Are For-Profits Manipulating Cohort Default Rates? How Student and School Characteristics Are Related to the Timing of Default

Erin Dunlop Velez[1,2]

Austin Lacy[1]

Johnathan Conzelmann[1]

March 2019

Abstract

Cohort Default Rates (CDRs) are meant to both sanction schools that produce a high fraction of defaulting students and warn students about their likelihood of default if they attend various institutions. However, there is evidence that institutions manipulate these rates by encouraging their students to enter into deferments and forbearances during the CDR period. While there is some descriptive evidence that private for-profit institutions are the main perpetrators of this manipulation, little empirical evidence is available due to previous data limitations. Using newly released student level data, we empirically measure whether CDR manipulation by for-profits is occurring and quantify the magnitude of the manipulation. We find evidence that for-profit institutions are manipulating their CDRs and the size of this manipulation is large. Furthermore, we investigate what other student and school characteristics are related to the timing of default. We find that students who are traditionally thought of as high defaulters – typically at-risk students – are mostly just high defaulters in the short-run. Different types of students – those who are typically less disadvantaged students – default later in their repayment careers.

[1] RTI International

[2] Corresponding author: evelez@rti.org

1. Introduction

The largest federal investment in postsecondary education involves the provision of grants and loans to students enrolling in Title IV eligible institutions. In the 2016-17 federal financial aid year, $137 billion dollars in federal grants and loans were distributed, with loans comprising 70% of this volume ($96 billion).[3] While recent metrics surrounding repayment have emerged, default – namely the Cohort Default Rate (CDR) – is the chief accountability mechanism for student loans. Default, in the context of student loans, has negative consequences both for borrowers and for postsecondary institutions. Borrowers face increased costs and negative effects on their credit when they default, among other difficulties. Postsecondary institutions face sanctions from the federal government, such as the loss of access to federal Title IV financial aid for their students, if their default rate exceeds 30 percent for three straight years, or 40 percent for one year.

One of the shortcomings of the cohort default rate as an accountability mechanism is that borrowers are measured at a single point in time – three years after entering repayment. As others have noted, these CDRs can be manipulated. The Department of Education itself recognized this saying “Repayment rates are generally considered more sensitive than default rates, which measure only the worst-case scenario for repayment outcomes and which can be manipulated through the use of allowable nonrepayment options like deferments and forbearances.”[4] Additionally, a Government Accountability Office report recently found that “to manage these 3-year default rates, some schools hired consultants that encouraged borrowers with past-due payments to put their loans in forbearance, an option that allows borrowers to

[3] Trends in Student Aid 2018 report, Table 1, The College Board: https://trends.collegeboard.org/student-aid/figures-tables/total-aid#Total%20Student%20Aid.
[4] https://collegescorecard.ed.gov/assets/FullDataDocumentation.pdf

temporarily postpone payments. While forbearance can help borrowers avoid default in the short-term, it increases their costs over time and reduces the usefulness of the 3-year default rate as a tool to hold institutions accountable."[5]

There is some descriptive evidence that for profits, specifically, are manipulating cohort default rates.[6] But this work is limited given the lack of student level repayment information that has historically been available. Last year the National Center for Education Statistics released the 2015 Federal Student Aid Supplement (FSA Supplement) to the 2004/09 Beginning Postsecondary Student Longitudinal Study (BPS:04/09). These new data enable researchers to examine student loan repayment and default, using a nationally representative sample, over a much longer time horizon than previously possible: up to 12 years after students first enter postsecondary education.

The first goal of this paper is to empirically model whether for-profits are manipulating CDRs, and if so, to measure the magnitude of this manipulation. We do this by estimating a difference-in-differences model, where we interact institution control (public, private nonprofit, private for-profit) with whether a student defaulted in a quarter directly after the CDR period ended. We find evidence that for-profits are manipulating cohort default rates, and the magnitude of the effect is large: the odds of for-profit students defaulting, relative to public institution students, in a quarter directly after the CDR period ends are about 10 times higher than the odds of for-profit students defaulting, relative to public students, in any other quarter.

The fact that we have shown students' likelihood of default is not constant over time begs another question: are the types of students who default early, for example within the first 3 years

[5] https://www.gao.gov/products/GAO-18-163.
[6] See The Institute for College Access and Success' (TICAS) Comments on Topics for Negotiated Rulemaking, for a summary: https://ticas.org/sites/default/files/pub_files/ticas_dtr_neg_reg_comments.pdf

after entering repayment, the same as the types of students who default later? With these new data, we have found that just over half (52 percent) of all students who defaulted within 12 years of beginning postsecondary education defaulted within 0 – 3 years of entering repayment, the time captured by CDRs. Another 32 percent of defaults occurred between 3 and 6 years and 16 percent of defaults occurred after 6 years in repayment. Before these data were available, it was assumed that most defaults happened within the first few years after students exited college. This is when individuals may struggle to find a job or be working in an entry level position. These recently released repayment data indicate that many students default well into their repayment careers.

