

Robert M.

La Follette School of Public Affairs

at the University of Wisconsin-Madison

Working Paper Series

La Follette School Working Paper No. 2010-023

<http://www.lafollette.wisc.edu/publications/workingpapers>

The (Un)Productivity of American Higher Education: From “Cost Disease” to Cost-Effectiveness

Douglas N. Harris

La Follette School of Public Affairs, University of Wisconsin-Madison

dnharris3@wisc.edu

Sara Goldrick-Rab

University of Wisconsin-Madison

srab@education.wisc.edu



**ROBERT M. LA FOLLETTE
SCHOOL OF PUBLIC AFFAIRS
University of Wisconsin-Madison**

1225 Observatory Drive, Madison, Wisconsin 53706

608-262-3581 / www.lafollette.wisc.edu

The La Follette School takes no stand on policy issues; opinions expressed
in this paper reflect the views of individual researchers and authors.

**The (Un)Productivity of American Higher Education:
From “Cost Disease” to Cost-Effectiveness**

Douglas N. Harris
Sara Goldrick-Rab

University of Wisconsin at Madison

November 30, 2010

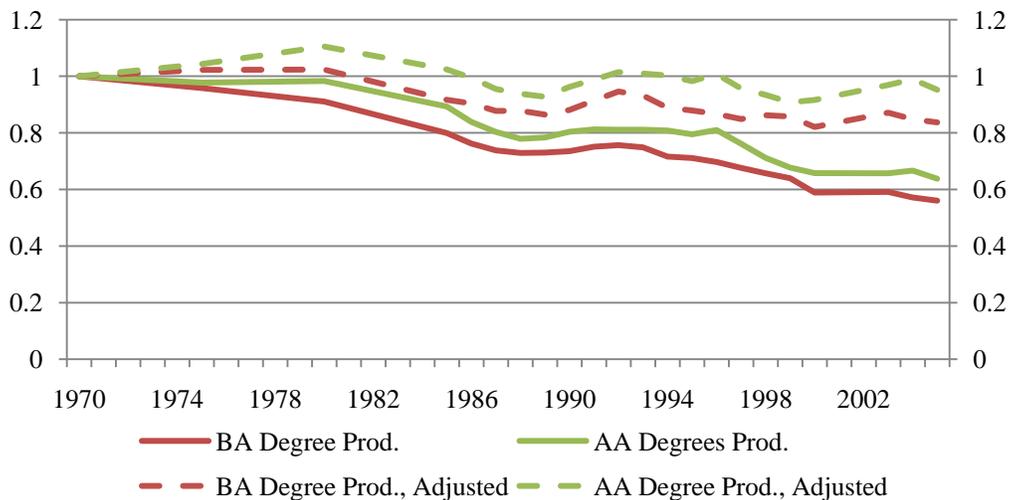
Abstract: Productivity in academic degrees granted by American colleges and universities is declining. While there is some evidence this is caused by an uncontrollable “cost disease,” we examine two additional explanations. First, few popular programs and strategies in higher education are cost-effective, and those that are may be underutilized. Second, a lack of rigorous evidence about both the costs and effects of higher education practices intersects with a lack of incentive to use cost-effectiveness as a way to guide decision-making. Rather than simply a “cost disease,” we argue that the problem is more a “system disease”—one that is partly curable.

Acknowledgements: We gratefully acknowledge financial support of the Lumina Foundation. For their useful advice, we wish to thank Charles Clotfelter, Kristin Conklin, Michael Olneck, Noel Radomski, David Weimer, John Wiley, and participants in symposia sponsored by the Lumina Foundation: Making Opportunity Affordable initiative (Indianapolis and Denver) and the Wisconsin Center for the Advancement of Postsecondary Education (WISCAPE), and a session at the 2010 annual meeting of the American Education Finance Association. For research assistance, we thank Gregory Kienzl, Regina Brown, Alan Nathan, Byoung-Ik Jeoung, and So Jung Park. All errors are our own.

The (Un)Productivity of American Higher Education: From “Cost Disease” to Cost-Effectiveness

The productivity of American colleges and universities, in terms of academic degrees granted, is declining. Since the early 1990s, real expenditures on higher education have grown by more than 25 percent, now amounting to 2.9 percent of the gross domestic product—greater than the percentage of GDP spent on higher education in any of the other G-8 countries (Clotfelter, 1996; Hauptman & Kim, 2009). Also, while the proportion of high school graduates going on to college has risen dramatically, the percent of entering college students *finishing* a bachelor’s degree has —at best—stagnated. A comparison of the class of 1972 and class of 1992 high school cohorts indicates that eight-year college completion rates declined by 4.6 percentage points (from 50.5% to 45.9%) during that time (Bound, Lovenheim, & Turner, 2009).

Figure 1: The Productivity Decline, 1970-2006¹
(Ratio of Degrees-to-Expenditures in Public Colleges; Base Year = 1970)



The combination of rising costs and declining or stagnating degree completion strongly suggests that productivity in academic degrees has declined. Figure 1 shows the trend during 1970-2006

¹ All calculations based on National Center for Education Statistics (NCES) Digest of Education Statistics. Expenditure data were not available beyond 2001 therefore revenue data were used in their place. Expenditure data are in real (inflation-adjusted) 2006 dollars. Data are available only every five years during the 1970s; the intervening years are interpolated. Data on private colleges is only sporadically available and therefore excluded.

expressed in terms of the ratio of degrees granted to total sector expenditures.² The decline is largest in the four-year sector where current productivity is less than half what it was 40 years ago. Even when adjusted for the growth in overall labor costs in the economy (see Figure 1's dashed lines) the decline in bachelor's degree production is nearly 20 percent. If these declines continue, maintaining the current rate of bachelor's degree production will cost an additional \$42 billion forty years from now.³ This means that even if state support for public higher education did not continue to decline, then tuition would have to increase by an average of \$6,885 per full-time equivalent student in public universities—almost doubling today's tuition.⁴ Expanding the number of degrees at current productivity rates would only add to the expense.⁵ Finding ways to improve the efficiency of the American higher education system is thus a top priority.

