Heterogeneity and Higher Education

Jeffrey Smith
Department of Economics
University of Michigan
econjeff@umich.edu

Version of August 14, 2007

Draft! Please do not cite without the author’s permission

---

1 I am grateful for discussions about higher education with John Bound, Paul Courant, Mike Lovenheim, Lock Reynolds and Sarah Turner that have shaped my thinking on the issue. Sarah Turner also provided detailed comments on an earlier draft and pointed me to the speech by Bill Bowen that I quote in the conclusion. Sadly, though inevitably, I retain all of the blame for both my outrageous opinions and for any errors.
Introduction

The higher education sector embodies a remarkable diversity of students, colleges and programs within colleges. This essay makes the case that this heterogeneity matters for how we think about success in higher education for individuals and colleges, for what research questions we ask and for how we go about answering them.

This paper uses the terms diversity and heterogeneity as synonyms to refer to the wide variation among students, colleges and programs within colleges. My conception of diversity at the student level goes well beyond its frequent use as shorthand for the representation of certain racial and ethnic groups. Instead, I have in mind diversity on many dimensions, including race and ethnicity but also ability and pre-college achievement in math, reading, writing, the arts and athletics, effort level, study skills, ambitions for life, religious background and so on. All these characteristics affect the optimal match between student and college and student and program. They also affect the nature and likelihood of success in college and the extent and nature of any effect of college on later life.

Colleges also exhibit extraordinary heterogeneity. They vary by the size of the student body, by expenditures per student (and available resources more broadly), by the quality of their faculty (on dimensions such as teaching, research and service), by the extent of the intramural and inter-collegiate athletic programs they offer, by location (contrast Columbia and Berea), by the amount of hand-holding they provide to students (or to helicopter parents!), by the student atmosphere they provide (contrast Reed and Hillsdale) and so on. Some colleges provide two-year degrees and others four-year degrees. Some specialize in the great books and others in engineering or science, like
Purdue and Cal Tech. While some college characteristics, such as selectivity, fit reasonably well on a simple numerical scale, no single scale or ranking can hope to capture the relevant information on all or even most of these dimensions.

Finally, programs within colleges and universities vary widely in availability, type and quality. Majors often vary widely in terms of number of students, faculty size and quality, and the type and amount of work required for successful completion. Some programs feature work opportunities as part of their curriculum. Others focus on writing and so on.

This amazing diversity among students, colleges and programs within colleges has implications for how we think about the outcomes associated with college attendance and completion and for how we think about related policy issues. To see this, consider a couple of quick examples.

First, suppose that a particular college decides to adopt an income-contingent student loan program. Some students will now apply this college who before did not apply to any college. These students will be differentially risk averse and differentially interested in majors that lead to jobs with low earnings, which are effectively subsidized under such schemes. Some existing students at the college will migrate toward majors that imply lower earnings as well. The students who migrate will generally be the ones with the lowest expected pecuniary and non-pecuniary benefits from the majors they leave (the marginal students, as economists like to say). Similarly, some particularly risk-averse current students will migrate toward majors that lead to jobs with a higher variance in earnings, which become more attractive due to the implicit insurance provided by the scheme.
Second, suppose that a college expands the number of majors graduating each year in its computer science program from 30 to 100 by lowering entry requirements for the major. The marginal students will likely come from other majors with weaker entry requirements but still with a lot of technical content, such as business, economics or applied math, and not from English literature or general studies. Classes in computer science will get bigger, and the material in them will be aimed lower. As a result, the 30 students who would have gotten in under the old regime will now likely learn less while the 70 students on the margin will have different, and likely more lucrative, career tracks.

In both examples, policy changes introduced in the context of heterogeneous students, colleges and programs leads to very specific changes in the behavior of certain decidedly non-representative groups of students. They also lead to different outcomes even for students who do not change their behavior. As a result, such policy changes require careful analysis to trace out who wins, who loses, and what the net result is for students and for the institutions involved. The following sections amplify this argument and apply it in a variety of higher education contexts.

