose credit card or mortgage balances exceed the ninety-ninth percentile in any year, leaving 539,269 borrowers. Next, I limit the sample to English speakers who answered at least one call within 140 days

[6] The online Appendix to Tang et al. (2011) provides a public-use list of common first names paired with the male-female proportions of New York City Facebook profiles with each name. LLS merged this list to first names in their borrower records at my request.

[7] Additional details concerning TransUnion data can be found in Dobbie, Goldsmith-Pinkham, and Yang (2017); Avery et al. (2003); and Finkelstein et al. (2012).

[8] While borrowers can hold loans from a mixture of FFEL, Direct, and private sources, the database I use only includes repayment information for FFEL borrowers. The analysis sample excludes borrowers with mixture of loans, so I can observe their complete repayment profile. Roughly 15 percent of LLS's 2012 FFEL borrowers also hold Direct loans, and fewer than 10 percent hold private student loans.

of falling delinquent, leaving 443,138 borrowers. Then, I remove borrowers who were already enrolled in IDR prior to their delinquency call, as they would not be eligible for call-induced IDR take-up. I also remove anyone with a previous IDR spell from the sample so that estimates can be interpreted as the effect of *initial* enrollment. From the remaining group of 402,043 borrowers, I keep only those who were "modeled-out"—borrowers who reported difficulty meeting their monthly payments and met other qualifications for monthly payment reduction—leaving 133,630 borrowers.[9]

To facilitate my empirical strategy, I use the sample of borrowers described above to create three balanced panels at the borrower-by-call level, centered around call dates.[10] For instrumental variables analysis of short-term repayment outcomes, I select all calls made from 2017 onward by agents with at least 100 total calls.[11] From the resulting sample of 78,050 calls, I create a balanced monthly panel of 49,775 calls with 20 leads and 10 lags. For the more speculative analysis of longer-term outcomes, I broaden the selection criteria to include calls from 2013 to 2016 and those made by small-cell agents. From this sample of calls, I create two additional balanced panels corresponding to the frequencies of outcome data: a *yearly* panel of 22,904 calls with 4 leads and 3 lags, and a *monthly* panel of 47,520 calls with 42 leads and 10 lags.

Table 1 provides summary statistics for samples of interest. The "Full sample" (column 1) is a random sample of 608,195 drawn from the population of LLS FFEL borrowers as of December 2012. The "Analysis sample" (column 2) is the entire subpopulation of borrowers selected according to the criteria described above. In the full sample, IDR has low take-up, with only 14 percent of borrowers enrolled in a plan. That share rises to 34 percent in the analysis sample, as it is constructed to include only borrowers who might benefit from the plan. Unsurprisingly, these borrowers have lower credit card limits, higher rates of bankruptcy and live in lower-income zip codes. Columns 3 and 4 of Table 1 break the analysis sample into "Enrolled" and "Not enrolled" groups, where "Enrolled" is defined as IDR enrollment within four months of answering an LLS delinquency call.[12] Baseline variables for enrolled borrowers are largely comparable to those for the non-enrolled group.

[9] Because I cannot observe eligibility criteria like income or employment status, I cannot directly observe whether a borrower qualifies for IDR payments. The "modeled-out" restriction serves as a proxy for these unobserved criteria because it removes borrowers that agents deemed ineligible for IDR early in the call. Nonetheless, I provide estimates for the pooled population of modeled and non-modeled borrowers and find qualitatively similar to those for the modeled-out sample (see Section IVA).

[10] Borrowers who receive multiple delinquency calls within the panel window can appear more than once in the data. To account for any within-borrower correlation in outcomes, I cluster standard errors at the borrower level. I also include controls for the number of prior calls a borrower has received to remove any influence that call history might have on outcomes.

[11] Removing agents with few calls reduces measurement error in the agent-score instrument because estimates of the mean taken over a small number of calls are highly imprecise. Restricting the sample to the post-2016 period removes any nonrandomly assigned calls placed by older auto-dialing systems.

[12] Note that IDR enrollment is defined at the call level, not the borrower level. For the borrower-level statistics reported in Table 1, the enrolled group consists of all borrowers with *any* calls resulting in IDR enrollment. Also note that 23 percent of the non-enrolled group does eventually enroll in IDR, though never within four months of a delinquency call included in the balanced panels.

Table 1—Summary Statistics

	Full sample	Analysis sample		
	Pooled	Pooled	Not enrolled	Enrolled
	(1)	(2)	(3)	(4)
Panel A. LLS data				
IDR	0.135	0.337	0.225	1
Female	0.592	0.693	0.686	0.737
Zip median income	61.36	53.21	53.45	51.77
Age	38.17	40.51	40.58	40.15
Amount borrowed	25.66	22.62	22.42	23.82
10+ days delinquent	0.375	0.796	0.803	0.756
90+ days delinquent	0.149	0.343	0.351	0.293
Days delinquent	43.40	89.59	91.57	78.03
Panel B. Credit data				
Credit score	679.8	594.8	594.5	596.6
Bankruptcy	0.0832	0.156	0.153	0.176
Derogatory rating	0.260	0.613	0.610	0.632
Number of credit cards	5.376	3.410	3.423	3.334
Credit card balances	4.231	1.588	1.618	1.417
Number of mortgages	1.242	0.789	0.809	0.674
Mortgage balances	68.96	29.29	30.70	21.14
Credit card limits	19.82	5.082	5.180	4.516
Number of auto trades	1.972	1.656	1.670	1.576
Observations	608,195	133,630	114,429	19,201

Notes: This table reports summary statistics at the borrower level. The full sample is a random sample of the population of borrowers in LLS's FFEL portfolio who carried a positive loan balance as of December 31, 2011 and hold no private or Direct loans. The analysis sample is a subsample from the same population, selected according to the following criteria outlined in Section II. "Enrolled" borrowers are those who enroll in IDR within four months of a delinquency call. "IDR" is an indicator for whether the borrower ever enrolled in IDR. "Female" is a measure of likelihood-female inferred from first name following Tang et al. (2011). "Zip median income" is the median 2010 income for the borrower's recorded 5-digit zip code. "Days delinquent" is the maximum number of days the borrower was ever past due on payments in the past year. IDR enrollment statuses reflect IDR enrollment histories through September 2019. All other LLS variables are taken from administrative records as of December 31, 2012. Credit scores, bankruptcies, derogatory ratings, credit card, mortgage, and auto loan information are taken from TransUnion credit bureau data collected in August 2012.

III. Empirical Strategy

Consider the following empirical model of borrower i's outcomes, t periods after receiving delinquency call c:

$$Y_{ict} = \beta_1 IDR_{ic} + \beta_2 \mathbf{X}_{ic} + \epsilon_{ict}, \tag{2}$$

where Y_{ict} denotes the outcome of interest, $\mathbf{X}_{ic}$ is a vector of borrower control variables (including call date and time fixed effects), IDR_{ic} is an indicator for IDR enrollment within four months of the call, and ϵ_{ict} is an error term.[13] To estimate β_1, my

[13] I fix IDR_{ic} to a specific month in order to capture the dynamic effects of IDR. Note, however, that a borrower's repayment plan as of month four need not reflect their repayment plan in later months. Indeed, attrition from IDR

primary identification strategy is an instrumental variables (IV) design that exploits the quasi-random assignment of servicing agents to calls. I compliment this IV strategy with estimates of longer-term differences-in-differences between IDR enrollees and non-enrollees before and after receiving delinquency calls. This second, more descriptive analysis is considerably more speculative; difference-in-differences results should be seen as providing only suggestive evidence of causal effects.