What accounts for declining productivity in this sector? In theory, it could stem from increases in the quality of higher education, but there is little evidence to suggest that is the case (Archibald & Feldman, 2008a). The economic returns to education have been rising, but this is more likely due to shifts in the demand for skilled labor (Goldin & Katz, 2008) rather than changes in the quality of degrees. Also, there is little evidence that contemporary college

² This is not a perfect measure of productivity because, for example, the resources included in the four-year (two-year) sector expenditures are not all supposed to go toward BA (AA) production, but the non-degree roles of colleges have not changed significantly over this time period, so this probably influences the productivity level, but not the trend.

³ Productivity in 2006 was 81 percent of 1970 levels. Current expenditures in 2007 on four-year colleges were \$196 billion (U.S. Department of Education, 2010a). Assuming the trend continues, productivity in 2050 will be 81 percent of 2010 levels and this will require an additional \$42 billion to generate the same degrees. This slightly understates the additional resources because the figure represents only 36 years rather than 40, so the extrapolated productivity would actually be slightly larger. These calculations exclude two-year colleges because productivity is largely unchanged in that sector. The calculations also exclude private colleges for which less data are available.

⁴ There were 6.1 million FTE undergraduates in public four-year colleges 2008 (U.S. Department of Education, 2010b). Dividing the \$42 billion by this number yields \$6,885. By comparison, according to the Delta Cost Project (2009), tuition at public research universities was \$6,741 (\$5,004) in public research (public master's) institutions in 2006.

⁵ Hanushek and Woessmann (forthcoming), Murray (2008), and Vedder (2007) argue that efforts to increase degrees are misguided.

students are gaining more than past generations of students in terms of higher-order thinking skills (Arum, Roksa, & Velez, 2008).

A number of other explanations have been given for this trend. The most commonly cited is “cost disease” (Baumol & Bowen 1966). That theory posits that productivity gains are more difficult to achieve in service sector, especially where the “quantity” of the service is defined in terms of the amount of time spent with customers. In American higher education, degrees are granted based on credit requirements, and credits are based on “seat-time.” In contrast, manufacturing enterprises can increase productivity by reducing the amount of labor hours spent in the production process; these productivity increases in turn lead to wage increases in non-service sectors, which the service sector, including universities, have to compete with by raising salaries for faculty and other staff. With the amount of time in the classroom fixed, and wages increasing, costs continue to rise while output remains unchanged, making it appear that productivity is constantly on the decline. This is why productivity declines are much smaller after taking into account the gradually increasing wages and productivity in the economy as a whole (see Figure 1). This problem is not limited to higher education—for example, legal services have seen larger cost increases than higher education while physician costs have risen at about the same rate (Archibald & Feldman, 2008b).

In addition to the postulated “cost disease,” some argue that increased access to higher education and corresponding increases in the enrollment of less-qualified and less-motivated students may reduce graduation rates, though there is little evidence to support this (Bound et al., 2009). A third explanation is that college quality is difficult to assess, leading students to use price as a proxy for quality—with the result that higher costs are viewed as a positive indicator of quality and something to be valued (Black and Smith, 2006; Smith, 2008; Zhang, 2005). A final and related concern is that the price of higher education (tuition) is disconnected from the cost (Winston, 1999); even with the recent growth in tuition, students at public institutions pay only 20 percent

of the total cost of education (Delta Cost Project, 2009). If higher tuition is equated with quality, and price does not reflect the total cost, then there is little reason for colleges to be concerned about cost containment. This may be why many colleges have raised tuition substantially and seen their application numbers soar (Glater & Finder, 2006).⁶ Circumscribing all of these explanations is the fact that most analysts emphasize the role of rising costs (Baumol & Blackman, 1995; Bowen, 1980; Breneman, 2001; Ehrenberg, 2000; Getz & Siefried, 1991; Jones & Wellman, 2009; Vedder, 2004; Weisbrod, Ballou, & Asch, 2008) or declining degree attainment (e.g., Bowen, Chingos, & McPherson, 2009) rather than the relationship between the two, i.e., productivity.⁷

While each of these explanations has some empirical support, together they also reinforce a common perception among college leaders and scholars that college productivity is impossible to control. That belief is articulated by Archibald and Feldman who write, “The problem in higher education is that productivity growth often is synonymous with lower quality. Adding more students to each class can diminish the benefit for each student, leading to diminished outcomes and lower graduation rates. Increasing the number of courses a professor teaches would reduce research or community service . . .” (2008a, p.270). Similarly, in a recent study of college presidents’ attitudes, a two-year president said, “I don't think there are any more efficiencies left to be squeezed out of public universities across the nation... There are no more efficiencies to be had” (Immerwahr, Johnson & Gasbarra, 2008).⁸ The clear implication is that institutional leaders are helpless to improve productivity without sacrificing something else of value.

⁶ Beyond the four explanations in the text, there are other explanations regarding high or rising college costs—program redundancy, the inefficiency of governmental and non-profit entities, and capital-skill complementarity (Archibald and Feldman, 2008b)—but these cannot easily explain declining productivity.

⁷ Exceptions include Massy (1996) and McPherson, Schapiro, Owen, and Winston (1993).

⁸ Direct quotations are not included in the cited paper, but were collected as part of the study and provided by the cited authors.

