Predicting Student Loan Delinquency and Default

Presentation at Canadian Economics Association Annual Conference, Montreal

June 1, 2013

SRDC SRSA

SOCIAL RESEARCH AND DEMONSTRATION CORPORATION

SOCIÉTÉ DE RECHERCHE SOCIALE APPLIQUÉE

Outline

- Introduction: Motivation and Research Questions
- Literature Review
- Methodology
- Selected Descriptive Analysis Results
- Results of the Multivariate Analysis
- Conclusions

Introduction - Motivation (1)

- Results from LAHS-BC suggested that students might not understand student financial aids adequately. Not all factors were taken into consideration when they borrowed.
- Canada Student Loan Program (CSLP) provides a mortgage style student loan (amortized to a period after study).
- The three-year default rate of the 2009-10 consolidation cohort was **9 per cent** among *university* graduates and **17 per cent** among *college* graduates (HRSDC, 2012).
- Serious consequences of a defaulted CSLP loan:
  - The loan is returned to CRA for debt collection.
  - The loan cannot be discharged in most cases.
  - The borrower is disqualified from further CSLP loans.
- Universal access: uniform “risk premium” within a cohort.

Introduction - Motivation (2)

- The Repayment Assistance Plan (RAP) provides relief to eligible borrowers by temporarily reducing required payments to an “affordable” level.
- Risk management relies on the identification of potential defaulters and delivering preventive interventions at the time of consolidation.
- Two main questions related to the efficacy of existing risk management:
  - How well do observable characteristics at the time of consolidation predict default?
  - Does RAP work?

Introduction - Research Questions

- What are the static factors that affect student loan repayment delinquency or default?
- What are the main characteristics of borrowers that predict missing loan repayment(s)?
- What are the main characteristics of borrowers that predict loan default?
- Are there significant provincial and regional differences in the likelihood of delinquency and default?
- Is it possible to create an indicator to assess borrowers’ risk of delinquency and default?

Literature Review - Theoretical

- The main theoretical model: A life-cycle model that students borrow to finance their postsecondary study and expect to repay the student loan in the future.
- Human capital investment is risky (Baum and Schwartz, 2006):
  - Not all graduates will be able to find remunerative jobs.
  - Unanticipated changes may derail a life plan.
  - The borrower may not be able to **maintain a living** with the **scheduled repayment of loans**.
- However, there is no consensus definition of manageable student debt (Hansen, 1991).
- Delinquency and defaults are some forms of unmanageable student debts.

Literature Review – Empirical (1)

- Three categories of determinants (Lochner, Stinebrickner, and Suleymanoglu, 2012):
  - lack of income and financial resources,
  - high student loan debt and other debt, and
  - beliefs about the cost of not repaying.
- There was two published Canadian studies using the CSLP admin data (Kapsalis, 2006; Lochner et al., 2013).
- Most empirical studies were from the United States.
- A few published Canadian studies relied on survey data. The measurements on payment difficulties were not necessary corresponding to actual student loan defaults.

Literature Review – Empirical (2)

- Some determinants found in the literatures:
  - During repayment: Income, unemployment, family characteristics (dependent children, single-parenthood, marriage dissolution), credit score, debt burden.
  - Before repayment: institutional characteristics (university, college, or private institution; subject of study), academic performance (graduation and grade).
  - Others: age (+), gender (not clear), socioeconomic background.
- Commonly used econometric model: logistic regression.

Methodology (1)

- Data: CSLP administrative databases
  - Repayment: designation files from Aug 2009 - July 2012
  - Characteristics before consolidation: designation files, loan disbursement files, Needs Assessment Records (NARS).
- Definitions:
  - 2009-10 Consolidation Cohort: borrowers who consolidated their student loans within the period from Aug 2009 - July 2010.
  - Delinquency: any borrower who did not repay the scheduled repayment completely on time.
  - Default: any borrower who was delinquent for 270 days or more.
  - Measurement: the loan became default between August 2009 and July 2012
  - **This research's 3-year delinquency rate is very different from the HRSDC published point delinquency rate**

Methodology (2)

- Descriptive analysis of the determinants of delinquency and default within three loan years of the 2009/10 loan cohort.
- Base model: Two logistic regressions of delinquency/default.
- Base model includes gender, marital status, age, last student loan application category, disability status, major field of study, type of educational institutions attended, level of the educational program, length of program, province, principal of the federal loans at consolidation, amount of monthly payment for federal loans, number of terms to repay for federal loans, receipts of CSG-LI, CSG-MI, and CAG-LI-FT, and family income at the time of loan application.
- Visible minority and aboriginal status (if reported).

Analysis

- Examining how delinquency / default vary with each determinant.
- Evaluating the model's prediction accuracy: true positive rates versus false positive rates.

- Did RAP reduce default?
  - Extended model #1: Base model with additional independent variables – the indicators of usages of RAP in 2009/10 and 2010/11.
- Would previous late payments predict default?
  - Extended model #2: Logistic regression of entering default during 2011/12 loan year - independent variables include an indicator of in delinquency for 3 or more months.

