2013

Propensity Score Analysis of an Honors Program's Contribution to Students' Retention and Graduation Outcomes

Robert R. Keller
Colorado State University

Michael G. Lacy
Colorado State University

Propensity Score Analysis of an Honors Program's Contribution to Students' Retention and Graduation Outcomes

Robert R. Keller and Michael G. Lacy

Colorado State University

INTRODUCTION

Honors directors and deans know or presume that retention and graduation rates of honors students substantially exceed those of non-honors students. In our research, we have attempted to better determine what portion of this success is attributable to the academic and other benefits of honors programs as opposed to the background characteristics of the students. Among the former, we would point to innovative and small classes, more individual attention for honors students from faculty and staff, residential learning communities, thesis experiences, and extra-curricular opportunities, all of which might be expected to make the college experience more engaging for honors students and thereby contribute to their success. Among the background characteristics, the superior academic achievement and ability enjoyed by honors students is a primary factor that determines retention and graduation (Cosgrove; Shushok; Slavin et al.), and other influences such as gender, in-state or out-of state residency, and family educational background are linked to both academic success and honors program participation. To better estimate the unique contribution of an honors program to retention and graduation outcomes, we have applied propensity score analysis (Guo and Fraser) to separate the effects of honors students' academic achievement and other background characteristics from the consequences of program participation.

SUMMARY OF RELATED RESEARCH

Past work on honors students' persistence published in the *Journal of the National Collegiate Honors Council* by Cosgrove, Shushok, and Slavin et al. has attempted to adjust for the differential academic background of honors students with regression analyses and by constructing pools of non-honors students whose past academic achievements and academic potential resemble those same characteristics in honors students.

Slavin et al. identified a group of high-achieving non-honors students that approximate the academic profile of honors students at the University of Maine, using a constructed group of non-honors students with SAT scores and class rank comparable to honors students. They also used logistic regression to compare non-honors and honors students in order ". . . to examine the relationship between honors participation and retention/graduation rates, statistically controlling for SAT scores and high school rank" (64). Using logistic regression analysis, they found that one-year retention rates were significantly higher for honors students than for non-honors students. For example, 94% of the 2006 honors class returned as compared to 85% of non-honors students (64). They also reported a corresponding odds ratio of 3.1, meaning that an honors student was 3.1 times more likely to return after one year than a non-honors student (an explanation of the "odds ratio" appears below). They found a four-year graduation rate, adjusted for high school rank, of 64% for honors students and 60% for non-honors students for the entering class of 2002, a difference that is not statistically significant (67). They speculated, however, that, in the context of the institution they studied, a trend toward development of a community identity among honors students might lead over time to higher four-year graduation rates for honors students.

Shushok selected a non-honors student group of the same size as the honors student group and with comparable characteristics at the University of Maryland, College Park. Shushok used "caliper matching" to define the acceptable level of comparability and was able to identify 86 comparable students in each of the non-honors and honors groups. Using this technique, he deemed two groups as comparable if their mean scores (e.g., GPA and SAT) were within 0.15 of a standard deviation of one another (87). With these adjustments, honors and non-honors groups had "identical" mean GPAs, mean SATs, percentages of females, percentages living on campus, and percentages White, Black, Asian, and Hispanic. Shushok determined that honors students had significantly higher cumulative GPAs at the end of the first year (honors 3.41 and non-honors 3.18) and higher one-year retention rates (honors 97% and non-honors 90%). However, both of these differences tended to disappear by the fourth year (93). He did not report four-year graduation rate comparisons for honors and non-honors students, but Shushok made an important contribution by examining some of the benefits of program participation. He also used a survey revealing that honors students, especially males, were more likely than non-honors peers to have met with faculty during office hours (odds ratio 2.5), to have discussed career plans and vocational aspirations with faculty (odds ratio 3.6), and to have participated in activities with an academic component (odds ratio 4.7) (92).

