The Effects of College Selectivity and Major Choice on Labor Market Outcomes

Valerie Bostwick
Department of Economics
University of California, Santa Barbara
June 14, 2012
Introduction

It has been well-documented in the economics literature that college graduates who major in natural sciences and engineering earn substantially more than humanities majors in the labor market. Research has further shown that while there is a significant amount of ability sorting across majors, those differences in ability only account for a small part of the earnings gap between majors. James et al. (1989) even go so far as to claim that, "while sending your child to Harvard appears to be a good investment, sending him to your local state university to major in Engineering, to take lots of math, and preferably to attain a high GPA, is an even better private investment."

These striking results on the returns to college major choice imply that some majors must provide human capital that is simply more valuable than that provided by other majors. However, data on recent graduates from Ivy League universities seem to point towards a very different result. In 2007 alone, 75% of Princeton University’s graduating class with full-time jobs were in the financial services or consulting sector. Comparably, 47% of Harvard’s 2007 graduating class went into these same industries.\footnote{Rampell, Catherine. (Dec. 21, 2011) Out of Harvard, and Into Finance. The New York Times. Retrieved June 13, 2012, from http://economixblogs.nytimes.com/2011/12/21/out-of-harvard-and-into-finance/} Presumably, these graduating classes encompass a wide variety of major choices. Yet these students are still achieving largely identical labor market outcomes.

This apparent contradiction begs the question: are the returns to major choice decreasing in college selectivity? This question is interesting even from a purely
descriptive point of view. While a significant amount of progress has been made in the economic literature on quantifying the returns to different majors and the returns to college selectivity, no one has yet considered college quality as a dimension along which major choices may vary. Documenting the changing patterns in the distribution of major choices and the distribution of ability across majors as college selectivity increases will be enlightening in and of itself. Given these patterns, I will then analyze the relationship between the returns to major choice and college selectivity level.

Furthermore, I hope to investigate potential causes for a decreasing relationship between these two variables. It may be the case that previous economic studies were merely using ineffective measures of students’ true ability levels. If we assume that students at top universities are more homogeneous in ability than students at other schools, we would then expect to see decreasing returns to major choice as selectivity increases. On the other hand, it may be the case that the returns to major choice reflect sorting across majors based on unobservable characteristics, such as work ethic or ambition. Then if we again assume that students at top schools are more homogeneous, we would once more expect to see the pattern of decreasing returns to college major choice. A third explanation is that teaching quality or the knowledge and skills imparted are more uniformly distributed across majors at top colleges.

Finally, it may be the case that there is some upper bound on the return to a college education. Suppose, for example, that above some set amount of school-based knowledge, the only way to increase labor market productivity is through on-the-job experience. Then a decreasing return to major choice could be evidence that students
at top schools are over-investing in their education.

**Relevant Literature**

The literature related to the returns to college selectivity and major choice is fairly extensive. One of the first papers to address this issue is Wise (1975). Wise asks whether academic achievement is a good predictor of job performance (as measured by salary). He models salary as a function of college degree, major, GPA, selectivity, and individual characteristics. Wise finds that both college selectivity and major choice have significant effects on labor market outcomes. Unfortunately, these results are limited by the fact that he does not control for ability sorting across both college and major.

Following this initial attempt, James et al. (1989) examine a model of the returns to selectivity and major that accounts for ability sorting. The authors model annual earnings as a function of college characteristics (including selectivity, control, and expenditure-per-student), college experience variables (GPA, major, math credits completed), and individual characteristics (SAT score, race, parental income, etc.). The authors find that there is a positive and significant effect of college selectivity on earnings, but that this effect is dwarfed by the much larger returns to major, GPA, and math credits.

A long line of research has since bolstered these early results, showing positive, significant labor market returns to both major choice and college selectivity (Grogger & Eide 1995, Loury & Garman 1995, Brewer, Eide & Ehrenberg 1999, Turner & Bowen 1999). However, in more recent work researchers have begun to question the
accuracy of these findings due to the potential for selection bias.

Black & Smith (2004) seek to address the main econometric difficulty in measuring the effects of college selectivity, which is that better students sort into better schools. The authors point out that the commonly used linearity assumption can hide the failure of the "common support" condition. That is, the outcomes of students who have a low likelihood of attending a high-quality college are used to identify the counterfactual outcomes for students who attend high-quality schools. For this reason, Black & Smith use propensity score matching methods to estimate the returns to college quality (as measured by average faculty salary, average freshman SAT score, and retention rate).

This method yields a positive estimate of the effect of college quality on wages, but with very large standard errors. This is likely due to their small sample size and the fact that very few individuals who attend low-quality schools have the characteristics that would give a high probability of attending a high-quality school. Unfortunately, this shortcoming would be very hard to address in any data set, making this method largely unfeasible.

