2021

2021-4 Innis Lecture: Return on Student Loans in Canada

Lance Lochner

Qian Liu

Martin Gervais

Innis Lecture: Return on Student Loans in Canada

by

Lance Lochner, Qian Liu, and Martin Gervais

Working Paper #2021-04 **July 2021**

Western

Centre for Human Capital and Productivity (CHCP)

Working Paper Series

Department of Economics
Social Science Centre
Western University
London, Ontario, N6A 5C2
Canada

Innis Lecture: Return on Student Loans in Canada*

Lance Lochner
University of Western Ontario

Qian Liu
Brock University

Martin Gervais
University of Georgia

July 30, 2021

Abstract

This paper uses new administrative data with detailed borrower information and lengthy repayment histories from the Canada Student Loans Program (CSLP) to measure rates of return on undergraduate student loans. We document substantial heterogeneity in returns based on information available at the time loans were disbursed, including province of residence, field of study, and institution of attendance. Field of study is a particularly important determinant of rates of return, explaining 22% of the variation in predicted returns across borrowers. We explore the implications of this variation for CSLP cross-subsidization across borrowers and potential risk-based loan limits. Given the variation in *ex ante* predicted returns across borrowers, using all available information at the time of loan disbursement, we study the implications of potential cream-skimming of high-return borrowers by private lenders.

JEL Codes: D14, H52, H81, I22, I28, J24

*Lochner thanks Francisco Ruge-Murcia and the Canadian Economic Association for the opportunity to deliver the 2021 Innis Lecture. We thank Dajin Li and Pierre Tourigny for their help with the CSLP administrative data, Rob Adams for providing CSLP data on collections and rehabilitation for borrowers who default, Mauricio Torres Ferro for excellent research assistance, and seminar participants at the 2021 CEA Annual Meeting and the University of Western Ontario's Center for Human Capital for their comments. We gratefully acknowledge support from the Social Sciences and Humanities Research Council of Canada.

1 Introduction

The rising costs of and returns to education over the past few decades have caused many countries to re-evaluate the structure of their higher education financing systems. The efficient structure of these systems, including government student loan programs, depends critically on the strength of various market frictions stemming from incomplete information and imperfect enforcement considerations (Lochner and Monge-Naranjo, 2016). This motivates a broad research agenda aimed at identifying and quantifying these market frictions (see, e.g., Bachas, 2019; Mueller and Yannelis, 2019; Karamcheva et al., 2020; Yannelis, 2020; Lochner et al., 2021).

This paper uses newly available administrative data on student borrowing and repayment from the Canada Student Loans Program (CSLP) to estimate the returns on undergraduate student loans to the government, focusing on heterogeneity in these returns across borrowers.[1] Variation in predictable returns across students raises concerns about the efficiency of student loan programs and implies *ex ante* cross-subsidization, where borrowers with high expected returns effectively subsidize those with low expected returns. This gives rise to concerns that private creditors may undercut government loan programs by cream-skimming, i.e. poaching profitable borrower types and leaving the government with unprofitable ones.[2] This can raise program costs, forcing governments to raise interest rates on student loans (driving more high-return borrowers away), reduce loan limits, or subsidize remaining pools of borrowers. While we examine return heterogeneity and cream-skimming based on easily observed differences across borrowers, unobserved differences would further exacerbate adverse selection concerns.

A major challenge to estimating returns on student loans is a lack of publicly available data on repayment records covering much of the repayment period. As a result, previous studies have typically examined student loan repayment status, such as default rates or the proportion of loan repaid, a few years after students leave school (Gross et al., 2009; Looney and Yannelis, 2015). Yet, Lochner and Monge-Naranjo (2015) show that knowing whether borrowers had ever defaulted or are currently in default as of some arbitrary date (especially early in the repayment process) need not provide an accurate picture of returns on student

[1] We make no effort to measure the welfare benefits of student loans (overall or to specific borrowers), which depends on the value of loans to students as well as any gains/losses to the government.

[2] This practice is already underway in the US by companies such as SoFi (Kosir et al., 2015). Using data from a student loan refinancing firm, Bachas (2019) provides a careful analysis of these issues in the U.S., including the impact of offering risk-based interest rates in the student loan market on borrower welfare and government revenue.

loans or how those returns differ across borrowers.

In this paper, we use CSLP data on loan disbursements, needs determination, and repayments available from 2003–2004 to 2015–2016, which provide extensive information on borrowers and nearly complete repayment streams for those entering repayment in the mid-2000s. Based on these data, we compute a measure of the realized return on each student loan disbursed to the cohort of upper-year undergraduate borrowers that attended public post-secondary schools in Canada in 2005. While our primary measure of returns reflects the net profit on a loan relative to the amount issued, we also report internal rates of return for various subgroups of borrowers.

We begin by estimating average returns to the CSLP for the full cohort of borrowers, as well as for different observable groups based on their demographic characteristics, province, field of study, and post-secondary institution—all factors known by the CSLP at the time loans are disbursed. Our estimates suggest that the CSLP lost about 7 cents for every loan dollar disbursed to upper-year undergraduates in 2005, with considerable heterogeneity across different types of borrowers. Heterogeneity across borrowers enrolled in different fields is substantial, implying considerable *ex ante* cross-subsidization from borrowers majoring in Health Sciences to those majoring in Arts/Sciences. We also show that borrowers who default during the first 3 years of repayment generate sizeable negative returns for the CSLP, while those who experience no early repayment problems earn the CSLP modest positive returns averaging 3%. Early default rates are also negatively correlated with average rates of return when looking across borrowers from different provinces, fields of study, or institutions.

Finally, we estimate borrower-specific predicted returns based on all information available at the time of loan disbursement, documenting the considerable variation in expected returns across borrowers. Altogether, the factors observed in CSLP records explain 12% of the variation in realized returns. Based on the empirical distribution of predicted returns, we estimate that aggressive poaching of high-return students by private lenders could exacerbate losses by the CSLP by 1 to 3 percentage points. If the government were to set field-specific loan limits to ensure non-negative average returns for borrowers in each field, our estimates suggest that loan limits would differ by as much as $20,000 between Arts/Sciences majors and Health Science majors.

We provide an overview of the CSLP in Section 2, while Section 3 discusses our measures of returns on student loans. Section 4 describes the CSLP data, as well as our sample selection criteria. Section 5 characterizes measured rates of return on CSLP loans, the estimated importance of observable factors for those returns, and the distribution of predicted rates of

return. Section 6 examines the relationship between measures of early repayment problems (i.e., default and enrolment in repayment assistance) and our measures of returns based on the full repayment cycle. Section 7 discusses cross-subsidization across borrowers in different fields of study, risk-based student loan limits, and the potential impacts of cream-skimming by private lenders. Section 8 concludes with some broader lessons and directions for future reseach.

2 Student Borrowing and Repayment in Canada

Working in collaboration with most provincial/territorial governments, the CSLP provides loans and grants to help Canadian students pay for post-secondary education based on their financial need.[3] In 2005–2006, the loan year we study, roughly 345,000 Canadian students (41% of all post-secondary students in participating provinces) borrowed nearly $2 billion ($5,600 per borrower) from the CSLP, while the average university student held a total loan balance of $14,700 (Office of the Chief Actuary, 2007).

As long as students remain in school, Canada student loans accrue no interest and need not be repaid. Six months after leaving school, all loans received during school are combined into a single repayment plan in a process known as consolidation, at which time repayment begins.[4] Consolidation establishes the repayment agreement, including the amortization period and interest rate. The standard repayment plan entails a constant debt-based payment with an amortization period of 9.5 years, which can be extended up to 14.5 years under special circumstances. Borrowers can choose a floating interest rate (the default option) or a fixed interest rate.[5] Nearly everyone chooses the floating rate (see, e.g., Office of the Chief Actuary, 2019).

Borrowers experiencing periods of financial hardship can apply for reduced payments through the CSLP's Repayment Assistance Plan (RAP) for a period of six months; however, they are not automatically enrolled in the program and must re-apply if they have continued

[3] All provinces and territories participate in the CSLP, except Quebec, the Northwest Territories and Nunavut, which operate their own financial assistance programs and receive alternative funding from the federal government.

[4] During our sample period, interest accrued during the grace period (i.e., between the end of schooling and consolidation), although this practice ended in November 2019.

[5] Prior to 2019, the floating interest rate was the prime rate plus 2.5 percentage points and the fixed rate was prime plus 5 percentage points. Effective November 2019, the floating interest rate was lowered to prime and the fixed rate was lowered to prime plus 2 percentage points. Interest rates have been set to zero during the Covid-19 pandemic.

need.[6] Under RAP, eligible borrowers whose family incomes fall below the minimum family size-based threshold (e.g., $25,000 for singles and $59,500 for families of four) need not make any payments, while those with family incomes above the threshold are expected to contribute an increasing fraction (ranging from 0 to 20 percent) of their incomes above the threshold toward their student loan (unless this income-based amount exceeds the standard debt-based payment set to ensure repayment within 15 years of consolidation, in which case they must only pay this amount). Among the 26% of borrowers on RAP in 2018–2019, the vast majority made no payment at all (Employment and Social Development Canada, 2020).

RAP entails two stages. During Stage 1 (Interest Relief), any reduced payment by the borrower goes toward principal first. The government pays all interest not covered by the borrower's payment, but the government does not pay any amount towards the loan principal. Eligible borrowers can receive this benefit for up to 60 months. In Stage 2 (Debt Reduction), which begins after 60 months in Stage 1 or 10 years after finishing school (whichever comes first), the government pays all interest and principal amounts not covered by the borrower's reduced payment in such a way that student debt is completely eliminated within 15 years of leaving school.[7]

Despite the availability of RAP, many student loans still end up in default (i.e., 9 months or more of consecutive missed payments).[8] Indeed, 3-year default rates have hovered around 10% for the last several borrower cohorts. When a loan enters default, it is sent to the Canada Revenue Agency (CRA) for collection through tax offsets, wage garnishments, judgments against borrowers, and other standard collection efforts.

Borrowers who default or enroll in RAP generate significant losses for the CSLP.[9] In March 2019, the government expected to lose 77% of all loan dollars in default and 25% of all dollars in RAP (Office of the Chief Actuary, 2019).

[6] RAP was introduced in 2009 and replaced two related repayment assistance programs known as Interest Relief (IR) and Debt Reduction in Repayment (DRR). Borrowers in financial difficulty could receive interest payments from the government for up to 54 months on IR; however, IR did not help with loan principal payments. After exhausting IR eligibility, borrowers continuing to face financial difficulties could receive up to 3 separate loan principal reductions (totaling up to $26,000) from DRR. See Appendix 1 of Office of the Chief Actuary (2019) for additional details. Throughout the paper, we do not distinguish between different forms of repayment assistance (i.e., RAP, IR, and DRR), simply referring to all as RAP.

[7] Borrowers with permanent disabilities skip Stage 1 and enter Stage 2 directly. Repayment assistance for these borrowers is such that student debt is eliminated within 10 years of leaving school.

[8] See Lochner et al. (2021) for a detailed analysis of RAP take-up decisions and an analysis of why many borrowers continue to make standard payments or default despite being eligible for RAP.

[9] A very small fraction of borrowers enter bankruptcy during their repayment period. In general, borrowers filing for bankruptcy are not shielded from student loan payments if they file within 7 years of leaving school. However, student loan debt is eligible for discharge, along with other debts, beyond the seven year point.

3 Measuring Returns on Student Loans

There are several ways to measure the profitability of a particular loan or set of loans issued by the government. We focus on two measures of nominal returns that we use throughout our analysis. We loosely refer to our first measure as the *realized return* on a particular loan. To calculate this return, we first compute the value that the lender would have attached to the *ex post* realized sequence of payments had it known that sequence *ex ante*. The realized return is then simply measured as the percent difference between that *ex post* government value and the initial loan amount. It reflects the *ex post* net profit on a loan relative to the amount issued and is analogous to the (net) benefit/cost ratio often used in analyses of early childhood and other educational programs.

The second measure we consider is the *Internal Rate of Return*, or *IRR*, which we compute for sets of loans issued to various subsets of the student population. This measure is commonly used in finance and, increasingly, in studies of educational investments which entail significant initial outlays and a lengthy period of returns. This section formally defines and discusses these measures.

Since we are particularly interested in understanding the extent to which returns can be predicted *ex ante* based on information available to the government—or lenders more generally—at the time loans are issued, this section also describes our approach for estimating predicted returns (based on realized returns). We discuss our data and several practical challenges in computing returns in Section 4.

3.1 Realized Return

Consider a loan issued to student i at date $t = 0$ for amount $L_{i,0}$. Over the life of the loan, the student makes a sequence of payments denoted $P_i = \{P_{i,t}\}_{t=1}^{T}$, where T is the horizon of the loan. In practice, these payments may reflect standard debt-based payments, payments while on RAP, or collections during default. If the government knew *ex ante* this stream of payments, it would value it at present discounted value

$$PDV(P_i, d) = \sum_{t=1}^{T} (1+d)^{-t} P_{i,t},$$

where d is the appropriate (constant) discount rate. One can think of this discount rate as the cost of funds for the government, which we measure by the cost of issuing debt of similar

horizons as the loans.[10] The realized return, then, is simply given by

$$R_i(L_{i,0}, P_i, d) = \frac{PDV(P_i, d) - L_{i,0}}{L_{i,0}}. \tag{1}$$

Loans that earn a net profit will have a positive realized return, while those entailing net losses will have a negative realized return.

For additional perspective on this return measure, note that if the interest rate on the loan is the same as the government's discount rate (i.e., $r = d$) and a student makes all the prescribed payments over the horizon of the loan, then the realized return on that loan is zero. This follows from the fact that standard debt payments are set so that $L_{i,0} = PDV(P_i, r)$.[11] With $d < r$, which is our benchmark, loans that are fully repaid are profitable for the government. In practice, these profits help offset losses from borrowers who fail to repay their loans in full.