In this paper, we test the hypothesis that productivity gains *are* possible, perhaps without losses to quality that might outweigh those gains. Our analysis follows an approach outlined by Harris (2009). Specifically, we examine the cost-effectiveness of higher education programs by drawing on evidence of impacts from prior studies and estimate costs by collecting additional data. Our analysis compares those impacts with the corresponding costs and the results suggest that widely-used programs (often seen as markers of quality) are quite costly and those costs are not matched by comparable impacts on degree attainment—that is, they seem to reduce productivity. We analyze separately the cost-effectiveness of programs targeted to disadvantaged students because helping these students is primarily a matter of equity, though we believe this evidence is helpful identifying efficient ways of improving equity.

Compared with other areas of public policy, costs are rarely considered in education research (Monk and King, 1993; Rice, 2002). One reason is that economists generally have shown much more interest in the methodological complexities of estimating (causal) impacts of programs than cost analysis. Educational research rooted in other disciplines has also paid relatively little attention to program costs. Harris (2009) describes this as a “catch-22”: “There have been few cost analyses because there has been no basis of comparison and no basis of comparison because there have been so few cost analyses” (p.3). Similarly, Weimer (2009) argues that education research could benefit from being more “policy analytic” (p.93), including greater utilization of cost-effectiveness and cost-benefit analyses. Our analysis suggests the problem is even worse in analyses of higher education programs perhaps because, even when costs are considered, the focus is usually on the cost of enrolling more students rather than helping them graduate (Hossler, 2004). Combined with the perception that productivity is

uncontrollable and the absence of cost-effectiveness evidence, this reinforces the conclusion that higher education has room for improvement in productivity.

We outline below the methodological and conceptual challenges involved in conducting a cost-effectiveness analysis in higher education and various ways to address them. We then describe the potential promise and pitfalls of a “productivity agenda” among higher education scholars.

Cost-Effectiveness Analysis: Methodology

We begin with a brief outline of cost-effectiveness analysis in general and how we apply it in our analysis. We discuss the meaning of costs, the calculation of effectiveness-cost ratios, which represent the primary metric in this analysis, the assumptions involved, and the types of prior impact evidence we incorporate into the larger analysis.

Opportunity and Budgetary Costs

We use a standard economics-based approach to cost-effectiveness rooted in opportunity costs that includes all costs—in this case, all costs borne by students, parents, and colleges themselves (Levin & McEwan, 2001). Opportunity costs are typically larger than budgetary costs. For example, some programs utilize volunteers who are not paid (i.e., not on budget), but whose time still represent economic resources. A second example is the cost of students’ own time, which can be quite large in relation to what colleges spend.⁹ Textbook costs do not show up in the college’s budget, or in most estimates of total higher education expenditures. Because

⁹ Suppose the average student FTE has 12 credit hours (12 hours in class) and spends an equal amount of time working on coursework outside of class, for a total of 24 hours per week. Excluding summers, this adds up to $30 \times 24 = 720$ hours per year. If the opportunity cost of student time is \$7 per hour, then this amounts to more than \$5,000 per year. With an average student-faculty ratio of 15 (Digest of Education Statistics, 2007, Table 237), the total student opportunity cost per faculty member is \$75,000, about as large as the opportunity costs of the faculty member (see later salary figures).

budgetary costs tend to under-state total economic costs, we focus more on the latter in order to capture all costs to society (Levin & McEwan, 2001).¹⁰ Since few studies estimate costs per student, we often have to estimate them ourselves making informed assumptions about each ingredient or type of economic resource and collecting data from other sources. All costs in this study are expressed in 2007 dollars, unless otherwise noted.

Economists typically “discount” costs and benefits (Levin and McEwan, 2001). This is due to the basic economic assumption, supported by research (e.g., Moore et al., 2004), that people value the present over the future. This means that the significance of a cost or benefit in the future is smaller than it is in the present and this is taken into account in cost-effectiveness analysis by reducing the value of future costs and benefits based on a “discount rate.” The programs we consider have costs arising within an 8-10 year window of time (from the beginning of high school to the end of college) and, with a standard discount rate of three percent¹¹, the potential effect of discounting is modest. Also, both the costs and the effects have to be discounted (Harris, 2009), so some of the influences of discounting cancel out. The largest influence of discounting reduces the ECR of one program by about 20 percent, but does not change the ECR ranking of any of the programs. Therefore, for simplicity, we report the results undiscounted and provide discounted ECRs in a technical appendix, available upon request.

Calculating and Standardizing Effectiveness-Cost Ratios

¹⁰ Economists also distinguish between marginal and average costs, though the distinctions are less important here than in some cost analyses. In cost analysis, we would typically want to measure the marginal cost, meaning the cost of doing something for one additional student. This can differ, sometimes substantially from the average cost, e.g., adding one additional student to a classroom can be essentially free (excluding the student’s own opportunity costs), but the average cost per student in the classroom is obviously much higher. However, the context of this analysis is a potentially large expansion of degrees being proposed by President Obama and others, which implies large changes. Therefore, in this case, the distinction between marginal and average costs is less important.

¹¹ A related reason that discounting would play a small role is that the generally accepted discount rate is only three percent (Lipscomb, Weinstein, & Torrance, 1996; Moore et al., 2004; and Muennig, 2002).

We employ the ratio of effects (E)-to-costs (C), or the effectiveness-cost ratio (ECR), as a measure of program productivity. Larger ECRs imply greater productivity. To standardize the measures across programs of varying sizes, we report the costs of implementing a program for an entering cohort of 100 students. For targeted programs (e.g., need-based financial aid or programs that serve only disadvantaged students), the cohort involves 100 students in the targeted population. Some programs or policies engage students prior to college entry, while others affect them once entry has occurred. The timing of the entering cohort aligns with the timing of the program—if the program begins in high school then we consider a cohort of 100 high school freshmen; if it begins in college then we consider a cohort of 100 college freshmen.