It is reasonable to hypothesize that the reasons that students default when they first enter repayment are different from the reasons they default later, and thus the types of students that default within the first 3 years after entering repayment are different than the types of students who default after being in repayment for 6 or more years. A second goal of this paper is to measure how student characteristics are related to the timing of default. We find that groups who we typically think of as defaulting – typically disadvantaged subgroups – are more likely to default quickly, but that students not typically associated with default – higher income and higher ability students – are the ones defaulting later.

The question of what types of students default when is important for two reasons. First, this speaks to the types of students currently being measured by CDRs. Currently CDRs are not just measuring which schools produce the most defaults, but which schools produce the most defaults quickly. If the types of students who default after the 3-year CDR period ends are different than the types of students who default within the 3-year CDR timeframe, we need to recognize that only certain students' defaults are being captured by CDRs. Second, there are many interventions aimed at reducing default. These interventions are either not targeted at all or targeted at students who default in the short-run (the only types of student defaults we knew about, until recently). More accurate information about the types of students who default when will help these interventions target the students in most need for help, when they most need it.

2. Literature Review

Interest in debt and repayment has spawned a growing body of research on default, both from the perspective of students and institutions. Literature where institutions are the unit of analysis has typically focused on the utility of cohort default rates, often choosing to compare the magnitude of parameters from a model of CDRs to models of alternative repayment rates (Belfield, 2013; Chou, Looney, and Watson, 2017; Kelchen and Li, 2017). A common finding among these papers is that student-body characteristics like racial profile and the proportion of low-income or first-generation students is associated with default and repayment. In addition, these papers tend to focus on the differences between the outcomes of private for-profit institutions and public or private nonprofit institutions. The former tends to exhibit higher default rates and lower repayment outcomes, even after controlling for enrolled student characteristics.

Though institution-level analyses can be useful for evaluating accountability policies like CDRs, they do not provide detailed enough information or estimates needed to understand default for borrowers. Many researchers have thus taken advantage of student-level data to analyze default and repayment, beginning as far back as the 1980s, when student loans had only recently become an issue of public concern. Early studies, often using public data like those from the National Postsecondary Student Aid Study (NPSAS), concluded that “default rates stem chiefly from students’ background characteristics rather than characteristics…of the institutions they attend" (Wilms, Moore, & Bolus, 1987). This statement was bolstered by other studies that found no or modest association between the type of institution a student attended and the likelihood of default (Knapp and Seaks, 1992; Volkwein et. al, 1998). Instead, background characteristics including family status, parental income, race, and citizenship were considered significant predictors of default. This early wave of research consistently found a negative relationship between degree completion and default as well (Greene, 1989; Knapp and Seaks, 1992; Volkwein et. al, 1998; Wilms, et. al, 1987), suggesting that students who obtained a degree were less likely to default than those who dropped out. At least one study did find institutions to have a significant effect on the probability of default, which was contradictory to the related literature at the time (Monteverde, 2000). However, the sample contained only law students that borrowed private, non-federal loans through the Law Access Loan program in 1991-92, calling into question the generalizability of the findings.

More recent studies of default emphasize institutions' effect on default and find these factors to be stronger and more significant than the earlier research, ushering in the second generation of default literature. These studies found students who attended for-profit institutions (Deming, Goldin, and Katz, 2012; Hillman, 2014; Looney and Yannelis, 2015) or publicly traded career and vocational schools (Woo, 2002) were significantly more likely to experience default. College selectivity was also found to be negatively related to default, where borrowers attending less selective schools experienced higher probabilities of default (Looney and Yannelis, 2015; Podgursky, Ehlert, Monroe, Watson, and Wittstruck, 2002).

Though the relationship between institution-level characteristics and default has evolved over time, many of the same student-level characteristics and their association to default have remained. For example, students who fail to complete their degrees have been consistently more likely to default than those who complete a credential of any kind (Deming, Goldin, and Katz, 2012; Greene, 1989; Hillman, 2014; Knapp and Seaks, 1992; Podgursky, Ehlert, Monroe, Watson, and Wittstruck, 2002; Volkwein, Szelest, Cabrera, and Napierski-Prancl, 1998; Wilms, Moore, and Bolus, 1987). Underrepresented minority borrowers (mainly Black borrowers) have a higher instance of default (Deming, Goldin, and Katz, 2012; Hillman, 2014; Lochner and Monge-Naranjo, 2014; Podgursky, et al., 2002), as do low-income borrowers (Christman, 2000; Hillman, 2014; Looney and Yannelis, 2015), holding all else constant. A student's ability, measured by SAT scores or GPA also has a negative relationship to default, where those with high test scores and GPAs are less likely to default than those with lower measures on these items (Podgursky, et. Al, 2002; Christman, 2000; Woo, 2002).

The default literature, though well-established, contains many limitations and areas for further exploration, often related to the data used. To analyze repayment outcomes like default requires measurements over a long time-horizon, which is typically absent from public data like those provided by the National Center for Education Statistics (NCES) at the U.S. Department of Education. Generally, studies using data from NCES are only able to measure outcomes after six years at the maximum for those using the Beginning Postsecondary Students Longitudinal Study (BPS) (e.g., Hillman, 2014). At least one study (Lochner and Monge-Naranjo, 2014) utilizes the 1993/03 Baccalaureate and Beyond Longitudinal Study to examine repayment and default. These data contain up to 10 years of repayment information, but this is a rare case.