Dale & Krueger (2002) also attempt to address the selection bias in measuring the effects of college quality, by controlling for both selection on observables and selection on unobservables. Dale & Krueger note that college admissions are based in part on student characteristics that are observable, such as SAT score and high school GPA, and in part on characteristics that are unobservable, such as perceived ambition, maturity, and dedication. If these unobservable characteristics are also correlated with labor market outcomes, then the standard OLS estimates that only
control for observables will over-state the returns to college quality.

Dale & Krueger attempt to control for this selection by matching each student to other individuals who applied to, and were accepted or rejected by, the same set of colleges. By including an indicator variable for each group of matched students in the returns to college quality regression, the authors hope to control for the unobservable characteristics common to each group. This method yields positive, but insignificant estimates of the returns to college selectivity. The main drawback to this technique is that it assumes that students make their enrollment decisions randomly. For this method to work, we must believe that the students who choose to attend lower-quality schools are not systematically different from their peers who received the same admissions decisions but chose to attend a higher-quality school.

Arcidiacono (2004) addresses a different aspect of the selection bias issue. He uses a structural model to investigate the source of observed ability sorting across majors. He examines whether the observed returns to major choice are driven by ability sorting or whether the causal relationship is reversed. Arcidiacono notes that high-ability students may self-select into majors that they know will yield higher wages.

Arcidiacono develops a dynamic model of college and major choice over three time periods. In the first period, individuals choose to either take a job offer or go to college. If they choose college, they must also select a major in the first period. In the second period, the student receives an information shock revealing their college ability. They then have the option to dropout, change schools, or change majors. Finally, in the third period every agent works and receives earnings based
on their educational choices. Arcidiacono calibrates this model and runs simulations to address the question of causality between returns to major choice and ability sorting. As in the previous literature, he finds a large return to major choice and smaller returns to college selectivity. The simulations reveal that ability sorting across majors is largely driven by preferences, rather than by the monetary returns to certain fields.

The existing literature has addressed, in depth, the effects of both college selectivity and major choice on labor market outcomes. Many of these papers also draw comparisons between the two effects and attempt to quantify their relative importance. My proposed research will fill an important gap in this literature by examining the interaction between college selectivity and major choice. This project will open up a new direction in the literature by considering college selectivity as a dimension along which major choices and their relative outcomes may vary.

Data

The data for this research will come from the National Center for Education Statistics’ Baccalaureate & Beyond Longitudinal Study. This study surveys a sample of students who completed bachelor’s degrees in 1993, 2001, and 2008. The survey cohorts include approximately 11,000, 10,000, and 19,000 students respectively. The data contain extensive information on the undergraduate experience, which includes the institution attended, GPA, major, and a number of financial aid variables. The survey focuses on post-graduation work experience, and therefore collects a variety of labor market outcome measures. Finally, the data also include the necessary
demographic background characteristics, the most important of which is SAT score. My primary concern with this data is whether it will prove to contain a wide enough range of college selectivity levels. Ideally, I hope to have a sizable subsample of students who attended colleges in the top 10% of the selectivity range.

**Estimation Strategy**

My preliminary estimation strategy will rely on a difference-in-differences approach. To keep the initial analysis simple, I will limit the comparison to only two majors (e.g. English and Economics) and two levels of selectivity (high and low). I will first estimate

\[
W_{ic} = \beta_0 + X_{ic}'\beta_1 + \beta_2 M_{1c} + \beta_3 S_{1c} + \beta_4 M_{1c} * S_{1c} + u_{ic},
\]

where \(W_{ic}\) are wages for individual \(i\) attending college \(c\), \(X_{ic}\) are individual characteristics (including ability measures), \(M_{1c}\) is an indicator for majoring in economics, and \(S_{1c}\) is an indicator that college \(c\) is highly selective. The coefficient \(\beta_4\) would then reveal the effect of college selectivity on the return to majoring in economics.

To extend the model to many majors and levels of selectivity, I will create a discrete variable measuring selectivity level and estimate

\[
W_{imc} = X_{imc}'\beta + \sum_{m=2}^{M} \delta_m D_m + \gamma S_c + \sum_{m=2}^{M} \theta_m D_m S_c + u_{imc}.
\]

Here \(D_m\) are indicators for each major and \(S_c\) is the selectivity level of college \(c\). \(\theta_m\) then reflects the effect of attending a college of one selectivity level higher on the
relative return to each major.

The main econometric challenge will be in accounting for the ability sorting across majors and selectivity levels. As discussed in Black & Smith (2004) and Dale & Krueger (2002), there may be selection bias issues that are not adequately addressed by simply controlling for ability as a covariate in my regressions. Moving forward, I will have to determine the extent of this potential bias and decide how to best address the selection issue.
References


