Calculating the Contribution of
Demographic Differences to Default Rates

Mark Kantrowitz

Publisher of FinAid.org and FastWeb.com

April 5, 2010. Updated April 29, 2010[1] and May 7, 2010.[2]

EXECUTIVE SUMMARY

Cohort default rates, as published by the US Department of Education and the White House Office of Management and Budget, differ significantly according to type of college and loan program. It is unclear, however, how much of the differences in default rates are due to demographic differences among the colleges and loan programs.

This paper presents a mathematical model for determining the contribution of demographic differences toward overall default rates. It factors overall default rates for a set of borrowers disaggregated by a demographic characteristic into a dot product of prevalence and default rate vectors. An adjusted default rate is calculated by substituting the prevalence vector for a different set of borrowers, showing what the default rate would have been at the first set of borrowers if it had the same demographic distribution as the second set of borrowers. The contribution of the demographic characteristic to the difference in default rates for the two sets of borrowers is calculated by dividing the difference between the overall default rate and the adjusted default rate by the difference between the overall default rates for the two sets of borrowers. This demonstrates how much of the difference in overall default rates is due to differences in the demographic distributions and how much is due to the inherent default rates within each demographic segment.

This model is then applied to analyzing the contribution of demographic characteristics toward differences in the overall default rates for two different sets of borrowers. The first set of borrowers is distinguished according to type of college, namely borrowers at public, non-profit and for-profit colleges. The second set of borrowers is distinguished according to loan program, namely borrowers in the federally guaranteed student loan program and the direct loan program.

Type of College

Risk factors that affect persistence and attainment account for 60.1% of the difference in default rates between for-profit and non-profit colleges and 38.6% of the difference in default rates between for-profit colleges and public colleges. This very significant result is caused by for-profit colleges serving a higher risk mix of students. The risk factors include delayed and part-time enrollment, working full-time while enrolled and single parent status, among other factors.

---

[1] Added a proof that the dot product of the default rate and prevalence vectors equals the overall default rate.
[2] Added clarification concerning the definition of the household size variable.

There are four other demographic factors that also have a significant impact on the differences in default rates. These include Pell Grant recipient status, parent educational attainment, dependency status and household size. (The contributions of two different demographic characteristics to the differences in default rates cannot be added, as the demographic characteristics may overlap significantly in their distribution of students. Calculating the combined impact involves a more complicated process.)

Pell Grant recipient status accounts for 32.9% of the difference in default rates between for-profit and non-profit colleges and 30.7% of the difference in default rates between for-profit and public colleges. Students at for-profit colleges are much more likely to have received a Pell Grant during their first year than students at non-profit and public colleges.

Parent educational attainment has an impact on the difference in default rates, but only with regard to whether the parent received at least a Bachelor's degree. Graduating or dropping out of high school has no impact. Whether the parents received at least a Bachelor's degree accounts for 18.1% of the difference in default rates between for-profit and non-profit colleges and 7.9% of the difference in default rates between for-profit and public colleges.

Dependency status accounts for 17.2% of the difference in default rates between for-profit and non-profit colleges and 14.3% of the difference in default rates between for-profit and public colleges. For-profit colleges enroll a much greater proportion of independent students than non-profit and public colleges. Independent student status also correlates with several of the risk factors that are included in the risk index, since single parents are automatically independent by virtue of having a dependent other than a spouse.

Household size accounts for 20.8% of the difference in default rates between for-profit and non-profit colleges and 16.6% of the difference in default rates between for-profit and public colleges. For-profit colleges enroll a much greater percentage of students from larger families than non-profit and public colleges. Default rates increase monotonically with increasing household size.

The contributions of the five demographic factors to differences in default rates are summarized by the following table.

Demographic Characteristic	Contribution to Differences in Default Rates	
	For-Profit vs. Non-Profit	For-Profit vs. Public
Persistence and Attainment Risk Index	60.1%	38.6%
Pell Grant Recipient Status	32.9%	30.7%
Household Size	20.8%	16.6%
Parent Educational Attainment	18.1%	7.9%
Dependency Status	17.2%	14.3%

Other demographic factors have a moderate impact on differences in default rates. These include adjusted gross income, enrollment status, race, the socioeconomic diversity index and whether the college's location close to home was a factor in the choice of college. The contributions of these demographic factors are summarized by the following table.

Demographic Characteristic	Contribution to Differences in Default Rates	
	For-Profit vs. Non-Profit	For-Profit vs. Public
Adjusted Gross Income	16.2%	7.8%
Socioeconomic Diversity Index	14.5%	10.6%
College Distance	13.7%	10.5%
Race	10.9%	10.6%
Enrollment Status	2.2%	10.8%
College Close to Home	2.1%	9.6%

Other demographic factors have a minimal impact on differences in default rates. These include gender, age when first enrolled, remaining need before loans, major, school/work conflicts, whether the student took remedial courses, whether the student has a goal of being well-off financially, whether the college's cost was a factor in the choice of college, and whether high tuition and fees or a personal or family crisis made it difficult for the student to stay in school. The contributions of these demographic factors are summarized by the following table.

Demographic Characteristic	Contribution to Differences in Default Rates	
	For-Profit vs. Non-Profit	For-Profit vs. Public
Goal: Well-Off Financially	7.1%	4.3%
Major	5.9%	-1.6%
Age at Enrollment	4.8%	3.0%
Gender	2.5%	3.0%
Difficulty Due to Personal/Family Crisis	1.9%	1.1%
Difficulty Due to High Tuition and Fees	1.3%	-0.2%
School/Work Conflicts	0.2%	-0.1%
Remaining Need Before Loans	-0.6%	1.4%
Choice Based on College Cost	-1.4%	-4.0%
Remedial Coursework	-2.0%	-4.8%

Loan Program

The following results demonstrate that most of the differences in default rates between the federally-guaranteed student loan program (FFEL) and the direct loan program (DL) are due to differences in the distribution of dollar loan volume according to institution type and loan type. While such a comparison may have been rendered moot by the passage of the Health Care and Education Reconciliation Act of 2010 (P.L. 111-152), which ends the FFEL program effective July 1, 2010, the method is still useful for comparing the performance of state loan agencies in servicing direct loan program loans. The reconciliation bill included an earmark that provides state loan agencies with a lock on servicing the loans from up to 100,000 borrowers in each state. This prevents an apples-to-apples comparison of state loan agency performance because it blocks the random assignment of borrowers to the state servicers. Without the methodology presented in this paper, it would not be possible to determine how much of the differences in default rates and customer service ratings among the state servicers are due to servicer

performance and how much are due to demographic differences among the states. The technique can be used to normalize the default rates among the servicers to a meaningful common scale by reweighting them according to the overall average distribution of borrowers.

The distribution of colleges according to level and control accounts for 44.7% of the difference in default rates between the FFEL and DL programs for the FY2007 cohort. The distribution of loan programs according to institution type contributed about all of the difference in default rates between the FFEL and DL programs for the FY2006 cohort (97.9%) and the FY2005 cohort (108.0%).

The distribution of dollar loan volume according to type of loan (e.g., subsidized Stafford, unsubsidized Stafford, PLUS and consolidation loan) accounts for 86.9% of the difference in long-term default rate projections between the FFEL and DL programs in FY2009 and 84.0% in FY2008. These results are mostly due to differences in the distribution between new and consolidation loans.

RECOMMENDATIONS

Since a primary purpose of federal student aid is to enable students from at-risk populations to pursue a college education, a college's eligibility for federal student aid should not be penalized because the college enrolls more at risk students.

Current law bases a college's eligibility for federal student aid on its cohort default rate. This is a short-term default rate measure equal to the percentage of all borrowers from a college who enter repayment one fiscal year who default by the end of a subsequent fiscal year.

Instead of basing eligibility on a single default rate, the default rate should be split into two default rates, one for at-risk students and one for low risk students. (An easier to implement and more practical alternative would involve calculating one default rate for students who received a Pell Grant and another for students who did not receive a Pell Grant, and basing institutional eligibility for federal student aid funds on the latter.) Disaggregating the default rates in this manner will allow the evaluation of a college's default rates to be focused more on effectiveness and institutional quality and not as much on the degree to which a college serves students from at-risk populations. Even with such a split, for-profit colleges still have higher default rates than non-profit and public colleges within each demographic category, and thus have much room for improvement. But they should not otherwise be penalized for fulfilling a public policy objective of enrolling and graduating students from at-risk populations, so long as their degrees are high quality with genuine value in the marketplace.

It would also be worthwhile to use a randomized case-controlled study to explore the extent to which the default rates can be reduced by targeting borrowers with two or more risk factors for intensive counseling and other interventions.

DATA SOURCES

This report is based on an analysis using the data analysis system for the 1996/2001 Beginning Postsecondary Students Longitudinal Study (BPS:96/01) conducted by the National Center for Education Statistics at the US Department of Education.[3] This study tracks a cohort of 12,000 students over a six year period from 1995-96 through 2000-01, starting with their initial enrollment in postsecondary education. It helps identify the factors that affect persistence and completion of a postsecondary education.

