Community College Graduation among Vulnerable Populations: Sexual Minority, Insecurely Housed, and Disabled Students

Janet Rosenbaum[1]

[1] University of Maryland, College Park, MD

Abstract

This research examines post-secondary degree attainment among three vulnerable populations of community college students: sexual minority, insecurely housed, and disabled students. The data was from Add Health respondents with high school diplomas enrolled in community college in 2001 (n=1310). Three post-secondary degree attainment outcomes were used: any credential/degree, AA or above, and BA or above. Poisson regression with robust standard errors estimated incidence rate ratios. Seven years later, 60% of community college students had attained a post-secondary credential/degree. Sexual minority students were more than 40% less likely to attain each of the three education levels, unless they were "out" to parents. Factors that predict not achieving post-secondary degrees include having an imprisoned fathers, having been in a group home, and stuttering. Subsequent analyses will use causal inference methods to minimize confounding and identify factors that mediate the community college graduation outcomes.

Objective

About half of community college students who left school after their first or second years left because of "personal reasons," but past research does not specify the personal reasons (1). This research tests the hypothesis that specific vulnerable populations are less likely to graduate community college than students without those factors, and investigates protective factors that help these vulnerable students graduate. The research investigates three vulnerable populations: sexual minority, insecurely housed, and disabled students. Improving the community college graduation rate is a central focus of the current administration's post-secondary education policy. About 15% of community college students finish an AA or certificate within 3 years, and 45% leave school with no credential (2). Half of former community college students cite "personal reasons" as their explanation for leaving school, about twice as many as cite "family" or "finances." (1). Personal reasons could include many risk factors not measured by standard education surveys, such as current and past drug use, unstable or abusive sexual and romantic relationships, or low self-efficacy (3). Figure 1 gives a general conceptual model to illustrate the relationship between generic risk and protective behaviors and college graduation.

Theoretical framework

Tinto's model of student retention posits that students who become well-integrated, both academically and socially, are more likely to persist in school than students who are not well-integrated (4). Students from marginalized populations are predicted by Tintos model to be less likely to graduate because they have lower levels of social integration. In most community colleges, students commute to college and have social

contexts outside of the college community, so social integration may play a different role for community college students than it does for 4 year college students.

Students from generally marginalized populations may also be at risk if they engage in risk behaviors that decrease their college attachment independent of their level of social integration in the college community. Marginalized populations generally engage in more risk behaviors, and these risk factors rather than their marginalized identities may be responsible for reducing graduation likelihood.

Using a survey with an unusually rich variety of indicators of risk behaviors, we hypothesize that integration may be affected by students identities and their risk behaviors, particularly ones that preceded college and have persisted in college. This study broadens our understanding of influences which are implicit in Tinto's model but rarely examined. The personal factors that community college students cite may impact either academic or social integration. The proposed research will test hypotheses about whether students from three vulnerable populations are less likely to graduate, and identify factors that mediate the influence of these risk factors on college attainment outcomes.

Methods

Bivariate and regression analyses

Data analysis used Stata 11 and R. Bivariate analysis used a chi-squared test of proportions. Multivariate analysis obtained relative risks from a general linear model Poisson regression with robust standard errors.

The predictors of interest are the three marginalized identity statuses: possible housing instability (e.g., having been kicked out of their parents home, having cohabited in late adolescence, history of homelessness or having lived in a group home), sexual minority status (identifying as lesbian/gay/bisexual, having same-sex attraction), and disability (e.g., hearing or sight impairment, asthma, depression diagnosis, stutter, limitations on activities of daily living.)

Outcome variables were measured at wave 4 (2008). Regressions used three dichotomous outcome variables: having earned any post-secondary credential, an AA or above, and a BA or above. The control variables were gender, Latino/Asian/Black race/ethnicity, age at wave 4, parent-reported educational level, log parent-reported household income, Peabody vocabulary test score percentile at wave 3 (singly imputed), age of high school graduation, and whether respondent enrolled in their current community college more than one year after high school graduation.