When $d \neq r$, the timing of payments matters. For example, the most profitable loan with $d < r$ is one that is fully repaid, with interest, in the final period, yielding a return to the government of $R_{max} = [(1 + r)/(1 + d)]^T - 1 > 0$. In practice, repayment streams are often erratic, with the return bounded between -1 (for loans with no payments made) and R_{max}.

3.2 Internal Rates of Return (IRR)

In multi-period settings, a standard measure of the return on an investment—a student loan in our case—with any given realized stream of payments is given by the Internal Rate of Return, IRR, defined as the discount rate which makes the present discounted value of payments equal to the original loan amount, i.e. $PDV(P, IRR) = L_0$.

For our purposes, the IRR has three undesirable properties as a measure of returns on individual loans. First, it need not be unique when loan repayment streams are erratic. Second, the IRR is not defined when payments are zero. Third, the interpretation of the IRR as a meaningful measure of returns on an investment implicitly assumes that payments are re-invested in the same "project" (i.e., borrower) at rate IRR, an unattractive assumption in the context of student loans. A more appealing assumption is that payments received from

[10] For a comprehensive view of the cost of issuing student loans in the U.S., see Lucas and Phaup (2010) and Lucas and Moore (2010).

[11] With a constant interest rate, students make a constant actuarially fair payment every period given by $P_t = P = \frac{r(1+r)^T}{(1+r)^T - 1} L_0$.

a portfolio of student loans are re-invested in the form of loans to similar students. Such a portfolio would tend to have strictly positive and less variable total repayment streams than individual loans subject to considerable idiosyncratic risk. For these reasons, we use the IRR as a measure of returns on sets of loans issued to subgroups of students who share important characteristics, such as post-secondary field of study or type of institution.

3.3 Average and Predicted Realized Returns

We are particularly interested in cross-subsidization across borrowers and the potential for cream-skimming by private lenders. Both depend critically on differences in *ex ante* expected returns across borrowers. Although we have no way of knowing what the government (or other lenders) actually expected the returns to be on any given loan or set of loans, we can estimate *ex post* expected returns on past loans.[12]

As noted above, we calculate IRR measures that characterize the returns on loans disbursed to subgroups of borrowers. It is also straightforward to estimate expected returns for different subgroups of borrowers based on R_i (see equation (1)), either using the simple average of returns or loan-weighted average returns (i.e., weighting borrowers by their loan amounts, $L_{i,0}$). While the simple average provides a measure of expected returns for borrowers in the specified group, the loan-weighted average measures expected returns on the total loan amount disbursed to borrowers in that group.

We can also estimate the expected returns for borrowers conditional on their observed characteristics X_i, $E(R_i|X_i)$, using a basic multivariate regression approach. Consider a simple model of returns $R_i = X_i'\beta + \varepsilon_i$, where the vector X_i includes observable characteristics such as the amount borrowed, demographic characteristics, field of study or major, and institution of attendance, while ε_i reflects idiosyncratic variation in returns conditional on X_i (with $E(\varepsilon_i|X_i) = 0$). In this case, regressing R_i on X_i yields the estimated coefficient vector $\hat{\beta}$, which can be used to calculate predicted returns $\hat{R}_i \equiv X_i'\hat{\beta}$ for a borrower with characteristics X_i. In the absence of aggregate risk, these predicted returns would provide an estimate of *ex ante* expected returns for lenders with rational expectations, i.e., $E(R_i|X_i)$.

[12] *Ex ante* and *ex post* expected returns may differ due to aggregate shocks that similarly affect many borrowers.

4 Data from the Canada Student Loans Program

In principle, it is straightforward to calculate the returns on loans given data on loan amounts and the full history of repayments. In practice, loan records are imperfect and full repayment histories are rarely available. In this section, we discuss the data and sample used to calculate returns and briefly describe our efforts to address these practical challenges.

4.1 Data Overview

We exploit newly available administrative data from the Education and Labor Market Longitudinal Platform (ELMLP) provided by Statistics Canada. This platform provides longitudinal data from the CSLP beginning with the 2003–2004 academic/loan year.[13] The version we use includes recipient-level longitudinal information for 13 years through the 2015–2016 loan year. These data include information on the loan application process, loan disbursement, and annual repayment amounts for each borrower. The CSLP data also contain information on each borrower's institution of enrolment, field of study, and basic demographics. As such, these data provide a unique opportunity to study the distribution of realized returns conditional on a rich set of observable borrower characteristics, enabling us to estimate expected returns for different individuals or subsets of individuals, as outlined in Section 3.

4.2 Sample Selection

Given our interest in rates of return on student loans at the undergraduate level, we limit the sample to individuals who received student loans for full-time undergraduate studies during the 2005–2006 loan year. Focusing on this "2005 cohort" of borrowers ensures a lengthy period of observed payments.[14] The longest repayment history we observe is 11 years (from 2005–2006 to 2015–2016), covering the full 9.5-year duration of standard debt-based repayment in Canada.

Our sample begins with all undergraduate borrowers attending Canadian public post-secondary institutions who were between the ages of 18 and 30 in 2005.[15] Because some

[13] The CSLP loan year starts August 1st and ends July 31th of the following calendar year. The ELMLP allows for linkage to other administrative data sources, including tax records from T1 Family Files (T1FF) and data from the Post-Secondary Information System (PSIS).

[14] Unfortunately, loan records prior to 2005 contain many inconsistencies (see Section 4.3).

[15] While the CSLP classifies law, dentistry, and medical students as undergraduates, nearly all of these students are older and have already completed a separate undergraduate program. We, therefore, treat

borrowers continue their undergraduate studies well after the 2005–2006 academic year, or go on to other levels of study (e.g., non-degree, Master's, PhD), the final loan consolidation year varies from borrower to borrower. To ensure that we observe repayment histories of at least 8 years after consolidation for each borrower (we impute payments beyond the sample period as discussed below), we impose two additional restrictions on our sample. First, we restrict our sample to upper-year undergraduates (i.e., students in their 3rd or 4th year of studies) in 2005: for reasons discussed below, returns are likely to be lower for 1st and 2nd year students. Second, we restrict our sample to borrowers who consolidate by their 6th year since beginning their undergraduate studies. This includes about two-thirds of all 3rd and 4th year undergraduates who borrowed in 2005.[16] On one hand, this restriction likely yields a negatively selected sample and depressed rates of return, because more students continue on to post-graduate studies (typically associated with higher earnings) than extend their undergraduate studies beyond 6 years. On the other hand, students attending post-graduate studies tend to borrow more and begin their repayment much later (with no interest accruing while in school), which likely lowers rates of return on their loans received in 2005. Appendix Table A1 shows that excluding borrowers who consolidate late has only modest effects on the distribution of borrower characteristics and early repayment behavior. We leave it to future work, when longer repayment histories become available, to examine the full implications of this assumption and broader distributions of returns for all undergraduates.

4.3 Inconsistent Loan Records

We address two main inconsistencies in the CSLP administrative data. First, we find that the accounting identity requiring that the amount of principal repaid plus amount of principal outstanding must equal the original loan consolidation amount is not always respected. In many cases, there is an obvious explanation (confirmed through conversations with CSLP officials), which we address as appropriate.[17] In other cases, there is no easy explanation:

them as graduate students and exclude them from our analysis. Due to extremely small sample sizes and confidentiality concerns, we also exclude students with a field of "trades" and 4th-year students at the University of Ontario Institute of Technology, which was newly founded in 2002.

[16] Specifically, we restrict our sample to borrowers who consolidate their loans by 2008 for 3rd year students and by 2007 for 4th year students. Among borrowers who received an undergraduate loan disbursement in their 3rd year in 2005 and consolidated between 2005 and 2015, 68% consolidated their loan by 2008 (1% in 2005, 14% in 2006, 35% in 2007, and 17% in 2008). Among those who were in their 4th year in 2005, 66% consolidated their loan between 2005 and 2007 (5% in 2005, 46% in 2006, and 16% in 2007).

[17] For example, some borrowers have multiple loan consolidations due to leaving and returning to school. The loan consolidation amount should be cumulative, reflecting the total outstanding balance at the time; however, sometimes the later loan consolidation amount clearly does not include outstanding amounts from previous consolidations. In these cases, a discrepancy arises between the repayment record and the reported

discrepancies can result from coding errors or changes in the reporting system over time and across service providers. Among those discrepancies that are not readily understood, we exclude borrowers for which the absolute discrepancy is at least 5% of the consolidation loan amount.

The second type of inconsistency observed in our data relates to the dynamics of cumulative loan repayment amounts, which should be non-decreasing. We observe a large share of decreasing cases for borrowers in British Columbia in loan-year 2011–12 associated with the integration of BC provincial loans with federal CSLP loans that year. These cases were easily resolved following consultation with CSLP officials. Across other provinces and years, inconsistencies in cumulative repayment amounts are rare, typically small in value, and have little impact on our analysis.

Altogether, we exclude about 4,400 borrowers due to sizeable unexplained discrepancies in initial loan balances and another 40 borrowers due to sizeable discrepancies in cumulative repayments over time. These exclusions have negligible impacts on the distribution of characteristics, choices, and outcomes we study. See Appendix A for further details.

4.4 Incomplete Loan Histories

In practice, we do not observe the entire repayment history for many borrowers. Although we observe payments for 8–11 years since consolidation in our sample, many borrowers have not fully repaid their loans within this time frame.

For borrowers still making regular payments or enrolled in RAP, we impute payments after our final sample period (2015) using information on their last observed loan status, payment amount, and outstanding balance. In particular, we use data from borrowers who consolidate in earlier years (and, therefore, have longer observed repayment histories) to estimate payment amounts in later years conditional on their earlier loan status, payment amount, and outstanding balance. Assuming similar repayment dynamics across borrowers who consolidate in different years, we use these estimates to impute payments beyond the sample period for all borrowers in our sample. See Appendix B for details.

Unfortunately, our data contain no information on collections or loan rehabilitation after someone enters default. We, therefore, acquired additional summary records from the CSLP on the average amounts collected by the Canada Revenue Agency (CRA) and total loan amounts rehabilitated each year after default, both conditional on the amount of loan

last consolidation amount, which we easily correct using repayment records.

entering default and years since consolidation when the default occurred. We also received information on average annual loan amounts paid during rehabilitation conditional on the amount rehabilitated. We use these summary records to impute payments for the first 8 years after default. See Appendix C for details.

4.5 Description of 3rd and 4th Year Undergraduate Borrowers

Our final sample consists of about 39,020 borrowers in 2005 who were enrolled in their 3rd or 4th year of undergraduate studies, with 54% in their 3rd year.[18] While we include borrowers ages 18–30, the average age in our sample in 2005 was 22.5.[19] Consistent with recent post-secondary enrolment patterns, 61% of borrowers in our sample were female. Most borrowers were either dependent (53%) or single independent (42%) students, while much smaller shares were married/common law (4%) or single parents (2%).

Figure 1(a) shows that more than half of all CSLP borrowers (56%) came from Ontario, followed by Alberta (12%) and British Columbia (11%). Figure 1(b) documents the fields of study in which students were enrolled at the time loans were disbursed in 2005. More than half of the borrowers in our sample majored in Arts/Sciences, an extremely broad category that includes Visual and Performing Arts, Physical and Life Sciences, Behavioural and Social Sciences, and Humanities.[20] Meanwhile, 14% of borrowers majored in Administration/Business, 10% each in Engineering/Technology and in Health Sciences, 8% in Community Services/Education, and only about 1% in Agriculture. Figure 1(c) reports the fraction of borrowers attending universities by their Macleans ranking.[21] The vast majority (98%) attended a university ranked among the top 49 Canadian institutions, with nearly 40% attending a top 10 university. Enrolment shares decrease as one progresses to lower-ranked institutions. Finally, Figure 1(d) reports the distribution of total CSLP loan amounts owed at the end of 2005 for our sample. Roughly half of all borrowers owed $10,000–$19,999, while about one-quarter owed less and another one-quarter owed more. Only 2% had accumulated over $30,000 in CSLP student debt. The average total amount owed at the end of 2005 was $14,800, while the average amount disbursed in 2005 was $5,500 (with a standard deviation

[18] As required by Statistics Canada, numbers of observations are rounded to the nearest 10 due to confidentiality concerns.

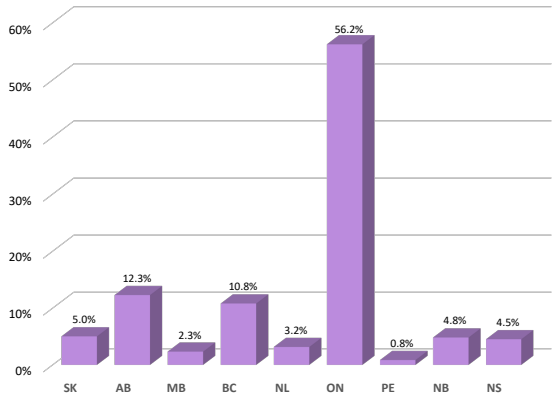
[19] We characterize borrowers by their situation in 2005, since we measure expected returns conditional on information available at the time loans are disbursed—2005 in our case.

[20] See Appendix D for descriptions of each field. Unfortunately, the CSLP data do not contain more detailed field breakdowns. As noted in Section 4.1, CSLP records can be linked to the PSIS, which provides a more detailed classification of fields but does not obtain universal coverage for all provinces until 2009.