Cosgrove added another important dimension to assess the benefits of honors programs/colleges by ". . . differentiating the honors experiences of students who complete all their honors requirements from those who do not" (46). For example, if entering honors freshmen drop out of the honors program after the first year but nevertheless persist until graduation, the honors graduation rate could be biased upward if the success of these "partial honors" students is not separated out but instead implicitly attributed to honors students who completed the program. To clarify this issue, Cosgrove presented retention and graduation data for three groups: students who completed the honors program, students who partially completed the honors program, and students who did not participate in honors at all. He concluded that, while the five-year graduation rate for program completers was (virtually by definition) 100%, the rate was 82% for partial honors students and 76% for non-honors students (61). The cumulative GPAs for the three groups were 3.71, 3.48, and 3.36. Cosgrove's study provides a significant comparison of GPAs among the three groups, but the exclusion of students who dropped out of the honors program virtually assured that the graduation rate for honors completers was 100% since a student must have remained in school to complete the honors program.

DATA AND METHODS

In the current study, we report data from students who initially entered Colorado State University (CSU) between fall 2005 and fall 2008, with outcomes tracked through summer 2012. All were recent graduates from high schools, most being freshmen but quite a few having sophomore standing because of college credits earned prior to matriculation. Although previous studies have examined longitudinal outcomes by year, our research pooled student retention and graduation data from 2005–8. As outcomes, we examined whether a student returned to Colorado State for the second fall semester and whether s/he graduated from CSU within a four-, five-, or six-year period from initial entry. In analyzing graduation rates, we included only students for whom the requisite number of years had passed so that, for example, six-year graduation was analyzed using only the 2005 and 2006 cohorts, with more cohorts available for analyzing four-year graduation. Graduation records reflect Colorado State University only and do not include students who transferred and graduated elsewhere. We classified honors program participation as simply present or absent and included as honors students all those who joined the program in their first year at CSU, whether or not they remained in it.

The current study statistically adjusted for a wider range of (confounding) background characteristics than has been typical in previous work. In addition to academic achievement in high school, we included information about ethnic

status, gender, in-state/out-of-state origin, first-generation college attendance, and academic unit at entry. Academic achievement was measured using the State of Colorado's college admission index, which is a composite encompassing high school grades or class rank and achievement test scores (ACT or SAT). The inclusion of high school performance as well as test scores has the advantage for our analysis of at least partly and implicitly including a measure of students' motivation and perseverance. Ethnic status was simplified to White-Anglo vs. any other status, with international students excluded from analysis. Academic unit at entry was defined as the college within the university (Agriculture, Engineering, Natural Sciences, Liberal Arts, Undeclared, etc.) listed for the student during the first semester. Each of these variables is related to honors participation (see Table 1) and has in previous institutional research at CSU (not reported here) been related to various graduation and retention outcomes.

To estimate the effect of honors program participation on the various outcomes, we chose "propensity score analysis," an analytic approach that has become increasingly popular for non-experimental (observational) studies of "treatment programs." We describe the basic idea of propensity score analysis here and recommend to interested readers a more detailed recent source, Guo and Fraser's *Propensity Score Analysis: Statistical Methods and Applications (Advanced Quantitative Techniques in the Social Sciences)*. Propensity score analysis rests on a "counterfactual" perspective: namely, that in examining the effects of a treatment such as honors program participation, we would in principle want to compare the outcomes of treated individuals to the outcomes they would have experienced had they *counter-to-fact* not participated in the honors program. While we can observe their outcomes under their actual "treatment" condition (in the honors program or not), we can only impute what would have happened had they not been "treated," based on outcomes experienced by similar persons who were not in the program. To impute the counterfactual outcome, the propensity scoring approach involves first applying a regression model to the full sample of participants, i.e., both honors and non-honors students in this case. This analysis is used to estimate, as a function of the various background factors (gender, admission index, etc), the probability that a student would have been in the program (honors). This probability or "propensity score" is the basis on which nonparticipants are judged to be similar to a participant in a given treatment program (honors).

Thus, for each program participant, one or more nonparticipants are chosen who are similar (matched) in propensity score and can therefore serve as comparable controls. The average outcome among these nonparticipant controls are used to impute the outcome that would have been expected for their matched treatment subject, had s/he counterfactually *not* received

treatment. The mean of the differences between each treatment subject and her/his controls serves as the measure of program effect and is commonly called the "average treatment effect among the treated" as it attempts to estimate only what effect the treatment (honors) had on the treated subjects (Guo and Fraser 46–47). As compared to more traditional statistical approaches (e.g., a regression model for the outcome, with the background variables as covariates), propensity scoring arguably reduces bias in estimating the effect of a treatment program.