There are several contributions of the current work. First, we use newly available loan repayment data that includes a much longer loan repayment time horizon than most previous studies. Second, we are able to test for manipulation of CDRs specifically, looking for deviations in the default rate right after the CDR period ends. Finally, we consider the time aspect of default. Until recently, default was considered to just be a binary outcome measured in the short run. We look at default at several periods of time and are able to measure whether the types of borrowers who default in the short-run are similar to borrowers who default in the long-run.

3. Data

The data used in this analysis come from the National Center for Education Statistics' 2015 Federal Student Aid Supplement (FSA Supplement) to the 2004/09 Beginning Postsecondary Student Longitudinal Study (BPS:04/09). BPS:04/09 surveyed a nationally representative sample of postsecondary students who began college in 2003-04. The students were surveyed approximately 1, 3 and 6 years after they first began postsecondary education. Rich information was collected on their postsecondary enrollment, experiences while enrolled, degree attainment, and post-college employment and other activities.

The FSA Supplement appends federal student loan records through June 2015 to the BPS:04/09 survey data. The FSA supplement data come from the National Student Loan Data System (NSLDS), a federal student loan database that includes historic information on every federal financial aid transaction. These new data allow researchers to examine student loan repayment and default for up to 12 years after students first began postsecondary education.

3.1 Descriptive Statistics

Table 1 describes first-time postsecondary students in 2003-04 who borrowed federal student loans. The variables included in Table 1 are the control variables used in the Difference-in-Differences and regression models, described below. Those covariates include: demographics (age, sex, race, parental education, family income, and marriage and dependents status), characteristics describing the individual's postsecondary experience (control of first institution attended, major, GPA, and degree attainment), characteristics describing individual’s student loan borrowing (cumulative amount borrowed, whether the student ever borrowed for graduate school, and the number of deferments for economic hardship or forbearances used), and characteristics describing the individual’s employment 6 years after first beginning postsecondary education (earned income).

The descriptive statistics in Table 1 indicate that federal loan borrowers who began college in 2003-04 were 61 percent female, 61 percent White, 37 percent from families with parents who had earned a bachelor’s degree or higher, and 82 percent were unmarried with no dependents when they began college. Some 62, 18, and 20 percent first began college at a public, private nonprofit, and private for-profit institution, respectively. Of these borrowers, 54 percent earned a degree within 6 years of beginning postsecondary education, they borrowed an average amount of $26,200 for their undergraduate education, and they took an average of 3 economic hardship deferments or forbearances within 12 years of beginning postsecondary education.

Some 73 percent of the sample did not default within 12 years of beginning postsecondary education, 14 percent defaulted in 3 years or less, 9 percent defaulted in 3 to 6 years, and 4 percent defaulted after more than 6 years.

3.2 Data Limitations

While the FSA Supplement provides a rich source of data for researchers, some limitations should be considered. First, although students are observed over a 12-year period after they first began postsecondary education, this does not mean they were in repayment for 12 years. Individuals enter repayment 6 months after exiting college, regardless of whether they dropped out or graduated. So even though students are observed for 12 years after they first enrolled, we observe fewer than 12 years of repayment, accounting of the time each student spent enrolled or in the grace period. Also, once a student finishes completely paying off his or her loans, the student is no longer in repayment and no longer included in our sample. Table 1 shows the distribution of quarters observed in repayment, with the median student being observed in our data in repayment for 31 quarters, about 8 years.

This presents two data challenges. First, we do not observe the same number of quarters in repayment for everyone in our sample. Students who are enrolled more quarters have shorter repayment histories in our data than other students. Given our long window of analysis – 12 years after students first begin postsecondary education – this creates less bias than in some data sets with shorter repayment windows, but it is still an important factor to consider.

Second, even though we use data with the longest observed repayment period available, our data are still censored. While we observe 12 years of data after students first enroll, we only observe an average of about 8 years of repayment. And while we capture twice as many defaults as 3-year CDRs, we are likely not capturing all the defaults that occur. Given this truncation of our outcome of interest, this introduces some bias in any OLS or logit regression model. As such, we also estimate a Cox Proportional Hazard model as a robustness check. One of the advantages of this model is that it accounts for censoring in the outcome.

4. Estimation Strategy

A complication in our analysis is that borrowers in our data entered repayment under both the 2-year and 3-year CDR periods. Borrowers in our data first entered repayment between January 2005 and June 2015. The federal government switched from calculating 2-year to 3-year CDRs starting in fiscal year 2009. This means everyone who entered repayment before October

2008 was involved in the calculation of 2-year cohort default rates, while everyone who entered repayment in October 2008 or later was involved in the calculation of 3-year cohort default rates.

While we could divide our sample among those who entered repayment under the 2-year CDR era or the 3-year CDR era, there are several reasons why we choose not to do this. First, there is nonrandom selection into who entered repayment during the 2-year versus 3-year CDR era – those who were enrolled longer are more likely to have entered under the 3-year era. Second, given the relatively small sample size among those who entered repayment during the 3-year CDR era (about 30 percent of our sample), dividing the sample means we will lose precision in the model for this group. Finally, we are not interested in comparing the relative size of the 2-year CDR manipulation to the 3-year CDR manipulation, we just want to document whether in general, manipulation is taking place. Given all these factors, we keep the sample pooled, and looked for manipulation of the CDR in the quarters directly after the end of a given CDR period.