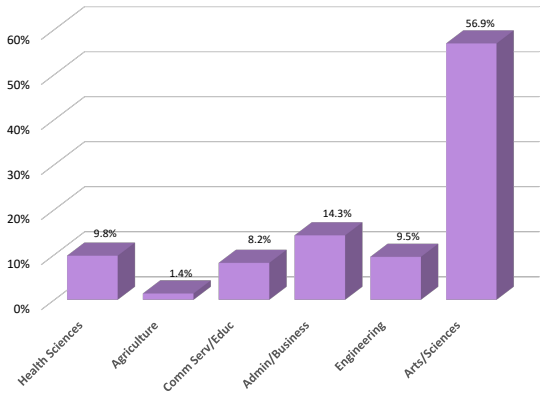
[21] See https://www.macleans.ca/education/canadas-top-school-by-reputation-2020/ for 2020 rankings.

Figure 1: Descriptive Statistics

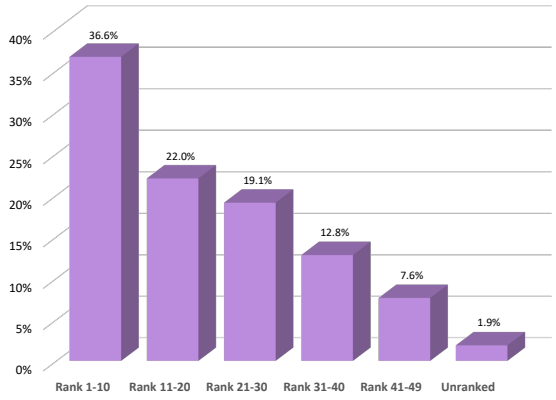
(a) Distribution of borrowers across Provinces



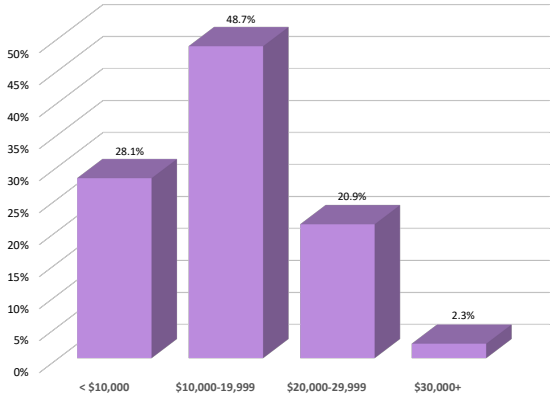
(b) Distribution of fields of study



(c) Distribution of institutional ranking



(d) Distribution of total amounts borrowed



Notes: Descriptive statistics of our baseline sample. Institutional ranks from Macleans ranking.

of $2,100).[22]

4.6 Average Realized Returns and the Overall IRR

We calculate the realized return, R_i, for each borrower as described in equation (1), using the discount rate of $d = 2.8\%$ to discount annual payments to the year when the undergraduate

[22] As required by Statistics Canada, all dollar values are rounded to 100 if greater than 1,000 or 10 if less than 1,000.

loan was disbursed, 2005.[23] This discount rate reflects the 10-year Canadian treasury rate averaged over 2005–2017.[24] For our sample of 3rd and 4th year undergraduates, the average realized return was -5.7%.

It is important to note that this average return does not reflect the return on all loan dollars disbursed to this cohort of students, because it does not differentiate between returns from students who borrowed $1,000 vs. those borrowing $10,000. Yet, this distinction is important for the CSLP in calculating the total return on its loan portfolio. To measure the return to the CSLP on the total amount disbursed, we can weight each borrower's realized return by the loan amount disbursed to that student in 2005. This *loan-weighted average return* is -6.9%, indicating that, on average, the CSLP lost 6.9 cents for every dollar it lent to this cohort. Consistent with this negative return, we obtain an overall IRR of only 1.6% for our sample, more than one percent less than the average 10-year treasury rate we use to calculate realized returns. Any discount rate above 1.6% would imply a negative loan-weighted average realized return.

5 Heterogeneity in Returns Across Borrowers

In this section, we discuss average (unweighted) returns and IRRs for different subgroups of students based on their demographic characteristics, home province, field of study, and post-secondary institution ranking. We then use multivariate regression to estimate the extent to which these factors help predict differences in realized returns across borrowers. These estimates are used to calculate predicted returns based on information observed by the government at the time student loans were disbursed. These predicted returns are critical for evaluating concerns about adverse selection and potential cream-skimming by private lenders as discussed below in Section 7.

[23] While we focus on the returns to loans disbursed in 2005, many borrowers that year also took out loans in other years. Our calculations assume that the returns on all loans (e.g., across years) consolidated by a student are the same. See Appendix A.3 for details.

[24] The 10-year horizon accords with the standard 10-year repayment period for CSLP loans (including the 6 month grace period after completing school). The years 2005–2017 cover this period for borrowers in our sample, who consolidate from 2005 to 2008. We also repeated our analysis with a discount rate of 4.1%, the government's estimated (nominal) cost of borrowing during the 2005–06 loan year as reported by the Office of the Chief Actuary (2006). Using this higher discount rate produces average realized returns that are about 6.5% less than our reported values using $d = 2.8\%$; however, the higher discount rate has little effect on the cross-borrower differences in returns that we emphasize.

Table 1: Average Return, IRR, Early Default, and Early RAP Enrolment by Borrower Characteristics

Population Subgroup	Realized Returns		IRR	Default within First 3 Years	RAP within First 3 Years
	Mean	Std. Error			
All	-5.7%	0.1%	1.6%	6%	30%
Gender					
Female	-5.4%	0.2%	1.7%	5%	30%
Male	-6.2%	0.2%	1.5%	7%	30%
Dependency Category					
Married/Common law	-8.8%	0.7%	1.0%	7%	33%
Single parent	-24.7%	1.4%	-1.0%	16%	53%
Single independent	-4.2%	0.2%	1.8%	6%	29%
Dependent	-6.0%	0.2%	1.6%	5%	30%
Year in Study					
Year 3	-7.9%	0.2%	1.3%	6%	31%
Year 4	-3.0%	0.2%	2.0%	5%	30%
Issue Province					
Alberta	0.2%	0.3%	2.7%	3%	19%
British Columbia	-4.3%	0.3%	1.7%	5%	22%
Manitoba	-3.8%	0.8%	1.8%	8%	20%
New Brunswick	-9.6%	0.7%	1.3%	7%	45%
Newfoundland and Labrador	-7.0%	0.8%	1.7%	9%	40%
Nova Scotia	-11.8%	0.8%	0.9%	12%	42%
Ontario	-7.0%	0.2%	1.3%	6%	32%
Prince Edward Island	-8.4%	1.6%	1.4%	10%	38%
Saskatchewan	1.3%	0.5%	2.9%	4%	22%

Notes: Columns (1) and (2) report the average realized return and its standard error (in %), respectively. Column (3) reports the IRR (in %) on all loans disbursed to the listed subgroup. Columns (4) and (5) report the percent of borrowers who ever defaulted or enrolled in RAP (or IR), respectively, during the first 3 years of repayment.

5.1 Average Returns and IRRs

Table 1 reports average returns (and their standard errors) by demographic characteristics, home province, field of study, and ranking of institution attended. On average, women generate a 0.8% higher return than men, which is somewhat surprising in light of the well-known disadvantage women face in the labour market (Fortin, 2019). Average returns are slightly lower for dependent students relative to single independent students; however, returns are nearly -25% for borrowers who were single parents during their upper years of undergraduate studies.

Borrowers in their 3rd year of university generate a 4.9% lower average return than those in their 4th year. About half of this difference can be explained by the fact that payments

typically begin one year later for 3rd year (relative to 4th year) students, coupled with the lack of interest charges during school.[25] The remaining gap is likely due to lower degree completion rates among 3rd year relative to 4th year students. Prior research shows that dropouts experience more repayment difficulties, likely due to diminished earnings prospects (see, e.g., Gross et al., 2009; Looney and Yannelis, 2015). Both of these factors suggest that returns are likely to be even lower for undergraduates in their 1st or 2nd years of university.

Looking across provinces, we observe substantial differences in returns, ranging from a high of 1.3% in Saskatchewan to a low of -11.8% in Nova Scotia. In general, returns appear to be lower in the Atlantic provinces and higher in Western Canada.

Table 2 shows that student choices about what and where they study also have important implications for student loan repayment. At the high end of the return distribution, Health Science majors generate modest positive returns of 0.8%, while those majoring in Arts/Sciences—more than half of all borrowers—yield a strong negative return (-8.9%) that is nearly 10 percentage points lower. Interestingly, returns are highest (though still negative) for students enrolled in institutions ranked 11–20 rather than those attending top-10 Canadian universities. Ironically, this is not because students attending top schools are more likely to experience repayment difficulties—see the last 2 columns of Table 2, which report the fraction of borrowers who default or enroll in RAP during the first 3 years of repayment.[26] Rather, many of these students repay their loans quickly, which results in fewer borrowers who generate large positive returns (due to $r > d$) to offset losses from those who do not repay in full.[27] We observe the lowest returns (-11.2%) for borrowers who attended unranked institutions.

Tables 1 and 2 also report IRRs computed separately for each subgroup of borrowers. Looking across different borrower types, IRRs are generally high among borrowers with high average realized returns; although, the correlation is not perfect. Across provinces, IRRs range from a low of 0.9% in Nova Scotia to a high of 2.9% in Saskatchewan. Arts/Sciences majors produce an IRR of only 1.0%, while Health Science and Agriculture majors generate an IRR of 2.9%. Consistent with the patterns for average realized returns, borrowers attending schools ranked 11–20 have the highest IRR (2.1%), while those attending unranked

[25] Discounting payments of 3rd year borrowers to their 4th year (rather than 3rd year) implies an adjusted realized return of $(1+d)(1+R)-1$. This adjustment would raise the average return from -7.9% to -5.4% for our sample of 3rd year students.

[26] Our measure of RAP enrolment includes those enrolled in IR prior to the introduction of RAP in 2009. Borrowers from the 2005 cohort would not typically have been eligible for DRR before 2009.

[27] For example, 43% of borrowers in our sample who attended top-10 institutions had fully repaid their loans during the first 3 years of repayment, compared to only 30% of borrowers from schools ranked 11–20.

Table 2: Average Return, IRR, Early Default, and Early RAP Enrolment by Fields of Study and Institution Ranking

Population Subgroup	Realized Returns		IRR	Default within First 3 Years	RAP within First 3 Years
	Mean	Std. Error			
Field of Study					
Administration/Business	-2.4%	0.3%	2.2%	3%	27%
Agriculture	0.4%	0.7%	2.9%	1%	20%
Arts/Sciences	-8.9%	0.2%	1.0%	8%	36%
Community Services/Educ.	-1.6%	0.4%	2.5%	3%	28%
Engineering	-2.7%	0.3%	2.2%	3%	21%
Health Sciences	0.8%	0.3%	2.9%	2%	14%
Institution Ranking					
Rank 1–10	-4.6%	0.2%	1.7%	4%	25%
Rank 11–20	-3.3%	0.2%	2.1%	5%	28%
Rank 21–30	-7.9%	0.3%	1.2%	7%	37%
Rank 31–40	-6.9%	0.4%	1.4%	8%	34%
Rank 41–49	-8.9%	0.5%	1.2%	8%	37%
Unranked Universities	-11.2%	1.2%	0.8%	10%	38%

Notes: Columns (1) and (2) report the average realized return and its standard error (in %), respectively. Column (3) reports the IRR (in %) on all loans disbursed to the listed subgroup. Columns (4) and (5) report the percent of borrowers who ever defaulted or enrolled in RAP (or IR), respectively, during the first 3 years of repayment.

institutions have an IRR of less than 1%.

5.2 Predicted Returns

We next examine the extent to which borrower returns can be predicted conditional on information available to the CSLP at the time loans were issued in 2005. As described in Section 3.3, we regress individual realized returns, R_i, on the total amount borrowed at the end of 2005 (allowing for nonlinearities with a linear spline function), available borrower characteristics (indicators for gender, family/dependency status, age, and year in undergraduate study), home province, field of study indicators, and institution indicators for all ranked universities (with all unranked schools the base/omitted category). Coefficient estimates (with standard errors) for this 'basic' specification are reported in column (5) of Appendix Table E1, which reports results for several other specifications as well. Here, we summarize key results by focusing on the roles of important determinants of estimated returns and the predictability of returns, noting that nearly all of the factors we include in our prediction

model are significantly related to borrower returns.[28]

Figure 2 summarizes the estimated effects of (a) home province, (b) field of study, and (c) institution of attendance on realized returns as estimated by our 'basic' prediction model, comparing these estimated effects with the raw differences in average returns reported in Tables 1 and 2. In panels (a) and (b), we sort estimated effects for provinces (relative to Ontario) and fields (relative to Arts/Sciences), respectively, from highest to lowest raw returns, whereas panel (c) sorts estimated effects by institutions (relative to unranked schools) according to their Maclean's ranking (from top to bottom ranked).[29] Panel (a) shows that differences in student borrowing, characteristics, and schooling choices account for some of the differences in returns across most provinces; yet, important differences remain even when conditioning on these factors. Focusing on the extremes, the raw difference of 13.1% in average returns between Saskatchewan and Nova Scotia drops to 9.1% after accounting for all other factors. Panel (b) shows that differences by field of study are not driven by student borrowing, characteristics, or institution of attendance. Appendix Table E1 further shows that, once one controls for borrower characteristics and province, whether or not one controls for the amount borrowed or institution of attendance has negligible impacts on estimated differences across fields. By contrast, panel (c) reveals that differences in returns across institutions are substantially affected by differences in student borrowing, characteristics, and field of study. Indeed, accounting for these other factors reduces the variance of estimated institution effects by 68%.[30] Panel (c) provides further evidence that institutional ranking provides little information about the returns on student loans, except to the extent that unranked institutions yield significantly lower returns than all but a few ranked schools.