**Students, Colleges, Programs and Outcomes**

We can think about college success in a number of ways: Does the student finish a degree? Does the student do well in the labor market relative to what he or she would have experienced without going to college? Does the student have better health outcomes as a result of attending college? Does the student contribute more to society in other ways, such as civic participation\(^2\), as a result of having attended college?

---

\(^2\) See the paper by Thomas Dee in this volume.
In each case, we can think about students having a variety of possible outcomes that depend on their own characteristics (measured and unmeasured), the particular college they attend and its characteristics (and the characteristics of its student body as a whole), and the particular program or programs they undertake. At the broadest level, each choice of school and program implies a different distribution of possible outcomes for each student and no single college represents the best match for all or even most students.

Economists and other social scientists know surprisingly little about the nature of the relationships that govern student outcomes. Most of what we do know represents mean effects for all students or for subgroups of students defined by a single variable, such as males and females. For example, we know for both men and women that starting at a two-year college rather than a four year college increases the probability of getting a two-year degree (no surprise there) but decreases the probability of getting a four-year degree, even among students who say they intend to do so before starting. This aggregate result, however, hides a wealth of variation relevant to both policy and to individual decision-making. For some students, the ability to live at home while attending a community college may improve school outcomes while for others it limits their ability to focus on their studies. Some two-year schools surely do a much better job of greasing the track to a four-year degree than others. Some students may learn in a two-year school that they really like particular programs offered at that level, and so on. Keeping heterogeneity in students, colleges and programs in mind in interpreting the aggregate result leads to more nuanced thoughts about how to react in terms of

---

institutional and governmental policies and also leads immediately to ideas for useful additional research to guide those policy (and individual) choices.

**Matching students to colleges (or no college)**

Economists sometimes talk about a mythical social planner with complete information and infinite processing capacity; in the land of economic theory, this planner makes optimal choices about resource allocation given the available inputs and known relationships between student, program and college characteristics and later life outcomes such as earnings and civic participation. Given the (potential or actual, depending on how one reads the literature) role of peer effects, which cause the best choice for each student to depend on the choices of all of the other students, this represents a daunting task indeed. Access to deeper knowledge concerning what programs and college characteristics fit best for particular types of students than we currently possess would complicate the allocation problem still more.

The United States, a country whose residents tend not to revere would-be social planners, implicitly takes the opposite approach via a highly decentralized system. Students apply to individual colleges. Each college makes admissions and aid decisions about individual students (though at many public universities most students face a fixed and known and highly subsidized price per credit or per term with little chance of aid). Information enters the system via commercial college guides, high school counselors, paid admissions consultants, parents, siblings, and many other sources. Students from families with more resources, either monetary or informational or both, may well make
better choices and some evidence suggests that students from disadvantaged backgrounds sometimes do not end up very well matched.\textsuperscript{4}

In general, the social science literature provides less systematic information than it should regarding the sources of information that students and colleges rely on in making their choices about where to apply or attend, in the case of the students, or whom to admit and how much financial aid to give them, in the case of the colleges. The literature on college quality does suggest that the various commercial rankings, such as those produced by \textit{Barron’s} and by \textit{U.S. News and World Reports}, provide valuable information, in the sense that these rankings (and the variables that underlie them) predict later labor market outcomes even after controlling for the non-random matching of students to colleges.\textsuperscript{5}

In contrast, I am aware of no formal studies of the value added in match quality by high school guidance counselors and admissions officers at selective colleges. The related literature on the value of caseworkers in assigning clients to particular treatments within the context of active labor market programs for the unemployed does not inspire much confidence. For example, Michael Lechner and I compare the performance of Swiss caseworkers in assigning unemployed individuals to different types of subsidized classroom and on-the-job training and find that they do about as well as random assignment.\textsuperscript{6} In contrast, an automated system that assigns each individual to a training

type based on a statistical model (what economists call a “statistical treatment rule”) and their individual characteristics performs better than random assignment.