If there was manipulation of CDRs, we would expect a spike in default directly after the CDR period ends. The way certain institutions manipulate CDRs is by encouraging their students to use forbearances and deferments which do not prevent default, but delay it. In the absence of manipulation, the probability of default in any given quarter would be smooth, with rises and falls in the rate over time. But with manipulation, we would likely see a mass of defaults right after the CDR period ends, the accumulation of all the defaults postponed through forbearance and deferment throughout the CDR period. Figure 1 shows just that, providing descriptive evidence that CDR manipulation is occurring.

Figure 1 shows the percent of borrowers who first default in a given quarter after entering repayment, by institution control. There is a spike in default rates for all institution types in quarter 5, the first quarter students can default after entering repayment. The default rates for public and private non-profit institutions are relatively flat in the subsequent quarters. The default rates for for-profits, however, show two additional noticeable spikes, one at quarter 9, the quarter directly after the 2-year CDR is measured and one at quarter 13, the quarter directly after the 3-year CDR is measured. This is descriptive evidence that for-profit institutions are manipulating the CDRs. In the quarter directly after the CDR period ends, there is a large spike in for-profit default rates that is not seen in other institution controls.

To empirically measure the size of this manipulation, we estimate a difference-in-differences model. We re-shaped our data so that it is at the borrower-quarter level, and the number of quarters each borrower is observed in the data is the number of quarters we observe them in repayment. None of our covariates vary over quarters, but our outcome of interest, first default, does. We estimate a logit model on the panel data where the outcome of interest is quarter of first default, and the key explanatory variables are first institution control (public, private nonprofit, private for-profit), whether the quarter is directly after a CDR period ends, and the interaction between the two variables. We count the 2 subsequent quarters after a CDR period ends, meaning we count default in quarters 9, 10, 13, and 14 as occurring directly after a CDR period ends. Out main empirical model can be expressed as:

$$Default_{it} = \beta_0 + Control_i\beta_1 + AfterCDR_t\beta_2 + Control_i * AfterCDR_t\beta_3 + X_i\,\beta_4 + \mu_i$$

where $Default_{it}$ is an indicator for whether respondent *i* first defaulted in quarter *t*; $Control_i$ is an indicator for whether respondent *i*'s first institution control was a public, private nonprofit, or private for-profit institution; and $AfterCDR_t$ is an indicator for quarter *t* being 9, 10, 13, or 14, the two quarters directly after the 2-year and 3-year CDRs are measured.

The second research question involves measuring what other characteristics of students vary with the timing of default. After all, if students at different institution types default at different times, perhaps the timing of default varies across other student characteristics as well. To answer this research question, we estimate two different models. The first in an OLS regression, where the outcome is years between entering repayment and first federal student loan default. Only students who default have a value for this outcome, so this model measures what characteristics predict the timing of default, among all defaulters. The second set of models predict whether a student first defaults in a given period of time: between 0 – 3 years, between 3 – 6 years, and in greater than 6 years. These models show what characteristics are related to defaulting later in one's repayment career, compared to earlier. Across all models, the same set of rich covariates is used.

5. Results

Overall, we find that students who attend for-profit institutions are significantly more likely to default in the quarters directly after the CDR period ends. In addition, we find that many student characteristics, such as sex, race, parental education, having dependents, major, GPA, and number of deferments and forbearances are related to the timing of default.

5.1 For-Profit Defaults Around CDR Periods

Table 2 reports odds ratios for the results of the difference-in-differences model that interacts institution control with whether the quarter is directly following the end of a CDR period. We find that overall, students who first attend private nonprofits have 12 percent lower odds of defaulting in a given quarter and students first attending private for-profits have 12 percent higher odds of defaulting in a given quarter, relative to students at public institutions, although these effects are only significant at the 0.10 percent level. There is no statistically significant effect for being in a quarter directly after a CDR period ends for the overall sample. However, when looking at the effects of control interacted with quarter, we see that private nonprofits are no more likely to default in the quarters directly after the CDR period ends, while for-profits are about 110 percent more likely.

Table 3 reports the combined odds of defaulting by quarter group (all quarters except those directly after when the CDR ends and quarters directly after when the CDR ends) and control. This table shows that relative to students who begin at public institutions, students who begin at for-profits are slightly more likely to default in a given quarter other than the quarters directly following when the CDR period ends: the odds are about 12 percent higher, but only significant at the 0.10 percent level. However, in the quarters directly after the CDR period ends, students who begin at for-profits are over twice as likely (about 140 percent as likely) to default. This pattern is not seen in private nonprofit institutions. These results demonstrate that for-profit institutions are manipulating cohort default rates and the size of that manipulation is large. The likelihood of students at for-profits defaulting, relative to students at publics, is more than 10 times higher in the quarters directly after the CDR period ends, than in other quarters.