A. *Instrumental Variables*

Using a sample of randomized delinquency calls made after 2016, my instrumental variables (IV) design estimates IDR's effect on monthly repayment outcomes within 20 months of enrollment. I instrument for IDR enrollment using "agent score," a leave-one-out measure of agents' ability to induce IDR enrollment, where post-call enrollment is residualized to account for the timing and ordering of delinquency calls. Specifically,

$$IDR^*_{ic} = IDR_{ic} - \gamma \mathbf{W}_{ic} \tag{3}$$

$$= Z^A_{icj} + \epsilon_{ic}, \tag{4}$$

where $\mathbf{W}_{ic}$ is a vector of call year-by-month, day-of-week, and hour-of-day dummies and Z^A_{icj} is agent score. I calculate the residualized rate of IDR take-up, IDR^*_{ic}, using OLS estimates of γ in equation (3). I then construct agent score Z^A_{icj} using the leave-one-out mean of this residualized rate,

$$Z^A_{icj} = \left(\frac{1}{n_j - 1}\right)\left(\sum_{k=0}^{n_j} IDR^*_{kcj} - IDR^*_{icj}\right), \tag{5}$$

where n_j denotes the number of calls made by agent j. The residualized agent-score distribution can be seen in Figure 1.[14] Variation in agent score can be driven by differences in agents' demeanor, clarity of instructions, or loan-servicing technologies. For example, in 2017, LLS adopted electronic signature technology ("e-sign") to a subset of call agents. In online Appendix Section B, I use e-sign status as an alternative instrument and find similar results to the agent-score IV specification.

In order for my two-stage least squares estimates to identify a local average treatment effect (LATE) of IDR take-up, the instrument must satisfy three conditions. First, IDR take-up must vary with agent assignment. To test this assumption, I estimate the first-stage relationship between the agent-score instrument and observed IDR enrollment:

$$IDR_{ic} = \alpha_1 Z^A_{ic} + \alpha_2 \mathbf{X}_{ic} + \epsilon_{ic}. \tag{6}$$

after the one-year recertification period will play an important role in interpreting my results. See online Appendix Figure A3.

[14] Note that while the two-stage least squares analysis is conducted on a balanced monthly panel of post-2016 calls, the agent-score instrument is calculated using the larger unbalanced panel of calls satisfying all other sample selection criteria in Section II. This sample includes calls from 203 different agents in four different call centers. Agents place 245 calls on average to borrowers in the sample, with a median of 155 calls.

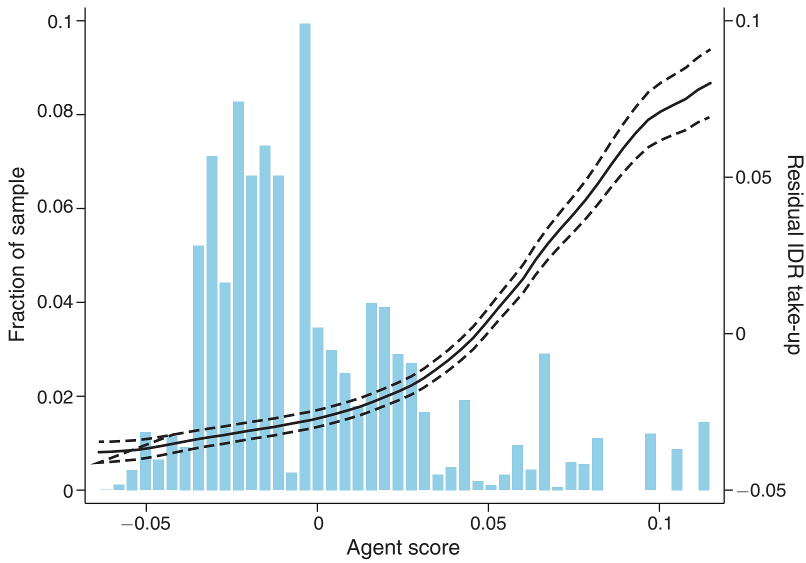

Figure 1. Agent-Score Instrument and IDR Enrollment

Notes: This figure reports first-stage effects and distribution of agent scores across delinquency calls, where agent score is the leave-out mean IDR take-up calculated using data from other calls made by the agent following the procedure described in Section III. The solid and dashed lines, plotted against the right axis, represent predicted means with 95 percent confidence intervals from a local linear regression of residualized IDR take-up on agent score. The histogram, plotted against the left axis, provides the distribution of agent scores across all delinquency calls in my analysis sample.

The first-stage OLS estimate of α_1 is 0.98, with or without borrower controls, and the F-statistic on a test of instrument significance is 167.57 (online Appendix Table A1). Graphical evidence of first-stage effects is provided by Figure 1, which plots a local linear regression of IDR take-up against the agent-score instrument.

The second identifying assumption is that agent assignment must correlate with borrower outcomes only through its effect IDR take-up. While randomized assignment of agents rules out many potential violations of this assumption, nonrandom selection of borrowers into the study sample could still cause concern.[15] For example, if agents experience differential rates of borrower hangup before reaching the "modeled-out" portion of the phone call, the sample would be selected based on agent-specific criteria that could potentially correlate with the instrument and bias my estimates. I address this concern by adjusting my main IV specification to include agent-induced sample-selection propensity, constructed as the leave-one-out mean "modeled-out" rate, Z_{ic}^{M} (Heckman 1979).[16] Balance tests confirm that, after

[15] Note that random assignment does not imply equal probability of assignment—an agent who makes shorter and more frequent phone calls will have a higher rate of availability during their shift. Any given delinquency call will therefore have a higher probability of being assigned to these "quicker" agents. The average call to which such agents are assigned, however, will nonetheless be no different from those calls assigned to relatively "slower" agents who make fewer calls per hour.

[16] I construct the leave-one-out mean "modeled-out" rate, Z_{ic}^{M}, among all calls assigned to the agent on a given call. I perform this calculation on the unconditional sample of calls and follow the same procedure as equations (3) through (5), replacing the treatment variable IDR_{ic} with $Modeled_{ic}$, an indicator for whether borrower i was

correcting for agent modeling propensity and call timing, borrowers do not vary systematically by agent-score (see online Appendix Table A2). I also present results from the unconditional sample with no sample-selection correction and find qualitatively similar results to my main specification (see Section IVA).

Even if agents are randomly assigned to borrowers, the exclusion restriction may still be violated if agents can influence borrower outcomes through channels other than repayment plan choice. If, for example, agents who induce high IDR take-up also convince borrowers to make timely payments, two-stage least squares estimates of IDR's effects on repayment would be biased upwards. While it is impossible to rule out agent effects through non-IDR channels, loan servicing practices suggest that such threats to validity are unlikely. LLS's delinquency calls are designed solely to provide borrowers with information on their repayment options. Agents provide no advice or counseling to borrowers, and follow a decision tree to present repayment alternatives.

The third identifying assumption requires monotonic agent effects across borrowers, so there can be no borrower for whom a higher-score agent decreases the likelihood of IDR take-up. I implement two partial tests of the monotonicity assumption. First, I estimate the first-stage relationship between my agent-score instrument and IDR take-up within subgroups of my monthly analysis sample. As online Appendix Table A3 shows, estimated coefficients are positive across a variety of subgroups. Second, I calculate a variety of *group-specific* agent-score instruments, capturing agents' average IDR inducement rates within observably different subsamples.[17] Online Appendix Figure A2 reports binned scatter plots and correlation coefficients for several pairwise comparisons of these group-specific instruments computed across the entire analysis sample. I find strongly positive correlations for each pair, suggesting agent inducement is similar across borrower characteristics.

B. *Difference-in-Differences*

I complement the instrumental variables design described above with a difference-in-differences (DD) design that compares pre-/post-call differences in outcomes between borrowers who take up IDR and borrowers who remain in standard repayment plans. While the self-selected nature of these groups makes any causal interpretation of DD estimates highly speculative, this design allows me to expand the study sample to earlier phone calls and investigate long-term trends in credit and employment outcomes.[18]

"modeled-out" during phone call c. I then include the sample selection measure Z_{ic}^{M} in my instrumental-variables regressions to ensure that assignment of Z_{ic} is conditionally random.