The core analysis in this report is based on the LOANEVDF variable, which tracks whether the student ever defaulted on a student loan, and the ITNPCT variable, which reports the control of the first institution in which the student enrolled in1995-96 (i.e., public, non-profit or for-profit). These variables are then crossed one-by-one with several variables relating to demographic characteristics, including:

- PRENRL2B, which indicates whether the student has attained or not yet attained a degree
- PELLAMT, which indicates whether the student received a Pell Grant in 1995-96 and the amount of the grant
- SBGENDER, which indicates whether the student is male or female
- SBRACE, which indicates the student's race
- SBDEP1Y1, which indicates the student's dependency status (dependent or independent) in 1995-96
- SBRSK1Y1, which is a risk index that measures risk factors known to affect persistence and attainment
- DISADVAN, which is an index that measures the degree of socioeconomic disadvantage
- SNEED6, which measures the remaining financial need after all student financial aid except the Stafford and PLUS loans[4]
- NETCST3, which measures the net cost of the college as the difference between the cost of attendance and all student financial aid (including loans)
- ICMILES, which measures the college's distance from the student's permanent residence
- RMANYY1, which indicates whether the student took one or more remedial courses in 1995-96
- ITNPLV, which indicates the level of the first institution in which the student enrolled in 1995-96 (4-year, 2-year or less-than-2-year)
- SEMAJ2Y1, which indicates the student's major during their first year
- DFWORKY1, which indicates whether conflicts between work and school made it difficult for the student to stay in school in 1995-96
- DFCRISY1, which indicates whether a personal or family crisis made it difficult for the student to stay in school in 1995-96
- DFTUITY1, which indicates whether the high cost of tuition and fees made it difficult for the student to stay in school in 1995-96

---

[3] While there is a more recent study begun in 2003-04 which tracks students through 2008-09, the data from the 2008-09 follow-up is not yet available through the data analysis system.

[4] The SNEED6 variable is equal to the cost of attendance minus the EFC and minus all aid except the Stafford and PLUS loans. Values less than $100, including negative values, are set to zero. Negative values occur mostly with higher-income students who may receive merit-based aid from the colleges in excess of financial need.

- ICCLOSE, which indicates whether the student chose the college in part based on a location close to home
- ICNUMCOS, which indicates whether the student chose the college in part because of cost
- PBEDHI2, which indicates the highest educational attainment of either parent, focusing on whether the parents did or did not graduate high school or had some college
- PBEDHI3, which indicates the highest education attainment of either parent, focusing on a more detailed differentiation of the degree of college education
- CAGI, which indicates the adjusted gross income of the student's parents (if dependent) and the student and spouse (if independent)
- PGFINC, which reports whether the student had a personal goal of being well-off financially
- SBHHSZB1, which reports the student's household size in 1998, including the student and spouse (if any)[5]
- SBAGFM, which reports the student's age during the first month enrolled in postsecondary education in 1995-96
- ENIPTT2B, which reports the student's enrollment status through 2001 as full-time, part-time or mixed

It was not possible to disaggregate the data by more than one variable at a time because of the small sample size of the BPS:96/01 study. Disjunctive combinations of multiple variables are possible using OR-filters.

Most of the results reported in this paper are restricted to students with debt, not all students, because of the emphasis on default rates.

This paper also uses cohort default rates disaggregated by type of college[6] and loan program[7] as published by the US Department of Education on the Default Prevention and Management web site,[8] as well as data included in the US Department of Education appendixes to the President's FY2010[9] and FY2011[10] budgets.

---

[5] This figure is not the same as household size as reported on the FAFSA. Per the variable description, a value of one was assigned if the student "was not married and had no dependents and either was under age 24 and had living parents or guardians or lived on campus year-round and was enrolled at the time of the interview". Effectively this treated most dependent students as having a household size of one.

[6] www.ed.gov/offices/OSFAP/defaultmanagement/instrates.html

[7] ed.gov/offices/OSFAP/defaultmanagement/3yrffelrates.pdf and ed.gov/offices/OSFAP/defaultmanagement/3yrdlrates.pdf

[8] ed.gov/offices/OSFAP/defaultmanagement/cdr.html

[9] www.whitehouse.gov/omb/budget/fy2010/assets/edu.pdf

[10] www.whitehouse.gov/omb/budget/fy2011/assets/edu.pdf

MATHEMATICAL MODEL OF THE CONTRIBUTIONS OF DEMOGRAPHIC DIFFERENCES TO OVERALL DEFAULT RATES

Assuming that the demographic characteristics do not overlap,[11] the overall default rate[12] is a weighted average of the default rates for each disaggregated characteristic weighted by the prevalence of the characteristic in the student population for the particular set of borrowers.[13] This can be represented as a dot product of a default rate vector with a prevalence vector, each of which is specific to the set of borrowers. In effect this factors each set's default rate into a combination of default rate vectors and prevalence vectors.

Let D be the default rate vector $(d_1, d_2, \ldots)$ and P be the prevalence vector $(p_1, p_2, \ldots)$ where d_i is the default rate for characteristic i and p_i is the prevalence of characteristic i in the student population such that $\sum_{i=1}^{n} p_i = 100\%$. The prevalence vector may also be referred to as the demographic distribution. Then the overall default rate is the dot product,

$$D \cdot P = \sum_{i=1}^{n} d_i \cdot p_i$$

For example,

$$D_{gender} \cdot P_{gender}$$
$$= (d_{male}, d_{female}) \cdot (p_{male}, p_{female})$$
$$= d_{male} \cdot p_{male} + d_{female} \cdot p_{female}.$$

Let D[*set*] and P[*set*] disaggregate the default rate and prevalence vectors for a particular *set* of borrowers, such as borrowers at a particular type of institution, by limiting the student population to just the students from among the particular set of borrowers. Thus we have the default rate and prevalence vectors for the various types of colleges

---

PROOF

The following proof demonstrates that the dot product of the prevalence and default rate vectors is the same as the overall default rate.

Let n_i be the number of students with characteristic i and let N be the total number of students, so that $\sum_{i=1}^{n} n_i = N$.

Let m_i be the number of students with characteristic i that defaulted on their loans and let M be the total number of defaulted students, so that $\sum_{i=1}^{n} m_i = M$.

Then $d_i = \frac{m_i}{n_i}$ and $p_i = \frac{n_i}{N}$.

Then the dot product of the prevalence and default rate vectors is the same as the overall default rate since

$$D \cdot P = \sum_{i=1}^{n} d_i \cdot p_i$$
$$= \sum_{i=1}^{n} \frac{m_i}{n_i} \cdot \frac{n_i}{N}$$
$$= \sum_{i=1}^{n} \frac{m_i}{N}$$
$$= \frac{\sum_{i=1}^{n} m_i}{N}$$
$$= M/N$$

---

[11] A characteristic is non-overlapping if it is categorical, where each student is assigned to one and only one category. For example, each individual student is reported as either male or female, not both male and female. Even continuous variables can be treated as categorical by using non-overlapping ranges.

[12] Some characteristics do not cover the full student population since some students did not respond to every survey question. In some circumstances responses that were specific to only one type of college were excluded to ensure an apples-to-apples comparison of the different college types. Accordingly, the weighted average default rate calculated relative to a specific demographic characteristic may vary slightly from the overall default rate for the student population.

[13] Sets of borrowers discussed in this paper include borrowers at a particular type of college or borrowers from the direct or federally-guaranteed student loan programs.

as D[*for-profit*] and P[*for-profit*]; D[*non-profit*] and P[*non-profit*]; and D[*public*] and P[*public*]. The overall default rate for a particular set of borrowers disaggregated by a particular demographic characteristic is then

$$D[set] \cdot P[set] = \sum_{i=1}^{n} d_i[set] \cdot p_i[set]$$

Let the *adjusted default rate* be the default rate for a particular set of borrowers disaggregated by a specified demographic characteristic, but substituting the prevalence vector for a different set of borrowers for the same demographic characteristic. The adjusted default rate shows what the default rate at the first set of borrowers would have been had it the same demographic distribution as the second set of borrowers. For example, the dot product D[*for-profit*] · P[*public*] is the default rate that for-profit colleges would have if they had the same demographic distribution as public colleges.[14] The difference between this and the actual default rate, the dot product D[*for-profit*] · P[*for-profit*], is the portion of the difference in default rates between for-profit and public colleges that is attributable to the differences in the prevalence vectors.

The degree to which the demographic differences contribute to the differences in default rates between the two sets of borrowers is the ratio of the difference between the original default rate and the adjusted default rate to the difference between the original default rate and the default rate for the other set of borrowers. This shows how much of the gap between the default rates for the two sets of borrowers is filled by the demographic differences. Let A and B be two sets of borrowers. Then the *contribution* of a demographic characteristic to the differences in the default rates for the two sets is

$$\text{Contribution}(A, B) = \frac{D[A] \cdot P[A] - D[A] \cdot P[B]}{D[A] \cdot P[A] - D[B] \cdot P[B]}$$

where the default rate and prevalence vectors are specific to the particular demographic characteristic.

Conversely, the remaining part of the difference in the default rates is attributable to the differences in the default rates within each demographic segment and not to differences in the distribution of the student population among the demographic segments, namely

$$\text{Remainder}(A, B) = 1 - \text{Contribution}(A, B) = 1 - \frac{D[A] \cdot P[A] - D[A] \cdot P[B]}{D[A] \cdot P[A] - D[B] \cdot P[B]} = \frac{D[A] \cdot P[B] - D[B] \cdot P[B]}{D[A] \cdot P[A] - D[B] \cdot P[B]}$$

This *remainder* is due to differences that are external to the distribution of the demographic differences within the sets of borrowers, such as differences in institutional quality for each type of college. Yet combinations of multiple demographic differences could still potentially account for part of the remainder. The prevalence pattern of each demographic characteristic has a different contribution to the differences in default rates. However, since different demographic characteristics may overlap in their contributions to the default rates, one cannot sum the contributions from each characteristic. For example, the dependency status and age demographic characteristics share mutual information since students who

---

[14] A similar approach may be used to normalize the default rates for a set of colleges to permit apples-to-apples comparisons by adjusting the default rate according to an overall average demographic distribution. So long as all the sets are adjusted according to the same distribution, the results will be comparable. Adjusting the default rates according to the overall average distribution yields results that are reflective of performance applied to the student population as a whole.

are age 24 or older are automatically independent.[15] Instead, measuring the combined contribution would require calculating prevalence vectors (as projections of multidimensional hyper-planes) for the cross products of the demographic variables. This is beyond the scope of this paper, as the sample size of the BPS:96/01 study is too small to permit such a fine-grained analysis to satisfy data quality standards. One can chop the data only so fine before it loses adequate statistical significance.