We make one other narrowed assumption in the service of pragmatism, and that is to focus mainly on productivity in *degree* completion. Degrees represent a central aim of students and policymakers across the political spectrum. We recognize that by focusing on a single outcome we reinforce a trend toward narrowing the functions of education only to those we can measure. This is borne of necessity—few higher education studies measure outcomes other than degree attainment, making it impossible to include them in a cost-effectiveness analysis. There are few studies with evidence of graduation impacts, as well as evidence about achievement effects. While we cannot report effectiveness-cost ratios in terms of achievement, we do still discuss the achievement evidence where it exists to see whether there is evidence of a trade-off among the outcomes.

A common problem in cost-effectiveness analysis is that different studies focus on different outcomes, requiring some type of conversion to a common metric. This is especially true of cost-benefit analyses that require conversion of effects on measures like achievement to monetary measures (e.g., Dynarski, 2008; Harris, 2007). In the present analysis, we review some

studies that report effects on graduation, while others report only persistence from one year to the next. We start by reporting the effects as the study authors reported them for the specific outcome they studied. If they studied persistence, then we report the effect on persistence. For the sake of comparability across studies, we also attempt to translate persistence effects into graduation effects based on evidence about the relationship between these two measures. While graduation requires more than persistence (for example the accumulation of good grades and credits earned), persistence is a necessary precondition for graduation. Therefore, using data from nationally representative studies, we calculate multipliers to translate impacts on persistence to impacts on graduation for studies that do not report graduation effects; and use a separate set of multipliers for the costs.¹² The multipliers have very little influence on cost-effectiveness comparisons among programs where persistence is the only outcome measured, though we are more cautious about making broader comparisons between persistence studies and those that measure effects on graduation directly.

Once costs and effects are calculated and applied, the effectiveness-cost ratio (ECR) is a straightforward computation: divide the (multiplier-adjusted) effect by the (multiplier) adjusted

¹² Impact multipliers are in the 0.30-1.00 range (and applied only to studies of initial enrollment and persistence conditional on enrollment). The size of the impact multipliers depends on the sector (two- versus four-year), student income status (different groups have different baseline persistent rates), and the number of years since the start of college that persistence is measured. In theory, impact multipliers could be greater than 1.00, depending on the baseline rate of persistence. If all four-year students who stay in college for four years also graduate at that time, then a baseline year-to-year persistence rate of 0.60 yields a multiplier of 1.00 (that is, a one percentage point increase in the persistence rate, to 0.61, yields a one percentage point increase in the graduation rate. If the baseline persistence rate is higher (lower) than about 0.60, then the multiplier is greater (lower) than 1.00. Data from the Beginning Postsecondary Survey (BPS) suggest that the baseline persistence rates are greater than 0.6 so that the multipliers should be greater than 1.00. Since it is common for the initial effects of programs to decline over time (even when the programs continue to be in place), we cap the multiplier at 1.00.

Cost multipliers are in the range 1.00-3.73 (and applied to all studies). The cost multipliers are greater than or equal to 1.00 because the unadjusted costs mentioned in the text and in Table 1 are on an annual basis, so the minimum cost is for one year; however, the impact estimates in many studies are based on the use of resources over more than one year. For example, if we were studying a program that is applied in every year that students are in college, and we assume a 100 percent participation rate in all five years, then we would have to multiply the annual costs by the cost multiplier of 5.00. Because some students will drop out before graduation and therefore require fewer resources, the actual cost multiplier is lower than 5.00.

cost. ECRs are usually positive, but they can be negative if a program has positive effects and also saves resources. A program with a ECR equal to 1.0 has a total cost of \$1,000 per each additional graduate. As we will see below, most ECRs are much smaller than one—in other words, it costs much more than \$1,000 to produce an additional college graduate, even for the most cost-effective programs.

The ECRs are based on a variety of assumptions. In particular, they implicitly assume that the costs and impacts of programs are independent of other programs (Harris, 2009). For any group of students “on the margin” of graduation, implementing a given program may propel them to graduation, but additional programs (beyond the initial one) may have smaller marginal effects—that is, there may be diminishing returns to programs in general. Since we know relatively little about even the “main effects” of most higher education programs, consideration of potential effect interactions is set aside as an issue for future research.

Other researchers have carried out cost-benefit analyses in which the benefits of higher education are translated into dollar terms (mainly based on the estimated returns to education in terms of future wages and salaries) and the question is whether the benefits exceed the costs. This approach has some advantages when there are multiple measured outcomes and when decision makers have control over the total resources, as in the case of state and federal government. The problem is that cost-benefit analyses are often limited to a single program, so that a recommendation about a given program is based solely on whether the benefits exceed the costs. However, policy analysis is inherently comparative so that the ratio of the benefits-to-costs has to be greater than it is for the alternative policies. This type of comparative cost-benefit-analysis is almost never carried out in education policy research. As we show later, the policy implications can be misleading using this standard cost-benefit approach.

A key goal of this analysis is to facilitate comparisons across programs, which in turn can help improve policy choices and, ultimately, student outcomes. However, it is worth noting three different types of comparisons and their relative usefulness: (1) comparisons across strategies; (2) comparisons across programs within strategies; and (3) comparisons within strategies but across student populations. We argue that the first two comparisons are important for improving productivity and, while still recognizing some of the necessary assumptions, are comparisons we wish to encourage (Harris, 2009). We try to avoid comparisons across student populations, however. Because some programs are targeted to specific groups of students (e.g., related to race and income), comparisons among programs and strategies can effectively pit groups against one another in competition for scarce resources. This issue arises with any comparison among programs, including the present one, where balancing multiple social goals, such as equity and efficiency, is inherent to the policy decision. We recommend a balanced consideration of the potential for such comparisons to improve outcomes for all groups, while accounting for the potential downsides of conflicts arising from policy decisions.