[17] Group-specific agent-score instruments are calculated as

$$Z_{icj}^{g} = \left(\frac{1}{n_j^g - 1}\right)\left(\sum_{k=0}^{n_j^g} IDR_{kj}^{*} - IDR_{ij}^{*}\mathbf{1}_{\{i \in g\}}\right).$$

For example, Z_{icj}^{men} is the residualized, leave-one-out propensity of agent j to induce men into IDR.

[18] Prior to 2016, LLS used a different autodialer to reach customers. While the frequency, timing, and content of calls during this period were unchanged, the details of how that system allocated calls between agents is not available.

Formally, the DD specification takes the following form:

$$(7)\quad Y_{ict} = \gamma_i + \gamma_t + \left[\sum_{\tau \neq -1} \delta_\tau \times IDR_{ic} \times \mathbf{1}\{t = \tau\}\right] + \beta_1 IDR_{ic} + \beta_2 \mathbf{X}_{ict} + \epsilon_{it},$$

where Y_{ict} denotes the outcome of interest, γ_i are individual fixed effects, γ_t are event-time fixed effects, IDR_{ic} is an indicator for IDR enrollment within four months of the call, $\mathbf{X}_{ict}$ is a vector of borrower control variables (including call date and time fixed effects), ϵ_{ict} is an error term, and δ_τ, the parameters of interest, are coefficients on IDR enrollment status which vary by event time. The specification omits γ_t and δ_τ terms at $t = -1$, so estimates can be interpreted relative to the baseline period of one month or year prior to the delinquency call.

Identification in the DD specification comes from variation in the propensity to take up IDR following a delinquency call. The identifying assumption is that, holding borrower-specific differences fixed, post-call trends in outcomes would be the same for enrolled and non-enrolled borrowers had neither group taken up IDR. Online Appendix Figures A3 through A6 plot mean outcomes for IDR enrollees and non-enrollees relative to call date and normalized by pre-call mean. Trends in pre-call outcomes appear similar between IDR and standard enrollees for several periods, diverging only after receiving the delinquency call. I also estimate IDR effects in an alternative differences-in-differences specification that controls for group-specific linear trends in months or years prior to call.[19]

Even if IDR and standard borrowers exhibit observably similar pretrends, DD estimates could be biased if enrolled and non-enrolled groups would have responded to delinquency calls differently in the absence of IDR. I address this concern with a placebo test designed to simulate this hypothetical scenario. Many enrolled borrowers receive one or more "non-converting" calls before their "enrolling call" (i.e., the call preceding their IDR enrollment). If, in the absence of IDR, enrolled and non-enrolled borrowers would have responded differently to their nth delinquency call, they would likely have had different responses to calls 1 through $n - 1$ as well. Online Appendix Figure A7 plots raw pre- and post-call repayment outcomes for non-IDR borrowers versus *eventual* IDR borrowers following these earlier "placebo calls" that did *not* induce IDR take-up within the following 12 months. Compared to enrolling calls in the main estimation sample, responses to non-converting calls track closely with calls for the non-IDR group, suggesting my DD estimates are not capturing a "call effect."

DD estimates could also be biased if IDR enrollees experienced a shock at the time of a delinquency call that induced them into IDR take-up and influenced outcome variables. However, delinquency calls are *outgoing*, so their incidence is determined by LLS and does not vary systematically between observably similar borrowers. If IDR borrowers were enrolling as a response to sudden shocks, outcomes should vary in the months immediately preceding the call. It is possible that

[19] Estimates from the specification including linear pretrends can be interpreted as IDR's impact on outcomes relative trend-predicted differences between groups. Formally, the model is given by (8):

$$(8)\qquad Y_{ict} = \gamma_i + \delta t \times IDR_{ic} \times \mathbf{1}\{t < 0\} + \left[\sum_{\tau \geq 0} \delta_\tau \times IDR_{ic} \times \mathbf{1}\{t = \tau\}\right] + \beta_1 IDR_{ic} + \beta_2 \mathbf{X}_{ict} + \epsilon_{it}.$$

some borrowers make IDR enrollment decisions based on expected *future* shocks to their financial well-being, though such forward-looking borrowers would likely enroll in IDR themselves rather than wait for a delinquency call from LLS. In any case, IDR benefits are strictly decreasing in income and available credit, so any potential bias created by forward-looking borrowers should be negative, attenuating any positive treatment effects of IDR.

IV. Results and Interpretation

A. *Short-Term Outcomes: Repayment and Balances*

Figures 2 through 4 plot agent-score instrumental-variables (IV) and difference-in-differences (DD) coefficients on minimum payments, loan balances, and indicators for more than 10, more than 90, and more than 270 days delinquent. Left-column graphs plot estimated coefficients on IDR take-up from separate two-stage least squares regressions in each month using the agent-score instrument. Right-column graphs plot estimated coefficients on the interaction between IDR take-up and months-since-call from the pooled DD specification given by equation (7).[20]

Minimum Payments and Re-enrollment.—The immediate effect of IDR enrollment on minimum payments is mechanical.[21] Nonetheless, estimating the IDR treatment effect on minimum payments can provide useful insight into the "first-stage" effects driving more downstream results.

Both instrumental variables and DD estimates of minimum payments effects suggest IDR provides borrowers with large but short-term increases to cash-on-hand. Agent-score IV estimates imply a 86 percent decline in monthly minimums immediately after enrollment, followed by a sharp rise 12 months later. The DD strategy finds very similar results. As the bottom panel of online Appendix Figure A3 illustrates, this pattern appears to be driven by a lack of re-enrollment. After one year on IDR, more than 60 percent of enrollees in my sample do not fulfill their income recertification requirements, resulting in a return of minimum payments to their pre-call levels.[22] This result is not restricted to my panel window or analysis sample. Online Appendix Figure A9 plots enrollment status for an expanded panel of IDR borrowers in the larger, representative sample in the months following their initial enrollment. While a small group of enrollees do eventually recertify, roughly

[20] IV point estimates for both agent-score and e-sign specifications are reported separately by three-month period in online Appendix Table A4, and corresponding difference-in-difference estimates are reported in online Appendix Table A5. All specifications include controls for call date and time, as well as amount borrowed, number of previous calls, inferred gender, and zip-code median income.

[21] Given adjusted gross income, family size, and debt balance, one could directly calculate IDR's effect on payment size using a standard loan amortization formula and equation (1). For the enrolled group in my sample, this effect is approximated by observed IDR payments minus payments in the month prior to receiving the delinquency call. Online Appendix Figure A8 provides a graphical illustration of this measured payment effect across the distribution of IDR enrollees in my analysis sample.

[22] Technically, standard payments might be higher after a year on IDR because unpaid interest is, under some circumstances, recapitalized into the principal amount.

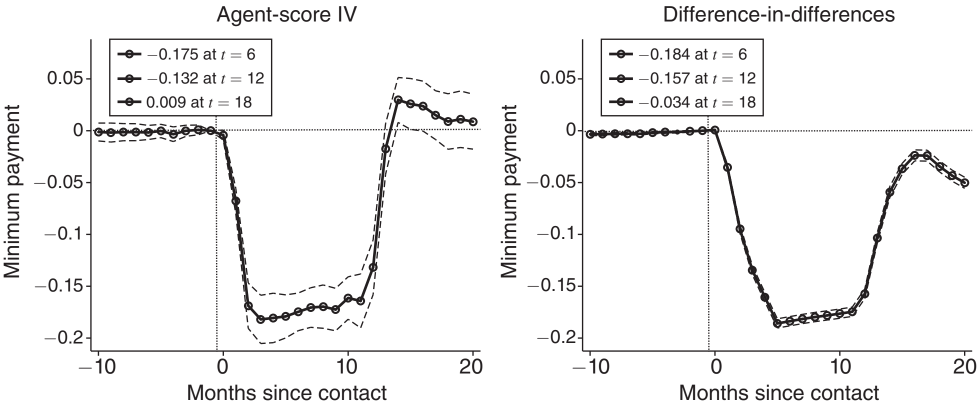

Figure 2. Estimates of the Effect of IDR Enrollment on Minimum Payments

Notes: This figure reports monthly agent-score two-stage least squares and difference-in-differences estimates for minimum monthly payments. Each point represents the estimated effect of post-call IDR status on minimum monthly payment at a given time period relative to the date of delinquency call. Relative months are plotted along the *x*-axis. Dashed lines represent 95 percent confidence intervals. Boxes list point estimates at selected months. All regressions include fixed effects for call year, month, day-of-week, and hour-of-day, as well as controls for initial amount borrowed, number of previous calls, inferred gender, pre-call debt balance, and pre-call zip-median income. IV estimates also control for agent modeling propensity (see Section III), and difference-in-differences regressions include individual fixed effects. Robust standard errors are two-way clustered at the borrower and agent levels for IV estimates and one-way clustered at the borrower level for difference-in-differences estimates.

one-half of initial enrollees have still not recertified by month 42. While some of this attrition may be driven by incomes rising above the reduced-payment-eligibility threshold, the more likely explanation is a behavioral response to the burdensome recertification process required under IDR (Cox, Kreisman, and Dynarski 2018).