These results indicate that students' field of study is more important than the school they attend in determining the returns on their student loans. Based on our 'basic' prediction model, we find that differences in returns across fields account for 22% of the variance in predicted returns, while differences across institutions account for less than half as much (9%).[31]

[28] It is important to note that, with one exception noted in Section 7.1, our analysis does not depend on any causal interpretation of these prediction estimates. We are only interested in quantifying the extent to which lenders might be able to predict returns and the factors that aid in that prediction.

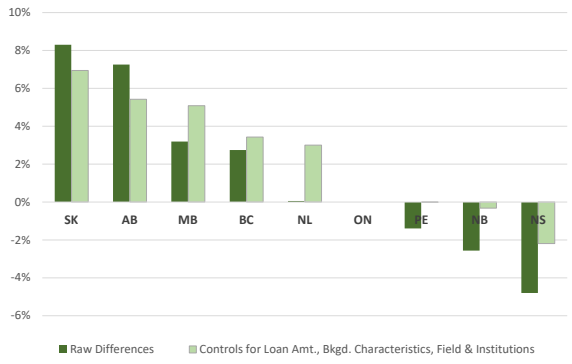
[29] Although we include indicators for all institutions in estimation, Statistics Canada does not allow public release of statistics at the institution level for universities in Manitoba, Prince Edward Island, or Newfoundland and Labrador. Estimates for 5 universities are, therefore, excluded from panel (c) of Figure 2.

[30] This reflects the reduction in the variance of borrower-level predicted returns when controlling only for institution indicators (i.e., raw differences in average returns) and when controlling for other factors (based on coefficient estimates from the specification reported in column (5) of Appendix Table E1).

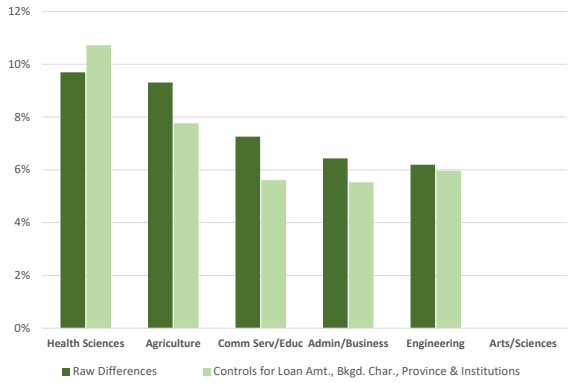
[31] These are based on the variance of predicted returns (from the specification reported in column (5) of Appendix Table E1) when holding all other factors constant except field of study or institution attended relative to total the variance of predicted returns from the model.

Figure 2: Estimated Differences in Expected Returns

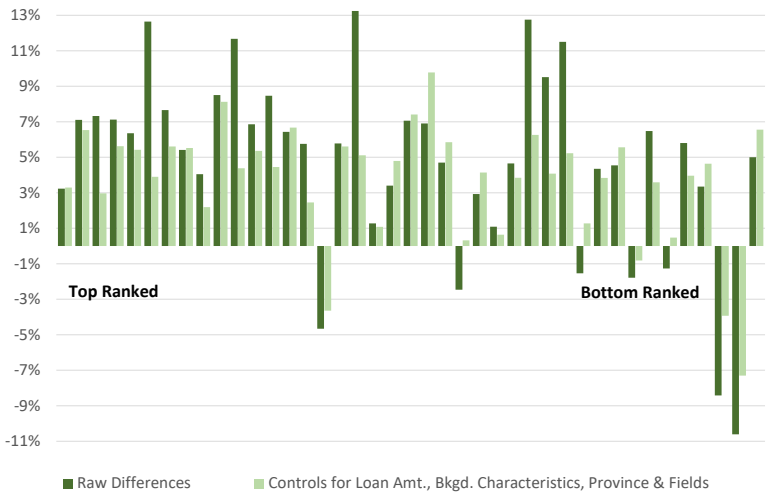
(a) by Province (Relative to ON)



(b) by Field of Study (Relative to Arts/Sciences)



(c) by Institution (Relative to Unranked)



Notes: Panel (a) reports raw differences in expected returns by province (relative to ON) and estimated coefficients on province indicators from the prediction model reported in column (5) of Appendix Table E1. Panel (b) reports raw differences in expected returns by field of study (relative to Arts/Sciences) and estimated coefficients on field indicators from the prediction model reported in column (5) of Appendix Table E1. Panel (c) reports raw differences in expected returns by institution (relative to unranked universities) and estimated coefficients on institution indicators from the prediction model reported in column (5) of Appendix Table E1; estimates for 5 schools in Manitoba, Prince Edward Island, and Newfoundland and Labrador are not shown by agreement with Statistics Canada.

Thus far, our 'basic' prediction model has accounted for differences in the level of returns across borrowers from different backgrounds, provinces, or studying in different fields or at different institutions, but it has assumed that returns for all borrowers are equally sensitive to student debt levels. To relax this assumption, we add to the 'basic' specification a set of interactions between province, field of study, and institution indicators with de-meaned loan amounts by year in undergraduate study. We refer to this extremely rich specification, reported in column (6) of Appendix Table E1, as our 'extended' prediction model, using it throughout the rest of the paper unless otherwise noted. The R^2 statistic for this specification suggests that information about borrowers at the time loans are disbursed explains 12.3% of the variation in their returns: the CSLP knows a lot about its borrowers that can help predict future payments.[32]

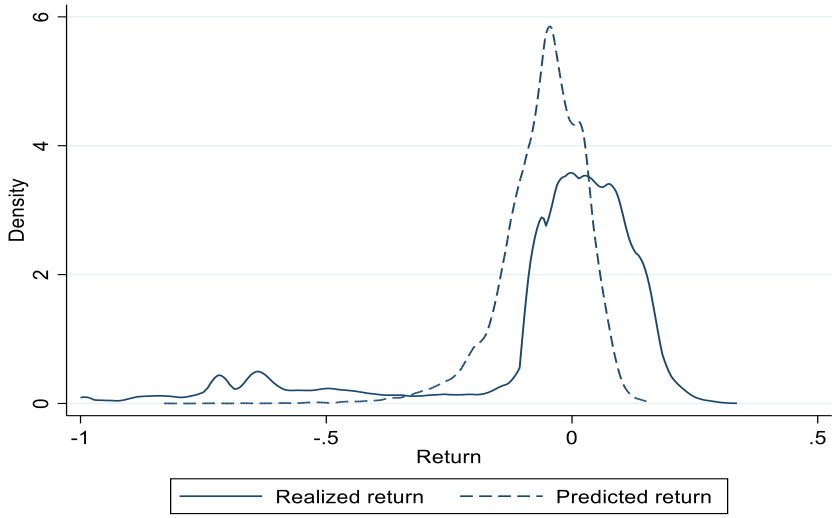
The same predictive information is also readily available to other private lenders, raising potential problems for the CSLP. To see why, note that perfect competition in this loan market would imply no predictable differences across borrowers, since any expected profits would be arbitraged away (e.g., through more attractive loan terms) while expected losses would be truncated (e.g., through higher interest rates or different loan limits). We discuss the potential for private cream-skimming of profitable borrowers below in Section 7.

The distributions of both realized and predicted returns are shown in Figure 3. The realized return distribution exhibits peaks at returns of less than -50% (associated with default) and a large hump over returns ranging from -10% to 20%. As a reference point, a 4th year (3rd year) borrower making all standard payments throughout the repayment period would generate a return of 10.8% (7.8%). The mass over higher returns indicates that a modest fraction of borrowers delays some payments while still repaying in full. The fact that most of the distribution lies below 8% indicates that many borrowers do not fully repay their student loans. The strong negative skewness in realized returns (associated mostly with default and RAP) means that the median realized return is notably higher than the mean return (-5.7%) and positive: 1.1%. Still, 46% of all upper-year undergraduates in our sample generate a negative return on their student loans.

The dashed line in Figure 3 shows the overall dispersion in predicted returns, with a sharp peak just below zero and clear negative skewness. Not surprisingly, we cannot perfectly predict which borrowers will default, so we do not see the same small peaks for predicted returns below -50% that are observed for realized returns; yet, the long left tail of the predicted return distribution reflects the fact that some borrowers predictably make very

[32]By comparison, our 'basic' prediction model explains 10.6% of the variation. More sophisticated prediction models can only improve upon the explanatory power of even our richest specification.

Figure 3: Distributions of Realized and Predicted Returns



Notes: Distributions based on kernel density estimation using the Epanechnikov kernel and optimal bandwidths of 0.012 for realized returns and 0.008 for predicted returns.

low payments due to default or prolonged enrolment in RAP. These low-return borrowers pull down the median predicted return (-4.7%). Moreover, nearly three-in-four upper-year undergraduates are predicted to yield a negative return.

6 Informational Content of Early Default and RAP

Both Canada and the U.S. use default rates over the first 3 years of repayment as performance indicators for government student loan portfolios. While Canada only publishes these default rates (nationally, provincially, and for many institutions), the U.S. goes a step further by sanctioning post-secondary institutions with sufficiently high default rates by withdrawing their eligibility for federal student aid. While not typically used as such, early RAP enrolment may also be a strong indicator of poor returns on a loan, since borrowers enrolled in RAP qualify for reduced payments, interest forgiveness, and eventually forgiveness of loan principal amounts—see Section 2 for details. If borrowers only enroll in RAP for a year or two, forgiven interest payments are unlikely to significantly reduce the returns on the loan; however, borrowers who spend several years on RAP may generate much lower returns. In this section, we explore the extent to which default or RAP enrolment in the first 3 years of repayment indicate poor expected returns, $E(R_i)$, over the life of the loan.

Table 3: Loan Status by Year of Repayment

Years in Repayment	Fraction Fully Repaid	Fraction Entering Default	Fraction Entering RAP
1	15.6%	0.1%	24.8%
2	25.0%	4.0%	3.7%
3	31.1%	1.5%	1.7%
4	37.0%	1.1%	1.0%
5	42.5%	0.7%	0.6%
6	48.5%	0.6%	0.4%
7	54.2%	0.5%	0.3%
8	60.5%	0.4%	0.2%

Notes: The first year in repayment corresponds to the year consolidation took place. Fraction fully repaid is cumulative. Fractions entering default and RAP refer to the first time a borrower enters that status. Some borrowers on RAP may subsequently enter default.

The official 3-year default rate for the cohort of all CSLP loans entering repayment in 2006–07 was 16%, while the default rate among university students was much lower at 10% (Human Resources and Skills Development Canada, 2009). In other words, 10 cents out of every loan dollar university students consolidated in 2006–07 was in default within 3 years of consolidation. Rather than focus on defaulted amounts, we focus on default rates among borrowers. Table 3 shows that in our baseline sample of upper-year undergraduates who borrowed in 2005, the fraction defaulting at any point during their first 3 years of repayment was around 6%, while 30% of all borrowers enrolled in RAP over that same period. The vast majority (82%) of borrowers who enrolled in RAP within 3 years of repayment did so the very first year. By contrast, since a loan is deemed in default when it is in arrears for more than 270 days, very few borrowers enter default during the first year of repayment. Interestingly, the fraction (first) entering default or RAP declines monotonically over time.

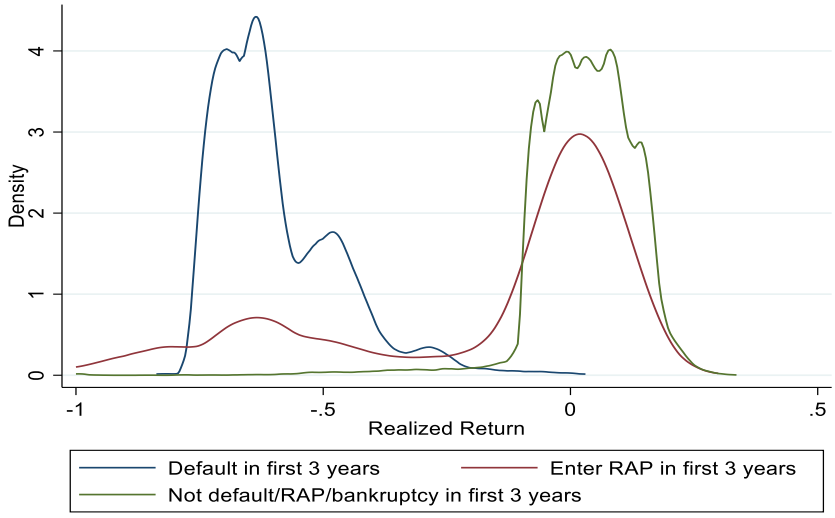
The last two columns of Tables 1 and 2 report early default and RAP enrolment rates for the subgroups of borrowers examined earlier. As with average returns, we see considerable heterogeneity across provinces, fields, and institutions. In most cases, early default and RAP enrolment rates are high for borrowers with low average returns. For example, borrowers from Saskatchewan generate the highest average returns (1.3%), very low default rates (4%), and modest RAP enrolment rates (22%), while borrowers from Nova Scotia yield the lowest average return (-11.8%) and have among the highest default (12%) and RAP enrolment (42%) rates. We observe a similar pattern across fields of study in Table 2, where we see that students majoring in Arts/Sciences are outliers with the lowest average returns (-8.9%)

and highest default (8%) and RAP enrolment (36%) rates. Consistent with their low average returns, students attending unranked universities have high early default and RAP enrolment rates of 10% and 38%, respectively.

Looking more closely at institutions, we calculate average returns, as well as early default and RAP enrolment rates, for each ranked institution in Canada plus all unranked universities combined. We find that, in general, institutions with high default rates also tend to have high RAP enrolment rates: the correlation coefficient is 0.65. More importantly, the correlations between average institutional returns and both early default and RAP enrollment rates are quite strong (-0.81 and -0.84, respectively). Altogether, these results suggest that 3-year default and RAP enrolment outcomes are both good, if imperfect, early indicators of lifetime expected returns on student loans.