Selected Descriptive Analysis Results (1)

	Delinquency / Late Payment Rate	Default Rate (Out of All Borrowers)
Overall Average	**62.4**	**13.4**
Last Student Loan Applicant Category		
Married / Common Law	61.5	12
Single Parent	**79.9**	**29.9**
Single Independent	61.6	13.1
Dependent	61.2	12.1
Last Reported Disability		
Without Disability	62.1	13.5
With Disability	**72.7**	**17.6**
Institution Type		
University	56.2	7.4
College	65.4	15.8
Private	**76.1**	**28.2**
Level of Study		
Less than Bachelor's Degree	**69.3**	**20.1**
Bachelor's Degree	56.7	7.7
Postgraduate Degree	50.4	4.8

Selected Descriptive Analysis Results (2)

	Delinquency / Late Payment Rate	Default Rate (Out of All Borrowers)
Principal at Consolidation		
First Quartile	55.4	11.5
Second Quartile	**65.3**	**17.8**
Third Quartile	64.1	14.7
Fourth Quartile	64.7	9.6
Scheduled Repayment		
First Quartile	54.6	10.5
Second Quartile	**66.8**	**17.4**
Third Quartile	64.6	14.9
Between 75th and 90th percentiles	62.4	10.9
Top 10 percentile	64.9	10.8
Number of Months in Loan Term		
Up to 18 months	36.9	3.4
19 to 42 months	60.3	11.8
43 to 66 months	58.3	13.1
67 to 90 months	**62.4**	**16.1**
91 to 114 months	*66*	14.6
More than 114 months	64.4	6.6

Selected Descriptive Analysis Results (3)

	Delinquency / Late Payment Rate	Default Rate (Out of All Borrowers)
Family Income (within Application Category)		
First Quartile	**67.4**	**17.7**
Second Quartile	64.7	15.3
Third Quartile	60.5	11.8
Fourth Quartile	57.6	9.8
Received CSG-LI		
No	62.0	14.8
Yes	**64.0**	**7.1**
Received CSG-MI		
No	62.6	13.9
Yes	**58.3**	**5.6**

Selected Descriptive Analysis Results (4)

	Number of cases	Delinquency / Late Payment Rate	Default Rate (Out of All Borrowers)
Applied for RAP before 2011-12			
No	117,634	48.9	10.5
Yes	**68,849**	**85.4**	**18.4**
Used RAP before 2011-12			
No	137,896	56.0	15.1
Yes	**48,587**	**80.4**	**8.7**

Selected Descriptive Analysis Results (5)

- There were 68,849 borrowers applied for the RAP in the loan years 2009-10 and 2010-11. The data also indicated that 48,587 had officially started the RAP.
- RAP seems to be effective in reducing loan defaults.
- Among those 20,262 who applied but not started the RAP:
  - High default rate;
  - The exact reason is unknown. The application might be in one of the following states :
    - Rejected (not likely)
    - Waiting for approval
    - Returning the loan to good standing (eligibility requirement)
    - Withdrawal

Results of Logistic Regressions (1)

- Patterns observed from the descriptive bivariate statistics mostly carry over to the logistic regressions of the base model.
- Exceptions:
  - Variations by province were mostly not statistical significant in the multivariate model.
  - Those who took a 4 year program had a higher default rate than others in the multivariate model when field of study, type of institution and level of study were controlled for.
  - Delinquency and default increased with the scheduled repayment amount, once the amount of principal at consolidation was controlled for.
  - Default rate increased with the number of terms until 114 months and decreased for longer terms.

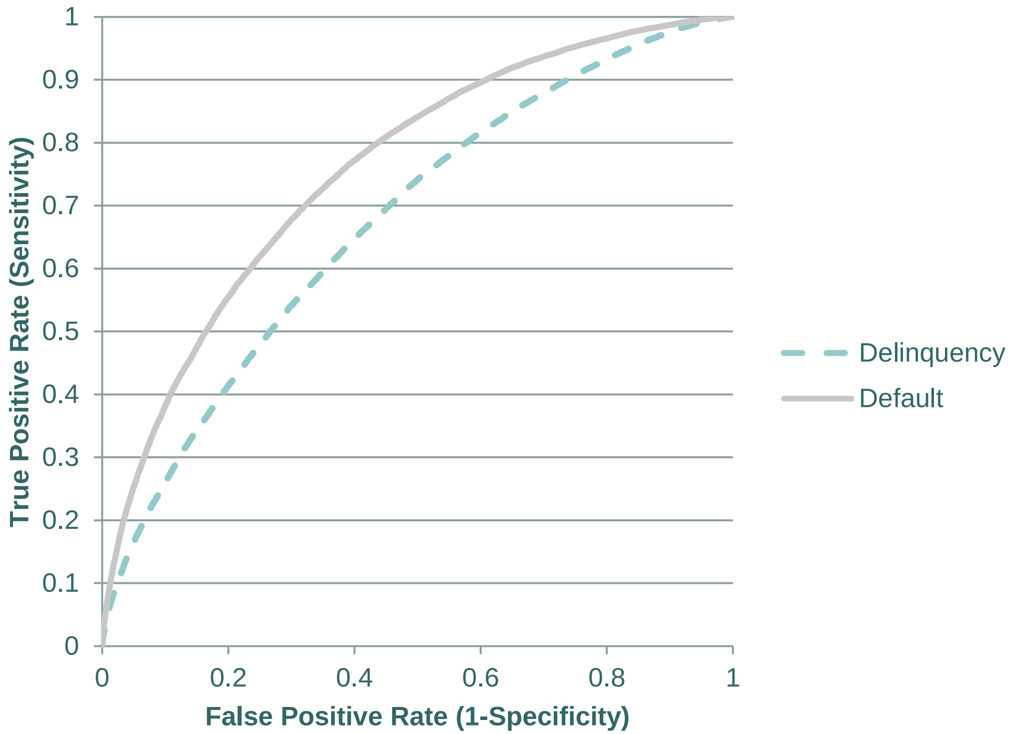
Results of Logistic Regressions (2)