Rigorous Research and “Break-Even” Impacts

In identifying promising practices and programs, we searched for studies of college programs or policies that used rigorous research methods and consulted with colleagues and experts in the field. Studies from outside the U.S. (except one from Canada) are omitted, as are studies from before 1980. We have attempted to be comprehensive in including studies that meet our criteria, but our general conclusions do not depend on comprehensiveness.

Next, we placed programs into categories of rigor using the standards similar to the U.S. Department of Education, What Works Clearinghouse (2010c). We consider levels of evidence

to be strong with randomized control trials that exhibit high internal validity, more moderate with quasi-experimental methods that have equivalence between treatment and comparison groups, and lower when relying on other types of evidence and/or theories based on expert opinion. The vast majority of research is in the low category. While we do not assume that all randomized trials exhibit high internal (or external) validity, we do include all experimental studies we found. The key is whether the authors can rule out alternative interpretations that would call into question the causal impacts (Shadish, Cook, & Campbell, 2002).¹³

Because we judge that most of the uncertainty probably has to do with the impacts rather than the costs, we also carry out a “break-even” analysis. Specifically, we calculate how large the impacts would have to be to equal the current average spending per degree. The details of many of these cost calculations, including more detail on the multipliers can be found in the separate technical appendix.

Results: Costs and Cost-Effectiveness

Typical Hallmarks of College Quality are Costly

We begin with two resource decisions that have important effects on budgets and are widely seen as key indicators of quality—student-faculty ratios and full-time faculty. While much of the debate about higher education has moved beyond these resources, they still comprise 13 percent of *U.S. News and World Report* undergraduate rankings.¹⁴ Below, we consider the costs of each and briefly summarize available evidence on impacts.

¹³ In addition to internal validity, impacts estimates are uncertain because of sampling variation. We report in the technical appendix ECR confidence intervals based on the standard errors of the impact estimates. None of the studies include information that would allow us to estimate confidence intervals for costs.

¹⁴ According to U.S. News (2010), the “faculty resources” component represents 20 percent of the total rating. Sixty-five percent of this portion is comprised of “percent faculty that is full time” (5 percent), “percent faculty with terminal degree in field” (15 percent), “student/faculty ratio” (5 percent), “class size 1-19 students” (30 percent), and “class size 50+ students” (10 percent).

Student-faculty ratio and class size. For a given faculty teaching load, a small student-faculty ratio means small class sizes. Faculty-student interaction outside the classroom may also be facilitated this way. Jacoby (2006) reports full-time faculty salaries of \$74,443 (\$58,041) and part-time faculty salaries of \$16,156 (\$12,174). (For brevity throughout the paper, when reporting data simultaneously for four- and two-year colleges, we report the two-year figure in parentheses immediately after the four-year figure, as in the previous sentence.)

Current student/faculty ratios (FTE basis) are 14.8 for four-year publics and 19.2 in two-year publics (Digest of Education Statistics, 2007, Table 237). Four-year (two-year) colleges currently already have 2/3 (1/3) of their courses taught by full-time faculty (USDOE, National Study of Postsecondary Faculty (NSOPF), 2004). Based on these ratios and (weighted) salaries, reducing the student-faculty ratio from 15 to 14 at four-year colleges (from 19 to 18 at two-year colleges) would therefore cost \$25,561 (\$8,064) per year for 100 students, excluding capital costs and fringe benefits.

Full-time faculty. One way colleges have attempted to reduce costs in recent decades is through hiring adjuncts or part-time faculty. We calculate the costs of this change based on data on percent part-time faculty and faculty salaries, as well as the precise number of courses taught by part-time/full-time faculty (NSOPF, 2004). The costs of switching from the above actual proportions to all full-time faculty would be \$30,425 (\$19,153). It might seem surprising that the two-year college costs are lower, given that more faculty have to switch to full-time status in the two-year sector in this hypothetical policy experiment. However, the part-time/full-time gap in salary per course is much larger in four-year colleges.

Summary. These cost estimates show that it is costly for four-year colleges to do what is necessary to generate a high national ranking. Of course, it makes no more sense to focus only

on the costs than it does to continue the tradition of reporting only impacts. Full-time faculty and small classes may well pay off for students in ways that justify the expense, a topic we take up later.

College Access Programs are Particularly Costly

Policymakers have focused for decades on increasing access to higher education by targeting disadvantaged middle and high school students. Some of the oldest and most researched access programs are GEAR UP, Talent Search, and Upward Bound.

GEAR UP. Gaining Early Awareness and Readiness for Undergraduate Programs (GEAR UP) is the umbrella name given to a loosely defined set of services for high school students. The purpose of GEAR UP is to foster increased knowledge, expectations, and preparation for postsecondary education among low-income middle school students and their families. GEAR UP projects provide services to students, parents and teachers at high-poverty schools with at least 50 percent of students eligible for free or reduced price lunch. Services include tutoring, mentoring, college field trips, career awareness, college-readiness counseling, and parent education about access to higher education, as well as college scholarships. GEAR UP chooses entire middle schools to participate and requires provision of services to an entire grade cohort, but individual student participation is voluntary. The funds are intended to support students through high school. Based on data collection on the costs of a single GEAR UP program, total costs, including personnel, facilities, and equipment, were \$264,000 for a cohort of 97 students (Albee, 2005).

Upward Bound. One of the original federal TRIO programs, aimed at increasing college access among low-income first-generation students, Upward Bound provides tutoring, SAT and

ACT test preparation, summer and after-school sessions aimed at improving language arts and math skills, as well as campus visits. These regular interactions with students make Upward Bound more costly. Cohort cost estimates range from \$480,000 (Myers, Olsen, Seftor, Young, & Tuttle, 2004) to \$516,000- \$677,000 (Albee, 2005). The former and lower figure is based on federal budgetary contributions, while the higher figure is based on opportunity costs in some specific sites, and this reflects the general observation that budgetary costs under-state total resources.