Matched sampling

Regression methods alone are probably insufficient for correcting for complex relationships between factors. The parametric assumptions of regression are rarely satisfied (5). After regression analyses, residual confounding by these and other factors has been found to remain (6–8). Subsequent analyses will use matched sampling to reduce confounding. Matched sampling methods are often able to create balance on many more factors than were present in the matching model (9–11). The subsequent analysis will use matched sampling methods (9–11), including a new method, coarsened exact matching (12–14). Traditional regression methods try to create comparability between possibly non-comparable individuals with arithmetic adjustments. Matched sampling methods reduce confounding by balancing groups on dozens of variables, so that the only observed difference between the groups is the predictor of interest. The potential impact of non-observed variables can be quantified using sensitivity analysis (15).

Matched sampling will occur as follows. We will identify potential confounders of the relationship between LGB status and college graduation, enter them in Table 4, fill in the means, and perform an appropriate test of difference. After matching, we test whether the differences between groups remain, and if not, we will fill in the remaining columns of Table 4. Otherwise, we will repeat the matching procedure until balance is achieved. Within the matched sample, we will repeat the above regression analysis.

Data sources

This study will use waves 1–4 of the National Longitudinal Study of Adolescent Health (Add Health). Add Health is a nationally representative sample of students who were grades 7–12 in 1995 (wave 1), who were surveyed again in 1996, 2001, and 2008. Add Health also includes high school transcripts, biomarkers, and surveys of parents and school administrators. Analysis was restricted to students who had a high school diploma, were enrolled in community college, and did not report having an associates degree in 2001 (n=1310).

Prior research on community college students and completion rates derives from the Beginning Postsecondary Students (BPS) Longitudinal Study, but like other Department of Education datasets BPS doesnt include risk behavior information.

Results

About 60% of all wave 3 community college students had earned some post-secondary credential at wave 4: 12% earned certificate(s), 28% earned AAs, and 20% earned a BA or above.

Sexual minorities

In bivariate analysis, identifying as lesbian/gay/bisexual (LGB) or having a same-sex attraction were associated with lower likelihood of earning any credential, an AA, and LGB identity was associated with lower likelihood of earning a BA.

The difference in graduation rates was lower for LGB students who were out to one or both of their parents. This reduction in graduation disparities by coming out to parents may suggest either that LGB students from supportive families or in places where LGB identity is less deviant are more likely both to be out to their parents and graduate; it may also suggest that being LGB while not being out to parents somehow reduces the likelihood of graduation such as by creating stress or cognitive dissonance or marginal identity. The change from statistically significant to non-statistically significant with coming out is partially a statistical artifact. As Gelman pointed out, the difference between statistically significant and non-statistically significant is not itself statistically significant: that is, small changes can alter a result from being statistically significant to non-significant (16). In this case, the change from significant to non-significant is partially attributable to statistical factors because such a small proportion of LGB students were out to their parents: 4% of students identify as LGB students, but only 1.5% and 1.1% of students are out to one or both parents, respectively. For all three education levels, however, the graduation rates of LGB students increased monotonically from all LGB to those out to at least 1 parent to those out to both parents.

In multivariate analysis, the same pattern of significance remained: LGB and same-sex-attracted students were significantly less likely to earn any postsecondary credential and AAs, and LGB students were less likely to earn BAs. This result suggests that, if parametric assumptions hold, the disparity between sexual minorities and non-sexual minorities is not attributable to educational background, socioeconomic status, or race/ethnicity. Further analysis will examine family and contextual factors because family and community attitudes about sexual minorities may mediate the relationship between sexual minority status and likelihood of earning a degree.

As in the bivariate analysis LGB students who are out to one or both parents are not less likely to earn degrees than non-LGB students, although again this non-significance is likely partially due to statistical factors such as higher standard error, which can be seen in the wider confidence intervals among the "out LGB students.

Housing/family instability

Having a father in prison, having been kicked out of home, having ever lived in a group home, and early cohabitation were associated with lower likelihood of graduation, but became less significant in multivariate analysis.

Disabilities

Limitations in either vigorous or any activities of daily living (ADL) were associated with lower likelihood of degree attainment in both bivariate and multivariate analysis, but less so at the BA level. Having limitations in moderate ADL was not associated with lower likelihood of degree attainment, which may be attributable to self-selection into community college. Limitations in moderate ADL represents a more severe form of disability than limitations in vigorous ADL, and the students who chose to attend community college were likely those who were more likely to succeed in college due to the nature of their disabilities, support system strength, or other factors.