It would be a very interesting and worthwhile exercise for one or more selective colleges to choose to admit some small fraction (say 10 percent) of their applicants in the following way: First, define a simple rule involving cutoff values on the SAT and on high school GPA. Second, randomly admit students who exceed the cutoff values to fill the slots available for this purpose. Then compare the academic and labor market outcomes of the two groups. Of course, one could also use more complicated rules in such an automated system; indeed, one way to obtain such a rule would consist of summarizing the behavior of the existing admissions system via a statistical model. In the event that the simple admissions rule does as well (or better) than existing admissions procedures the college gains access to additional instructional resources. In the event that it does not, the college has produced some useful knowledge (and earned the undying gratitude of its admissions officers) at a fairly low cost.

Our understanding of the matching process would also benefit from further systematic qualitative and quantitative research on how counselors and admissions officers do what they do, and on how students combine the information from counselors and other sources to produce choices about application and acceptance. In addition, comparative work on alternative, more centralized application and allocation systems such as those in some European countries would help to illustrate alternatives and would highlight the costs and benefits of the present system.
College as an experience good

Many students know very little about college in general, and even less about their best match among the wide variety of colleges and programs available to them. College life, particularly when it involves living away from home, involves major changes in social networks, responsibility, supervision and academic intensity. Because of all the uncertainty surrounding them, college in general as well as specific colleges and programs represent what economists call “experience goods”. Students cannot really know how well they match to a particular school or program without enrolling and experiencing it, at least for a while.

Thinking about colleges and programs as experience good leads naturally to thinking about how students learn. They begin at one particular college with some idea about programs of interest. Over time, as their studies continue, they learn about how well their initial choices of college and program of study suit them, and likely gain more information about the characteristics and outcomes associated with other colleges and, even more so, other programs of study. At some point, they may acquire enough information to cause the perceived benefits of changing college or changing program to exceed the costs. All else equal, we would expect the amount of such switching to decrease in proportion to the quality and quantity of information underlying the initial choice and to increase in proportion to the number of choices available. Note that switching as used here includes dropping out (i.e. “switching” from college to work). Thinking about college as an experience good suggests an optimal dropout rate different from zero (though perhaps still smaller than the one we observe) and also suggests useful ways to reduce both the dropout and transfer rates.
We know that both switching schools and switching majors has increased in recent decades, and we know that students from backgrounds likely to imply less information about where to go and what to study do more switching as well\(^7\), though some (or most or all) of that may result from financial rather than informational issues. Beyond that, we do not know very much. For example, we do not know how much of the increasing time to degree documented by Sarah Turner elsewhere in this volume results from increases over time in the extent of college and program shopping. Nor do we know how the earlier lower levels of switching or the current higher levels of switching correspond to the optimal level of switching. Nor do we know how policy choices at the university, system, state and federal levels regarding admissions, tuition, loans, grants, program requirements, and other issues affect the amount of switching and its value (or lack thereof) to the student and to society.

**Optimal institutional mix**

Up to this point in the paper I have treated the existing set of colleges, and their sizes and characteristics, as fixed. But, of course, in the longer run government policymakers, along with administrators, alumni and other funders of private colleges (as well as the firms that operate for-profit colleges) can affect the available mix. Thinking about this wider problem in the context of trying to find the best matches of heterogeneous students to heterogeneous colleges and programs raises some interesting questions indeed. The fact of government provision (combined with indirect government funding via research grants and student loans and grants at private universities) complicates the picture even

---

\(^7\) Goldrick-Rab, op. cit.
further. Some existing research examines ability sorting⁸, but in the main we know very little.

I consider two questions here. First, do we have the right amount of heterogeneity among colleges? Casual empiricism suggests that, with a few exceptions such as Evergreen State College in Washington State, with its liberal arts flavor, or engineering schools such as Cal Tech, the public sector provides a relatively homogeneous product on several important dimensions. Certainly, public colleges do differ in important ways on certain dimensions, such as selectivity, expenditures per student, local environments, research quality and so on, but in other ways they all seem quite similar: secular, large, and impersonal, with little monitoring of students and no central vision or theme (other than, in some cases, success at football or basketball).