Talent Search. A second of the original federal TRIO programs—and the largest in terms of the number of students served—Talent Search provides a combination of academic support, career development activities, and financial aid assistance to high school students (Constantine et al., 2006). Specific services include test-taking, study skills, academic advising, course selection, college orientation, college campus visits, referrals, counseling, financial aid counseling and workshops, FAFSA assistance, and scholarship searches. Federal contributions amounted to \$392 per participant in 2009, which is considerably lower than other TRIO programs. We investigated costs further by searching for web sites of specific Talent Search programs. Talent Search programs do not apparently make use of volunteers or receive substantial institutional resource contributions therefore the budgetary costs appear to be a reasonable approximation of total (opportunity) cost. The implied cohort cost is therefore \$39,200.

Bridges to Opportunity. The Bridges to Opportunity intervention is a series of courses aimed at improving English skills for non- or poor-English speakers. The three interventions highlighted were: Vocational ESL, Workplace Basics, and Technology Career. All three programs focus on skill development (English language communication and applied mathematics

and applied computer applications) as well as language acquisition. Jenkins (2002) reports costs of: Vocational ESL \$684, Workplace Basics \$684, and Technology Career, \$4,791 (per enrollee).

Summary. GEAR UP and Upward Bound are widely regarded pillars of college access, but they are also costly. At more than a half-million dollars annually, the Upward Bound costs are equivalent to hiring nine full-time faculty at a four-year college for one year. Increasing access to higher education for disadvantaged students is a worthy goal, but continuing these programs is still questionable if there are other more cost-effective ways to reach that goal. Talent Search appears to be much less costly (though less cost information is available).

Program Costs Vary Widely

A basic principle of policy analysis, and especially cost-effectiveness analysis, is that good decisions cannot be made without comparing potential policies to the plausible alternatives (Weimer and Vining, 2005). In this section, we provide cost information about a wide variety of other programs that fall into three general categories: student services, financial aid, and instruction.

Student counseling. As part of the MDRC Opening Doors initiative, low-income students who were just starting college, and who had histories of academic difficulties, were provided additional counseling and given a small stipend of \$300 per semester when they used those services in two Ohio community colleges; the average stipend was \$210. Counselors had a much smaller than usual caseload (119 versus 1,000 in the control group) because of the expectation that they would be spending more time with each student; students also were given a designated contact in the financial aid office. Students did use counseling and financial aid services at

greater rates than control group students (who also had access to standard campus services). Based on the number of counselors involved and Bureau of Labor Statistics (BLS) data on average counselor salaries, we estimate counselor costs of \$340 per year per student. Adding the time of counselors to the student stipends, the unadjusted cohort cost is \$54,898.

Call centers. An alternative student service approach, the call center, involves literally making phone calls to students who apply but do not register, register but do not show up for class, show up for class initially but then stop attending, and so on. When a student is reached by phone, the college representative tries to learn the reasons why the student is not progressing and to direct the student to other services or assistance to help. It could be that the services and assistance were important to students or that the effort made by the college representatives gives students a closer connection to the campus and motivated them to return. We estimated cohort costs of \$200-\$500 based on information from call centers in other sectors and our own back-of-the-envelope calculations. However, these costs exclude any additional ancillary services students might receive if, for example, the call leads students to contact financial aid officers or counselors to seek additional assistance. To be conservative, we therefore double the direct cost of the calls themselves.

Financial aid (no services). Tuition is the heavily subsidized price of college paid by students. The cost of these subsidies, as well as grants to students, is essentially the face value of the subsidy or grant. Some grants and scholarships have “merit” requirements based on courses and grades. The situation is more complicated with loans. The U.S. federal government, and to a lesser degree state governments, have long used loans as a means of making college accessible.

Since loans have to be paid back, grants would likely have a larger impact than loans, albeit with much greater costs to the government. Dynarski (2003b) estimates that the

government subsidy for Stafford loans, in which all interest is paid by the government while the student is in school and the interest rates are subsidized after students leave college. She concludes that the “subsidy value of a [subsidized] loan is about a third of its face value” (p.21). However, this figure is apparently based on what students with high credit ratings would obtain and this probably over-states the credit situations of the average student. We estimate that the more typical subsidy is probably closer to 43 percent of face value.¹⁵ We therefore assume the cost to the government for a \$1,000 loan is \$430.

Financial aid combined with services. Some financial aid programs are bundled with other student services. MDRC’s Opening Doors also included an experiment that combined services with performance-based financial aid. The most obvious costs of Opening Doors in New Orleans are the costs for performance-based scholarship, as well as advising and counseling students. Based on data from two community colleges, Richburg-Hayes et al. (2009) report that average total scholarship payment per student over two semesters was \$1,246. Based on the number of counselors employed and the national average salaries of these workers, we estimate the costs of the counselors to be \$340 per student so that the total average cohort cost is $\$1,246 + \$340 = \$1,586$ (per year for two years).