5.2 Robustness Checks in the For-profit Manipulation Model

As mentioned earlier, a shortcoming of using a logit model to measure default is that default is censored in any data. While we observe all defaults that occur within 12 years of students first enrolling in postsecondary education, it is possible that some defaults occur outside of this window. The fact that we do not observe these future defaults causes some bias in our estimation.

To understand the magnitude of this potential bias, we estimate an Event History Analysis (EHA) model. EHA is a group of regression-like models that go by several other names (i.e. survival analysis, duration analysis). These models predict the probability that a binary event will occur at time *t*, given that it has not occurred in the time period before (*t-1*). There are several advantages to this class of models, but one main advantage is that the models explicitly incorporate the fact that the outcome may be censored. The specific model we estimate is the Cox Proportional Hazard model. This specific model is well suited to social science research because it allows researchers to examine the effect of various covariates without specifying a parametric form for the baseline hazard rate.

While the details of the Cox Proportional Hazard model are somewhat different from a logit, our general estimation strategy was the same. We interacted institution control (public, private nonprofit, private for-profit) with the quarter being directly after the CDR period ends. We found that in both significance and magnitude, our EHA results were very similar to our original logit model results. This finding gives us more confidence in the results we found in our main specification and also indicates that the bias introduced from the censoring of default is minimal.

A second robustness check we ran was limiting the sample to students we observed for a given period of time. As mentioned above, the number of quarters we observe borrowers in repayment varies. Our models capture difference in repayment rates across various quarters. We want to verify that these differences are not being driven by different students being observed in each quarter. As such, we re-ran all models, limiting the sample to students we observe for at least 12 quarters and students we observe for at least 20 quarters. In both subsamples, the results were similar in magnitude and significance to our original model.

5.3 Student Characteristics and the Timing of Default

The results in Table 4 show that different types of students default at different times. Female defaulters first default about a quarter of a year earlier than male defaulters (although only significant at the 0.10 percent level). In addition, females have odds that are about 37 percent lower for defaulting in years 3-6. Black students have higher odds of defaulting in the earlier time periods – 123 percent higher odds of defaulting in years 0-3 and 83 percent higher odds of defaulting in years 3-6, relative to White students, but no statistically significant higher odds of defaulting in the latest time period.

Students whose parents did not earn a bachelor’s degree are more likely to default sooner, with 34 percent higher odds of defaulting in years 0-3, but no more likely to default in the later time periods. Students with higher family incomes are less likely to default in 0-3 years, but again, not statistically less likely to default in the later time periods. Students with dependents are more likely to default sooner as well, with 48 – 90 percent higher odds of defaulting in years 0 – 3.

Defaulters who first attended a private nonprofit institution defaulted about a third of a year later on average than students who first attended a public institution, all else equal. Additionally, the odds of private nonprofit students defaulting were about 33 percent higher in years 3 – 6 and 48 percent higher in years 6+, relative to students who began at a public institutions. Defaulters who began at private for-profit institutions defaulted about three quarters of a year later on average than defaulters who began at public institutions, all else equal. Additionally, the odds of for-profit students defaulting in all time periods were higher than the odds of public students defaulting, although the odds of students at for-profits defaulting relative to publics increased from 40 percent higher to 190 percent higher, over the time periods examined. These results indicate that current 3-year CDRs disproportionally do not capture defaults that occur at private nonprofit and private for-profit institutions. Note, this is different than these schools manipulating their CDRs. We see manipulation by looking for default in specific quarters. These finding just mean that all else equal, students at private nonprofit and for-profit institutions take longer to default than students at public institutions.

Defaulters who majored in education default about 0.85 years later on average, compared to humanities majors. Additionally, education majors have 37 percent lower odds of defaulting in years 0 – 3 and 110 percent higher odds of defaulting in years 6+. Those who complete a degree also default later. Defaulters who completed a degree default about a third of a year later on average. Additionally, those who complete a degree have 54 percent lower odds of defaulting in years 0 – 3, 38 percent lower odds of defaulting in years 3 – 6, and no statistically significantly different odds of defaulting in the latest time period.

Finally, students who took more economic hardship deferments or forbearances default later. Among defaulters, for each deferment or forbearance, students default about 0.14 years later. And among all students, the number of deferments and forbearances is associated with a lower probability of defaulting in years 0 – 3 and a higher probability of defaulting in years 3 – 6 and 6+. This demonstrates the mechanism though which for-profits manipulate cohort default rates. By encouraging students to take deferments and forbearances, they lower the likelihood of defaulting in the short-run, but the students are no less likely to default overall, so this increases the likelihood of defaulting later.

Overall, we found that students who are female, black, low parental education, and have dependents are likely to default sooner. Students who attend private nonprofit and for-profit institutions, are education majors, have high GPAs, attained a credential, and used economic hardship deferments and forbearances are likely to default later.

6. Discussion and Conclusion

There has been much discussion of the shortcomings of CDRs, chiefly because they can be manipulated and as such, are not effective as an accountability metric and are also not effective for warning students away from high default schools. While there has been some descriptive evidence that for-profit institutions are the main perpetrators of this manipulation, little empirical evidence has previously been available, due to a lack of data suited to this type of analysis.