Delinquencies.—I measure IDR's impact on delinquencies using the likelihood of falling more than 10 days delinquent, the likelihood of falling more than 90 days delinquent, and the likelihood of falling more than 270 days delinquent. These three benchmarks reflect points of increased delinquency penalties: at 11 days past due, borrowers begin to accrue late fees for delinquent loans. At 91 days past due, borrowers are reported to credit bureaus. At 271 days past due, a borrower becomes eligible for default.

Monthly DD and IV estimates, shown in Figure 3, indicate a large negative effect of IDR enrollment for all three delinquency measures in the short term, but attenuate or reverse direction after the 12-month recertification period. I find that IDR reduces 10-day delinquency measures by 19 percentage points and 90-day delinquencies by −8 percentage points. IV and DD results are very similar in magnitude, although the DD results are more precise. I find no significant impact on 270-day defaults. While IDR leads to small increases in delinquencies in later months, these increases are short-lived, so the net impact on delinquencies is negative.

Balances.—In theory, IDR could affect balances on student loans in either direction. IDR borrowers face lower monthly minimums payments than those on standard plans, increasing relative balances among those who stay current on their loans.

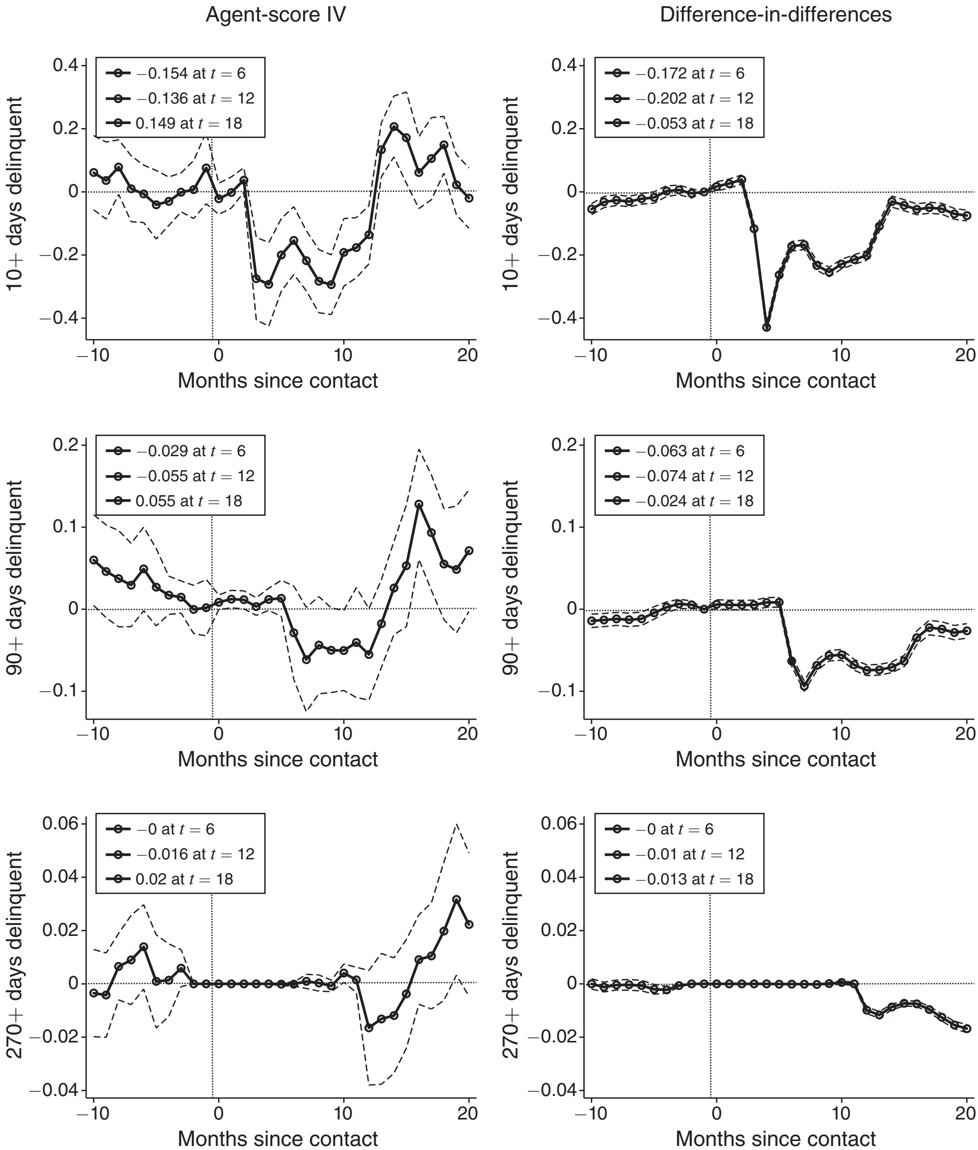

Figure 3. Estimates of the Effect of IDR Enrollment on Delinquencies

Notes: This figure reports monthly agent-score two-stage least squares and difference-in-differences estimates for borrower delinquencies. Each point represents the estimated effect of post-call IDR status on the likelihood of being more than 10, more than 90, and more than 270 days delinquent at a given time period relative to the date of delinquency call. Relative months are plotted along the x-axis. Dashed lines represent 95 percent confidence intervals. Boxes list point estimates at selected months. All regressions include fixed effects for call year, month, day-of-week, and hour-of-day, as well as controls for initial amount borrowed, number of previous calls, inferred gender, pre-call debt balance, and pre-call zip-median income. IV estimates also control for agent modeling propensity (see Section III), and difference-in-differences regressions include individual fixed effects. Robust standard errors are two-way clustered at the borrower and agent levels for IV estimates and one-way clustered at the borrower level for difference-in-differences estimates.