- Family Income
  - Family income at the final student loan application was negatively related to delinquency and default when all other factors were controlled for.
  - Delinquency and default rates of those who received CSG-LI and CSG-MI were negative even in the multivariate model.
- Efficacy of RAP

	Delinquent vs. Not Delinquent		Default vs. Not Default	
	Coefficients	S.E.	Coefficients	S.E.
Previous Usages of RAP				
Used RAP in 2009-10	1.105	(0.016)	-0.873	(0.022)
Used RAP in 2010-11	1.94	(0.035)	-1.069	(0.04)

Prediction Accuracy – Base Model (1)

- Concordant Percentages:
  - 67.2% for Delinquency
  - 75.5% for Default
- Receiver Operating Characteristic (ROC Curve)



SRDC SRSA

Prediction Accuracy – Base Model (2)

- Suppose the logistic regression of default is used to predict the propensity to default. Those with a propensity above a certain threshold will receive a prevention intervention.
- If the threshold is 13.4 per cent (the average default rate):
  - True positive rate: 69.0 per cent. i.e. 69.0 per cent of defaulters (17,255 borrowers) were above the threshold.
  - False positive rate: 31.2 per cent. i.e. 31.2 per cent of non-defaulters (50,339 borrowers) were above the threshold.
  - 3 out of every 4 borrowers who received the intervention are not defaulters.
  - 31.0 per cent of defaulters would not be identified to receive the intervention.

SRDC SRSA

Prediction Accuracy – Base Model (3)

- If the threshold is 75 per cent (to identify high risk group):
  - True positive rate: 0.4 per cent. i.e. 0.4 per cent of defaulters (109 borrowers) were above the threshold.
  - False positive rate: 0.02 per cent. i.e. 0.02 per cent of non-defaulters (34 borrowers) were above the threshold.
  - Only 1 out of every 4 borrowers who received the intervention did not need it.
  - But 99.6 per cent of defaulters would not be identified to receive the intervention.
- The low prediction accuracy of the model makes it difficult to achieve better efficiency while maintaining equity.

Improving Prediction Accuracy

- Payment history is commonly used in credit scoring.
- Exercise: to predict the risk of entering default during 2011-12, with an indicator of in delinquency for 3 or more months (during 2010-11) as an additional predictor.
- Comparison of accuracy by concordant percentages:
  - Without the 3-month delinquency indicator: 67.5%.
  - With the 3-month delinquency indicator: 85.5%.
- A complete hazard model is potentially accurate in predicting default.
- Dynamic model involves regular periodic monitoring of payment history. It requires communication with those who have elevated risk of default during repayment.

Conclusions (1)

- Many static factors and borrower characteristics that affect student loan default found in the literatures are confirmed in this study.
- Two key findings regarding default prevention:
  - First, loans with longer than the standard 114 month term were less likely to be defaulted, likely because of the lower repayment amount per month. Extending repayment periods (to up to 15 years) is the first change suggested to borrowers who find it difficult to repay their loan.
  - Second, the Repayment Assistant Plan was effective in reducing loan default. A substantial number of borrowers who applied for RAP but were not found eligible to start the program were more likely to default on their loans.

Conclusions (2)

- A value function on the false positives and false negatives is needed to determine whether a statistical model to profile defaulters is "better" than other ways of identifying default prevention intervention.
- However, the prediction accuracy using only the information collected at the time of consolidation is mediocre (many false positives or false negatives).
- If preventative measures could be delivered during the repayment period (with regular delinquency monitoring), targeting accuracy could be substantially improved.
- An effective mortgage style student loan program requires:
  - borrowers to have sufficient financial literacy;
  - sophisticated monitoring to prevent costly loan defaults.

SRDC SRSA