The recent release of the FSA Supplement to BPS:04/09 allows researchers to have a much clearer understanding of default in the long run. We find that about half of defaults (or more), occur outside of the 3-year CDR period. These data also allow us to empirically measure whether institutions are manipulating CDRs by looking for spikes in default rates in the quarters directly after the CDR period ends.

We find evidence that for-profit institutions are manipulating CDRs and the size of the manipulation is large. Students at for-profit institutions are about 10 times more likely to default, relative to public institution students, in the quarters directly after the CDR period ends, compared to the likelihood of for-profit students defaulting, relative to public students, in any other quarter.

We also investigate whether the types of students who default changes over time. We find that the populations who are generally considered by the literature to be high defaulters – minority, low SES, low ability students – are most likely to default in the short-run, but not necessarily in the long run. Over 3 – 6 and 6+ years, different types of students default, students who are generally less disadvantaged.

The findings in this paper suggest that changes to the CDR need to be made to make it less susceptible to manipulation. One possibility for this would be to not count students who spent the majority of their first 3 years in repayment in forbearance or deferment as successful repayers. Making changes to the CDR would both help make it a more effective tool for accountability purposes, and also make it more useful to students, as a better indicator of which institutions produce the most defaults. Alternatively, replacing the CDR with another metric all together, such as one that measures progress towards repayment, as opposed to just the worst-case outcome of default, would also be beneficial.

Additionally, we need to be more thoughtful when describing the types of students who default. The types of students who default in the short run (until recently, the only types of defaulters we could measure) are not the same as the types of students who default in the long run. When targeting interventions at borrowers to reduce the risk of default, we need to be more specific about which types of borrowers we target when, to make these interventions most effective.

Defaulting on federal student loans in harmful for students, costing them financially and having lasting effects on their credit. In order to reduce student loan default rates, we need to do a better job of holding institutions accountable for defaults. We also need to do a better job of warning students away from high default producing schools. Finally, we need to understand that students default well into their repayment careers, and the reasons and solutions related to defaulting earlier may be different from those related to defaulting later.

References

Belfield, C. R. (2013). Student loans and repayment rates: The role of for-profit colleges. *Research in Higher Education*, *54*(1), 1-29.

Chou, T., Looney, A., & Watson, T. (2017). *Measuring loan outcomes at postsecondary institutions: Cohort repayment rates as an indicator of student success and institutional accountability* (No. w23118). National Bureau of Economic Research.

Christman, D. E. (2000). Multiple realities: Characteristics of loan defaulters at a two-year public institution. *Community College Review*, *27*(4), 16-32.

Deming, D. J., Goldin, C., & Katz, L. F. (2012). The for-profit postsecondary school sector: Nimble critters or agile predators?. *The Journal of Economic Perspectives*, *26*(1), 139-163.

Greene, L. L. (1989). An economic analysis of student loan default. *Educational Evaluation and Policy Analysis*, *11*(1), 61-68.

Hillman, N. W. (2014). College on credit: A multilevel analysis of student loan default. *The Review of Higher Education*, 37(2), 169-195.

Kelchen, R., & Li, A. Y. (2017). Institutional accountability: A Comparison of the predictors of student loan repayment and default rates. *Annals of the American Academy of Political and Social Science, 671*(1), 202-223.

Knapp, L. G., & Seaks, T. G. (1992). An analysis of the probability of default on federally guaranteed student loans. *The review of economics and statistics*, 404-411.

Lochner, L. J., & Monge-Naranjo, A. (2014). *Default and repayment among baccalaureate degree earners* (No. w19882). National Bureau of Economic Research.

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.

Monteverde, K. (2000). Managing student loan default risk: Evidence from a privately guaranteed portfolio. *Research in Higher Education*, *41*(3), 331-352.

Podgursky, M., Ehlert, M., Monroe, R., Watson, D., & Wittstruck, J. (2002). Student loan defaults and enrollment persistence. *Journal of Student Financial Aid, 32*(3), 27-42.

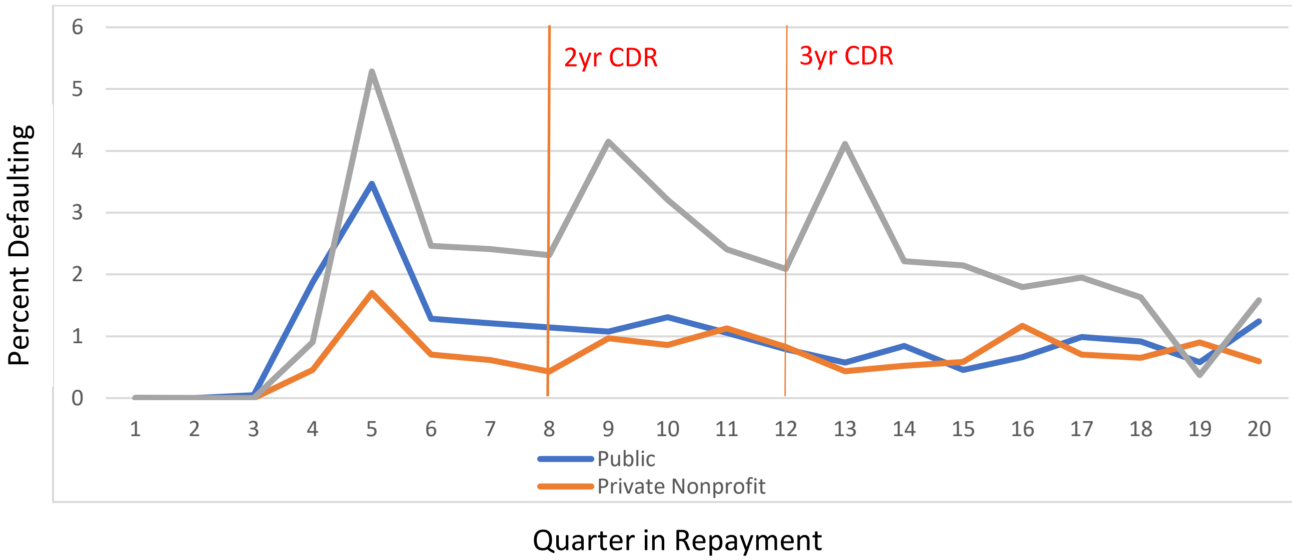
Volkwein, J. F., Szelest, B. P., Cabrera, A. F., & Napierski-Prancl, M. R. (1998). Factors associated with student loan default among different racial and ethnic groups. *The Journal of Higher Education*, *69*(2), 206-237.