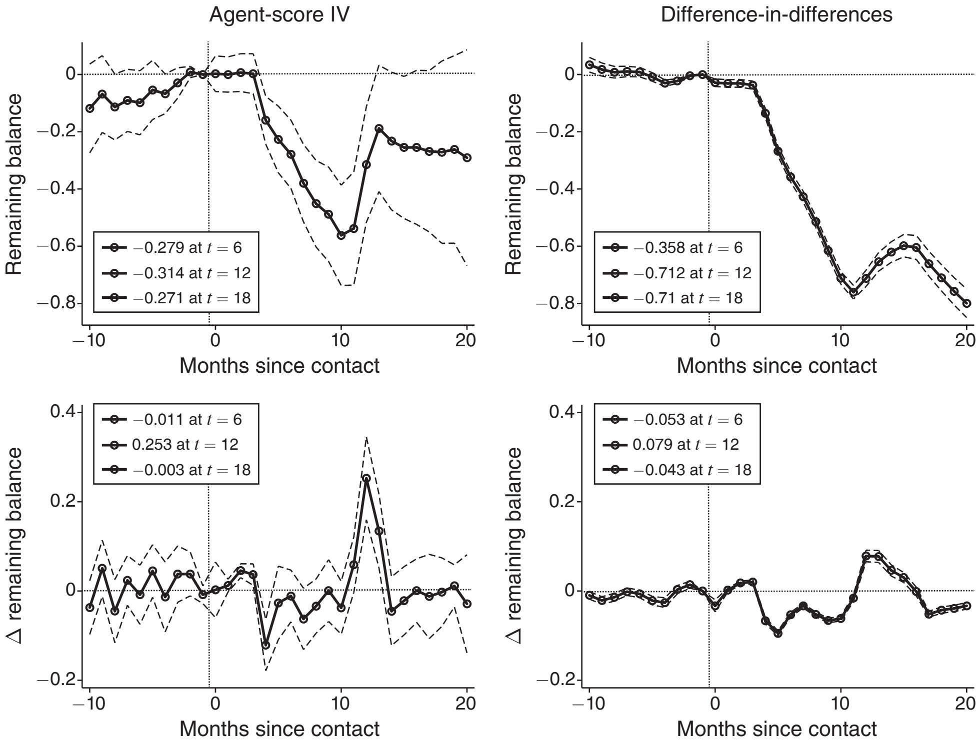

Figure 4. Estimates of the Effect of IDR Enrollment on Balances

Notes: This figure reports monthly agent-score two-stage least squares and difference-in-differences estimates for borrower balances. Each point represents the estimated effect of post-call IDR status on borrowers' month-to-month balance and change in debt balances, respectively, at a given time period relative to the date of delinquency call. Relative months are plotted along the *x*-axis. Dashed lines represent 95 percent confidence intervals. Boxes list point estimates at selected months. All regressions include fixed effects for call year, month, day-of-week, and hour-of-day, as well as controls for initial amount borrowed, number of previous calls, inferred gender, pre-call debt balance, and pre-call zip-median income. IV estimates also control for agent modeling propensity (see Section III), and difference-in-differences regressions include individual fixed effects. Robust standard errors are two-way clustered at the borrower and agent levels for IV estimates and one-way clustered at the borrower level for difference-in-differences estimates.

However, IDR borrowers are also more likely to actually *make* their monthly payments, a consideration that is often ignored in fiscal projections of IDR. Figure 4 reports estimated coefficients for student loan balances and monthly changes in balance.[23] In months 6 through 8, IDR borrowers pay down more debt each month ($46 for both DD and $36 for IV), but much of those gains are lost by months 12 through 15, when their balances begin to *increase* relative to non-IDR borrowers by a monthly average of $67 for DD and $114 for IV.

My results suggest the effect of reduced minimums on loan balances is dominated by more timely repayment, at least in the short term. While the cumulative effect on

[23] Note that, depending on the specific plan and minimum payment amount, IDR borrowers can sometimes receive partial forgiveness on accumulated interest. While effects on balance levels partially reflect these forgiveness provisions, my measure of change in balances removes any interest forgiveness.

balance levels remains negative throughout the panel window, the sharp reversal in effects on changes in balances at the 12-month mark points once again to the negative influence of the recertification process on repayment likelihood. Importantly, however, estimated effects on balances are relative, not absolute. On average, neither standard nor IDR borrowers are decreasing their total balances over the entire period (see online Appendix Figure A5).

Note that for all monthly repayment outcomes, estimates from the difference-in-difference approach are broadly consistent with the more plausibly identified IV estimates, especially in the first ten months following take-up. In later months, however, estimates diverge, with IV estimating less favorable effects for borrowers than DD. While some of this divergence could be caused by estimation error or biased DD estimates, the pattern is consistent with expected differences in the local-average treatment effects identified by IV and DD strategies. In particular, one might expect agent-score IV compliers—those who are induced into IDR by slight variations in agent-specific factors—to be less likely to recertify than those who enroll in response to more general call-specific factors.

Robustness Checks.—These results are robust to a number of alternative samples and specifications. First, I conduct my analysis for a subsample of borrowers with predicted IDR payments greater than zero. This exercise should attenuate the influence of mechanically lower default rates among IDR borrowers with monthly minimums equal to zero.[24] Realized IDR payments are nonzero for more than 80 percent of enrolled individuals in this subsample, yet the repayment effects of IDR persist. Online Appendix Table A6 reports delinquency results for this subsample. IV results are noisy yet qualitatively similar to main results, while DD results continue to find a large and significant effect on repayment rates.

Second, I extend my analysis sample to include borrowers who were not "modeled-out" during their delinquency calls. Many of these borrowers are effectively ineligible for IDR because they qualify for deferments instead or have incomes that are too high relative to their remaining debt balance to qualify them for reduced payments under IDR. Online Appendix Figure A10 and Table A7 report estimates for this pooled population of modeled and non-modeled borrowers. IV results are similar but show larger standard errors because the inclusion of non-modeled borrowers weakens agent-score instrument's first stage—one cannot be induced into IDR by their agent if they do not qualify for the program. DD estimates are also similar, though slightly attenuated. This attenuation might occur because IDR-ineligible borrowers, now in the no-IDR comparison group, are more likely to expect post-call improvements in earnings or financial health than the eligible borrowers who opt into the program.

Third, I estimate effects under alternative specifications for the IDR enrollment variable. To account for administrative lags, the main specification designates IDR status based on their repayment plan as of the fourth month following their

[24] While this mechanical effect could still be characterized as a liquidity effect under a neoclassical model, it may be driven in part by psychological frictions or "hassle costs" if borrowers facing payments of $\epsilon > 0$ dollars would face higher delinquency rates than zero-payment borrowers.

delinquency call.[25] Online Appendix Tables A8 and A9 report estimates after I redefine the IDR variable to be enrollment within three and five months of the call, respectively. Estimates are similar under both specifications.

Finally, I estimate results under two alternative call-inclusion specifications. In online Appendix Table A10, I estimate results after expanding the analysis sample to include pre-2016 calls. While some of these calls may be nonrandom and contaminate the instrument, I nonetheless find very similar results, only much more statistically significant. In online Appendix Table A11, I estimate IV results using an alternative instrument construction that excludes pre-2016 calls when constructing agent score. These results closely resemble those from the main specification.

B. *Long-Term Outcomes: Credit Scores*, *Mortgages*, *and Zip-Median Income*

To investigate IDR's potential impact on long-term financial health, I turn my attention to credit scores, mortgage-holding rates, and zip-median income.[26] To investigate effects on these long-term outcomes, I shift my focus to calls made in 2014 and 2015, a period when some calls may not have been randomized. I therefore rely solely on the difference-in-difference strategy to estimate effects on these outcomes. For reasons stated above, these estimates carry considerably stronger caveats than IV results, and should been seen only as suggestive evidence of causal effects.

Figure 5 plots DD estimates of the effect of IDR on credit scores and mortgages. Plotted points represent the estimated coefficients on IDR in consecutive years from the pooled regression specified in equation (7), beginning with the year of the delinquency call ("Year 0"), while dashed lines represent corresponding 95 percent confidence intervals.[27] Relative to those who remained in standard repayment, borrowers who enrolled in IDR experienced a statistically significant 6.65-point increase in credit scores within one year of the delinquency call off of a pre-call mean of 596.5 points, an increase that persisted for the following four years. IDR's effects on the likelihood of holding a mortgage are also effectively zero in the year of the call, but rise to 1.9 percentage points by year 4, an increase of 9 percent off of the pre-call mean.

My data do not include direct measures of income. I can however, construct a proxy using the median income among households in each borrower's reported zip code, taken from the 2006–2010 American Community Survey (US Census Bureau 2010). ACS zip-median income was taken from the Michigan Population Studies Center (MPS 2017). While zip codes are reported at a monthly frequency, this value is self-reported by borrowers and usually only updated during contact between the borrower and LLS. Standard-plan borrowers receive more follow-up delinquency

[25] It typically takes one or two months following contact to process and enroll borrowers in IDR, and eventual enrollees often forgo making payments until IDR enrollment is complete. The relative timing of successful enrollment, next payment due date, and data collection date at the end of the calendar month adds further lag time before IDR effects can be realized.