In the current application, with a binary outcome, propensity scoring analysis also has the advantage of permitting a single estimate of effect that is expressed as a percentage difference, something often more intuitively understandable than odds ratios associated with logistic regression models for binary outcomes. As an illustration to help readers in translating between percentage comparisons and corresponding odds ratios: If hypothetically 90% of honors students and 75% of non-honors students graduated, the odds for graduating among honors students would be 90/(100 - 90) = 9.0 while the odds for non-honors students would be found as 75/(100 - 75) = 3.0, giving an odds ratio of 9/3 = 3.0. Thus, an honors student would have been three times more likely to graduate than a non-honors student with "likely" judged on the scale of odds. For many readers, odds ratios can seem large in relation to a comparison based on percentages.

We applied propensity scoring to our data as follows: For all persons with data available on a given outcome, we used the background variables described above to estimate a logistic regression model for whether or not a student would have been in the honors program, yielding for each person a propensity score (between 0 and 1) describing the similarity of her/his background profile to an honors program participant. For each honors program participant, the five non-honors students closest in propensity score were chosen as controls. Eligible controls were restricted to non-honors students having a propensity score falling within a so-called "caliper" distance of 0.25 standard deviations of the propensity score of the matched honors program participant; this is a commonly suggested criterion of sufficient closeness of propensity scores of controls to treatment subjects (Guo and Fraser, 147), and the use of five controls is a common compromise between statistical precision and ease of finding matching controls. We used the Stata (Stata Corp) add-on program "psmatch2" (Leuven and Sianesi) to conduct all aspects of this analysis.

HONORS AT COLORADO STATE UNIVERSITY

Colorado State University is a land-grant institution and Carnegie Research University with approximately 26,000 resident students, of whom approximately 22,500 are undergraduates. The University Honors Program began in

1957 with several honors seminars and approximately fifteen students. Today the program has nearly 1,600 honors students, including 70% women, 10% ethnic minorities, and 65% from within Colorado. The university-wide honors option consists of four interdisciplinary honors seminars that fulfill four general education requirements, two honors courses in the major, and a senior honors thesis; departmental honors programs require honors courses in the major and a senior honors thesis. Other features include an honors learning-community residence for freshman students that includes the honors office and space for honors faculty, a first-semester honors seminar with an orientation component, and rigorous supplemental advising. Other honors opportunities include an Honors Undergraduate Research Scholars program, ample opportunities for extracurricular activities through the Honors Student Association, and a special honors scholarship of $1,000 per year; student surveys and anecdotal evidence suggest that honors program students highly value these features of the program.

RESULTS

Table 1 displays summary information on program participation and background variables for the combined sample of all cohorts used in the analysis. While these figures differed slightly in absolute terms across cohorts, in the interest of brevity we present only this summary.

As might be anticipated, given program admission restrictions, honors participants had substantially higher admission index scores than others, with honors program participants having a mean about 20 points (2 standard deviations) higher than other students. The mean admission index of 132 for honors students corresponds to a 3.9 GPA or 8th percentile class rank and a combined critical reading and mathematics score of 1340 on the SAT or 30 on the ACT. The mean admission index of 111 among non-honors students corresponds to a 3.5 GPA or 25th percentile class rank and an 1120 SAT or 25 ACT.

Considering demographic factors, we found that women were more than twice as likely as men to participate in honors and that students from outside Colorado were similarly much more likely to participate. From a different perspective, these numbers indicate that the proportion of women in honors is much higher than the proportion of women in the university (70% versus 54%) and that non-residents are a significantly higher proportion of honors students than the student body as a whole. Substantial differences across academic units also prevailed. For example, persons entering the College of Veterinary Medicine/Biomedical Sciences participated at more than three times the rate of students in general while undeclared students were represented at only about one-fourth the average percentage.

Table 1: Honors Program Participation and Background Variables, Fall 2005–Fall 2008 Cohorts (N = 15,821)

Categories	Percent Honors Program Participants within Category
Total Sample	6.8
Females/Males	8.8/4.3
White-Anglo/Minority	7.2/5.0
In-state resident/non-resident	6.0/10.2
First generation students/others	4.4/7.7
College major at entry	
Agriculture	10.0
Applied Human Sciences	3.6
Business	5.9
Engineering	11.3
Undeclared	1.5
Liberal Arts	7.3
Natural Sciences	11.3
Vet. Medicine/Biomedical Sciences	22.4
Natural Resources	10.1
Admission Index, Mean (SD) of Honors Participants/ Others	132.1(4.7)/ 111.4(9.9)

Clearly, these and the other background factors show sharp differences between honors and non-honors students that would be expected to significantly affect honor students' retention and graduation rates regardless of any program effects. Considering those differences, the results in Table 2 are not surprising.