Wilms, W. W., Moore, R. W., & Bolus, R. E. (1987). Whose fault is default? A study of the impact of student characteristics and institutional practices on guaranteed student loan default rates in California. *Educational Evaluation and Policy Analysis*, *9*(1), 41-54.

Woo, J. H. (2002). Factors affecting the probability of default: Student loans in California. *Journal of Student Financial Aid*, *32*(2), 5-23.

Figures

Figure 1: The Probability of First Default, by Quarter and Control of First Institution



Tables

Table 1: Descriptive Statistics of Federal Student Loan Borrowers

Characteristic	Proportion/Mean
Age	20.8
Sex	
Male	39.4%
Female	60.6%
Race/ethnicity	
White	60.6%
Black	16.9%
Hispanic	13.7%
Asian	3.7%
Other race	5.0%
Highest Parental Education	
Less than BA	62.8%
BA or higher	37.2%
Percentile of Family Income	47.1
Marriage and Dependents Status	
Unmarried, no dependents	82.2%
Married, no dependents	1.5%
Unmarried, dependents	10.6%
Married, dependents	5.7%
Institution Control	
Public	62.1%
Private nonprofit	18.1%
Private for-profit	19.8%
Major When Last Enrolled	
Humanities	11.7%
Social science	10.2%
STEM	14.6%
Education	6.1%
Business	15.1%
Health	14.2%
Other major	28.2%
GPA When Last Enrolled	
Below 1.24	2.2%
1.25-1.74	1.7%
1.75-2.24	4.4%
2.25-2.74	15.5%
2.75-3.24	22.3%
3.25-3.74	35.2%

3.75 and above	18.7%
Degree Attainment and Persistence	
Did not attain a degree within 6 years	46.4%
Attained a degree within 6 years	53.6%
Cumulative Federal Amount Borrowed	$26,244.1
Took Out Graduate-level Loan	17.3%
Number of Hardship Deferments or Forbearances	3.0
Earned Income 6 Years Later	$30,837.1
Timing of Default	
Did not default	72.7%
First defaulted in 3 years or less	14.3%
First defaulted between 3 and 6 years	8.6%
First defaulted in more than 6 years	4.3%
Number of Quarters Observed	
10th percentile	9
25th percentile	19
50th percentile	31
75th percentile	39
90th percentile	44

Source: 2004/09 Beginning Postsecondary Student Study (BPS:04/09)

Table 2. Difference-in-Difference Logit Model of Control and Quarter on Default

	(1)
Private nonprofit	0.877*
	(0.0612)
Private for-profit	1.122*
	(0.0725)
Quarter is directly after CDR period ends	1.080
	(0.129)
Private nonprofit x Quarter is directly after CDR period ends	1.073
	(0.231)
Private for-profit x Quarter is directly after CDR period ends	2.092***
	(0.359)
Age	1.002
	(0.00409)
Female	0.914*
	(0.0474)
Black	1.639***
	(0.0996)
Hispanic	1.148*
	(0.0842)
Asian	0.754
	(0.141)
Other race	1.326***
	(0.124)
Highest parental education is less than BA	1.189***
	(0.0703)
Percentile of family Income (10s)	0.901***
	(0.0300)
Squared of percentile of family income	1.004
	(0.00386)
Married, no dependents	0.924
	(0.202)
Unmarried, dependents	1.268***
	(0.0926)
Married, dependents	1.429***
	(0.160)
Major when last enrolled was social science	0.952
	(0.104)
Major when last enrolled was STEM	0.917
	(0.0918)
Major when last enrolled was education	0.987

	(0.127)
Major when last enrolled was business	0.943
	(0.0894)
Major when last enrolled was health	1.005
	(0.0986)
Major when last enrolled was other	0.978
	(0.0816)
GPA when last enrolled	0.950***
	(0.0144)
Attained a degree within 6 years	0.609***
	(0.0328)
Cumulative federal amount borrowed	0.980***
	(0.00218)
Squared cumulative federal amount borrowed	1.000***
	(7.25e-06)
Took out graduate-level loan	0.591***
	(0.0856)
Earned income 6 years later	0.998
	(0.00204)
Constant	0.0253***
	(0.00414)
Observations	284,600
Number of ID	10,000