There will also be different contributions depending on which type of college is used as the baseline for comparison. For example, Contribution(*for-profit*, *public*) will differ from Contribution(*for-profit*, *non-profit*) because public and non-profit colleges have different prevalence and default rate vectors as well as different overall default rates.

Although the examples so far have calculated the contribution of demographics to differences in the default rates of one type of college versus another (or one type of loan program versus another), nothing prevents this methodology from being applied to a comparison of an *individual* college with a *type* of college, including the same type of college as the subject college. It's just a matter of calculating the default rate and prevalence vectors for the individual college as opposed to a set of colleges. If one compares a college with the same type of college it will indicate how well the college fits within the type, a measure of the college's *typicality*.

This approach assumes that the default rate within a demographic segmentation is independent of the prevalence of the characteristic within the student population. This is a necessary assumption for the present mathematical analysis, as one cannot factor the default rate into the dot product of default rate and prevalence vectors without it, but it is not an entirely accurate assumption. [16] The relative size of two social groups can affect the performance of the two groups, as has been demonstrated in the theory of cross-cutting social circles.[17] When a low-performing student is surrounded by high-performing peers, performance improves, and when a high-performing student is surrounded by low-performing students, the student coasts and often fails to achieve his or her full potential. Individuals within a smaller social circle have greater exposure to out-group individuals than in-group individuals and so are more likely to adopt the behaviors of the out-group unless an insulating factor isolates the smaller social circle. For example, a recent study found that among students who intend to obtain a Bachelor's degree, those who start at a community college are 14.5% less likely to obtain a Bachelor's degree within nine years.[18] Even so, the lack of true independence most likely causes the demographic-specific default rates of interest to be higher, not lower, so this paper's methodology for calculating the impact of demographics on default rates yields estimates that are probably lower bounds. The degree of potential error is also relatively small. Thus the independence assumption yields a reasonable approximation.

---

[15] 85.7% of independent undergraduate students are age 24 or older and 14.3% are under age 24. 11.3% of undergraduate students under age 24 are independent (6.6% with dependents other than a spouse, 2.4% married, and 2.2% unmarried with no dependents).

[16] To illustrate, the independence assumption would assume that the default rates for female students are the same regardless of whether the women are enrolled in single-sex or coeducational institutions, all else being equal.

[17] See, for example, Peter M. Blau and Joseph E. Schwartz, *Crosscutting Social Circles: Testing A Macrostructural Theory of Intergroup Relations*, Academic Press, 1984.

[18] Bridget Terry Long and Michal Kurlaender, *Do Community Colleges provide a Viable Pathway to a Baccalaureate Degree?*, National Bureau of Economic Research, NBER Working Paper Number 14367, September 2008. www.nber.org/papers/w14367 The study accounted for potential selection bias.

THE IMPACT OF A PERSISTENCE AND ATTAINMENT RISK INDEX

The BPS:96/01 longitudinal study rates each survey respondent according to a risk index as of 1995-96. This risk index counts the number of risk factors from a set of 7 risk factors that are known to affect persistence and attainment:

- Delayed enrollment
- Lack of a high school diploma
- Part-time enrollment
- Financial independence
- Having dependents other than a spouse
- Single parent status
- Working full-time while enrolled (35 or more hours)

The risk index is reported on a scale from 0 (no risk) to 7 (high risk). For the following analysis the four highest risk index values were lumped together to ensure statistical significance at the non-profit colleges, which have very low enrollments among high risk students.

Since failure to graduate correlates strongly with student loan defaults, a higher risk index is predictive of a higher default rate, as is demonstrated by the following chart.

6-Year Default Rate

Risk Index 1995-96	Default Rate
None	5.6%
1	11.4%
2	12.5%
3	16.3%
4, 5, 6, 7	18.6%
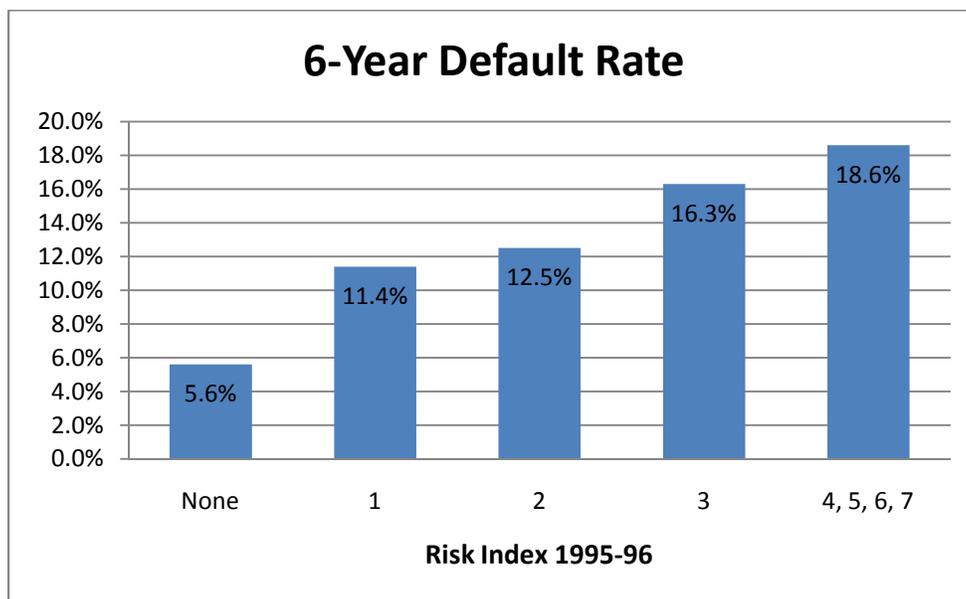
The following table demonstrates that the demographic distribution of students at for-profit colleges is shifted toward higher risk students while the demographic distributions of students at public and non-profit colleges are shifted toward lower risk students. The more selective admissions policies at non-profit colleges contribute to a preference for no-risk students.

BPS:96/01	Risk Index 1995-96				
Prevalence	None	1	2	3	4, 5, 6, 7
Public Colleges	55.8%	19.3%	8.1%	6.6%	10.2%
Non-Profit Colleges	79.6%	12.2%	2.8%	1.9%	3.4%
For-Profit Colleges	16.1%	15.9%	16.5%	18.0%	33.5%

The following table illustrates how the default rates vary based on risk. The overall average is the weighted average of the data reported in the table. This differs slightly from the average default rates of 7.3%, 7.0% and 24.6% for borrowers at public, non-profit and for-profit colleges because not all students provided sufficient data to calculate their risk index.

BPS:96/01	Risk Index 1995-96					Overall
Default Rates	None	1	2	3	4, 5, 6, 7	Average
Public Colleges	5.3%	8.2%	8.3%	10.7%	11.3%	7.1%
Non-Profit Colleges	5.5%	11.0%	9.4%	20.8%	9.2%	6.7%
For-Profit Colleges	10.2%	27.9%	22.3%	24.4%	29.7%	24.1%

The overall default rates are due to a combination of the prevalence and default rate vectors. Using the methodology presented in this paper and the data in these two tables, 60.1% of the difference in default rates between for-profit and non-profit colleges and 38.6% of the difference in default rates between for-profit and public colleges are due to for-profit colleges serving a higher risk mix of students. The rest of the difference in overall default rates is due to for-profit colleges having higher default rates within each risk segment of the population.

While this does not excuse the higher default rates at for-profit colleges, it isolates the portion of the overall default rate that is due to demographic differences from the portion that is due to other factors, such as institutional quality. For example, since the risk index accounts for 60.1% of the difference in default rates, the remainder of the difference in default rates, 39.9%, is not due to demographic differences.

Accounting for the differences due to demographics yields default rates at for-profit colleges that are much closer to the default rates at non-profit and public colleges, albeit still higher. The default rate at for-profit colleges would have been 13.6% instead of 24.1% if their student population had the same risk profile as non-profit colleges. Conversely, the default rate for non-profit colleges would be 11.0% instead of 6.7% if their student population had the same risk profile as for-profit colleges. The default rate at for-profit colleges would be 17.5% instead of 24.1% if their student population had the same risk profile as public colleges. Conversely, the default rate for public colleges would be 9.2% instead of 7.1% if their student population had the same risk profile as for-profit colleges.