Figure 4 shows dramatic differences in the distribution of realized returns for borrowers who default early vs. those who neither default nor enter RAP during the first three years of repayment. Returns are almost always negative for borrowers who default early, while they are mostly positive for those who experience no repayment difficulties early in the repayment process. The distribution of returns for borrowers who enroll in RAP during the first 3 years of repayment looks like a mixture of the return distributions for those who default and those with no early repayment problems. These patterns are well-summarized by the average returns, which are quite negative (-60%) for those who default early, higher but still quite negative (-17%) for those who enroll in RAP early on, and positive (3%) for those who experience no early repayment problems. Although losses of 60% on borrowers with early default are substantial, they also indicate that the government continues to collect modest sums from borrowers even after they default.

Figure 4: Realized return, kernel density



Notes: Distributions based on kernel density estimation using the Epanechnikov kernel and optimal bandwidths of 0.024 for the early default group, 0.043 for the early RAP group, and 0.011 for those with no early repayment difficulties.

7 Cross-Subsidization, Cream-Skimming and Risk-Based Loan Limits

Our results point to considerable heterogeneity in expected returns based on *ex ante* observable characteristics of student borrowers. Indeed, our basic prediction model (see Section 5.2) implies that the expected return on a $15,000 loan issued to a 3rd year student attending the lowest return school in Nova Scotia majoring in Arts/Sciences is -20%, while the same loan issued to a 4th year student attending the highest return school in Saskatchewan majoring in Health Sciences is over 15%. These vast differences in returns indicate a high degree of cross-subsidization within the federal student loan system.

While there is nothing inherently wrong about the idea of some students subsidizing others, there are several reasons one might want to characterize this phenomenon. First, it is important to know whether implicit subsidies to different students are consistent with overall government objectives. Second, there may be more efficient ways to subsidize some students than through reduced debt repayments. Third and perhaps more importantly, cross-subsidization across *ex ante* identifiable groups of borrowers is largely incompatible with competitive private markets (see, e.g., Puelz and Snow, 1994). This raises concerns that private lenders could exploit the cross-subsidization by cream-skimming *ex ante* high-

return borrowers, leaving an uncomfortable choice for the CSLP between subsidizing loans to all students or facing a deteriorating pool of borrowers—options that could threaten the viability of the program itself.

Below, we characterize the extent of cross-subsidization using field of study as an example. We also use our model of predicted returns from Section 5.2 to compute borrowing limits that could ensure non-negative expected returns for each field of study. Finally, we show how the pool of student loans would be affected if the private sector were to cream-skim individual borrowers with particularly high predicted returns.

7.1 Cross-Subsidization and Borrowing Limits across Fields of Study

Table 4 reports average loan-weighted returns across fields of study. As discussed earlier, returns differ widely across majors. Given our coarse measure of fields, the majority of borrowers are in Arts and Sciences who collectively borrowed over $120M in 2005 and inflicted over $12M in losses to the program. The meager revenues from Health Science and Agriculture borrowers are clearly insufficient to cover these losses. Nevertheless, these numbers highlight the fact that students across fields of study do not contribute equally to the CSLP's bottom line.

We further illustrate this point by looking at net revenues/losses, after netting out the overall cost of the program (thereby normalizing average returns for the entire program to zero). This framing is consistent with the notion that the federal government knowingly subsidizes each borrower by about $375 through the CSLP and focuses squarely on the cross-subsidization within the pool of borrowers that arises explicitly from differential returns. As one might expect, these numbers suggest that borrowers in Arts and Sciences are cross-subsidized to various degrees by borrowers enrolled in all other fields: per borrower, each Arts and Sciences student receives about $200 from students in other fields, who themselves give between $200 and $400 each. Cross-subsidization is prevalent even at this broad level of aggregation, underlining the scope for cream-skimming as discussed below.

One instrument that is commonly used to mitigate lending risk consists of a limit on the amount of credit extended to certain borrowers.[33] The last column of Table 4 reports one potential set of loan limits: the maximum loan amount consistent with a non-negative

[33]CIBC, for example, advertises terms (loan limits as well as interest rates) on student loans that are explicitly contingent on field of study, although it emphasizes professional programs such as law, medicine, and MBA's.

Table 4: Cross-Subsidization across Fields of Study

Field of Study	Avg Loan-Weighted Return	Total Amount Issued	Total Revenue or Loss	Net Revenue or Loss	Loan Limit
Health Sciences	0.6%	$23,241,000	$131,535	$1,562,369	$20,300
Agriculture	0.4%	$3,024,000	$12,054	$214,849	$17,500
Community Serv./Educ.	-2.1%	$18,502,000	-$388,407	$809,588	$14,000
Engineering	-3.2%	$20,034,000	-$639,517	$753,762	$12,200
Admin./Business	-3.5%	$28,964,000	-$1,001,963	$1,089,833	$11,300
Arts/Sciences	-10.7%	$119,880,000	-$12,767,544	-$4,430,401	$0

Notes: 'Total Amount Issued' reflects the total value of all loans issued in 2005 to 3rd and 4th year undergraduates in our sample. 'Total Revenue or Loss' reflects the 'Avg Loan-Weighted Return' multiplied by the 'Total Amount Issued' to borrowers in that field. 'Net Revenue or Loss' is a normalized version of 'Total Revenue or Loss' that sums to zero across all fields; it nets out total program losses by effectively subsidizing each borrower the average loss of $375. 'Loan Limit' reflects the loan amount that would yield an average return of zero for borrowers within each field of study given the characteristics of borrowers in that field (based on estimated coefficients from column (5) of Appendix Table E1).

expected return for the average borrower in each field of study. In calculating these limits, we use the 'basic' model of expected returns estimated in Section 5.2 and the fact that predicted returns are monotonically decreasing in student loan amounts.[34] A student with average characteristics borrowing up to these limits would be predicted to net a zero return, whereas the same student borrowing less than the limit would generate a positive return.[35]

The results in Table 4 show that these risk-based loan limits vary considerably across fields of study. Notably, Arts and Sciences majors would be precluded from borrowing altogether under this scheme, whereas students in other fields would face debt limits varying from $11,300 in Administration/Business to $20,300 in Health Sciences.

[34]Our 'basic' prediction model implies that expected returns conditional on choice of major m can be written as $E(R(L)|m) = E(X_B\beta_B + \delta_I|m) + \Gamma(L)$, where X_B reflects observable borrower characteristics (i.e., gender, dependency category, age, year in study, province), δ_I reflects institution effects, and $\Gamma(L)$ reflects the (non-linear) effects of loan amounts. Since estimated returns are monotone in the amount borrowed—$\Gamma(L)$ is a strictly decreasing function—we can use the estimated coefficients from column (5) of Appendix Table E1 to compute the amount of debt that would imply a zero average return on all loans issued to students in each field of study. If some current borrowers continued to borrow less than the calculated limits, the limits would imply positive average returns.

[35]We note two important caveats. First, these loan limits assume a causal relationship from loan amounts to returns; otherwise, reductions in borrowing induced by tighter loan limits need not raise returns. Second, these calculations ignore any behavioral responses in terms of field of study or who borrows.

7.2 Cream-Skimming

Given the prevalence of cross-subsidization within the student loan program, one concern is that private lenders could poach students with particularly attractive expected returns. Table 5 displays back-of-the-envelope calculations for several potential levels of cream-skimming by private lenders. Imagine that private lenders had access to our data and used our most sophisticated model to predict returns based on information available at loan origination. If private lenders managed to successfully attract all student borrowers with an expected return of at least 5% at the time of origination, they would poach around 7% of all upper-year undergraduate borrowers. The pool of borrowers in the CSLP would then consist of the bottom 93% of borrowers ranked by expected returns. Naturally, that would reduce the return on loans issued by the CSLP, exacerbating program losses by nearly 1 percentage point. If private lenders could manage to poach all borrowers who generate a positive return, overall CSLP losses would rise to nearly 10%.[36]

We highlight two important limits of these simple calculations. On one hand, they assume that students would be willing to switch to private lenders at the time of origination, which is far from obvious given interest grace periods and potential future payment reductions through RAP.[37] On the other hand, our calculations also assume that poaching only occurs at origination. In practice, private lenders can go after student borrowers at various points in time as the information set changes: at loan origination; at degree completion; at consolidation; after first employment; and so on.[38] As information flows both to private lenders and student borrowers, a market with even more targeted borrowers could emerge, further deteriorating the quality of the pool of borrowers remaining in the program.

[36] These cream-skimming figures implicitly assume that private lenders can discriminate on all observable factors, including some that are prohibited by law, like age and gender. Because these characteristics are relatively weak predictors of loan returns, eliminating them from our prediction model would lead to only minor adjustments in these calculations.

[37] While nothing prevents private lenders from offering contracts that mimic that provided by the CSLP, their more limited enforcement capabilities likely reduce the profitability of such an effort.

[38] In the U.S., for example, SoFi has been aggressively advertising their student loan refinancing program, which emphasizes various post-graduation outcomes such as job offers, employment, or steady income.

Table 5: Average Return and IRR If Excluding High Expected Return Borrowers

	% of Borrowers	Avg. Loan-Weighted Return	IRR
All borrowers	100%	-6.9%	1.6%
Exclude predicted returns > 10%	99%	-7.0%	1.6%
Exclude predicted returns > 5%	93%	-7.7%	1.4%
Exclude predicted returns > 3%	87%	-8.4%	1.3%
Exclude predicted returns > 1%	78%	-9.4%	1.2%
Exclude predicted returns > 0%	74%	-9.9%	1.1%

Notes: Exclusion of borrowers with high predicted returns is based on estimates from column (6) of Appendix Table E1.

8 Conclusions

Using newly available administrative records from the CSLP, covering roughly 10 years of student loan repayment histories, this paper offers a first look at lifetime returns on student loans as earned by the Canadian government. We draw 3 main conclusions from our analysis: (1) the aggregate return on student loans is quite low; (2) there is an enormous amount of heterogeneity in individual realized returns; and (3) a significant portion of this heterogeneity is readily predictable at the time of loan origination.

The fact that the aggregate return is low should not be too surprising: if the program were highly profitable, the private sector would simply take over. As it stands, losses of about 7 cents for every dollar disbursed and an aggregate internal rate of return of 1.6% for the upper-year undergraduates we study, the program *as a whole* is not particularly attractive to private lenders.

Despite the negative aggregate return, the median borrower generates a slightly positive return as the distribution of returns is negatively skewed due to default and prolonged episodes of repayment assistance. More generally, realized returns differ markedly across all individual and institutional characteristics: returns are higher for females than for males; in Western relative to Atlantic provinces; at highly ranked relative to low/unranked schools, for 4th relative to 3rd year students, and for all majors relative to those in Arts and Sciences. To put these differences in perspective, aggregate losses for the CSLP program (for upper-year undergraduates) amount to a $375 implicit subsidy to each borrower, while differential returns across fields implies a subsidy of about $200 to each Arts and Sciences borrower paid

for by students in other fields at costs of between $200 and $400 each.

This type of cross-subsidization is prevalent within the CSLP: a simple prediction model of student loan returns—based exclusively on information observed by the CSLP at the time of loan disbursement—explains 12% of the variation in individual returns. This *ex ante* heterogeneity is generally inconsistent with a competitive market, which would tend to truncate negative returns and arbitrage away positive returns through more attractive contract terms. And while average returns are negative, the presence of many borrowers with high predicted returns raises concerns about cream-skimming by private lenders. Taking a very simple approach, we show that if private lenders could effectively poach all upper-year borrowers earning returns of 5% or more, the CSLP would see its average return fall by nearly 1 percentage point.

Several challenges remain open for future work. For example, the cream-skimming issue clearly calls for a theory of private lending and dynamic cream-skimming, accounting for degree completion, post-school earnings, and partial repayment histories. The value of the latter as an opportunity for private cream-skimming is highlighted by our results documenting the strong relationship between early repayment outcomes and long-run returns. A still more ambitious agenda would consider the implications of unobserved differences across borrowers and the extent to which this further exacerbates adverse selection problems for government programs and the student loan market more generally. From a purely statistical view, advances in more sophisticated prediction models based on machine learning could point to more severe problems down the road. Finally, while we characterize returns for the 2005 cohort of borrowers—largely due to data availability—it will prove interesting to study variation across cohorts, for its own right but also in relation to aggregate conditions, especially since the Great Recession came and went over the period covered by the data.

References

Bachas, N. (2019). The impact of risk-based pricing in the student loan market: Evidence from borrower repayment decisions. Working Paper.

Employment and Social Development Canada (2020). Canada Student Loans Program: Annual Report 2018–2019.

Fortin, N. M. (2019). Increasing earnings inequality and the gender pay gap in canada: Prospects for convergence. *Canadian Journal of Economics 52*(2), 407–440.

Gross, J. P., O. Cekic, D. Hossler, and N. Hillman (2009). What matters in student loan default: A review of the research literature. *Journal of Student Financial Aid 39*(1), 19–29.

Human Resources and Skills Development Canada (2009). Canada Student Loans Program: Annual Report 2008–2009.

Karamcheva, N., J. Perry, and C. Yannelis (2020). Income-driven repayment plans for student loans. Unpublished manuscript.

Kosir, M., J. Danette, P. Li, and T. Sanford (2015). Assessing the feasibility of creating a self loan refinancing program. Minnesota Office of Higher Education.

Lochner, L. and A. Monge-Naranjo (2015). Default and repayment among baccalaureate degree earners. In B. J. Hershbein and K. Hollenbeck (Eds.), *Student Loans and the Dynamics of Debt*, Chapter 8, pp. 235–284. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Lochner, L. and A. Monge-Naranjo (2016). Student loans and repayment: Theory, evidence, and policy. In E. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education*, Volume 5, pp. 397–478. Elsevier.

Lochner, L., T. Stinebrickner, and U. Suleymanoglu (2021). Parental support, savings, and student loan repayment. *American Economic Journal: Economic Policy 13*(1), 329–71.