Table 2 displays naive (unadjusted) comparison of retention and graduation rates, which show that 93% of honors students returned for the second fall as compared to 82 percent of all other students. Four-, five-, and six-year graduation percentages were respectively about 29, 24, and 27 points higher among honor students than others, translating to odds ratios in the range of about 3 to 5. These differences reflect both the effects of the background characteristics typical of honors students and whatever effects participation in the program might have had.

Table 2: Unadjusted Outcome Comparisons, Honors Participants vs. Others

	Returned for Second Fall	Graduated in Four Years	Graduated in Five Years	Graduated in Six Years
Honors participants	92.9%	64.2%	81.9%	88.6%
All other students	81.8%	35.2%	57.6%	61.9%
Difference	11.1%	29.0%	24.3%	26.7%
Odds Ratio for Outcome, Honors vs. Others.	2.9	3.3	3.3	4.8
N for honors/ all others	2,071/ 26,115	1,081/ 14,740	796/ 10,946	466/ 7,229

The first step of our analysis was to obtain the propensity scores used to adjust for these background variables. Table 3 displays the results of the logistic regression analysis used to obtain propensity scores for the cohort of students available for analysis of four-year graduation. (We include this table for illustration but, in the interest of brevity, omit it for other outcome periods/ cohorts.)

As might be expected from the bivariate results of Table 1, the results in Table 3 show that admission index, first-generation status, residency, and initial college major had strong relationships with honors participation. When adjustment was made for these variables, however, ethnicity was not strongly connected to honors participation. For the reader unfamiliar with logistic regression results, we should note that, in Table 3, positive coefficients indicate that increases in variables are associated with being more likely to be an honors student while negative ones indicate the reverse.

Table 4 displays the final results, after using the propensity scores to find matched controls. It shows that, as compared to what would have been expected had honors students not participated in the program, there was a gain of 5 percentage points in second-year returnees and increases in four-, five-, and six-year graduation estimated at 8.4 percentage points, 12.3 percentage points, and 14 percentage points respectively. (For comparability to previous work using logistic regression, note that these translate to odds ratios of 1.8, 1.4, 2.0, and 2.7.)

Table 3: Logistic Regression Coefficients for Honors Program Participation as a Function of Background Variables

	b(se)
College major at entry[1]	
Agriculture	0.855 (0.237)**[2]
Applied Human Sciences	0.596 (0.220)**
Business	0.423 (0.217)
Engineering	0.078 (0.205)
Liberal Arts	0.734 (0.193)**
Natural Sciences	0.960 (0.185)**
Vet. Medicine/Biomedical Sciences	1.289 (0.207)**
Natural Resources	1.070 (0.297)**
Female gender	0.509 (0.105)**
Colorado resident	-0.377 (0.105)**
White-Anglo ethnicity	-0.112 (0.143)
Admission index	0.347 (0.0099)**
First generation status	-0.231 (0.114)*
Constant	-46.359 (1.296)**
N	15,821
Likelihood Ratio Chi-Squared, $df = 13$	4615**
McKelvey-Zavoina R^2	0.83

[1]Reference category is Undeclared (intra-university) college major.
[2] * $p<0.05$; ** $p<0.01$

All of the effect estimates in Table 4 were statistically significant at conventional levels. Comparing them to the unadjusted figures of Table 2, one sees that, of the total (unadjusted) difference of 11.1% in second-year return of honors students versus others, about half can be attributed to the program itself. This interpretation observation rests on presuming that the unadjusted difference of 11.1% reflects the combined effects of participation in the program and any gains reflecting the background variables (gender, academic ability, etc.)

Table 4: Propensity Score Adjusted Comparison of Outcomes among Program Participants vs. Controls

	Returned Second Fall	Four Year Graduation	Five Year Graduation	Six Year Graduation
Honors	92.9%	64.2%	81.9%	88.9%
Controls	87.9%	55.8%	69.6%	74.9%
Difference (se)[1]	5.0% (1.7)**	8.4% (3.1)**	12.3% (3.1%)*	14.0% (3.1)**[2]
Total N in cohort	28,186	15,821	11,742	7,588[3]
N of Honors Students Analyzed/ Expected Increase in Numbers with Outcome	2,071/ 104	1,081/ 91	796/ 98	459/ 64[3]
Odds Ratio, Honors vs. Non-Honors	1.80	1.42	1.98	2.69

[1] Average treatment effect among the treated, with standard error in parentheses.