Stuttering was associated with reduced likelihood of attaining a BA in both bivariate and multivariate analysis, but not associated in multivariate analysis with reduced likelihood of attaining other degrees. Community college students who stutter were less likely to earn an AA, but after adjusting for background factors in regression, the difference did not remain.

Asthma, depression diagnosis, and hearing impairments were not associated with lower likelihood of degree attainment in either bivariate or multivariate analysis.

Further analysis

Further analysis will investigate these relationships in greater depth, and use more rigorous statistics to reduce confounding further. The regression can reduce confounding to the extent that the confounders are log-linearly associated with the outcome, but these parametric assumptions are rarely true. Non-parametric methods that reduce confounding including matched sampling can reduce confounding beyond what is possible with regression.

Additional analyses can also attempt to identify explanations for the results presented here, such as the extent to which coming out helps LGB community college students succeed versus the extent to which LGB students who are more likely to succeed are also more likely to come out.

Significance

The most common reason that former community college students give for leaving school is "personal reasons" (1), but these reasons likely encompass a diversity of situations. This research will begin to describe how the factors identified in preliminary analysis may impact community college students' graduation likelihood, with policy implications as diverse as the factors.
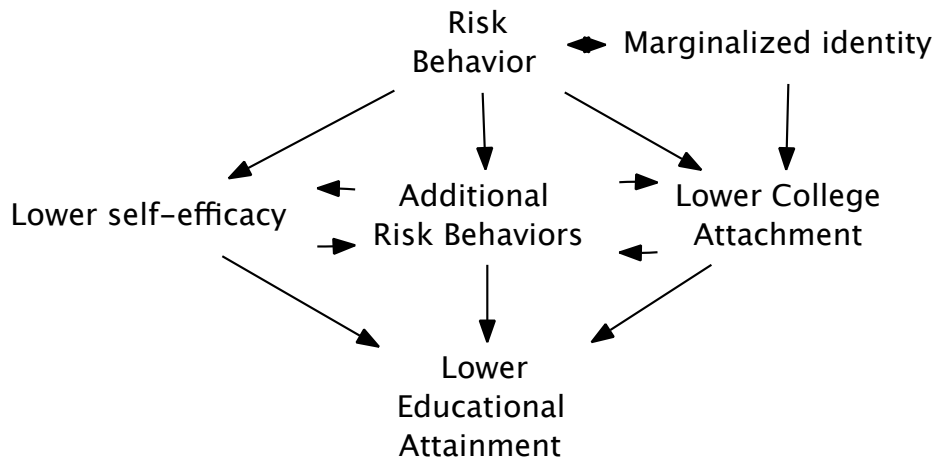
Figure 1: Hypothesized relationship between marginalized identities, health risk behaviors, and educational achievement. Marginalized students may engage in more risk behaviors, and risk behaviors may cause students to feel further marginalized, like they are not part of the mainstream of society. Marginalization may induce lower college attachment. Simultaneously, risk behaviors are associated with lower self-efficacy, additional risk behavior, and lower college attachment, which are also associated with each other. All three predict lower educational attainment.

Table 1: Wave 4 (2008) educational attainments of wave 3 (2001) community college students who reported a high school diploma but no AA at wave 3. Educational attainment is defined as self-reported highest level of education.

Degree	Number	Survey-weighed %
No post-secondary degree	510	40.2
Certificate(s)	171	12.3
AA(s)	357	27.9
BA(s)	236	16.3
Above BA	36	3.4
Total	1310	100

Table 2: Graduation rates with and without the listed traits, and proportion with traits. Bivariate analysis of individual attributes associated with receiving post-secondary credentials 7 years after community college enrollment (n=1310). All individual attributes are measured at wave 3, in 2001. Educational credentials are measured at wave 4, in 2008. Yes are those with the trait, and no are those without the trait. The average raw difference summarizes the difference for each factor, and is defined as the average difference in the unprotected sex percentages across the three waves between those with and without the factor. The results are sorted in order of the average raw difference.