Looney, A. and C. Yannelis (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*, 1–89.

Lucas, D. and D. Moore (2010). Guaranteed versus direct lending: The case of student loans. In D. Lucas (Ed.), *Measuring and Managing Federal Financial Risk*, Chapter 7, pp. 163–205. University of Chicago Press.

Lucas, D. and M. Phaup (2010). The cost of risk to the government and its implications for federal budgeting. In D. Lucas (Ed.), *Measuring and Managing Federal Financial Risk*, Chapter 3, pp. 29–54. University of Chicago Press.

Mueller, H. and C. Yannelis (2019). Reducing barriers to enrollment in federal student loan repayment plans: Evidence from the Navient field experiment. Working Paper.

Office of the Chief Actuary (2006). Actuarial Report on the Canada Student Loans Program as at 31 July 2005. https://www.osfi-bsif.gc.ca/Eng/Docs/cslp05.pdf.

Office of the Chief Actuary (2007). Actuarial Report on the Canada Student Loans Program as at 31 July 2006.

Office of the Chief Actuary (2019). Actuarial Report on the Canada Student Loans Program as at 31 July 2018. https://www.osfi-bsif.gc.ca/Eng/Docs/CSLP_2018.pdf.

Puelz, R. and A. Snow (1994). Evidence on adverse selection: Equilibrium signaling and cross-subsidization in the insurance market. *Journal of Political Economy 102*(2), 236–57.

Yannelis, C. (2020). Strategic default on student loans. Unpublished manuscript.

Appendix A Variable Construction and Data Inconsistencies

This appendix describes the construction of our main variables (loan amount, payments, and returns), the data inconsistencies we encounter, and how we address those inconsistencies. To facilitate our discussion, we introduce the following variables and their definitions from the CSLP data codebook:

balstudyend: Balance of student loan at the end of the study period. It does not include any amount repaid before the end of study.

graceint: Interest accrued over the 6-month non-repayment period following study end date.

paidprin: The total amount of principal that has been repaid.

paidint: The total amount of interest that has been repaid.

outprin: The principal amount that is not yet repaid.

A.1 Loan Amount

In the data, borrowers may have multiple loan consolidations due to leaving and returning to school. For each consolidation, the consolidation amount is the sum of *balstudyend* and *graceint*, which should be cumulative and reflect the total outstanding balance at the time.

To construct the loan amount variables, we begin with the last consolidated loan and its repayment. In theory, *outprin* + *paidprin* should equal *balstudyend* + *graceint* if no payments were made prior to consolidation. We can use either for the last consolidation amount, denoted L_l, if this equality holds. In many cases, this equality is not respected. One common case is that for borrowers who have multiple loan consolidations, the case officer may find it convenient to close the earlier loan with a "paid in full" status, despite the fact that the loan is not fully repaid, and combine the unpaid portion into the later loan. When the later loan consolidation amount does not include the outstanding amounts from previous consolidations, a discrepancy arises between the repayment record (*paidprin* + *outprin*) and the reported last consolidation amount (*balstudyend* + *graceint*). In this case, we use *paidprin* + *outprin* for the last consolidation amount L_l.

Another reason *outprin* + *paidprin* sometimes differs from *balstudyend* + *graceint* for the last consolidated loan is that *paidprin* includes payments made prior to the last loan consolidation, while *balstudyend* is net of payments made before consolidation. As a result, *outprin* + *paidprin* is greater than *balstudyend* + *graceint*. In this case, we track all payments made since the year of loan disbursement (2005) and set the last consolidation amount L_l to *balstudyend* + *graceint* for the last consolidated loan.

In other cases with discrepancies, there is no clear explanation. For these unexplained cases, we exclude about 4,400 borrowers from our sample for whom the absolute discrepancy between *paidprin* + *outprin* and *balstudyend* + *graceint* is at least 5% of *balstudyend* + *graceint*; otherwise, we set $L_l = paidprin + outprin$ for the last consolidated loan.

Because borrowers may continue their education beyond the undergraduate level, the final consolidated loan does not always reflect undergraduate borrowing alone. We, therefore, create another loan variable, L_u, which represents the amount of the last consolidated undergraduate loan. For borrowers whose last level of study was undergraduate, we have $L_u = L_l$. For those who attend other levels after their undergraduate studies, we calculate L_u in the same way as L_l based on the last consolidated undergraduate loan (i.e, $L_u = paidprin + outprin$ at the time of last undergraduate consolidation if no payments were made prior to that date); otherwise, we set $L_u = balstudyend + graceint$. The use of L_l and L_u to calculate returns is explained further in Section A.3.

We use two additional loan amount variables in our analysis: the total loan amount as of 2005, which we use in prediction model regressions, and the loan amount disbursed in 2005, which we use as weights in computing average loan-weighted returns. The total loan amount as of 2005 equals the total loan amount borrowed as of the end of undergraduate study less grace period interest and any additional amount disbursed after 2005. The loan amount disbursed in 2005 is directly taken from disbursement records in 2005.

A.2 Payments

To calculate annual payments, P_t, we use *paidprin* and *paidint*: $P_t = P_t^{prin} + P_t^{int} = (paidprin_t - paidprin_{t-1}) + (paidint_t - paidint_{t-1})$, where P_t^{prin} (P_t^{int}) denotes payment towards principal (interest) in year t. We favor using *paidprin* over *outprin* to calculate principal payments, because the change in *outprin* includes not only borrower payments but also amounts paid by the government. When a borrower is in RAP stage 2, for example, *outprin* declines even when the borrower makes no payment; however, *paidprin* does not

change in this case, reflecting the borrower's own payment amount.

Because a separate record is created whenever the loan status changes or at the end of a loan year, there may be multiple values for *paidprin*, *paidint*, and *outprin* in any given loan year. The highest values for *paidprin* and *paidint* for each loan year are used as the end-of-year amounts, while the lowest value of *outprin* for each year is used as the end-of-year amount.

Issues with *paidprin*/*paidint*

Because *paidprin* and *paidint* are both cumulative amounts, they should not decrease over time. However, we observe significant declines in *paidprin* and *paidint* for many borrowers from British Columbia (BC) from the 2010–11 loan year to the 2011–12 loan year. Conversations with the CSLP indicate that these cases are associated with BC becoming an integrated province in 2012, which led to known loan reconciliation issues in their records.

To solve this problem, we use *outprin* when appropriate. For those in standard repayment (not on RAP), we set $P_t^{prin} = outprin_{t-1} - outprin_t$, since all payments should be made by borrowers themselves. If $paidint_t - paidint_{t-1}$ is also problematic, we impute the interest payment P_t^{int} using the fact that payments are made towards interest first, in which case we should generally expect that

$$paidint_{t-1} - paidint_{t-2} = outprin_{t-2} \times r, \tag{2}$$

and

$$paidint_t - paidint_{t-1} = outprin_{t-1} \times r, \tag{3}$$

for any interest rate r. Using equation (2) to substitute in for r in equation (3), we impute interest payments as follows:

$$P_t^{int} = (paidint_{t-1} - paidint_{t-2}) \times \frac{outprin_{t-1}}{outprin_{t-2}}. \tag{4}$$

Fortunately, this imputation provides a very good approximation to interest payments for cases in other years/provinces that were unaffected by the 2012 BC loan integration process.

If the borrower is on RAP (not in Stage 2 or permanent disability, PD), we set $P_t^{prin} = outprin_{t-1} - outprin_t$. Because the government covers interest when the borrower is on RAP, we assume that $P_t^{int} = 0$ if $paidint_t < paidint_{t-1}$, (i.e., $P_t^{int} = \max\{paidint_t - paidint_{t-1}, 0\}$).

If the borrower is on RAP Stage 2 or RAP–PD, when the government covers both interest and principal amounts, we assume $P_t^{prin} = 0$ and $P_t^{int} = \max\{paidint_t - paidint_{t-1}, 0\}$. This reflects the fact that the vast majority of borrowers enrolled in RAP make zero payments.

There are several hundred additional cases (for borrowers not from BC in 2011–2012) for which *paidprin* or *paidint* decreases. In most cases *paidprin* (*paidint*) decreases by less than 5% of the consolidation amount, in which case we adjust it by setting $paidprin_t$ ($paidint_t$) equal to the maximum of its current or any earlier reported values, so it is non-decreasing. This ensures that annual payments are non-negative. Roughly 40 borrowers experience a decrease in *paidprin* that is 5% or more of the consolidation amount. We exclude these borrowers from our sample.

A.3 Returns

Given the loan amounts and payments, we calculate realized returns using equation (1). For those who did not make any payments prior to their last consolidation, the calculation is straightforward. We discount all payments made since the last consolidation to 2005, divide the sum of discounted payments by the last consolidation amount less any grace period interest (i.e., $L_0 = L_l - graceint$) and then subtract 1. We use the same measure of returns for borrowers whose last loan was not for undergraduate studies under the assumption that the returns are identical for all of a borrower's loans. This is equivalent to assigning payments after leaving school proportionally to the amounts owed for each loan.

For borrowers who made payments prior to their last consolidation, if their last study level was undergraduate, we discount all payments beginning in 2005 back to 2005. The loan amount L_0 equals $L_l - graceint$ plus any payments made between 2005 and the last consolidation. If their last study level was not undergraduate, we assume that repayments are proportional to amounts owed from different levels of study. We discount payments made towards the undergraduate loan beginning in 2005 back to 2005. The loan amount L_0 equals the amount of the last consolidated undergraduate loan, L_u, less the associated grace period interest plus payments made between 2005 and last undergraduate loan consolidation.

A.4 Effects of Sample Exclusions

Our analysis is based on 3rd and 4th year undergraduates who borrowed in 2005–2006. To ensure observation of sufficient repayment periods, our sample excludes borrowers who

consolidated more than 6 years after starting their undergraduate studies. We also exclude borrowers whose repayment records show sizeable unexplained discrepancies, as discussed in Sections A.1 and A.2.

We explore the implications of our sample restrictions for early repayment behavior and borrower characteristics in Table A1. Column (1) reports descriptive statistics for full-time undergraduate students who borrowed in the 2005–06 loan year, subject to our age, major, and institution restrictions but not the consolidation year or data consistency. The sample in column (2) excludes those who consolidate after their 6th year since beginning their undergraduate studies. This has little effect on the average 2005 loan amount or early repayment behavior. Changes in the sample composition are generally small except that excluding those who consolidate late produces a sample with 9% fewer Arts/Sciences majors and modest increases in the share of borrowers from other fields. Column (3) is our baseline sample, which excludes borrowers with data inconsistencies (in loan amount or *paidprin*) or missing key variables (payments, loan amount, or return) from those in column (2). Comparing columns (2) and (3) shows that excluding borrowers with inconsistencies has negligible impacts on the composition of our sample.

Table A1: Descriptive Statistics for Different Samples

	All 3rd and 4th Year Borrowers (1)	Excludes Late Consolidators (2)	Baseline Sample (3)
Sample size	66,990	43,620	39,020
Average loan disbursed in 2005	5,500	5,500	5,500
Fraction default in first 3 years	5%	6%	6%
Fraction enter RAP in first 3 years	33%	31%	30%
Fraction paid in full in first 3 years	28%	31%	31%
Average age	22.4	22.6	22.5
Gender			
Female	61%	60%	61%
Male	39%	40%	39%
Dependency Category			
Married/Common law	4%	4%	4%
Single parent	2%	2%	2%
Single independent	40%	44%	42%
Dependent	54%	50%	53%
Issue Province			
Alberta	12%	13%	12%
British Columbia	11%	11%	11%
Manitoba	2%	2%	2%
New Brunswick	5%	5%	5%
Newfoundland and Labrador	3%	3%	3%
Nova Scotia	5%	5%	5%
Ontario	56%	55%	56%
Prince Edward Island	1%	1%	1%
Saskatchewan	5%	5%	5%
Year in Study			
Year 3	54%	54%	54%
Year 4	46%	46%	46%
Field of Study			
Administration/Business	11%	14%	14%
Agriculture	1%	1%	1%
Arts/Sciences	65%	56%	57%
Community Services/Educ.	7%	9%	8%
Engineering	9%	10%	10%
Health Sciences	8%	10%	10%
Institution Ranking			
Rank 1-10	38%	37%	37%
Rank 11-20	21%	22%	22%
Rank 21-30	19%	19%	19%
Rank 31-40	12%	13%	13%
Rank 41-49	7%	8%	8%
Unranked Universities	2%	2%	2%

Notes: Column (1) consists of the full sample of 3rd and 4th year undergraduate borrowers ages 18–30 during the 2005–2006 loan year. Column (2) restricts the sample to borrowers who consolidate within 6 years after starting their undergraduate studies. Column (3), our baseline sample, excludes borrowers from column (2) with significant loan amount or repayment inconsistencies, as well as borrowers missing key variables.

Appendix B Post-Sample Payment Imputations

We do not always observe the entire repayment history for borrowers. Therefore, we use data from borrowers who consolidate earlier and have longer observed repayment periods to impute payments after the observed period for borrowers who consolidate later. This appendix describes our imputation approach.

While our analysis generally uses data from 2005 onward, we incorporate data back to 2003 in order to impute payments after the observed sample period for our sample of 2005 borrowers. Incorporating these earlier years of data allows us to use repayment patterns for borrowers consolidating in 2003 (2004), who we observe for 13 (12) years of repayment, to predict payments in year 13 (12) of repayment conditional on earlier loan status, repayment amount, and outstanding debt.[39]

Let c be the last consolidation year for a borrower. For $c \in \{2003, 2004, ..., 2008\}$, the longest observable payment period is $t' = (2016 - c)$ years. For borrowers who last consolidated in 2003, we observe up to 13 years of repayment, while we only observe up to 8 years of repayment for borrowers who last consolidated in 2008.