Source: 2004/09 Beginning Postsecondary Student Study (BPS:04/09)

Note: Coefficients are reported as odds ratios. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Combined Estimates for Control and Quarter from Difference-in-Difference Model of Default

	Public	Private nonprofit	Private for-profit
All quarters except those directly after when the CDR period ends	0	0.877*	1.122*
Quarters directly after when the CDR period ends	0	0.941	2.347***

Source: 2004/09 Beginning Postsecondary Student Study (BPS:04/09)
Note: Coefficients are reported as odds ratios. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Results of the Effect of Student Characteristics on the Probability of Default in a Given Number of Years After Entering Repayment

	(1) Years Between Entering Repayment and First Federal Loan Default	(2) First Default Occurred within 0-3 Years of Entering Repayment	(3) First Default Occurred within 3-6 Years of Entering Repayment	(4) First Default Occurred within 6+ Years of Entering Repayment
Age	-0.0105	1.008	0.978*	1.011
	(0.0120)	(0.00940)	(0.0119)	(0.0116)
Female	-0.249*	1.030	0.666***	0.776
	(0.133)	(0.106)	(0.0840)	(0.136)
Black	-0.211	2.227***	1.830***	1.261
	(0.160)	(0.302)	(0.264)	(0.266)
Hispanic	0.0897	0.930	1.295	1.217
	(0.188)	(0.150)	(0.208)	(0.256)
Asian	0.443	0.393***	1.050	0.411
	(0.325)	(0.130)	(0.394)	(0.246)
Other race	-0.263	1.632**	1.063	1.489
	(0.265)	(0.328)	(0.266)	(0.402)
Highest parental education is less than BA	-0.152	1.339**	1.183	1.130
	(0.172)	(0.165)	(0.180)	(0.191)
Percentile of family Income (10s)	0.119	0.842**	1.082	0.866
	(0.0911)	(0.0583)	(0.0834)	(0.0824)
Squared of percentile of family income	-0.00737	1.003	0.989	1.007
	(0.00953)	(0.00729)	(0.00827)	(0.0102)
Married, no dependents	0.861	0.860	1.066	1.726
	(0.610)	(0.344)	(0.460)	(0.922)
Unmarried, dependents	-0.184	1.475**	1.297	1.018
	(0.193)	(0.229)	(0.231)	(0.241)
Married, dependents	-0.0400	1.901**	2.074***	1.316
	(0.243)	(0.501)	(0.541)	(0.402)

Private nonprofit	0.369**	0.918	1.332**	1.478**
	(0.146)	(0.121)	(0.175)	(0.273)
Private for-profit	0.750***	1.380***	2.319***	2.906***
	(0.164)	(0.170)	(0.282)	(0.583)
Major when last enrolled was social science	-0.251	1.141	0.648	1.274
	(0.236)	(0.239)	(0.179)	(0.432)
Major when last enrolled was STEM	0.202	0.814	0.794	1.130
	(0.249)	(0.170)	(0.183)	(0.392)
Major when last enrolled was education	0.843**	0.640**	1.045	2.105**
	(0.419)	(0.139)	(0.292)	(0.771)
Major when last enrolled was business	0.167	0.838	0.811	1.345
	(0.223)	(0.148)	(0.170)	(0.420)
Major when last enrolled was health	0.364*	0.742	1.393	1.082
	(0.215)	(0.142)	(0.319)	(0.365)
Major when last enrolled was other	0.182	0.806	0.963	1.174
	(0.184)	(0.124)	(0.184)	(0.330)
GPA when last enrolled	0.0720	0.873***	0.917***	0.987
	(0.0455)	(0.0276)	(0.0296)	(0.0562)
Attained a degree within 6 years	0.302**	0.450***	0.618***	0.851
	(0.133)	(0.0501)	(0.0765)	(0.131)
Cumulative federal amount borrowed	-0.00428	0.982***	0.989**	0.994
	(0.00635)	(0.00460)	(0.00487)	(0.00610)
Squared cumulative federal amount borrowed	4.42e-05	1.000***	1.000	1.000
	(3.17e-05)	(1.35e-05)	(3.45e-05)	(1.68e-05)
Took out graduate-level loan	0.373	0.371***	0.910	0.546*
	(0.283)	(0.0801)	(0.215)	(0.177)
Number of Hardship Deferments or Forbearances	0.142***	0.927***	1.030***	1.081***
	(0.0238)	(0.0148)	(0.0117)	(0.0148)
Earned income 6 years later	-0.00827	0.999	0.995	0.980***
	(0.00506)	(0.00395)	(0.00529)	(0.00757)
Constant	2.481***	1.007	0.234***	0.0489***

	(0.435)	(0.317)	(0.0998)	(0.0230)
Observations	2,400	10,000	10,000	10,000

Source: 2004/09 Beginning Postsecondary Student Study (BPS:04/09)

Note: *** p<0.01, ** p<0.05, * p<0.1.