[2] * $p<0.05$; ** $p<0.01$

[3] Of 466 honors participants, 7 were excluded because suitable matched controls could not be found.

that characterized honors students. The matched comparison serves to remove the latter effect so that the 5.0 percentage point difference would estimate the gain attributable to the program alone out of the total 11.1% difference. Similarly, the portion of the difference attributable to the CSU honors program itself is one-fourth the size of the raw four-year graduation percentage difference and about half of that for five-year and six-year graduation.

Presuming that these adjusted differences are valid, we can apply them to the total number of honors students in each cohort to estimate what they imply regarding actual numbers of additional students retained or graduated. We find that the 5.0 percentage point increase in second-year retention resulted in 104 more honors students out of 2,071 returning for their second year at CSU while the corresponding figures for graduation were an additional 91 graduates for an honors cohort of 1,081 (four-year graduation), 98 out of a cohort of 796 (five-year), and 64 out of a cohort of 459 (six-year).

CONCLUDING REMARKS

The preceding study conducted at Colorado State University has shown that participation in the honors program was associated with meaningful increases in the proportion of these students who returned for their second year at the university and in the proportion of them who graduated within a four-, five-, or six-year period. These estimates come from a comparison of outcomes among honors students to outcomes among individually matched controls, students who were similar but did not participate in the CSU honors program. Measured in percentage-point gains, these increases in success among honors participants were larger for outcomes of longer duration, even when considered relative to the base rate of success among their matched peers.

The particular method for this matched comparison is known as propensity score analysis and offers relatively rigorous adjustment for achievement factors (admission index) and non-academic characteristics such as gender and non-residency in analyzing the contribution of an honors program to the retention and graduation rates of its students. Adjusting for the influence of such background factors gave results indicating that the effects of the program itself were much more modest than indicated by the raw comparison, but those effects were still found to be relatively large and statistically significant. For example, Table 4 indicates that the rate of five-year graduation for honors students was 81.9% versus the 69.6% that would have been expected among these students had they not participated in the program; this shows a much smaller gain for honors participation difference than would be implied by a comparison to the unadjusted honors graduation percentage of 57.6% (Table 2) and demonstrates the potential importance of an adjusted comparison in describing retention and graduation outcomes among honors participants. The adjusted comparisons of effects on five and six-year graduation, however, show much larger differences than on four-year graduation. Presuming that other universities' experiences would resemble those of CSU, honors administrators who want to demonstrate a positive effect of program participation will find comparison of longer-term outcomes more useful.

The adjusted differences in retention and graduation rates between honors students and comparable non-honors students may be ascribed to benefits offered by the honors program such as innovative and small classes, residential learning communities, and extracurricular activities that engage these bright students. However, our work does not indicate which parts of an honors program produce these improved outcomes, and future researchers might address this issue by combining a focus on our outcomes with more processual factors along the lines initiated by Shushok.

REFERENCES

Cosgrove, John. 2004. "The Impact of Honors Programs on Undergraduate Academic Performance, Retention, and Graduation." *Journal of the National Collegiate Honors Council,* 5(2), 45–53.

Guo, Shenyang and Mark W. Fraser. 2010. *Propensity Score Analysis: Statistical Methods and Application*. Los Angeles: Sage Publications.

Leuven, Edwin and Barbara Sianesi. 2003. "PSMATCH2: Stata module to perform full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing." Version 4.0.5. <http://ideas.repec.org/c/boc/bocode/s432001.html>.

Shushok, Frank. 2006. "Student Outcomes and Honors Programs: A longitudinal Study of 172 Honors Students, 2000–2004." *Journal of the National Collegiate Honors Council,* 7(2), 82–96.

Slavin, Charlie, Theodore Coladarci, and Phillip Pratt. 2008. "Is Student Participation in an Honors Program Related to Retention and Graduation Rates?" *Journal of the National College Honors Council*, 9(2), 59–69.

StataCorp. 2013. *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.

****

The authors may be contacted at

Michael.Lacy@colostate.edu.