[26] I also investigate bankruptcies, credit cards, auto loans, and unemployment deferments. These outcomes are reported in online Appendix Figures A11 through A13.

[27] All specifications include controls for call date and time, as well as amount borrowed, number of previous calls, inferred gender, and zip-code median income. Online Appendix Table A12 provides these estimates alongside estimates from a regression which omits pre-call month dummies and includes a linear time trend.

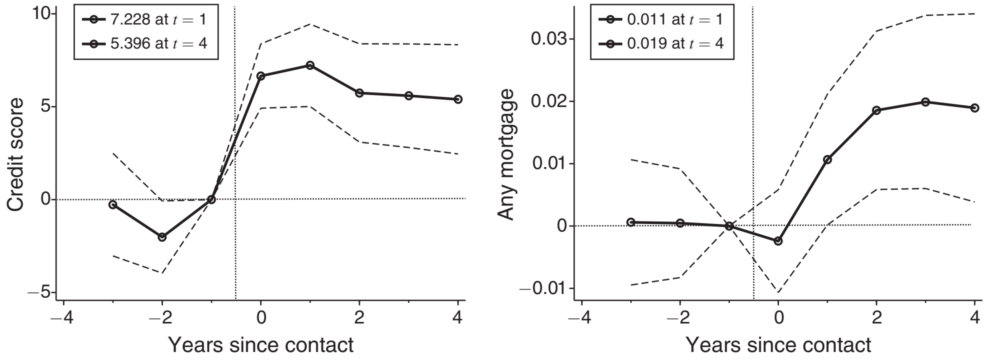

FIGURE 5. ESTIMATES OF THE EFFECT OF IDR ENROLLMENT ON LONG-TERM OUTCOMES

Notes: This figure reports annual difference-in-differences estimates for credit scores and mortgages. Each point represents the estimated effect of post-call IDR status on credit score or mortgage-holding status at a given time period relative to the date of delinquency call. Relative years are plotted along the x-axis. Dashed lines represent 95 percent confidence intervals. Boxes list point estimates at selected years. All regressions include fixed effects for call year, month, day-of-week, and hour-of-day, as well as individual fixed effects. Regressions also control for initial amount borrowed, number of previous calls, inferred gender, pre-call debt balance, and pre-call zip-median income. Robust standard errors are clustered at the borrower level.

calls than IDR borrowers for 15 months following the initial call, giving them more opportunity to update their zip codes during this period. Such borrowers may have higher *reported* incomes, biasing income effects downward.[28] To address this concern, I restrict attention to effects in months 18 and onward, when recertification periods have passed and enrolled and non-enrolled borrowers are equally likely to have had recent contact with LLS.[29] While the timing of potential biases is difficult to determine, more than 95 percent of borrowers have already recorded at least one change in zip code as of month 40, so late-month estimates likely to reflect effects beyond the potential bias period.

Results for zip-median income are reported in Figure 6.[30] In month 42, zip-median income shows a small increase of 0.7 percent off a pre-call mean of 3.9, and borrowers are 1.8 percentage points more likely to move to a higher-income zip code. These estimates suggest the positive effects of IDR overcome any zip-code-reporting bias, which should be negative if zip-median incomes are rising in general, though it should be emphasized that results for both outcomes should be interpreted with caution given the measurement concerns outlined above.

[28] Results for unemployment deferments, reported in online Appendix Figure A13, are susceptible to the same reporting bias.

[29] Online Appendix Figure A14 plots average number of additional points of contact for each month relative to the reference call. As expected, rates of contact for IDR borrowers spike during the initial enrollment and re-enrollment periods, differing considerably from non-IDR borrowers during that time period. In later periods, however, contact rates converge, suggesting both groups are equally likely to provide updated zip-code information to LLS during these months.

[30] All specifications include controls for call date and time, as well as amount borrowed, number of previous calls, and inferred gender. Estimates for selected months are reported in online Appendix Table A13.

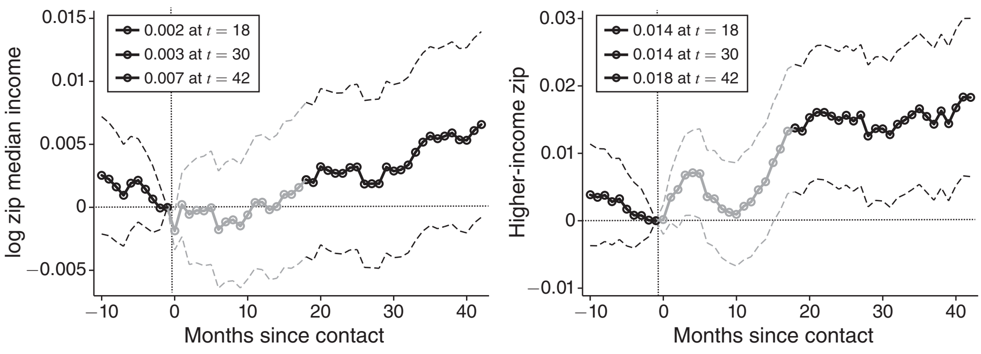

Figure 6. Estimates of the Effect of IDR Enrollment on Zip-Median Income

Notes: This figure reports monthly difference-in-differences estimates for zip-median income. Each point represents the estimated effect of post-call IDR status on log median income in a borrower's zip-code and a dummy for whether zip-median income exceeds its pre-call level at a given month relative to the date of delinquency call. Relative months are plotted along the *x*-axis. Results are estimated using an expanded monthly panel of 42 leads and 10 lags. Dashed lines represent 95 percent confidence intervals. Gray portions of the plot represent periods during which uneven rates of contact with LLS may bias estimates (see discussion in Section IVB). Boxes list point estimates at selected months. All regressions include fixed effects for call year, month, day-of-week, and hour-of-day, as well as individual fixed effects. Regressions also control for initial amount borrowed, number of previous calls, inferred gender, pre-call debt balance and pre-call zip-median income. Robust standard errors are clustered at the borrower level.

C. *Interpretation and Alternative Mechanisms*

Results for short-term repayment outcomes suggest IDR has a large and immediate increase to cash-on-hand, improving the balance sheets of liquidity-constrained borrowers. In addition to their potential fiscal benefits for the government, increased repayment rates provide one channel for welfare improvements for borrowers, as non-repayment can severely impact credit and employment prospects. These results also speak to the long-standing debate over the determinants of default. Increased repayment following a reduction in minimum payments is suggestive of a liquidity motive for default rather than a strategic motive, as lower monthly payments should not influence strategic default decisions.

Re-enrollment results highlight potential importance of behavioral barriers and re-certification rules in IDR design. The steep increase in payments 12 to 14 months following the delinquency call corresponds to the one-year recertification period when borrowers are required to provide updated proof of income or revert to standard payment levels. High attrition from IDR during this period implies many borrowers have failed to recertify. Either their incomes have increased above the level which would make them eligible for reduced payments, the "hassle-costs" of recertification exceed the expected benefit of continuing IDR, or behavioral phenomena like inattention or myopia prevented them providing proof of income.

Delinquency, repayment, and re-enrollment results highlight the importance of considering counterfactual repayment behavior when evaluating the budgetary implications of student loan reforms. Programs that offer lower monthly payments and potential debt forgiveness seem expensive, but may be budget neutral or even generate revenue depending on their repayment effects. In the case of IDR, back-of-the-envelope fiscal simulations suggest the costs of loan forgiveness may be small, and the savings from fewer defaulted loans may be large.[31]

While monthly payments results show that the increase to cash-on-hand through IDR is short-lived, long-run results suggest its effects may be quite persistent. The positive estimated effects of IDR on credit scores, mortgage-holding rates, and zip-median income are suggestive of long-lasting welfare improvements to liquidity-constrained borrowers through two channels—a direct response to the immediate increase in cash-on-hand, and an indirect effect through the increased credit access associated with higher credit scores. This indirect credit channel may be an important one, as prior research finds a ten-point increase in credit scores can increase credit card balances by more than \$500 one year later (Dobbie et al. 2016).