THE IMPACT OF INSTITUTION LEVEL

These differences in risk profiles may be due in part to the differences in the prevalence of the different institution levels at each type of college. As the following chart demonstrates, enrollments at non-profit colleges are mostly at 4-year institutions, about one third of enrollments at public colleges are at 4-year institutions, almost two thirds of enrollments at public colleges are at 2-year institutions, almost one third

of enrollments at for-profit colleges are at 2-year institutions and almost two-thirds of enrollments at for-profit colleges are at less-than-2-year institutions.
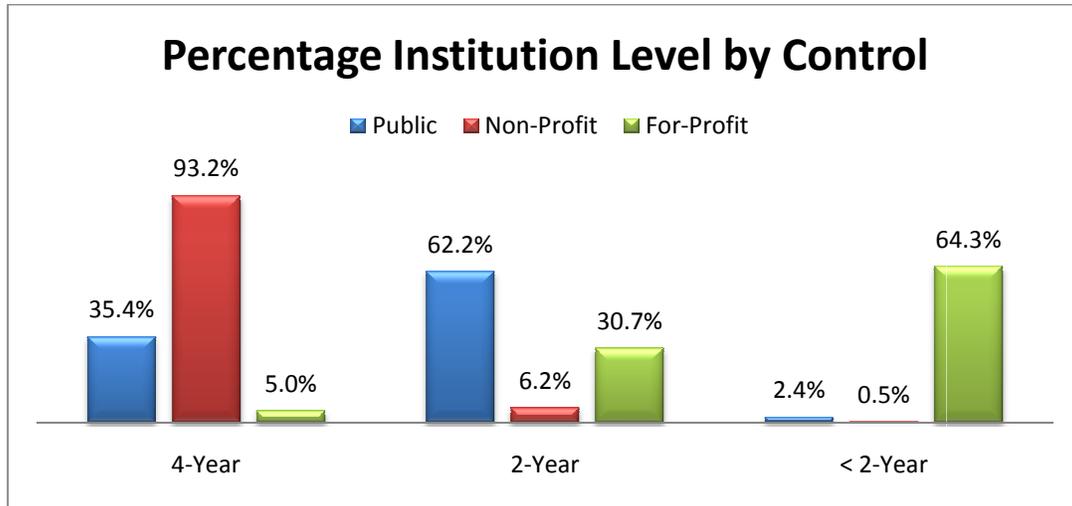
If the risk analysis is restricted to students at two-year colleges,[19] all of the difference in default rates between for-profit and non-profit colleges (105.9%) and about one-third of the difference in default rates between for-profit and public colleges (32.7%) are due to differences in the risk distribution of the student population. Three-fifths of the students at non-profit 2-year colleges have no risk factors and two-fifths have one risk factor, with no students having two or more risk factors. In fact, the default rate at for-profit colleges would be 15.2% instead of 19.9% if their student population had the same risk profile as non-profit colleges. This is less than the actual 15.5% default rate at non-profit 2-year colleges. The distribution of students according to risk at public 2-year colleges is much more similar to that of for-profit 2-year colleges.

Note that these results are due to the stratification of risk within a particular level of institution and not because of the prevalence of the different levels of institution. The following tables contain the prevalence and default rate vectors for the institution level demographic. Based on this data none of the difference in default rates between for-profit and non-profit colleges and only 8.7% of the difference in default rates between for-profit and public colleges are due to the differences in distribution of students among level of institution. It is only when one further stratifies the demographics according to the risk index that differences become apparent.

BPS:96/01 Prevalence	Level 1995-96		
	4-Year	2-Year	< 2-Year
Public Colleges	52.6%	46.5%	0.9%
Non-Profit Colleges	94.2%	5.8%	0.0%
For-Profit Colleges	55.2%	35.8%	9.1%

---

[19] A similar analysis was not performed for less-than-2-year institutions because there was insufficient data at non-profit colleges for statistically significant results. Likewise, a similar analysis was not performed for 4-year institutions because there was insufficient data at for-profit colleges for statistically significant results. The results for 4-year and less-than-2-year institutions would likely be much different than the results for 2-year institutions.

BPS:96/01 Default Rates	Level 1995-96			Overall Average
	4-Year	2-Year	< 2-Year	
Public Colleges	6.2%	8.5%	9.3%	7.3%
Non-Profit Colleges	6.3%	17.3%	NA	6.9%
For-Profit Colleges	24.8%	21.0%	26.8%	24.5%

THE IMPACT OF PELL GRANT RECIPIENT STATUS

The following table shows the demographic distribution of students according to whether or not they received a Pell Grant during their first year in college. For-profit colleges enroll a much greater proportion of students who receive Pell Grants than non-profit and public colleges.[20]

BPS:96/01 Prevalence	Pell Grant Recipient Status	
	No	Yes
Public Colleges	65.4%	34.6%
Non-Profit Colleges	68.4%	31.6%
For-Profit Colleges	33.1%	66.9%

Notice how the percentage of students receiving a Pell Grant during their first year in college at for-profit colleges is about double the percentages at non-profit and public colleges.

The following table shows how the default rates vary based on the student's receipt of a Pell Grant during their first year in college. Pell Grant recipients are more likely to default than non-recipients.

BPS:96/01 Default Rates	Pell Grant Recipient Status		Overall Average
	No	Yes	
Public Colleges	3.8%	14.0%	7.3%
Non-Profit Colleges	4.4%	12.6%	7.0%
For-Profit Colleges	13.6%	30.0%	24.6%

Using the methodology presented in this paper, 32.9% of the difference in default rates between for-profit and non-profit colleges and 30.7% of the difference in default rates between for-profit and public colleges are due to students at for-profit colleges being much more likely to have received a Pell Grant their first year than students at non-profit and public colleges. These figures represent significant contributions toward the differences in default rates.

THE IMPACT OF PARENT EDUCATIONAL ATTAINMENT

Analysis of two different variables that measure the highest educational attainment of either parent yields some interesting results concerning the relative importance of the parents completing high school versus

---

[20] The higher proportion of Pell Grant recipients at for-profit colleges may be due in part to 95.4% of students at for-profit colleges submitting the FAFSA, compared with 71.7% of students at non-profit colleges and 52.1% of students at public colleges. It may also be due to non-profit colleges admitting a wealthier mix of students who do not demonstrate financial need. See Mark Kantrowitz, *FAFSA Completion Rates by Level and Control of Institution*, October 14, 2009, which is available online at www.finaid.org/educators/20091014fafsacompletion.pdf.

the parents receiving a Bachelor's degree in college. The first variable includes a more detailed differentiation of the parent's educational attainment in high school but lumps educational attainment in college into a single category. The second variable lumps educational attainment in high school into a single category but uses a more fine-grained differentiation of educational attainment in college.

The following table shows the demographic distribution of students according to the highest educational attainment of either parent, with a more detailed differentiation of the parents' educational attainment at the high school level or below.

BPS:96/01 Prevalence	Parents Highest Education		
	High School Dropout	Completed High School	Some College or More
Public Colleges	3.8%	30.4%	65.8%
Non-Profit Colleges	2.4%	25.3%	72.3%
For-Profit Colleges	12.9%	49.6%	37.5%

The following table shows how the default rates vary based on the highest educational attainment of either parent, with a more detailed differentiation of the parents' educational attainment at the high school level or below.

BPS:96/01 Default Rates	Parents Highest Education			Overall Average
	High School Dropout	Completed High School	Some College or More	
Public Colleges	15.6%	7.8%	6.3%	7.1%
Non-Profit Colleges	17.0%	8.7%	5.9%	6.9%
For-Profit Colleges	20.0%	23.8%	22.6%	22.9%

Using the methodology presented in this paper, 0.1% of the difference in default rates between for-profit and non-profit colleges and 0.0% of the difference in default rates between for-profit and public colleges are due to for-profit colleges enrolling a greater proportion of students whose parents did not pursue a college education. Differences in whether the student's parents dropped out of high school, graduated high school or attended college have no impact on the differences in default rates according to type of college.

The following table shows the demographic distribution of students according to the highest educational attainment of either parent, with a more detailed differentiation of the parents' educational attainment at the college level.

BPS:96/01 Prevalence	Parents Highest Education		
	High School Diploma or Less	Some College	Bachelor's Degree or Higher
Public Colleges	37.4%	24.2%	38.4%
Non-Profit Colleges	29.8%	17.8%	52.4%
For-Profit Colleges	68.0%	17.8%	14.1%

Notice how not only are students at for-profit colleges less like to have parents with college degrees, but when their parents do have some college experience, they are much less likely to have attained at least a

Bachelor's degree as compared with students at non-profit and public colleges whose parents have some college experience.

The following table shows how the default rates vary based on the highest educational attainment of either parent, with a more detailed differentiation of the parents' educational attainment at the college level.

BPS:96/01 Default Rates	Parents Highest Education			Overall Average
	High School Diploma or Less	Some College	Bachelor's Degree or Higher	
Public Colleges	9.3%	7.4%	5.9%	7.5%
Non-Profit Colleges	9.1%	12.2%	4.1%	7.0%
For-Profit Colleges	23.9%	34.5%	15.6%	24.6%

Notice how the default rates are much higher for students whose parents had some college but not a Bachelor's degree than for students whose parents obtained at least a Bachelor's degree.

Using the methodology presented in this paper, 18.1% of the difference in default rates between for-profit and non-profit colleges and 7.9% of the difference in default rates between for-profit and public colleges are due to differences in the demographic distribution of students according to their parents' educational attainment in college.

These results demonstrate that Bachelor's degree attainment by a student's parents, not their graduation from high school, contributes to the difference in default rates at for-profit colleges versus non-profit and public colleges.

THE IMPACT OF DEPENDENCY STATUS

The following table shows the demographic distribution of students according to dependency status. For-profit colleges enroll a much greater proportion of independent students than non-profit and public colleges.

BPS:96/01 Prevalence	Dependency Status	
	Dependent	Independent
Public Colleges	82.2%	17.8%
Non-Profit Colleges	92.6%	7.4%
For-Profit Colleges	38.9%	61.1%

The following table shows how the default rates vary based on the student's dependency status.

| BPS:96/01 Default Rates	Dependency Status		Overall Average
	Dependent	Independent	
Public Colleges	6.2%	11.3%	7.1%
Non-Profit Colleges	6.2%	12.6%	6.7%
For-Profit Colleges	20.7%	26.3%	24.1%

Using the methodology presented in this paper, 17.2% of the difference in default rates between for-profit

and non-profit colleges and 14.3% of the difference in default rates between for-profit and public colleges are due to for-profit colleges enrolling a greater proportion of independent students than non-profit and public colleges. These figures represent significant contributions toward the differences in default rates.