		Any credential			Associates+			BA+			Mean raw
	%	No	Yes	p	No	Yes	p	No	Yes	p	Difference
Sexuality											
Identify as LGB	4.0	62.0	39.6	***	48.8	28.3	**	21.3	7.5	*	18.9
LGB and out	1.5	61.3	42.1	+	48.3	31.6	+	20.8	15.8		13.6
LGB out to both parents	1.1	61.2	46.7		48.2	33.3		20.8	20.0		10.1
Same sex attraction	9.8	62.2	50.8	**	49.2	37.5	**	21.2	16.4		9.2
Family/housing stability											
Ever in group home	1.7	61.3	45.5	+	48.3	31.8	+	21.1	0	*	17.8
Father in prison	14.6	62.9	50.3	***	50.2	35.1	****	21.8	14.7	*	11.6
Ever homeless	3.2	61.5	47.6	+	48.4	35.7	+	20.9	16.7		10.3
Ever kicked out of home	10.7	62.1	52.9	*	49.4	36.4	**	21.5	15.0	+	9.6
Early cohabitation	30.4	62.4	58.0	+	51.1	41.0	***	23.0	15.6	**	7.3
Disability											
Stutter	7.3	61.3	58.3		48.8	38.5	*	21.7	8.3	**	8.9
Vigorous ADL limits	22.8	62.7	55.5	*	49.9	41.8	**	21.9	17.1	+	6.7
Any ADL limits	27.1	62.8	56.3	*	50.1	42.5	*	22.2	16.9	*	6.5
Long-term ADL limits	18.6	62.0	57.0	+	49.2	43.0	+	21.5	17.6		5.0
Moderate ADL limits	4.5	61.2	59.3		48.3	42.4		21.0	15.3		4.5
Hearing impairment	4.8	61.1	60.3		48.3	42.9		21.1	14.3		4.3
Asthma	17.2	61.5	59.1		48.5	45.8		20.8	20.4		1.8
Depression diagnosis	10.4	61.2	60.3		48.1	47.1		21.1	17.6		1.8

Graduation rate

+ $p \leq 0.1$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$

ADL= Activities of daily living.
The mean raw difference is the arithmetic average of the differences between categories.

Table 3: Multivariate analysis of individual attributes associated with receiving post-secondary credentials 7 years after community college enrollment (n=1276). All individual attributes are measured at wave 3, in 2001. Educational credentials are measured at wave 4, in 2008. Incidence rate ratios are computed from Poisson regression with robust standard errors, controlling for demographic and educational factors.

	%	any credential	p	AA+	p	BA+	p	Mean IRR
Sexuality								
Identify as LGB	4.0	0.59 (0.41, 0.83)	**	0.57 (0.37, 0.88)	**	0.34 (0.13, 0.88)	*	0.49
LGB and out	1.5	0.62 (0.35, 1.10)	+	0.71 (0.38, 1.31)		0.82 (0.29, 2.32)		0.71
Same sex attraction	9.8	0.77 (0.64, 0.92)	**	0.76 (0.60, 0.95)	*	0.75 (0.50, 1.14)		0.76
LGB out to both parents	1.1	0.68 (0.37, 1.25)		0.75 (0.38, 1.46)		1.03 (0.37, 2.84)		0.81
Family/housing stability								
Ever in group home	1.7	0.65 (0.39, 1.07)	+	0.57 (0.30, 1.08)	+	na	na	0.61
Father in prison	14.6	0.84 (0.73, 0.97)	*	0.77 (0.63, 0.93)	**	0.71 (0.50, 1.03)	+	0.77
Ever kicked out of home	10.7	0.88 (0.75, 1.02)	+	0.76 (0.61, 0.95)	*	0.75 (0.50, 1.13)		0.79
Ever homeless	3.2	0.77 (0.54, 1.08)		0.78 (0.51, 1.20)		0.90 (0.46, 1.76)		0.81
Early cohabitation	30.4	0.93 (0.85, 1.03)		0.86 (0.75, 0.99)	*	0.78 (0.60, 1.01)	+	0.85
Disability								
Stutter	7.3	0.99 (0.83, 1.18)		0.86 (0.67, 1.10)		0.41 (0.21, 0.81)	**	0.70
Hearing impairment	4.8	0.99 (0.82, 1.20)		0.88 (0.66, 1.16)		0.68 (0.37, 1.25)		0.84
Vigorous ADL limits	22.8	0.88 (0.79, 0.98)	*	0.84 (0.73, 0.97)	*	0.80 (0.61, 1.06)	+	0.84
Any ADL limits	27.1	0.89 (0.80, 0.98)	*	0.85 (0.75, 0.98)	*	0.79 (0.61, 1.02)	+	0.84
Depression diagnosis	10.4	0.92 (0.80, 1.06)		0.90 (0.75, 1.09)		0.75 (0.51, 1.11)		0.85
Long-term ADL limits	18.6	0.90 (0.80, 1.01)	+	0.86 (0.73, 1.00)	*	0.82 (0.61, 1.10)		0.86
Moderate ADL limits	4.5	0.97 (0.78, 1.20)		0.90 (0.66, 1.22)		0.80 (0.43, 1.50)		0.89
Asthma	17.2	0.93 (0.82, 1.04)		0.89 (0.77, 1.04)	+	0.90 (0.68, 1.19)		0.91