Finally, the discussion above interprets IDR treatment effects as operating through liquidity effects, transferring cash-on-hand *within* borrowers, but debt forgiveness provisions under IDR provide a potential alternative mechanism. If borrowers expect their loans to be forgiven, they may increase repayment to try and qualify for forgiveness or raise short-term consumption out of increases to their expected lifetime wealth. Evidence for such wealth effects would also raise concerns about moral hazard, as borrowers may distort their labor supply decisions if they expect income-contingent loan forgiveness. However, persistently low recertification rates suggests most borrowers in my sample should not expect future loan forgiveness.[32] It's possible that post-2018 improvements in the IDR program increase this likelihood (US Department of Education 2021), but borrowers would have to anticipate these policy changes for wealth effects to play a meaningful role in behavior during my sample period.

D. *Generalizability and Selection into IDR*

This study aims to identify the average treatment effect of IDR among those targeted by the policy—borrowers who would qualify from reduced payments and plausibly benefit from the program. In this section, I consider how well my estimates might generalize to this population.

First, I compare both my analysis sample and the "full" LLS-representative sample from which it is drawn to corresponding subsamples in a separate, nationally representative dataset from the 2008/2012 Baccalaureate and Beyond Longitudinal Study (B&B).[33] Ideally, the full sample would be representative of the student borrowing population, and the analysis sample would be representative of borrowers eligible for lower payments under IDR. Columns 1 and 3 of online Appendix Table A14 provide summary statistics for the full and analysis samples in the LLS data, restricted to include only 2008 graduates. Columns 2 and 4 report the

[31] In online Appendix Section C, I use my estimates to create simulations of the budgetary impacts of IDR, accounting for re-enrollment and repayment effects. While projections are speculative and dependent on future policies, I find considerably lower fiscal cost estimates relative to existing studies (Lucas and Moore 2010; Di and Edmiston 2017), which generally assume perfect repayment and zero attrition from the program.

[32] Online Appendix Section C expands upon this argument with projections of future loan forgiveness.

[33] Provided by the National Center for Education Statistics (NCES), the B&B data include restricted-use administrative loan and financial aid records linked to survey responses for a representative sample of four-year US college graduates in the spring of 2008, followed up in 2011–2012 (NCES 2016).

corresponding statistics for two comparable subsamples of the B&B data. The first sample includes all B&B borrowers who took out federal loans. The second sample includes the "IDR eligible" borrowers in the B&B data—those whose reported 2012 incomes and loan balances would have qualified them for reduced payments under IDR.[34] Mean values for variables common to the two data sources are very similar in both comparison samples, suggesting my study sample is largely representative of the policy-relevant population.

Next, I investigate differences between IDR and non-IDR borrowers in both full and analysis samples. The comparison provides a descriptive sense of the types of borrowers driving my estimates and where they fall in larger distribution of student borrowers. Column A of online Appendix Figure A15 plots histograms of 2013 credit scores, loan balances, and zip-median incomes for IDR enrollees and non-enrollees in the full sample of borrowers. IDR borrowers have lower credit scores, higher debt balances, and live in lower income zip codes than their non-enrolling counterparts. While there is negative selection into IDR in the full sample of borrowers, that selection pattern all but disappears in the analysis sample. Column B of Figure A15 plots the same histograms as column A restricted to only those in the analysis sample. As of 2013, IDR enrollees and non-enrollees in this sample have similar distributions of credit scores, debt balances, and zip-median incomes. The observed negative selection into IDR is almost entirely captured by selection into the analysis sample. This suggests LLS's outgoing delinquency calls, combined with their "modeling-out" procedure and IDR eligibility requirements, effectively target the "right" individuals for IDR—financially distressed borrowers who might gain from the program.

Despite evidence that analysis sample represents a policy-relevant population for IDR, there remain some limits to the generalizability of my results. First, individuals in my analysis sample are restricted to those with loans originating prior to 2010. This selection criterion removes many borrowers for whom we would expect IDR to be most effective, as younger borrowers typically have higher debt-to-income ratios. Second, I estimate effects among those who enrolled in IDR between 2014 and 2018. Since this period, IDR has since expanded to many more borrowers, so the marginal IDR enrollee has likely changed. Third, I estimate effects of a specific variant of IDR known as Income-Based Repayment (IBR). While IBR is the largest IDR plan in the United States and shares most features with alternatives like Pay-As-You-Earn (PAYE), results may not extend to international IDR plans or hypothetical repayment schemes of policy relevance.[35] In fact, in the years following my sample period, IDR sign-up and recertification procedures have been simplified through the FUTURE Act (US Department of Education 2021).

Finally, note that the treatment effects I estimate are relative to a counterfactual that includes a multiple outside options. A borrower who declines IDR may instead

[34] Note that I cannot construct the analogous "IDR-eligible" subsample in the LLS data because they do not contain income information. Instead, column 3 of online Appendix Table A14 reports summary statistics for the analysis sample, which was constructed to approximate the population of borrowers eligible for reduced payments under IDR (see Section IIA).

[35] Estimates are specific to the Income-Based Repayment (IBR) because my sample is comprised exclusively of FFEL borrowers (see footnote 8), who are ineligible for alternative IDR plans like PAYE and REPAYE. PAYE and REPAYE reduce payments to 10 percent of discretionary income and forgive remaining balances after 20 years. A full description of each plan's eligibility rules and repayment terms can be found at www.studentaid.gov.

put their loans in forbearance, which temporarily pauses their monthly payments. They may even convert their loan into a Direct Loan to qualify for an alternative IDR plan like REPAYE. While these options are available to both non-IDR and IDR borrowers, a differential take-up between the two groups means part of my treatment effect could be driven through a forbearance or consolidation channel. Online Appendix Figure A16 plots the incidence of forbearance between IDR and non-IDR borrowers. Most IDR and non-IDR borrowers take up forbearance in the months immediately following the delinquency call. By month 6, however, most IDR borrowers have opted out of forbearance while roughly one-third of non-IDR borrowers remain enrolled. It is therefore important to interpret IDR effects as relative to a forbearance-optional counterfactual, which are likely smaller than IDR effects relative to a strict repayment regime. Consolidation, by contrast, does not appear to be a relevant outside option for non-IDR borrowers in my sample. Only eight total delinquency calls in my sample (0.02 percent) are followed by loan consolidation.

V. Conclusion

In this paper, I use administrative student loan servicing data to estimate the causal effect of IDR enrollment on borrower outcomes. Exploiting quasi-random assignment of loan-servicing agents to delinquency calls, I find that IDR lowers monthly minimum payments by $172 within eight months of take-up and reduces delinquencies by 22 percentage points. Despite facing lower monthly minimums, IDR borrowers pay down $36 *more* debt each month than standard borrowers during this period. Difference-in-differences estimates find that IDR enrollees are 2.0 percentage points more likely to hold mortgages and 1.8 percentage points more likely to move to a higher-income zip code than non-enrollees three years after enrollment.

These results do not appear driven by borrower responses to expected loan forgiveness. Instead, they suggest IDR improves borrower welfare principally through a liquidity channel, providing short-term increases to cash-on-hand during periods of financial distress. Indeed, despite its persistent effects on long-run outcomes, the period of reduced payments under IDR is remarkably short, largely because most IDR borrowers fail to recertify their incomes after one year.

The policy implications of this study are twofold. First, it illustrates the benefits of flexible student loan contracts. Relative to standard, flat repayment plans, IDR helps borrowers smooth consumption, invest in homes, and avoid default during periods of financial distress. For many borrowers, these liquidity benefits appear inaccessible through private lending markets, leaving considerable scope for other policies that improve contracts for financing college, particularly those that implicitly extend credit or insurance to the student borrowing population (Herbst and Hendren 2021).