If the data is restricted to students at 2-year institutions, 25.3% of the difference in default rates between for-profit and non-profit colleges and 12.2% of the difference in default rates between for-profit and public colleges are due to 2-year for-profit colleges enrolling a greater proportion of independent students than 2-year non-profit and 2-year public colleges. While the percentage distribution of independent students are closer at the 2-year institutions – 50.4% for-profit, 31.2% non-profit and 27.4% public – and so yield less of an impact, the default rates at 2-year for-profit colleges are somewhat lower and the default rates at 2-year non-profit colleges are much higher. In other words, the difference in default rates is lower, so the demographic differences have a much greater relative contribution.

THE IMPACT OF HOUSEHOLD SIZE

The following chart shows that overall default rates increase with increasing household size.[21]
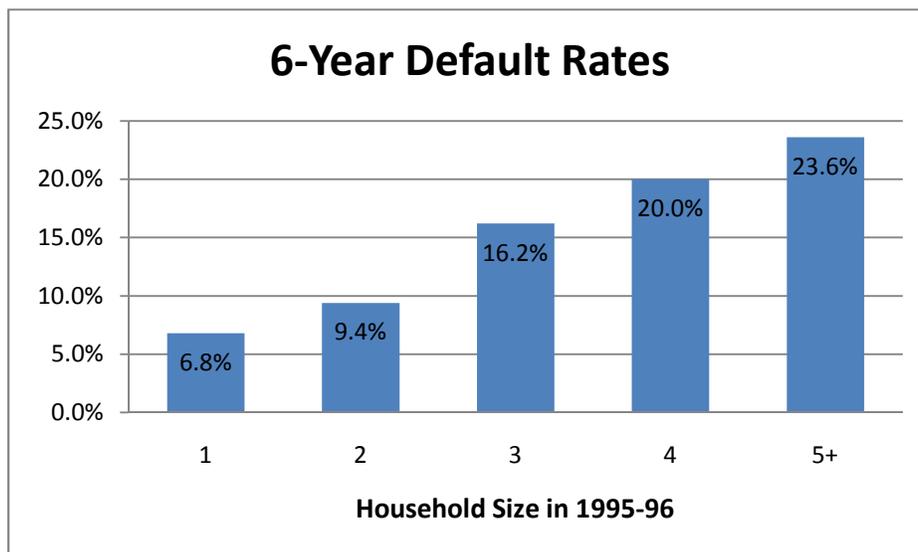
The following table shows the distribution of students according to household size. For-profit colleges enroll a much greater percentage of students from larger families than non-profit and public colleges.

BPS:96/01 Prevalence	Household Size in 1995-96				
	1	2	3	4	5+
Public Colleges	74.9%	8.2%	6.0%	5.1%	5.9%
Non-Profit Colleges	87.6%	3.8%	3.7%	2.8%	2.0%
For-Profit Colleges	31.0%	18.6%	14.4%	16.8%	19.2%

---

[21] Household size is reported as of the first follow-up interview in 1998. It is not necessarily the same as household size as reported on the FAFSA, as dependent students could potentially be reported as having a household size of one, per the SBHHSZB1 variable definition.

The following table shows how the default rates vary based on household size.

BPS:96/01 Default Rates	Household Size in 1995-96					Overall Average
	1	2	3	4	5+	
Public Colleges	6.1%	4.4%	8.3%	16.7%	16.6%	7.3%
Non-Profit Colleges	5.5%	7.7%	21.4%	20.6%	26.4%	7.0%
For-Profit Colleges	20.0%	19.6%	28.3%	24.3%	32.5%	24.2%

Using the methodology presented in this paper, 20.8% of the difference in default rates between for-profit and non-profit colleges and 16.6% of the difference in default rates between for-profit and public colleges are due to for-profit colleges enrolling a much greater percentage of students from larger families than non-profit and public colleges.

DEMOGRAPHIC FACTORS WITH A MODERATE IMPACT ON THE DIFFERENCES IN DEFAULT RATES

Demographic factors with a more moderate impact on the differences in default rates include adjusted gross income, enrollment status, race, the socioeconomic diversity index and whether the college's location close to home was a factor in the choice of college.

Adjusted Gross Income

The following table shows the demographic distribution of students according to adjusted gross income.

BPS:96/01 Prevalence	AGI (Continuous)		
	≤ $50,000	$50,001 to $100,000	> $100,000
Public Colleges	79.4%	19.0%	1.6%
Non-Profit Colleges	62.4%	31.4%	6.2%
For-Profit Colleges	93.9%	6.1%	0.0%

The following table shows how the default rates vary based on adjusted gross income.

BPS:96/01 Default Rates	AGI (Continuous)			Overall Average
	≤ $50,000	$50,001 to $100,000	> $100,000	
Public Colleges	8.3%	3.3%	2.1%	7.2%
Non-Profit Colleges	9.5%	2.8%	2.7%	7.0%
For-Profit Colleges	25.3%	15.0%	NA	24.7%

Since there are almost no students with family income over $100,000 at for-profit colleges, it is not possible to calculate a default rate for these students. Instead, the methodology presented in this paper is applied to just the two income bands, ≤ $50,000 and $50,001 to $100,000, renormalizing the prevalence vector. This indicates that 16.2% of the difference in default rates between for-profit and non-profit colleges and 7.8% of the difference in default rates between for-profit and public colleges are due to for-profit colleges enrolling a greater proportion of lower income students than the other types of colleges. Lower income students are more likely to default than middle and upper income students.

The choice of $50,000 and $100,000 as thresholds was arbitrary. Using $35,000 and $70,000 as thresholds (and renormalizing to exclude family AGI over $70,000 due to a lack of such students at for-profit colleges) yields contributions to the default rates of 25.5% of the difference in default rates between for-profit and non-profit colleges and 10.4% of the difference in default rates between for-profit and public colleges due to for-profit colleges enrolling a greater proportion of lower income students than the other types of colleges.

The following table shows the median, 75th percentile and 90th percentile AGI for students with debt at public, non-profit and for-profit colleges. It demonstrates that students at for-profit colleges represent a much lower income mix than students at public and non-profit colleges.

BPS:96/01 Default Rates	AGI (Continuous)			
	Median	75th Percentile	90th Percentile	Average
Public Colleges	$30,763	$52,725	$72,094	$37,991
Non-Profit Colleges	$43,765	$66,302	$90,468	$50,400
For-Profit Colleges	$15,613	$30,280	$47,366	$24,118

Enrollment Status

The following table shows the demographic distribution of students according to enrollment status.

BPS:96/01 Prevalence	Enrollment Status through 2001		
	Full-Time	Part-Time	Mixed
Public Colleges	50.4%	0.0%	49.6%
Non-Profit Colleges	70.2%	0.0%	29.8%
For-Profit Colleges	67.5%	7.2%	25.4%

The following table shows how the default rates vary based on enrollment status.

BPS:96/01 Default Rates	Enrollment Status through 2001			Overall Average
	Full-Time	Part-Time	Mixed	
Public Colleges	7.7%	NA	7.1%	7.4%
Non-Profit Colleges	6.8%	NA	7.6%	7.0%
For-Profit Colleges	26.4%	27.2%	19.0%	24.6%

While default rates are not available for students enrolled only part-time at public and non-profit colleges, this does not affect the use of the methodology presented in this paper for calculating the contribution of enrollment status on the difference in default rates. Using this methodology, 2.2% of the difference in default rates between for-profit and non-profit colleges and 10.8% of the difference in default rates between for-profit and public colleges are due to differences in the distribution of students according to enrollment status.

Race

The following table shows the demographic distribution of students according to race. For-profit colleges are much more likely to enroll minority students, especially Black and Hispanic students, than non-profit and public colleges.

BPS:96/01 Prevalence	Race			
	White	Black	Hispanic	Other
Public Colleges	70.4%	12.4%	10.1%	7.0%
Non-Profit Colleges	70.4%	12.0%	10.4%	7.2%
For-Profit Colleges	53.9%	19.7%	22.3%	4.0%

The following table shows how the default rates vary based on race. The default rates for minority students are much higher than for non-minority students.

BPS:96/01 Default Rates	Race				Overall Average
	White	Black	Hispanic	Other	
Public Colleges	4.4%	18.2%	11.2%	11.2%	7.3%
Non-Profit Colleges	4.3%	22.1%	10.2%	3.4%	7.0%
For-Profit Colleges	19.5%	41.3%	22.3%	22.9%	24.6%

Using the methodology presented in this paper, 10.9% of the difference in default rates between for-profit and non-profit colleges and 10.6% of the difference in default rates between for-profit and public colleges are due to differences in the racial distribution of the student population.

Socioeconomic Diversity Index

The socioeconomic diversity index measures the degree of socioeconomic disadvantage based on three factors:

- total family income as a percentage of the 1994 federal poverty level
- the highest educational level completed by either parent, and
- the percentage of the student body in the student's high school eligible for the free or reduced-price lunch program in 1994-95

The following table shows the demographic distribution of students according to the socioeconomic diversity index. For-profit colleges are much more likely to enroll socioeconomically disadvantaged students than public and non-profit colleges.

BPS:96/01 Prevalence	Socioeconomic Diversity Index		
	Not Disadvantaged	Minimally Disadvantaged	Moderately or Highly Disadvantaged
Public Colleges	47.1%	35.4%	17.5%
Non-Profit Colleges	58.4%	31.5%	10.1%
For-Profit Colleges	18.9%	45.7%	35.5%

The following table shows how the default rates vary based on the socioeconomic diversity index. The default rates for disadvantaged students are much higher than for non-disadvantaged students.