Perhaps this large and (on some dimensions) rather homogeneous public sector keeps us away from a lot of potential gains that would arise from a more heterogeneous system costing the same amount of money but doing a better job of matching the specific needs of diverse students. To what extent does the subsidized tuition available at public universities deter students from attending their best match? Would a system that subsidized only students and not schools change the mix of colleges and the sorting we observe a little or a lot? Theorizing about and estimating the private and social benefits of additional college diversity and the additional sorting it would imply represents a tough research challenge; it represents an interesting and important one as well.

---
Another (not unrelated) question concerns the effects of tuition differentials at state universities between in-state and out-of-state residents. Do these differentials steer students away from the colleges that would yield the highest value-added? It seems likely that they deter some matching on ability that would otherwise occur, as some high-ability students in states with relatively weak flagship schools stay in their home state because of the price difference. It seems possible to use variation among states in the amount of available choice and in tuition levels and differences to pin this down. Canada, which (largely) lacks these differentials, also provides scope for a comparative analysis.

The effects of policy reforms

Much of the discussion of policy reforms in higher education focuses on average effects of various sorts. For example, discussions of policies designed to increase access to higher education often focus on the simple “return to schooling” studied by economists, which (under certain rather implausible assumptions) gives the average percentage increase in earnings for an additional year of schooling. Policies that affect individuals close to the margin on a particular choice often get analyzed with evidence that applies to everyone making a particular choice, as when large earnings effects for current computer science majors get cited to justify an increase in the number of computer science majors. More generally, mean effects nearly always get treated (often implicitly) as common effects, which is to say that it is assumed that every individual experiences the mean

---

effect, rather than that the mean provides a crude summary of a wide variety of positive and negative effects.

Most policy reforms aim to move individuals not presently going to college to some sort of college or else they aim to change the college or program choices, or the rates of college completion, of very specific subgroups of students. The diversity of students, colleges and programs emphasized in this paper suggests that the effects on outcomes of policies that affect the choices only of particular subgroups of individuals in specific ways may differ quite substantially from the average effect over all students.

To make things concrete, consider a particular policy reform, namely, a program providing modest scholarships to poor kids presently on the margin of college attendance. What would happen in response to this reform depends on whether or not the marginal students have avoided college in the past due to an inability to borrow to cover tuition (what economists call “credit constraints”) or because of a lack of information about the college process or because they perceive the costs to exceed the benefits in the absence of the scholarship (where it matters whether those perceptions were, on average, correct or incorrect). In the first case, we might expect the students receiving the scholarships to experience larger labor market effects of college than other students as this group may include students for whom college represents a very good investment but who were previously unable to secure financing for that investment. In contrast, if the issue consists solely of information we might expect the affected group to have labor market

\[\text{\ldots} \]


effects of college attendance similar to those of the broader population of students, while if the issue consists of the affected students correctly (on average) estimating a low value-added to attending college, then we would expect relatively small impacts on their labor market outcomes (of course, if the students were mistaken in their beliefs regarding the value of college attendance, we would expect larger effects).

In each case, the reality of a modest scholarship, combined with the similarly modest test scores typically associated with being on the margin of college attendance, suggest that those affected would differentially attend community colleges or public non-selective four-year schools near enough to allow them to live at home. This in turn suggests that they may not end up optimally matched to a college or program, which suggests lower than average labor market effects from college attendance. The broader literature on college quality suggests a similar pattern. Moreover, we would expect to observe the scholarship recipients differentially sorting into programs with relatively low requirements in terms of technical preparation in high school (i.e. sociology and not engineering) as well as into programs more common in two year colleges, such as law enforcement.