While we are generally focused on U.S.-based results, we make one exception with the Canada STAR study. This is the only randomized trial of financial aid to our knowledge that occurred at a university, although another is ongoing. In addition to the control group, there

¹⁵ We are aware of no direct evidence on the credit histories of students or their parents (who usually co-sign on the loans). Dynarski’s one-third subsidy figure assumes that the market rate is seven percent, interest rate which she describes as the rate for borrowers with excellent credit histories. She writes that the rate for borrowers with poor credit histories was nine percent. For market rates of seven, eight, and nine percent, and a (subsidized) Stafford loan interest rate of six percent, the value to students of the reduced interest rate is: 30 percent (matching Dynarski’s calculation), 43 percent, and 57 percent (respectively). These estimates incorporate both the lower interest rate and the fact that the government pays all interest while students are in school. We assumed an 8-year repayment schedule and two years of time in college with complete government subsidy. We use the middle figure as our cost estimate, which we believe best represents the average student; however, the correct rate certainly varies across individuals.

were multiple treatments: (a) services-only, such as facilitated study groups; (b) scholarship money-only; and (c) a combination of (a) and (b). The costs for these options are reported as \$302, \$366, and \$739, respectively (Angrist et al., 2009).

Early commitment programs. In contrast to the above programs that target students who have already entered college, promise and early commitment programs try to influence students in high school or earlier by promising future college funding. There are dozens of such programs around the country, so we simply describe the Indiana 21st Century Scholars program as a case in point. Established in 1990, this program provides need- and performance-based aid and support services such as mentoring and organizing college visits. The program promises middle school students who qualify for the federal school lunch program eight semesters of full tuition at an Indiana public college or university, or a like amount at one of the state's private schools. Students remain eligible must maintain a C average in high school. In 2008, the maximum awards were \$5,172 for public state colleges and \$10,014 for in-state private ones. Since tuition at every Indiana university appears to exceed these amounts (The National Center for College Costs, 2010), and most students attend public institutions, we use \$5,172 as the typical payout, for a total cohort cost of \$517,200. While we do not have cost estimates for the services that go along with these payments, they appear quite small in comparison to the above tuition subsidies.

Emergency financial aid. The Dreamkeepers program attempts to reduce attrition among community college students by providing funding for those emergencies that arise and threaten the financial security of enrolled students (Geckeler et al., 2005). For low-income students who may already be struggling to meet their financial obligations an unexpected expense (for

example, auto repair, rent increase, eviction, etc.) can sometimes be the catalyst for delaying or severing their chance at a diploma.

Eleven community colleges participated in the Dreamkeepers pilot program. Each institution was required to match grant funding after the first two years in order to receive subsequent funding. Community colleges taking part in the program varied in their location, size and setting and developed their own criteria and structures for distributing the funds. More than 1,500 students received emergency funds ranging from \$11 to more than \$2,000 with an average of \$393 per recipient.

The Angel Fund Program is intended to meet similar emergency funding needs as those of the Dreamkeepers program, but to a specific population—Native American students. In the United States there are thirty-two colleges that are accredited to serve the needs of a primarily Native American student population. Most of these colleges award two-year degrees, although some award bachelor's degrees. Like students in Dreamkeepers colleges, students attending these tribal colleges have considerable financial needs; most students come from backgrounds of high levels of poverty. The Angel Fund program operated with the same specifications as the Dreamkeeper program. In 2006 data indicates that more than 600 awards were distributed to 587 students. Awards varied from a low of \$15 to a high of \$2,055; the average award was \$266.

Online and distance learning. So far, we have considered programs that attempt to influence students indirectly—by changing the general faculty resources available (student-faculty ratios and adjuncts) providing various forms of services (e.g., counseling and mentoring), and financial aid. But this means we have ignored what is arguably the core activity of colleges: instruction. We therefore conclude this section by discussing evidence about online learning, other modes of instruction, and remediation.

Early incarnations of distance education included correspondence courses, education television, and video-conferencing. Of greater interest here are more advanced online learning using the internet to deliver multi-media instruction. Supporters of online instruction point out the possibility of improving instructional quality and student engagement, as well as increasing convenience and expanding the reach of colleges to serve non-traditional populations.

We have chosen to spend more time below discussing the costs of online education both because online education is of such great public interest and because the cost structure is somewhat distinctive. There are considerable fixed costs to online education. While it might be aimed at teaching the same content, online instruction requires very different instructional techniques and materials which take time to develop. Online education also requires computer and related equipment, including special software licenses (e.g., Blackboard), servers, and maintenance personnel to keep these systems running.

There remains some question whether, after accounting for these fixed costs¹⁶, teaching an online course requires more time of the instructor and/or students. And any additional costs to both groups might be offset by reduced transportation costs and the benefits of increased convenience, all of which, to our knowledge, are yet to be quantified.

Some colleges pay faculty a premium to develop and teach online courses. To the degree that these fees are intended to compensate for course development time, they should be attributed to that cost category. But the fees might also reflect additional time needed to teach online courses, an issue that again remains unresolved. Anecdotally, online courses that attempt to transfer existing in-person courses to online formats do seem to require more faculty time. So,

¹⁶ These costs are not strictly fixed because they vary to some degree based on the number of students (e.g., Blackboard licenses might be priced on a per-student basis). In these situations, it is perhaps more accurate to refer to the costs as “lumpy.”

while online education no doubt expands access to courses and appears to increase quality, it seems more expensive.

The difficulties of measuring costs have received little attention in the literature on online learning (Gordon, He, and Abdous, 2009). To provide at least some basis for analysis, we estimated the costs of online instruction from information used in a study of Marshall University (Morgan, no date), from which we estimate that the additional cost per student is roughly \$100 per student per course. For an entire online program, for a student attending online full-time (eight courses per year), this would amount to \$800 per year, for a cohort cost of \$80,000.