Second, my findings demonstrate the importance of considering behavioral phenomena in the design of such contracts. While my first-stage estimates add to existing evidence of psychological frictions in student-loan repayment (Cox, Kreisman, and Dynarski 2018; Abraham et al. 2018b; Dynarski et al. 2018; Marx and Turner 2017), my re-enrollment findings highlight how the *persistence* of such frictions can compound these behavioral effects when borrowers are confronted with an onerous recertification process. If policymakers want IDR to provide more than just short-term increases to cash-on-hand, IDR re-enrollment must be streamlined or automated.

IDR represents the largest change to higher education financing in more than 50 years. Measuring its impact requires many considerations—the positive externalities of college, the redistributive impact of subsidies, the welfare gains from insuring earnings, and the distortionary costs of income-contingent benefits. While many of these questions remain unanswered, this study provides a crucial first step. These findings speak not only concerns of existing student loan policy, but also to the larger question of how society can best finance investments in human capital.

REFERENCES

Abraham, Katharine G., Emel Filiz-Ozbay, Erkut Y. Ozbay, and Lesley J. Turner. 2018a. "Behavioral Effects of Student Loan Repayment Plan Options on Borrowers' Career Decisions: Theory and Experimental Evidence." NBER Working Paper 24804.

Abraham, Katharine G., Emel Filiz-Ozbay, Erkut Y. Ozbay, and Lesley J. Turner. 2018b. "Framing Effects, Earnings Expectations, and the Design of Student Loan Repayment Schemes." NBER Working Paper 24484.

Avery, Robert B., Paul S. Calem, Glenn B. Canner, and Raphael W. Bostic. 2003. "An Overview of Consumer Data and Credit Reporting." *Federal Reserve Bulletin* 89: 47–73.

Barr, Nicholas, Bruce Chapman, Lorraine Dearden, and Susan Dynarski. 2017. "Getting Student Financing Right in the US: Lessons from Australia and England." Unpublished.

Bleemer, Zachary, Meta Brown, Donghoon Lee, Katherine Strair, and Wilbert Van der Klaauw. 2017. "Echoes of Rising Tuition in Students' Borrowing, Educational Attainment, and Homeownership in Post-Recession America." FRBNY Staff Report 820.

Britton, Jack, Laura van der Erve, and Tim Higgins. 2019. "Income Contingent Student Loan Design: Lessons from around the World." *Economics of Education Review* 71: 65–82.

Britton, Jack W., and Jonathan Gruber. 2019. "Do Income Contingent Student Loan Programs Distort Earnings? Evidence from the UK." NBER Working Paper 25822.

Chapman, Bruce. 2006. "Income Contingent Loans for Higher Education: International Reforms." In *Handbook of the Economics of Education,* Vol. 2, edited by E. Hanushek and F. Welch, 1435–1503. Amsterdam: Elsevier.

Chapman, Bruce, and Andrew Leigh. 2009. "Do Very High Tax Rates Induce Bunching? Implications for the Design of Income Contingent Loan Schemes." *Economic Record* 85 (270): 276–89.

Cox, James C., Daniel Kreisman, and Susan Dynarski. 2018. "Designed to Fail: Effects of the Default Option and Information Complexity on Student Loan Repayment." NBER Working Paper 25258.

Di, Wenhua, and Kelly D. Edmiston. 2017. "Student Loan Relief Programs: Implications for Borrowers and the Federal Government." *ANNALS of the American Academy of Political and Social Science* 671 (1): 224–48.

Dobbie, Will, Paul Goldsmith-Pinkham, Neale Mahoney, and Jae Song. 2016. "Bad Credit, No Problem? Credit and Labor Market Consequences of Bad Credit Reports." NBER Working Paper 22711.

Dobbie, Will, Paul Goldsmith-Pinkham, and Crystal S. Yang. 2017. "Consumer Bankruptcy and Financial Health." *Review of Economics and Statistics* 99 (5): 853–69.

Dynarski, Susan, C. J. Libassi, Katherine Michelmore, and Stephanie Owen. 2018. "Closing the Gap: The Effect of a Targeted, Tuition-Free Promise on College Choices of High-Achieving, Low-Income Students." NBER Working Paper 25349.

Field, Erica. 2009. "Educational Debt Burden and Career Choice: Evidence from a Financial Aid Experiment at NYU Law School." *American Economic Journal: Applied Economics* 1 (1): 1–21.

Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, and Oregon Health Study Group. 2012. "The Oregon Health Insurance Experiment: Evidence from the First Year." *Quarterly Journal of Economics* 127 (3): 1057–1106.

Friedman, Milton. 1955. "The Role of Government in Education." In *Capitalism and Freedom*, 85–107. Chicago: University of Chicago Press.

Gicheva, Dora, and Jeffrey Thompson. 2015. "The Effects of Student Loans on Long-Term Household Financial Stability." In *Student Loans and the Dynamics of Debt*, edited by Brad J. Hershbein and Kevin M. Hollenbeck, 287–316. Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (1): 153–61.

Herbst, Daniel. 2023. "Replication data for: The Impact of Income-Driven Repayment on Student Borrower Outcomes." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.38886/E150261V1.

Herbst, Daniel, and Nathaniel Hendren. 2021. "Opportunity Unraveled: Private Information and the Missing Markets for Financing Human Capital." NBER Working Paper 29214.

Johnston, Alison, and Nicholas Barr. 2013. "Student Loan Reform, Interest Subsidies, and Costly Technicalities: Lessons from the UK Experience." *Journal of Higher Education Policy and Management* 35 (2): 167–78.

Large Loan Servicer (LLS). 2018. "Anonymous Large Loan Service Borrower Rcords." Unpublished data. Accessed June 27, 2018.

Lucas, Deborah, and Damien Moore. 2010. *Costs and Policy Options for Federal Student Loan Programs. A CBO Study*. Washington, DC: Congressional Budget Office.

Marx, Benjamin M., and Lesley J. Turner. 2017. "Student Loan Nudges: Experimental Evidence on Borrowing and Educational Attainment." NBER Working Paper 24060.

Mezza, Alvaro, Daniel Ringo, Shane Sherlund, and Kamila Sommer. 2020. "Student Loans and Homeownership." *Journal of Labor Economics* 38 (1): 215–60.

Michigan Population Studies Center (MPS). 2017. "Zip Code Characteristics: Mean and Median Household Income." https://www.psc.isr.umich.edu/dis/census/Features/tract2zip (accessed June 27, 2017).

Mueller, Holger M., and Constantine Yannelis. 2019. "Reducing Barriers to Enrollment in Federal Student Loan Repayment Plans: Evidence from the Navient Field Experiment." Unpublished.

National Center for Education Statistics (NCES), Department of Education. 2016. "2008/12 Baccalaureate and Beyond (B&B) study data." Restricted-use data. Accessed October 11, 2016.

Tang, Cong, Keith Ross, Nitesh Saxena, and Ruichuan Chen. 2011. "What's in a Name: A Study of Names, Gender Inference, and Gender Behavior in Facebook." In *Database Systems for Advanced Applications,* edited by Jiangliang Xu et al., 344–56. New York: Springer.

US Census Bureau. 2010. "Selected Characteristics of the Native and Foreign-Born Populations: 2006–2010 American Community Survey 5-Year Estimates." US Department of Commerce, Economics and Statistics Administration.

US Department of Education. 2020a. "Consequences of Default." US Department of Education. https://studentaid.gov/manage-loans/default (accessed February 13, 2020).

US Department of Education. 2020b. "Federal Student Loan Portfolio." US Department of Education. https://studentaid.gov/data-center/student/portfolio (accessed September 20, 2021).

US Department of Education. 2020c. "Repayment Plans." US Department of Education. https://studentaid.gov/manage-loans/repayment/plans (accessed February 13, 2020).

US Department of Education. 2021. "Data Sharing for a Better Customer Experience." US Department of Education. https://studentaid.gov/sites/default/files/future-act-fact-sheet.pdf (accessed September 17, 2021).