$+ p \leq 0.1$, $* p \leq 0.05$, $** p \leq 0.01$, $*** p \leq 0.001$, $**** p \leq 0.0001$

ADL= Activities of daily living.
The mean incidence risk ratio is the geometric mean of the incidence risk ratios. The table is sorted in order of average IRR.

Table 4: Background characteristics of community college students with risk factor X ("With X") and without the characteristic ("Without X"), before and after matching. Examples of characteristic X could be being a sexual minority or having ever been kicked out of their parents' house. P gives the significance of the test for difference of the two groups. All characteristics are dichotomous and the cell entries are the percent of those with X who satisfy the characteristic (e.g., male, black, have a mother who graduated from high school). In the actual paper, this table could have 50–100 lines. Differences between groups can also be displayed visually in a "Love Plot" pioneered by Tom Love at Case Western (17).

	Before matching			After matching		
	With X	Without X	P	With X	Without X	P
Male						
Black						
Asian						
White						
Mother high school graduate						
Mother college graduate						
Father high school graduate						
Father college graduate						
....						

References

[1] Laura Horn. On track to complete? a taxonomy of beginning community college students and their outcomes 3 years after enrolling: 2003-04 through 2006. Technical report, National Center for Educational Statistics, July 2009.

[2] Stephen Provasnik and Michael Planty. Community colleges special supplement to the condition of education 2008. Technical Report NCES 2008-033, National Center for Educational Statistics, 2008.

[3] Elliott Currie. *Road to Whatever: Middle-Class Culture and the Crisis of Adolescence*. Howard Holt, New York, 2004.

[4] Vince Tinto. Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45:89–125, 1975.

[5] D Ho, K Imai, G King, and E Stuart. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, 15(3):199–226, 2007.

[6] Donald B Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688–701, 1974.

[7] R Lalonde. Evaluating the econometric evaluations of training programs. *American Economic Review*, 76(4):605–620, 1986.

[8] HR Dehejia and S Wahba. Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association*, 94(448):1053–1062, 1999.

[9] SL Morgan and C Winship. *Counterfactuals and causal inference: Methods and principles for social research*. Cambridge University Press, New York, 2007.

[10] Andrew Gelman and Jennifer Hill. *Data Analysis using regression and multilevel/hierarchical models*. Cambridge University Press, New York, 2007.

[11] Donald B Rubin. *Matched Sampling for Causal Effects*. Cambridge University Press, New York, 2006.

[12] Matthew Blackwell, Stefano Iacus, Gary King, and Giuseppe Porro. cem: Coarsened exact matching in stata. *Stata Journal*, 9(4):524–46, 2009.

[13] Stefano M. Iacus, Gary King, and Giuseppe Porro. Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, Forthcoming, 2011.

[14] Stefano M. Iacus, Gary King, and Giuseppe Porro. Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association*, Forthcoming, 2011.

[15] Paul R Rosenbaum. *Observational Studies*. Springer-Verlag, New York, 2 edition, 2002.

[16] Andrew Gelman and Hal Stern. The difference between "significant" and "not significant" is not itself statistically significant. *American Statistician*, 60(4):328–331, 2006.

[17] Thomas E Love. Displaying covariate balance after adjustment for selection bias. In *Section on Health Policy Statistics*, New York, August 11 2002. Joint Statistical Meetings.