The general point is that the extensive diversity of students, colleges and programs combined with a complicated relationship linking them to labor market outcomes means that the effects of particular reforms that affect certain, non-random groups of students in certain colleges and certain majors require both better evidence and more careful reasoning than they receive in policy discussions. The reasoning in this section also suggests the value of more carefully targeted, and thus more easily evaluated, policy reforms at all levels: school, state and federal.

12 See, e.g., Dan Black, Kermit Daniel and Jeffrey Smith, op. cit.
Conclusions and Recommendations

I begin this section with a quotation from a speech by former Princeton University and Mellon Foundation President William Bowen gave recently at the University of North Carolina:

“I think (hope) we can agree that accountability is key in connection with essentially everything we have been discussing: admissions policies, financial aid structures, pedagogy, graduation rates, and so on. Fortunately, there seems to be much more willingness today than in times gone by to look with a cold eye at the actual evidence concerning outcomes, and to be willing to make adjustments, or even to shift directions entirely, if something is not working. It will not do simply to assume that what appears to be a good idea is in fact a good idea.”

I am not so sure that I see the time trend that Bowen sees (though I have a shorter time series of personal observations to work with), but I certainly agree with his call for more systematic evaluation of all aspects of what universities do. Those of us at universities preach about the value of research, the accumulation of knowledge and its use in policymaking and openness about research findings to our students, our governments and our society at large, and yet often universities themselves practice none of these when managing their own affairs.

In the course of this paper, I have suggested a number of valuable and often relatively untouched research agendas related to higher education. The central argument
of this paper suggests that research in higher education will have complications that the study of some other institutions may not, due to the wide heterogeneity of students, institutions and programs. This heterogeneity in turn has implications for the design of research initiatives, as we are likely to learn the most from highly targeted initiatives that affect one or just a few margins of choice, rather than from broad policy reforms that change every aspect of the sorting of students among universities and programs.

I can add a few other items to the many research ideas already scattered throughout the text. Systematic ethnographic and survey-based studies of college choice remain too scarce. Thoughtful comparative research on alternative systems in other jurisdictions for matching students, colleges and programs would also add real value. The economics literature offers many examples of studies of specific assignment mechanisms\textsuperscript{13} but little in the higher education context. Such research would include both centralized application systems, as in Ontario, centralized allocation systems and systems in which students apply not just to a college, but to a specific program within a college, as at the graduate level in the United States. Indeed, a consideration of why we organize this process differently at the graduate and undergraduate levels in the U.S. would not be amiss. In short, as this list suggests, the opportunities for valuable research on higher education motivated by thinking in terms of heterogeneity and its implications seem nearly endless.

Creating the evidence on which to base evidence-based policymaking in higher education requires data – more of it than we currently have available. Indeed, I would argue that lack of good data represents one of the main reasons that research in higher

education lags behind that in primary and secondary education in both quality and quantity. Thus, I end with some thoughts on how we can improve the data available.

First, it would be a great boon to have reliable data on the characteristics of all colleges, two-year and four-year, public and private, good and bad. The existing Integrated Post-Secondary Education Data System (IPEDS) represents a good start but it has three key problems: it suffers from a lot of item non-response; that is, many schools do not report all of the requested variables, it has very poor coverage for two-year colleges and it lacks a number of variables that would aid in analysis, such as more detailed information about the size and admission requirements for particular programs.

Second, and more important, colleges could also contribute by making their data readily available to serious researchers studying reasonable topics. Several states, including Texas and North Carolina, have made detailed individual level data on all public school students in certain grades available to the research community (with appropriate safeguards for student confidentiality). The research undertaken with these data has added quite a lot to our understanding of public primary and secondary schools and their effects. I have little doubt that a similar bounty of knowledge about higher education will follow upon the availability of complete individual level data from one or more state university systems. The many fine papers and books that have relied on the “College and Beyond” data hint at the possibilities. Broader data, that included the large segments of the college world missed by the “College and Beyond” data, and perhaps allowed researchers to capture almost all of one or more entering classes in a state where most students stay in-state, would allow researchers to paint a more complete picture.