BPS:96/01 Default Rates	Socioeconomic Diversity Index			Overall Average
	Not Disadvantaged	Minimally Disadvantaged	Moderately or Highly Disadvantaged	
Public Colleges	4.7%	7.9%	13.3%	7.3%
Non-Profit Colleges	4.9%	8.3%	14.9%	7.0%
For-Profit Colleges	19.6%	25.1%	26.6%	24.6%

Using the methodology presented in this paper, 14.5% of the difference in default rates between for-profit and non-profit colleges and 10.6% of the difference in default rates between for-profit and public colleges are due to differences in the socioeconomic diversity of the student population.

College Location Close to Home

The following table shows the demographic distribution of students according to whether their choice of college was based on the college's location close to home. These results are not consistent with other data that indicate that students at for-profit colleges tend to choose the college based on convenience factors such as distance from home. It is possible that the students misinterpreted the question as referring to their parents' home and not their own home.

BPS:96/01 Prevalence	College Located Close to Home	
	No	Yes
Public Colleges	53.8%	46.2%
Non-Profit Colleges	79.7%	20.3%
For-Profit Colleges	87.3%	12.7%

The following table shows how the default rates vary based on whether the student's choice of college was based on the college's location close to home.

BPS:96/01 Default Rates	College Located Close to Home		Overall Average
	No	Yes	
Public Colleges	7.5%	7.5%	7.5%
Non-Profit Colleges	7.4%	5.4%	7.0%
For-Profit Colleges	25.3%	20.4%	24.7%

Using the methodology presented in this paper, 2.1% of the difference in default rates between for-profit and non-profit colleges and 9.6% of the difference in default rates between for-profit and public colleges are due to for-profit colleges enrolling a smaller proportion of students who chose their colleges based on the college's location close to home. The demographic differences do not contribute much to the difference in default rates because the default rates do not differ by much in each demographic segment.

Further analysis of the contribution of the distance of the college from the student's permanent residence demonstrates a result that is more consistent with other analyses. The following table shows the

demographic distribution of students according to the distance of the college from their permanent residence.

BPS:96/01 Prevalence	College Distance from Permanent Residence			
	≤ 5 miles	6-25 miles	26-50 miles	> 50 miles
Public Colleges	27.6%	30.4%	13.8%	28.2%
Non-Profit Colleges	25.8%	15.2%	10.8%	48.2%
For-Profit Colleges	44.6%	34.0%	11.2%	10.1%

The following table shows how the default rates vary based on the distance of the college from the student's permanent residence.

BPS:96/01 Default Rates	College Distance from Permanent Residence				Overall Average
	≤ 5 miles	6-25 miles	26-50 miles	> 50 miles	
Public Colleges	9.4%	6.6%	5.5%	6.8%	7.3%
Non-Profit Colleges	9.3%	8.3%	5.2%	5.7%	7.0%
For-Profit Colleges	29.6%	20.3%	23.1%	18.6%	24.6%

Using the methodology presented in this paper, 13.7% of the difference in default rates between for-profit and non-profit colleges and 10.5% of the difference in default rates between for-profit and public colleges are due to for-profit colleges enrolling more students who have a permanent residence closer to the college. The default rates are higher for students who live within 5 miles of the college, suggesting that students who choose a college primarily based on convenience are more likely to default. This is consistent with other studies that have shown that students who choose a college based on convenience are more likely to drop out.[22]

DEMOGRAPHIC FACTORS WITH A MINIMAL IMPACT ON THE DIFFERENCES IN DEFAULT RATES

Demographic factors with a minimal impact on the differences in default rates include gender, age when first enrolled, remaining need before loans, major, school/work conflicts, whether the student took remedial courses in 1995-96, whether the student has a goal of being well-off financially, whether the college's cost was a factor in the choice of college, and whether the high tuition and fees or a personal or family crisis made it difficult for the student to stay in school in 1995-96.

Gender

Gender should not have much of a contribution to the differences in default rates, except that women tend to have lower salaries than men. Since for-profit college enrollments are tilted more heavily to women, there should be some impact.

---

[22] Jean Johnson and Jon Rochkind, *With Their Whole Lives Ahead of Them*, Public Agenda, December 2009. www.publicagenda.org/theirwholelivesaheadofthem

The following table shows the demographic distribution of students according to gender.

BPS:96/01 Prevalence	Gender	
	Male	Female
Public Colleges	45.9%	54.1%
Non-Profit Colleges	44.1%	55.9%
For-Profit Colleges	36.0%	64.0%

The following table shows how the default rates vary based on gender.

BPS:96/01 Default Rates	Gender		Overall Average
	Male	Female	
Public Colleges	6.4%	8.0%	7.3%
Non-Profit Colleges	6.4%	7.4%	7.0%
For-Profit Colleges	21.2%	26.5%	24.6%

Using the methodology presented in this paper, 2.5% of the difference in default rates between for-profit and non-profit colleges and 3.0% of the difference in default rates between for-profit and public colleges are due to for-profit colleges enrolling a greater proportion of women than the other types of colleges. The margin of error is 1.8%, so these differences are negligible.

Age When First Enrolled

The following table shows the demographic distribution of students according to age when first enrolled.

BPS:96/01 Prevalence	Age When First Enrolled		
	≤ 21	22-26	≥ 27
Public Colleges	87.1%	6.5%	6.4%
Non-Profit Colleges	94.7%	2.4%	3.0%
For-Profit Colleges	57.8%	18.6%	23.5%

The following table shows how the default rates vary based on the student's age when first enrolled.

BPS:96/01 Default Rates	Age When First Enrolled			Overall Average
	≤ 21	22-26	≥ 27	
Public Colleges	7.1%	11.1%	5.8%	7.3%
Non-Profit Colleges	6.6%	13.5%	14.9%	7.0%
For-Profit Colleges	23.6%	35.9%	18.0%	24.6%

Using the methodology presented in this paper, 4.8% of the difference in default rates between for-profit and non-profit colleges and 3.0% of the difference in default rates between for-profit and public colleges are due to for-profit colleges enrolling a greater proportion of older nontraditional students than the other types of colleges.

Remaining Need Before Loans

The following table shows the demographic distribution of students according to remaining financial need before the awarding of Stafford and PLUS loans.

BPS:96/01 Prevalence	Remaining Need Before Loans				
	None	$1 to $2,500	$2,501 to $5,000	$5,001 to $10,000	> $10,000
Public Colleges	32.9%	22.1%	27.5%	15.3%	2.2%
Non-Profit Colleges	18.9%	9.6%	26.1%	30.4%	15.1%
For-Profit Colleges	7.1%	6.6%	13.2%	45.0%	28.1%

The following table shows how the default rates vary based on remaining need before Stafford and PLUS loans.

BPS:96/01 Default Rates	Remaining Need Before Loans					Overall Average
	None	$1 to $2,500	$2,501 to $5,000	$5,001 to $10,000	> $10,000	
Public Colleges	4.0%	8.6%	8.6%	9.2%	14.5%	7.3%
Non-Profit Colleges	5.4%	6.8%	6.6%	7.5%	8.8%	7.0%
For-Profit Colleges	12.5%	30.9%	31.1%	29.2%	15.6%	24.6%

Using the methodology presented in this paper, -0.6% of the difference in default rates between for-profit and non-profit colleges and 1.4% of the difference in default rates between for-profit and public colleges are due to students at for-profit colleges having a greater remaining need before Stafford and PLUS loans. These differences are within the margin of error.

Major

While all types of colleges offer majors in STEM, business and health fields, there are some differences in other fields of study. For example, public and non-profit colleges offer majors in humanities, social sciences and education while for-profit colleges do not. Likewise, for-profit colleges offer degrees in vocational and technical fields while public and non-profit colleges generally do not. These other fields of study are excluded from the following analysis, which considers only the distribution of students within the STEM, business and health fields.

The following table shows the demographic distribution of students according to major.

BPS:96/01 Prevalence	Major		
	STEM	Business	Health
Public Colleges	41.9%	30.7%	27.4%
Non-Profit Colleges	48.2%	29.6%	22.1%
For-Profit Colleges	29.8%	49.2%	20.9%

The following table shows how the default rates vary based on major.

BPS:96/01 Default Rates	Major			Overall Average
	STEM	Business	Health	
Public Colleges	5.8%	4.4%	9.2%	6.3%
Non-Profit Colleges	4.2%	7.6%	6.5%	5.7%
For-Profit Colleges	16.9%	25.0%	45.1%	26.8%

Using the methodology presented in this paper, 5.9% of the difference in default rates between for-profit and non-profit colleges and -1.6% of the difference in default rates between for-profit and public colleges are due to differences in the distribution of students according to majors in STEM, Business and Health fields of study.

This analysis excludes an "Other" category for students majoring in other fields of study because it would skew the results due to the mismatch between majors at the different types of colleges. It would not represent an apples-to-apples comparison. Vocational and technical majors at for-profit colleges are not the equivalent of majors in humanities, social sciences and education at non-profit and public colleges. However, excluding an "Other" category also affects the results, since 42.4% of students at public colleges, 46.6% of students at non-profit colleges and 16.8% of students at for-profit colleges pursue majors in other fields of study. Note that students enrolled in vocational and technical fields at for-profit colleges have a lower default rate, 18.2%, than students majoring in business and health. The default rates for students majoring in humanities, social sciences and education are 10.4% at public colleges and 10.2% at non-profit colleges, higher than the default rates for students majoring in STEM, business and health fields.

School/Work Conflicts

The following table shows the demographic distribution of students according to whether school/work conflicts made it difficult for the student to stay in school in 1995-96.

BPS:96/01 Prevalence	School/Work Conflicts Made It Difficult to Stay in School	
	No	Yes
Public Colleges	91.0%	9.0%
Non-Profit Colleges	95.9%	4.1%
For-Profit Colleges	92.6%	7.4%

The following table shows how the default rates vary based on whether the student took remedial coursework in 1995-96.