For each c in $\{2003, 2004, ..., 2007\}$, we use borrowers who last consolidated in c (i.e., $c_i = c$) to regress their last observed payment, $P_{i,t'}$, conditional on payment $P_{i,t}$, outstanding loan balance $L_{i,t}$, and status $S_{i,t}$, at t years after consolidation for $8 \leq t < t'$. Assuming repayment dynamics are similar across cohorts, estimates from the fitted models are then used to predict payments t' years after consolidation, $\hat{P}_{i,t'}(P_{i,t}, L_{i,t}, S_{i,t})$, for borrowers from later cohorts who are last observed t years after consolidation, i.e., who last consolidated in $2016 - t$ ($c_i = 2016 - t > c$).[40] For example, we use borrowers who last consolidated in 2003 and have $t' = 13$ years of observed payments to regress their payments at $t' = 13$ on $(P_{i,t}, L_{i,t}, S_{i,t})$ at $t = 8$ to predict payments at $t' = 13$ for those who last consolidated in 2008.

Our prediction model used to impute future payments relies on two indicator variables related to loan amounts outstanding. First, we create an indicator variable for whether outstanding debt exceeds \$3,000: $D_{1,i,t} = \mathbb{1}(L_{i,t} > 3000)$. Second, we create an indicator variable for whether continuing to make the same payment P_t for each post-consolidation year $t+1$ to

[39] We address data inconsistencies in these earlier years, which are more common than from 2005 onward, in the same way as with our main sample.

[40] We also explored specifications that conditioned on borrower characteristics (e.g., gender, field of study), but these characteristics had very low (generally insignificant) predictive power after conditioning on $(P_{i,t}, L_{i,t}, S_{i,t})$.

t' would cover the outstanding debt plus interest: $D_{2,i,t,t'} = \mathbb{1}\left(\frac{(1+r)^{t'-t}-1}{r} P_{i,t} > (1+r)^{t'-t} L_{i,t}\right)$, where r is the applicable interest rate on the loan. The first indicator is useful in differentiating repayment behavior for smaller and larger remaining balances, while the second enables us to differentiate between loans that would or would not be paid off by period t' if the borrower continued making the same annual payment amount. The latter is quite important (especially when t' is only a few years later than t), since many borrowers have reasonably stable repayment patterns several years into their repayment period. To see why, note that a borrower making a payment of \$1000 in year t and who only owes \$500 in year $t+1$ is likely to repay $P_{t+1} = \$500$ that year, while that same borrower is likely to make a payment of roughly $P_{t+1} = \$1,000$ if he owes more than that amount.

Recognizing that payments cannot be negative nor exceed the amount owed (with interest), we estimate Tobit regressions with a lower limit of zero and an upper limit of the outstanding balance at post-consolidation year t plus interest, i.e., $(1+r)^{t'-t} L_{i,t}$, where $r = 5.5\%$. We estimate separate regressions for borrowers in repayment at t vs. on RAP at t.

For borrowers in repayment in post-consolidation year t, we regress $P_{i,t'}$ on $L_{i,t}$, $P_{i,t}$, $L_{i,t} \times D_{1,i,t}$, $P_{i,t} \times D_{1,i,t}$, $D_{2,i,t,t'}$, $L_{i,t} \times D_{2,i,t,t'}$, and $P_{i,t} \times D_{2,i,t,t'}$. When creating $D_{2,i,t,t'}$, we use $r = 5.5\%$ for these borrowers.

We have much smaller sample sizes of borrowers on RAP. To avoid overfitting predicted payment amounts, we use simpler specifications for some of these borrowers.[41] For borrowers on RAP in post-consolidation year t, when $c = 2003$ (i.e., $t' = 13$) and $10 \leq t \leq 12$, we regress $P_{i,t'}$ only on $L_{i,t}$ and $P_{i,t}$, since we have fewer than 100 observations. When $c = 2003$ and $t < 10$, we have more observations, so we regress $P_{i,t'}$ on $L_{i,t}$, $P_{i,t}$, $D_{2,i,t,t'}$, $L_{i,t} \times D_{2,i,t,t'}$, and $P_{i,t} \times D_{2,i,t,t'}$. For $c \geq 2004$, sample sizes are larger still, and we estimate the same specifications as for borrowers in repayment.

Table B1 reports estimated coefficients from these Tobit specifications for all (t, t') when using the 2003 and 2005 cohorts, with Panel (A) reporting results for borrowers in repayment in year t and Panel (B) reporting results for borrowers enrolled in RAP in year t.[42] Columns report results for different specifications (separately by year t repayment status) depending on the cohort c and post-consolidation years t and t' for which payments in t' are predicted conditional on payments and loan amounts from t. The results are qualitatively similar across specifications and suggest that higher year t outstanding loan amounts and payments (L_t, P_t) predict higher payments in subsequent years t'. As expected, coefficients on $P_{i,t} \times D_{2,i,t,t'}$ are

[41] Because the government covers unpaid interest when borrowers are on RAP, we use $r = 0$ when creating $D_{2,i,t,t'}$ indicators for borrowers on RAP at t.

[42] Results for other cohorts are similar.

typically significantly negative. For t close to t', we generally observe estimated coefficients on P_t of roughly 1 and coefficients on $P_{i,t} \times D_{2,i,t,t'}$ of roughly -1. This is consistent with predicted future payments $P_{t'}$ being very similar to earlier payments P_t, except when continuing to make the same payment P_t each year would more than fully repay the loan by period t'.

Based on these prediction models, we impute payments up to 13 years after consolidation for every cohort. Since almost no borrowers continue to make payments more than 15 years after consolidation, we use the imputed payment and remaining balance in $t = 13$ to impute payments for 2 more years. First, consider borrowers in repayment when last observed. If continuing annual payment $\hat{P}_{i,13}$ over the next two years would fully cover their debts, we assume that they continue making this payment (or reduced amounts once the debts are covered). If continued payments of $\hat{P}_{i,13}$ would not cover their debts, we assume that $\hat{P}_{i,14} = \hat{P}_{i,15}$ are raised so that the loan is fully repaid. Second, if the borrower is on RAP when last observed, we assume that he will continue to make the same annual payment $\hat{P}_{i,13}$ until his time on RAP ends (after 2 more years) or the loan is fully repaid, whichever occurs sooner.

Table B1: Tobit Regressions for Imputations

Cohort c	2003	2003	2003	2003	2003	2005	2005	2005
t'	13	13	13	13	13	11	11	11
t	8	9	10	11	12	8	9	10
(A) Borrower in repayment at t								
Estimated coefficients								
$L_{i,t}$	0.01	-0.02	0.09	-0.11	-0.04	0.33*	0.50*	0.10
$P_{i,t}$	0.25	0.63*	0.48*	0.96*	0.95*	-0.47*	0.18	1.02*
$L_{i,t} \times D_{1,i,t}$	0.03	0.10	-0.04	0.15	0.04	-0.25*	-0.42*	-0.06
$P_{i,t} \times D_{1,i,t}$	0.05	-0.37	0.08	-0.36	-0.01	0.97*	0.60*	0.07
$D_{2,i,t,t'}$	-1533.88*	-1326.93*	-894.77*	-634.79*	-172.48	-1726.05*	-1657.99*	80.77
$L_{i,t} \times D_{2,i,t,t'}$	0.50*	0.73*	0.62*	1.29*	1.08*	1.09*	1.34*	0.91*
$P_{i,t} \times D_{2,i,t,t'}$	-1.61*	-1.78*	-1.42*	-2.07*	-0.94*	-2.58*	-1.89*	-1.03*
Constant	-86.96	-45.46	-59.96	171.52	129.50	-154.14	-276.50*	-138.03
Pseudo R^2	0.143	0.163	0.12	0.063	0.045	0.091	0.086	0.025
Sample size	3,160	2,470	1,090	520	360	6,270	4,890	2,600
(B) Borrower enrolled in RAP at t								
Estimated coefficients								
$L_{i,t}$	0.06*	0.05*	0.09*	0.05	0.06	0.09	-0.05	0.01
$P_{i,t}$	0.60*	0.60*	0.44	0.37	0.46*	0.52	1.02	1.62*
$L_{i,t} \times D_{1,i,t}$						-0.06	0.08	0.00
$P_{i,t} \times D_{1,i,t}$						0.58	0.43	-0.11
$D_{2,i,t,t'}$	-289.65	-44.04				-766.26	-184.21	-679.80
$L_{i,t} \times D_{2,i,t,t'}$	0.23	0.39				0.90	0.30	0.75
$P_{i,t} \times D_{2,i,t,t'}$	-1.35	-1.62				-2.47*	-1.65*	-1.05
Constant	-388.07*	-424.04	-740.26*	-299.51	-352.94	-881.89*	-966.15*	-1008.47*
Pseudo R^2	0.019	0.015	0.012	0.007	0.014	0.009	0.013	0.015
Sample size	150	130	100	80	70	990	850	700

Note: Table reports coefficient estimates from Tobit specifications predicting $P_{t'}$ conditional on repayment status and reported regressors in year t. Each column in each panel reflects a separate Tobit regression for (c, t, t'). * significant at 5% level.

Appendix C Collections and Loan Rehabilitation after Default

When a loan enters default, it is "Returned to Government" from CSLP's third-party Service Provider, and the Canada Revenue Agency (CRA) takes responsibility for loan collection. When a loan is rehabilitated, loans are sent back to the Service Provider for additional repayment. Unfortunately, loan histories in our data are broken whenever a default occurs, so we cannot generally track collections by the CRA or repayments in rehabilitation. We, therefore, impute payments following any default using summary data on repayment in collections and in rehabilitation provided to us by the CSLP.

These summary data contain the following statistics based on all borrowers who entered default or rehabilitated their loan between April 1, 2010, and March 31, 2011:

$\Phi_{Coll,t}(L_D, T_D)$:	Fraction of student loan debt entering default, L_D, in year T_D collected from CRA in year t since default.
$\Phi_{Rehab,t}(L_D, T_D)$:	Fraction of student loan debt entering default, L_D, in year T_D rehabilitated in year t since default.
$\theta_{Rehab,\tau}(L_R)$	Fraction of the amount rehabilitated, L_R, repaid during year τ since rehabilitation.

These fractions are based on 6 bins for L_D or L_R: \$1–3,000; \$3,001–6,000; \$6,001–10,000; \$10,001–15,000; \$15,001–20,000; and \$20,001 or more. T_d reflects 1, 2, 3, 4, and 5+ years. Payment rates are available annually for 8 years after default or rehabilitation. We impute payments in year $t = 1, ..., 8$ since default, $P_t(L_D, T_D)$, as a function of the loan amount at the time of default, L_D, and year of default (since consolidation), T_D, based on the following:

$$P_t(L_D, T_D) = L_D \left[\Phi_{Coll,t}(L_D, T_D) + \sum_{\tau=1}^{t} \Phi_{Rehab,t-\tau+1}(L_D, T_D)\theta_{Rehab,\tau}(L_D) \right].$$

The first term in brackets reflects the fraction of debt initially defaulted on that is collected from CRA each post-default year t, while the second term reflects the fraction of debt initially defaulted on that is collected through rehabilitation in post-default year t. The latter is calculated as sum of the fraction of debt in default that was rehabilitated each year up to post-default year t multiplied by the fraction of that rehabilitated debt that is repaid in post-default year t.[43]

[43]This approach implicitly assumes that borrowers who rehabilitate their student loans do not exit rehabilitation and that they hold similar loan balances when they enter rehabilitation as they had when they initially defaulted. This and other reasonable assumptions have little impact on post-default payment amounts P_t,

See Tables C1–C3 for $\Phi_{Coll,t}(L_D, T_D)$, $\Phi_{Rehab,t}(L_D, T_D)$, and $\theta_{Rehab,\tau}(L_R)$ as provided by the CSLP.

since loan rehabilitation rates, $\Phi_{Rehab,t}(L_D, T_D)$, are extremely low (less than 3% for nearly all (L_D, T_D) combinations).