BPS:96/01 Default Rates	School/Work Conflicts Made It Difficult to Stay in School		Overall Average
	No	Yes	
Public Colleges	7.4%	5.1%	7.2%
Non-Profit Colleges	6.6%	14.1%	6.9%
For-Profit Colleges	24.7%	25.8%	24.8%

Using the methodology presented in this paper, 0.2% of the difference in default rates between for-profit and non-profit colleges and -0.1% of the difference in default rates between for-profit and public colleges are due to demographic differences in school/work conflicts making it difficult for the student to stay in school. This is partly because the default rates at the for-profit colleges do not differ by much according to the impact of the school/work conflicts.

Remedial Coursework

The following table shows the demographic distribution of students according to whether the student took remedial coursework in 1995-96. Surprisingly, a smaller percentage of students at for-profit colleges participated in remedial coursework.[23]

BPS:96/01 Prevalence	Remedial Coursework 1995-96	
	No	Yes
Public Colleges	77.8%	22.2%
Non-Profit Colleges	87.2%	12.8%
For-Profit Colleges	94.3%	5.7%

The following table shows how the default rates vary based on whether the student took remedial coursework in 1995-96.

BPS:96/01 Default Rates	Remedial Coursework 1995-96		Overall Average
	No	Yes	
Public Colleges	7.4%	8.2%	7.6%
Non-Profit Colleges	6.5%	9.0%	6.8%
For-Profit Colleges	23.7%	28.5%	24.0%

Using the methodology presented in this paper, -2.0% of the difference in default rates between for-profit and non-profit colleges and -4.8% of the difference in default rates between for-profit and public colleges are due to demographic differences in students taking remedial coursework. The negative contribution rates mean that the demographic differences do not contribute to differences in the default rates.

Goal: Well-Off Financially

The following table shows the demographic distribution of students according to whether the student has a goal of being well-off financially.

BPS:96/01 Prevalence	Goal: Well-Off Financially	
	No	Yes
Public Colleges	19.5%	80.5%
Non-Profit Colleges	25.9%	74.1%
For-Profit Colleges	10.6%	89.4%

---

[23] This may be due to some for-profit colleges using just-in-time remediation as opposed to separate remedial courses.

The following table shows how the default rates vary based on whether the student has a goal of being well-off financially.

BPS:96/01 Default Rates	Goal: Well-Off Financially		Overall Average
	No	Yes	
Public Colleges	6.0%	8.0%	7.6%
Non-Profit Colleges	6.0%	7.2%	6.9%
For-Profit Colleges	16.8%	24.6%	23.8%

Using the methodology presented in this paper, 7.1% of the difference in default rates between for-profit and non-profit colleges and 4.3% of the difference in default rates between for-profit and public colleges are due to demographic differences in student goals to be well-off financially.

College Choice Based on College Cost

The following table shows the demographic distribution of students according to whether their choice of college was based on the college's cost.

BPS:96/01 Prevalence	Cost was a Factor in Choice of College	
	No	Yes
Public Colleges	64.1%	35.9%
Non-Profit Colleges	80.3%	19.7%
For-Profit Colleges	89.6%	10.4%

The following table shows how the default rates vary based on whether the student's choice of college was based on the college's cost.

BPS:96/01 Default Rates	Cost was a Factor in Choice of College		Overall Average
	No	Yes	
Public Colleges	7.1%	8.3%	7.,5%
Non-Profit Colleges	7.4%	5.5%	7.0%
For-Profit Colleges	24.4%	27.1%	24.7%

Using the methodology presented in this paper, -1.4% of the difference in default rates between for-profit and non-profit colleges and -4.0% of the difference in default rates between for-profit and public colleges are due to for-profit colleges enrolling a smaller proportion of students who chose their colleges based on the college's cost. The negative contributions indicate that the default rates would be higher if more of the students were focused on cost in their choice of college, since the default rate was higher within the segment that considered colleges based on cost. Even so, the magnitude of the contribution is minimal.

Impact of High Tuition and Fees

The following table shows the demographic distribution of students according to whether high college tuition and fees made it difficult for the student to stay in school in 1995-96.

BPS:96/01 Prevalence	High Tuition and Fees Made It Difficult to Stay in School	
	No	Yes
Public Colleges	93.3%	6.7%
Non-Profit Colleges	89.7%	10.3%
For-Profit Colleges	92.9%	7.1%

The following table shows how the default rates vary based on whether high college tuition and fees made it difficult for the student to stay in school in 1995-96.

BPS:96/01 Default Rates	High Tuition and Fees Made It Difficult to Stay in School		Overall Average
	No	Yes	
Public Colleges	7.4%	5.3%	7.3%
Non-Profit Colleges	6.7%	8.7%	6.9%
For-Profit Colleges	25.3%	17.8%	24.8%

Using the methodology presented in this paper, 1.3% of the difference in default rates between for-profit and non-profit colleges and -0.2% of the difference in default rates between for-profit and public colleges are due to whether high tuition and fees made it difficult for the student to stay in school in 1995-96. These differences fall within the margin of error.

Impact of Personal or Family Crisis

The following table shows the demographic distribution of students according to whether a personal or family crisis made it difficult for the student to stay in school in 1995-96.

BPS:96/01 Prevalence	Personal/Family Crisis Made It Difficult to Stay in School	
	No	Yes
Public Colleges	94.9%	5.1%
Non-Profit Colleges	96.2%	3.8%
For-Profit Colleges	93.4%	6.6%

The following table shows how the default rates vary based on whether a personal or family crisis made it difficult for the student to stay in school in 1995-96.
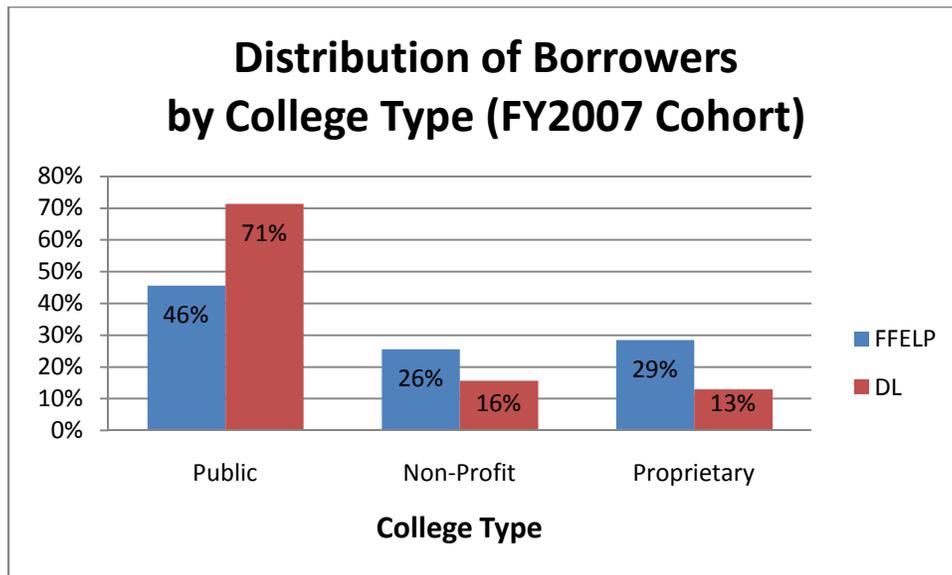
BPS:96/01 Default Rates	Personal/Family Crisis Made It Difficult to Stay in School		Overall Average
	No	Yes	
Public Colleges	7.1%	10.2%	7.3%
Non-Profit Colleges	6.7%	11.4%	6.9%
For-Profit Colleges	24.0%	36.1%	24.8%

Using the methodology presented in this paper, 1.9% of the difference in default rates between for-profit and non-profit colleges and 1.1% of the difference in default rates between for-profit and public colleges

are due to whether a personal/family crisis made it difficult for the student to stay in school in 1995-96. These differences fall within the margin of error. The differences are minimal despite much higher default rates for students who were negatively affected by the personal/family crisis because the demographic distribution was similar for for-profit, non-profit and public colleges.

DIFFERENCES IN DEFAULT RATES BETWEEN THE FFEL AND DL PROGRAMS

A similar approach can be used to calculate the contribution of institution type to the differences in default rates between the federally-guaranteed student loan program (FFEL)[24] and the direct loan program (DL).[25] The FFEL program is more heavily weighted toward non-profit and for-profit colleges and the DL program is more heavily weighted toward public colleges, as is demonstrated in the following chart.

Distribution of Borrowers by College Type (FY2007 Cohort)

College Type	FFELP	DL
Public	46%	71%
Non-Profit	26%	16%
Proprietary	29%	13%

The following table shows the distribution of the DL and FFEL loan programs according to control of institution.

FY2007 Prevalence	Institution Control		
	Public	Non-Profit	For-Profit
DL	71.3%	15.7%	13.0%
FFEL	45.7%	25.6%	28.5%

---

[24] See www2.ed.gov/offices/OSFAP/defaultmanagement/3yrffelrates.pdf
[25] See www2.ed.gov/offices/OSFAP/defaultmanagement/3yrdlrates.pdf

The following table illustrates how the default rates vary based on control of institution. Default rates are higher at for-profit colleges and lower at non-profit colleges, so the distribution will contribute to the differences in default rates.[26]

FY2007 Default Rates	Institution Control			Overall Average
	Public	Non-Profit	For-Profit	
DL	4.2%	3.5%	10.6%	4.9%
FFEL	6.8%	3.9%	11.1%	7.2%

Using the methodology presented in this paper and the data in these two tables, 16.0% of the difference in default rates between the FFEL and DL programs for the FY2007 cohort is due to differences in the distribution of the loan programs according to institution control. Similarly, the differences in the distribution of the loan programs according to institution control contributed 17.2% of the difference in the default rates between the FFEL and DL programs for the FY2006 cohort.