Table C1: $\Phi_{Coll,t}(L_D, T_D)$

		L_D					
		\$1–3,000	\$3,001–6,000	\$6,001–10,000	\$10,001–15,000	\$15,001–20,000	\$20,001 or more
$t=1$	$T_D=1$	24.6%	12.5%	8.0%	6.4%	5.5%	3.7%
	$T_D=2$	28.1%	16.1%	9.7%	7.8%	6.6%	5.0%
	$T_D=3$	30.8%	17.9%	11.6%	8.6%	7.4%	5.8%
	$T_D=4$	31.9%	17.0%	12.4%	8.9%	6.5%	4.9%
	$T_D=5+$	27.9%	16.3%	11.0%	7.9%	6.7%	3.0%
$t=2$	$T_D=1$	15.5%	10.3%	7.9%	5.8%	4.8%	3.9%
	$T_D=2$	15.4%	13.0%	9.7%	7.4%	8.0%	4.8%
	$T_D=3$	15.0%	13.3%	11.1%	7.3%	6.3%	5.2%
	$T_D=4$	14.2%	13.0%	11.1%	7.9%	7.4%	4.7%
	$T_D=5+$	11.9%	11.0%	9.4%	7.1%	4.0%	5.0%
$t=3$	$T_D=1$	10.1%	9.5%	7.3%	5.7%	5.1%	3.7%
	$T_D=2$	10.9%	11.2%	8.8%	7.6%	7.5%	4.6%
	$T_D=3$	12.7%	9.4%	10.2%	8.2%	6.4%	4.5%
	$T_D=4$	9.2%	10.4%	9.7%	8.4%	6.9%	5.8%
	$T_D=5+$	9.3%	8.9%	7.8%	6.6%	4.8%	4.2%
$t=4$	$T_D=1$	8.0%	8.0%	7.0%	5.7%	5.1%	3.7%
	$T_D=2$	8.6%	9.5%	7.7%	6.6%	6.9%	4.7%
	$T_D=3$	9.4%	10.2%	9.4%	6.7%	6.9%	4.7%
	$T_D=4$	8.8%	7.8%	8.2%	7.5%	5.2%	5.6%
	$T_D=5+$	6.5%	6.4%	6.3%	6.4%	4.8%	4.1%
$t=5$	$T_D=1$	7.6%	7.1%	6.5%	5.4%	5.1%	3.7%
	$T_D=2$	7.8%	7.3%	7.3%	7.0%	7.0%	4.6%
	$T_D=3$	6.0%	8.1%	7.8%	6.7%	6.2%	5.5%
	$T_D=4$	5.4%	7.1%	7.6%	5.9%	5.5%	5.4%
	$T_D=5+$	5.4%	5.8%	5.4%	6.0%	3.7%	4.1%
$t=6$	$T_D=1$	3.0%	4.2%	4.1%	4.0%	4.5%	3.5%
	$T_D=2$	2.7%	5.0%	5.4%	5.4%	5.6%	4.4%
	$T_D=3$	3.9%	5.5%	4.8%	5.9%	5.0%	4.6%
	$T_D=4$	2.6%	4.8%	5.7%	6.0%	6.4%	4.8%
	$T_D=5+$	3.8%	3.6%	4.1%	4.9%	4.3%	3.8%
$t=7$	$T_D=1$	1.1%	2.4%	2.5%	3.3%	3.1%	3.3%
	$T_D=2$	1.6%	3.0%	3.8%	4.2%	5.2%	4.0%
	$T_D=3$	2.3%	3.0%	3.6%	4.1%	4.9%	3.9%
	$T_D=4$	2.6%	4.0%	3.8%	3.7%	4.5%	4.4%
	$T_D=5+$	2.0%	2.7%	3.2%	2.8%	3.6%	2.7%
$t=8$	$T_D=1$	0.7%	1.4%	2.0%	2.6%	3.3%	3.0%
	$T_D=2$	1.3%	2.1%	2.7%	3.3%	4.2%	3.1%
	$T_D=3$	0.8%	2.0%	2.2%	3.9%	4.5%	3.8%
	$T_D=4$	0.9%	2.3%	3.6%	2.8%	4.8%	3.5%
	$T_D=5+$	1.6%	2.0%	2.2%	3.1%	3.5%	2.5%

Notes: These statistics are based on data provided by CSLP from newly defaulted loans established in Departmental Accounts Receivable System (DARS) from April 1, 2010 to March 31, 2011.

Table C2: $\Phi_{Rehab,t}(L_D, T_D)$

		L_D					
		\$1–3,000	\$3,001–6,000	\$6,001–10,000	\$10,001–15,000	\$15,001–20,000	\$20,001 or more
$t=1$	$T_D=1$	1.5%	1.7%	1.3%	1.7%	2.1%	2.1%
	$T_D=2$	2.0%	2.3%	2.1%	2.5%	2.5%	3.4%
	$T_D=3$	1.2%	1.7%	2.1%	3.3%	1.7%	2.4%
	$T_D=4$	0.6%	1.7%	1.8%	2.1%	2.0%	2.8%
	$T_D=5+$	0.3%	1.2%	1.9%	2.8%	2.8%	0.9%
$t=2$	$T_D=1$	1.9%	2.0%	1.7%	2.6%	2.3%	1.9%
	$T_D=2$	1.7%	1.7%	2.1%	3.1%	2.1%	2.6%
	$T_D=3$	1.2%	1.3%	1.4%	2.2%	2.9%	2.7%
	$T_D=4$	0.1%	1.3%	1.6%	1.9%	2.3%	1.9%
	$T_D=5+$	0.4%	1.1%	0.7%	1.2%	2.0%	0.9%
$t=3$	$T_D=1$	1.0%	1.3%	1.1%	1.1%	1.6%	1.5%
	$T_D=2$	1.1%	1.2%	1.3%	1.4%	1.2%	1.1%
	$T_D=3$	0.5%	0.7%	0.6%	1.2%	1.5%	1.1%
	$T_D=4$	0.3%	0.6%	0.5%	1.5%	1.8%	2.1%
	$T_D=5+$	0.2%	0.1%	0.8%	0.4%	1.1%	2.0%
$t=4$	$T_D=1$	0.5%	1.0%	1.1%	0.6%	1.4%	1.4%
	$T_D=2$	0.3%	0.6%	0.8%	0.7%	1.4%	1.9%
	$T_D=3$	0.7%	0.6%	0.4%	0.4%	1.5%	1.4%
	$T_D=4$	0.5%	0.2%	0.3%	0.2%	0.7%	0.6%
	$T_D=5+$	0.0%	0.3%	0.2%	0.3%	0.5%	0.5%
$t=5$	$T_D=1$	0.3%	0.5%	0.8%	1.0%	0.5%	0.8%
	$T_D=2$	0.2%	0.7%	0.3%	1.3%	0.5%	0.4%
	$T_D=3$	0.2%	0.7%	0.7%	0.4%	0.7%	1.5%
	$T_D=4$	0.1%	0.3%	0.4%	0.4%	0.5%	0.6%
	$T_D=5+$	0.4%	0.0%	0.2%	0.6%	0.6%	0.0%
$t=6$	$T_D=1$	0.3%	0.3%	0.5%	0.6%	0.9%	0.9%
	$T_D=2$	0.3%	0.2%	0.4%	0.8%	1.2%	0.8%
	$T_D=3$	0.0%	0.1%	0.4%	0.6%	0.7%	0.4%
	$T_D=4$	0.1%	0.0%	0.3%	0.5%	0.6%	0.8%
	$T_D=5+$	0.1%	0.0%	0.6%	1.2%	0.7%	1.0%
$t=7$	$T_D=1$	0.1%	0.4%	0.4%	0.8%	0.7%	0.9%
	$T_D=2$	0.1%	0.4%	0.4%	0.7%	1.8%	1.0%
	$T_D=3$	0.1%	0.4%	0.7%	0.5%	1.1%	0.5%
	$T_D=4$	0.5%	0.4%	0.4%	0.7%	0.8%	2.1%
	$T_D=5+$	0.1%	0.1%	0.4%	0.7%	0.9%	0.2%
$t=8$	$T_D=1$	0.4%	0.2%	0.3%	0.7%	0.2%	0.8%
	$T_D=2$	1.0%	0.5%	0.3%	0.3%	0.8%	0.7%
	$T_D=3$	0.0%	0.2%	0.3%	0.5%	0.5%	0.6%
	$T_D=4$	0.0%	0.1%	0.3%	0.3%	0.6%	0.8%
	$T_D=5+$	0.0%	0.3%	0.2%	0.4%	0.9%	0.4%

Notes: These statistics are based on data provided by CSLP from newly defaulted loans established in Departmental Accounts Receivable System (DARS) from April 1, 2010 to March 31, 2011.

Table C3: $\theta_{Rehab,\tau}(L_R)$

	L_R					
	\$1–3,000	\$3,001–6,000	\$6,001–10,000	\$10,001–15,000	\$15,001–20,000	\$20,001 or more
$\tau = 1$	16.6%	8.5%	4.9%	2.9%	6.5%	5.0%
$\tau = 2$	13.3%	8.4%	6.6%	6.0%	3.1%	8.1%
$\tau = 3$	12.9%	7.5%	6.4%	6.4%	5.5%	4.6%
$\tau = 4$	14.2%	9.7%	6.3%	5.8%	4.7%	4.8%
$\tau = 5$	14.1%	7.7%	5.4%	6.9%	6.9%	5.3%
$\tau = 6$	14.3%	8.6%	8.2%	5.8%	7.2%	3.4%
$\tau = 7$	14.9%	10.0%	6.9%	4.7%	5.2%	6.6%
$\tau = 8$	12.0%	4.7%	7.6%	4.8%	7.0%	3.2%

Notes: These statistics are based on data provided by CSLP from loans that were rehabilitated from Departmental Accounts Receivable System (DARS) from April 1, 2010 to March 31, 2011.

Appendix D Details on Field of Study

The following table lists detailed areas of study for our CSLP field of study classification.

Table D1: CSLP Field of Study Classification

Field of Study	Area of Study
Administration/Business	Business Administration; Accounting; Commerce; Communications; Court Reporter; Hotel Administration; Hospitality; Marketing; Media Resources; Municipal Government; Public Relations; Realty Appraisal; Stenography; Programming; Systems Analysis
Agriculture	Agriculture; Veterinary Medicine; Environmental Sciences; Forestry; Horticulture; Dairy Farming; Fish and Wildlife
Arts/Science	Archaeology; Astronomy; Biology; Chemistry; Ecology; Economics; Fine Arts; Geography; Geology; History; Linguistics; Material Science; Mathematics; Philosophy; Physics; Psychology; Sociology; Statistics; Physiology; Library Science; Theatre; Visual Arts; Crafts; Interpreting Journalism
Community Service/Education	Community Services; Counselling; Social Work; Rehabilitation; Visiting Bookmaker; Outdoor Recreation
Engineering/Technology	Engineering; Electronics; Instrumentation; Kinetics; Pilot Training; Architecture; Urban/Regional; Planning; Navigation; Marine Engineering; Waste Technology
Health Sciences	Dietetics; Home Economics; Chiropractice; Chiropody; Podiatry; Registered Nursing; Registered Assistant Nurse; Optometry; Pharmacy; Laboratory Techonology; X-Ray Technology; Audiology; Speech Therapy; Medical Secretary; Food Products; BIO-Medical Sciences; Environmental Food; Medical Radiography

Appendix E Rate of Return Regression Estimates

Table E1: Estimating Predicted Realized Returns

	(1)	(2)	(3)	(4)	(5)	(6)
Loan amt (in $10,000)	-0.008		-0.012*		-0.009	-0.074***
(Loan amt - 10,000) × 𝟙(amt > $10,000)	-0.052***		-0.047***		-0.052***	-0.052***
(Loan amt - 20,000) × 𝟙(amt > $20,000)	-0.038***		-0.042***		-0.040***	-0.054***
(Loan amt - 30,000) × 𝟙(amt > $30,000)	-0.020		-0.014		-0.016	-0.018
Gender						
Female	(base)		(base)	(base)	(base)	(base)
Male	-0.015***		-0.012***	-0.010***	-0.013***	-0.013***
Dependency category						
Married/Common law	-0.035***		-0.039***	-0.038***	-0.040***	-0.038***
Single parent	-0.149***		-0.158***	-0.195***	-0.152***	-0.147***
Single independent	0.003		0.001	0.006	0.001	0.003
Dependent	(base)		(base)	(base)	(base)	(base)
Age						
18	-0.097**		-0.105**	-0.070	-0.095**	-0.125***
19	-0.039***		-0.047***	-0.039***	-0.044***	-0.046***
20	0.000		0.002	0.007*	0.001	0.002
21	(base)		(base)	(base)	(base)	(base)
22	0.006		0.005	-0.001	0.005	0.006
23	-0.002		-0.004	-0.017***	-0.003	-0.001
24	-0.017***		-0.020***	-0.039***	-0.018***	-0.019***
25	-0.005		-0.010	-0.031***	-0.009	-0.012*
26	-0.014*		-0.020**	-0.045***	-0.018**	-0.022***
27	-0.009		-0.015*	-0.043***	-0.013	-0.016*
28	-0.016		-0.025***	-0.056***	-0.023**	-0.028***
29	-0.019*		-0.027**	-0.061***	-0.025**	-0.028**
30	-0.009		-0.020	-0.050***	-0.017	-0.025**
Year in study						
3	(base)		(base)	(base)	(base)	(base)
4	0.065***		0.066***	0.050***	0.065***	0.087***
Issue province						
Alberta	0.055***		0.055***	0.063***	0.054***	0.049***
British Columbia	0.040***		0.027***	0.022***	0.034***	0.029***
Manitoba	0.057***		0.026***	0.037***	0.051***	0.044***
New Brunswick	0.010		-0.026***	-0.046***	-0.003	-0.012
Newfoundland and Labrador	0.054***		-0.006	-0.020***	0.030**	0.023
Nova Scotia	-0.012		-0.036***	-0.059***	-0.022**	-0.021*
Ontario	(base)		(base)	(base)	(base)	(base)
Prince Edward Island	0.013		0.005	-0.032**	0.000	0.000
Saskatchewan	0.077***		0.079***	0.075***	0.069***	0.060***
Issue province × De-meaned loan amount						
Alberta						0.041***
British Columbia						0.014
Manitoba						0.047**
New Brunswick						0.003
Newfoundland and Labrador						0.017
Nova Scotia						-0.005
Ontario						(base)
Prince Edward Island						0.002
Saskatchewan						0.053***
Field of study						
Administration/Business		0.064***	0.055***	0.061***	0.055***	0.056***
Agriculture		0.093***	0.084***	0.086***	0.078***	0.077***
Arts		(base)	(base)	(base)	(base)	(base)
Community Services/Educ.		0.073***	0.055***	0.057***	0.056***	0.058***

Engineering		0.062***	0.062***	0.059***	0.060***	0.061***
Health Sciences		0.097***	0.113***	0.104***	0.107***	0.092***
Field of study × De-meaned loan amount						
Administration/Business						0.038***
Agriculture						0.071***
Arts						(base)
Community Services/Educ.						0.040***
Engineering						0.059***
Health Sciences						0.099***
Constant	-0.100***	-0.089***	-0.066***	-0.101***	-0.107***	-0.011
Additional Controls:						
Institution	Yes	No	No	No	Yes	Yes
Institution × De-meaned loan amount	No	No	No	No	No	Yes
R^2	0.087	0.024	0.099	0.067	0.106	0.123

Notes: All specifications based on regressions of realized returns on reported characteristics using the full sample of 39,020 undergraduate borrowers in 3rd or 4th year of university. *** significant at 1% level; ** significant at 5% level; * significant at 10% level.