The following table shows the distribution of the DL and FFEL loan programs according to institution type (level and control) for the FY2007 cohort. It demonstrates significant differences in the distribution of the loan programs among borrowers at 4-year public, non-profit and for-profit colleges and at 2-year public colleges.

FY2007 Prevalence	Institution Type								
	Public < 2 yr	Public 2-3 yr	Public 4yr	Non-Profit < 2yr	Non-Profit 2-3 yr	Non-Profit 4yr	For-Profit < 2yr	For-Profit 2-3 yr	For-Profit 4yr
DL	0.1%	9.3%	62.0%	0.3%	0.1%	15.3%	5.0%	5.0%	3.0%
FFEL	0.3%	15.9%	29.4%	0.1%	0.5%	25.0%	3.5%	8.6%	16.4%

The following table illustrates how default rates vary according to institution type.

FY2007 Default Rates	Institution Type								
	Public < 2 yr	Public 2-3 yr	Public 4yr	Non-Profit < 2yr	Non-Profit 2-3 yr	Non-Profit 4yr	For-Profit < 2yr	For-Profit 2-3 yr	For-Profit 4yr
DL	6.7%	9.1%	3.4%	13.2%	7.3%	3.3%	10.6%	12.0%	8.1%
FFEL	7.7%	10.1%	5.0%	12.1%	8.2%	3.8%	12.7%	12.7%	9.9%

Using the methodology presented in this paper and the data in these two tables, 44.7% of the difference in default rates between the FFEL and DL programs for the FY2007 cohort is due to differences in the distribution of the loan programs according to institution type. However, the differences in the distribution of the loan programs according to institution type contributed 97.9% of the difference in the default rates between the FFEL and DL programs for the FY2006 cohort and 108.0% of the difference in the default rates between the FFEL and DL programs for the FY2005 cohort.

---

[26] Rounding errors in the default rates presented in the tables on the US Department of Education web site have been corrected.

The smaller contribution of demographics to the difference in overall default rates for the FY2007 cohort may be caused by changes in the distribution of colleges in the DL and FFEL loan programs and changes in the default rates at each type of college in the two programs from the FY2006 to FY2007 cohorts.

The following table shows the distribution of the DL and FFEL loan programs according to institution type (level and control) in FY2006. The distribution is shifted somewhat toward lower default rate colleges, especially in the FFEL program, as compared with the distribution for the FY2007 cohort.

FY2006 Prevalence	Institution Type								
	Public < 2 yr	Public 2-3 yr	Public 4yr	Non-Profit < 2yr	Non-Profit 2-3 yr	Non-Profit 4yr	For-Profit < 2yr	For-Profit 2-3 yr	For-Profit 4yr
DL	0.1%	9.4%	60.3%	0.3%	0.2%	15.9%	5.3%	4.9%	3.7%
FFEL	0.2%	14.3%	31.6%	0.0%	0.5%	29.1%	3.1%	7.3%	13.3%

The following table illustrates how default rates vary according to institution type. The default rates in the DL and FFEL programs are much closer in the FY2006 cohort than in the FY2007 cohort, and in some cases the FFEL default rates are lower than the corresponding DL default rates.

FY2006 Default Rates	Institution Type								
	Public < 2 yr	Public 2-3 yr	Public 4yr	Non-Profit < 2yr	Non-Profit 2-3 yr	Non-Profit 4yr	For-Profit < 2yr	For-Profit 2-3 yr	For-Profit 4yr
DL	6.8%	8.3%	3.2%	10.9%	10.5%	3.3%	9.7%	11.5%	7.8%
FFEL	6.4%	8.4%	3.4%	8.7%	5.7%	2.3%	11.5%	11.1%	8.4%

It is unclear why there is such a big shift in the default rates for the FFEL program but not for the DL program. The most likely explanation is the impact of the early repayment status loophole, which allowed borrowers to consolidate their loans while they were still in school. This loophole affects the cohort in which a borrower is included. It also eliminates the six month grace period before repayment begins. Borrowers in the FFEL program were more likely to take advantage of the early repayment status loophole than borrowers in the DL program. This is reflected in the changes in the total number of borrowers, with much greater volatility in the FFEL numbers than in the DL numbers. The number of borrowers in the FFEL program decreased 17.6% from 3.2 million in the FY2006 cohort to 2.6 million in the FY2007 cohort while the number of borrowers in the DL program decreased 2.1% from 797,000 in the FY2006 cohort to 781,000 in the FY2007 cohort. During the previous year the number of FFEL borrowers increased by 15.8% and the number of DL borrowers decreased by 1.6%.

Similar results can be obtained using long-term default rate projections as published in the education appendix to the President's FY2011 budget. The following table shows the distribution of dollar loan volume in the FFEL and Direct Loan programs according to type of loan.

FY2009 Prevalence	Type of Loan			
	Subsidized Stafford	Unsubsidized Stafford	PLUS	Consolidation
DL	26.8%	31.2%	12.3%	29.6%
FFEL	38.0%	48.4%	13.3%	0.3%

The following table illustrates how default rates vary according to type of loan.

FY2009 Default Rates	Type of Loan				
	Subsidized Stafford	Unsubsidized Stafford	PLUS	Consolidation	Overall
DL	16.44%	16.25%	7.93%	22.24%	17.05%
FFEL	15.88%	16.51%	6.68%	16.16%	14.96%

Using the methodology presented in this paper and the data in these two tables, 86.9% of the difference in long-term default rate projections between the FFEL and DL programs in FY2009 is due to differences in the distribution of the loan programs according to type of loan. Further analysis shows that this is mostly due to the differences in distribution between new and consolidation loans and not as much due to the differences in the distribution of the various types of new loans. Differences in default rates for consolidation loans are largely responsible for the differences in the overall projected default rates for the DL and FFEL programs. The long-term projected default rates for just the new loans are 14.87% for the DL program and 14.96% for the FFEL program. If the DL program loans had the same distribution as the FFEL program loans for new and consolidation loans, the DL program would have a projected long-term default rate of 15.21%, not much different than the 14.96% projected default rate for the FFEL program. The results for FY2008 are similar, with differences in the distribution of the loan programs according to type of loan contributing 84.0% of the difference in the default rates between the FFEL and DL programs.

Although comparing the performance of the FFEL and DL programs may have been rendered moot by the elimination of the FFEL program, this type of analysis will remain useful. The Student Aid and Fiscal Responsibility Act (SAFRA), as incorporated into the Health Care and Education Reconciliation Act of 2010 (P.L. 111-152), includes a $1.5 billion earmark that gives state loan agencies a servicing lock on up to 100,000 borrowers in each state. This prevents the US Department of Education from evaluating state loan agency servicer performance (especially default aversion rates) on an apples-to-apples basis through the random assignment of borrowers to servicers. Default rates vary from state to state for several reasons, including differences in unemployment rates and the mix of colleges according to level and control of institution. Default rates are higher in southern states than in Massachusetts because more higher-income students tend to enroll in New England colleges and because the mix of college types and costs vary considerably. Without the methodology presented in this paper, it would be difficult to determine how much of the servicer performance is due to differences in geography-specific demographics and how much due to the effectiveness of each state servicer.

FURTHER RESEARCH

It is clear that service by for-profit colleges to at-risk populations that are underserved by non-profit and public colleges is responsible for a significant part of the difference in the default rates between for-profit colleges and non-profit and public colleges.

This result calls into question whether non-profit and public colleges are adequately fulfilling their charitable missions by enrolling and graduating enough students from at-risk populations. If a college admits mainly higher-income students, is it still a charitable organization? The greater selectivity of non-profit colleges is in effect acting as a source of bias against higher-risk students. Although six dozen of the more elite institutions have adopted "no loans" financial aid policies that substitute grants for loans in the need-based financial aid package, none of these colleges have adopted admissions preferences for low-income students. When generous financial aid packages are awarded to moderate and upper income students too, the competition for admission intensifies, making it more difficult for low-income students to get in. Low-income students are often the primary wage-earners for their families and do not have the opportunity to participate in athletics or other extracurricular activities. But the most elite institutions seem to be more impressed by a student who volunteers at a soup kitchen (and plays a musical instrument and letters in a sport) than by a student who has benefited from the soup kitchen, even though a student who has succeeded academically despite adversity is much more impressive than a student who was born with a silver spoon in his or her mouth. Some college admissions offices claim to give extra credit to students who have overcome adversity, but the weighting is rarely sufficient to compensate for the more favorable treatment accorded to more "well-rounded" students.

On the other hand, is the greater prevalence of high-risk students at for-profit colleges due to open enrollment policies, better customer service, aggressive counseling and higher FAFSA completion rates? Or are the for-profit colleges exploiting a vulnerable population? This paper does not address these questions, as this would depend on the causes of the differences in the prevalence and default rate vectors and on an independent and objective evaluation of the quality of the degree programs at each type of college.

Another area of possible future research involves analysis of combinations of demographic variables, such as the risk index and Pell Grant recipient status, by forming disjunctions of these variables. For example, a risk index of 2 or higher could be combined with Pell Grant recipient status using an OR-filter, and a risk index of 0 or 1 could be combined with Pell Grant non-recipient status with an AND-filter (i.e., NOT (A | B) is the same as NOT(A) & NOT(B)). This approach would avoid the sample size issues inherent in using prevalence vectors that result from a cross-product of multiple variables. This might help better determine how much of the differences in default rates are due to all of the demographic characteristics.

Alternately, one could explore combinations of the four demographic factors that have a significant impact on default rates with the risk index, perhaps by adding Pell Grant recipient status, household size, parent educational attainment of at least a Bachelor's degree and independent student status to the list of risk factors. This would require the addition of a new derived variable to the data analysis system for the BPS longitudinal studies.