On the other hand, the National Center for Academic Transformation (NCAT), led by Carol Twigg, created the Program in Course Redesign (PCR) with the hope of redesigning instructional approaches using technology to achieve cost *savings* as well as quality enhancements. It was conducted from 1999 to 2004 and during this time 30 two- and four-year colleges and 50,000 students per year participated. Redesign projects focused on large-enrollment, introductory courses, which have the potential of impacting large numbers of students. Content is delivered entirely online and students have nearly round-the-clock online access to instructors. As this is a competency-based model, successful students can complete the courses quickly and require few resources. PCR leaders asked each site to provide cost savings information and we reviewed available documentation from the PCR web site for each of the program sites where implementation was deemed “fully successful.”¹⁷ The simple average was a 44 percent reduction in reported costs. It appears that the majority of savings in most cases was from the replacement of small sections with large lecture (sometimes delivered online) and the reduction in the number of teaching assistants. Cost methodology was not reported, but since

¹⁷ http://www.thencat.org/PCR/Proj_Success.html

cost savings are coming from reduced faculty and teaching assistant time, and there are no obvious effects on off-budget resources, these may be reasonable estimates of opportunity costs.

These examples reinforce the varied nature of online learning and associated variation in costs (or cost savings). Some are entirely online while others blend online and in-person instruction; some change not only delivery method, but also the structure of the curriculum. This complicates the comparisons being made and the ability to make generalizations about the effect of “online learning.”

Other modes of instruction. Online instruction is not the only alternative to face-to-face instruction. The form or mode of instruction has been subject to a great deal of research over many decades. It is also one of the very few topics for which an extensive cost-effectiveness analysis has already been undertaken. Brown and Belfield (2002) in an extensive review report costs of a variety of instructional methods (relative to the cost of lecture): Discussion/inquiry (-\$454), Independent Study (-\$482), Personalized (\$2,867), and Other modes (\$1,320). That is, the negative costs imply that the discussion/inquiry and independent study were cheaper than lectures, while personalized and other modes were more expensive.

Remediation. A growing concern is that students who enter college are not adequately prepared for college-level work. While this is partly seen as a flaw of high school preparation, many colleges try to address the issue through remediation programs. Placement in remediation is often based on scores on standardized tests. Texas spends \$172 million per year on remediation programs (McFarlin and Martorell, 2007) that educated 162,597 (mostly four-year) students in 2006 (Terry, 2007). This translates to \$1,057 per student, or \$105,700 per cohort.

Some have expressed concern, however, that the quality of the typical remediation programs is relatively low and that more extensive, high-quality programs would have a positive

impact. Dowd and Ventimiglia (2008) estimate the costs of a high-quality remediation program, Pathways, which includes a combination of math and language arts. They estimate total costs of \$1,700 per student session. This yields an unadjusted cohort cost of \$170,000 per year which, as expected, is somewhat higher than the cost of standard remediation reported above.

Summary and break-even analysis. The cohort costs of these programs vary widely, from as low as \$1,000 for call centers to more than 600 times the amount for Upward Bound. From a productivity standpoint, call centers could produce tiny impacts and still be worthwhile. Likewise, expensive programs may be cost-effective, but only if they generate very large impacts on student outcomes.

Since we do not have convincing estimates of causal impacts on graduation for the majority of the programs considered above, we begin the comparison of programs by summarizing the cost of each program and calculating break-even impacts—that is, the impact necessary so the ECR equals current productivity rates, as measured by the current average cost per graduate.

To our knowledge, the best evidence on current cost per degree comes from Johnson (2009). He uses multiple sources of data and accounting methods to calculate costs, relying mainly on detailed data from Florida. He estimates that the opportunity cost for a bachelor's degree (what he calls "full cost attribution") is \$40,000. We believe this is an underestimate for two reasons: (1) Florida spends about one-third less than the national average on higher education compared with other states¹⁸; and (2) these estimates include only 73% of total spending on undergraduate education. Also, graduate education in some ways subsidizes

¹⁸ We compiled data from the National Center for Education Statistics, though we could not find full-time equivalent data by state and total enrollment figures for the same year as expenditure data. National college expenditures were roughly \$170 billion in 2001 with enrollment of 4.5 million in 2006 (combining part-time and full-time), for an average of \$37,700. The equivalent figures for Florida were \$5.4 billion and 234,000 students, for an average of \$23,000, or 61% of the national figure (same years).

undergraduate education in many universities by creating a captive pool of low-cost graduate students to teach courses. Therefore, we argue that a more realistic cost estimate is \$60,000. Johnson also estimates costs for two-year degrees of approximately \$20,000. In this case there are fewer costs excluded and we therefore adjust this figure up by a smaller percentage to \$25,000.

These average cost per degree figures are useful because they are, in essence, estimates of the current rates of college productivity. By simply inverting these figures (and multiplying by \$1,000 to express the effectiveness-cost ratio in thousands), we obtain baseline four-year (two-year) ECRs of $1/60=0.017$ ($1/25=0.04$). In Table 1, we estimate what impact for each program would be necessary to generate these same ECRs and thus break-even compared with current costs per degree.

The calculations could also be carried out using comparisons other than current cost per degree, but the main point here is to show the differences in required effects of each program relative to the others in the list. The break-even effects are important because they provide a guide for future research that might further explore the effects of these programs.

Notice in Table 1 that there is a very wide range of costs and therefore a very wide range of break-even impacts. The cheapest program—call centers—is also the one with the smallest break-even impact (0.09). This means that spending \$1,000 on call centers need only produce 1/14 of one graduate out of a cohort of 100 in order to match the productivity of existing spending. At the other end of the spectrum, Upward Bound costs \$677,000 per year for 100 students. It would have to produce almost 17 new college graduates in order to match current costs per degree. This reinforces the conclusions of the previous section and shows the wide range of costs across a variety of programs.

Table 1 distinguishes programs by whether the evidence pertains to two-year students (“(2y)”) and/or disadvantaged students (“disadv”). In cases where there is no indication about two-year or student disadvantage status, the evidence pertains to average students in four-year colleges. Again, the cost per degree is lower in the two-year sector, so this distinction is important for establishing the basis of comparison. Likewise, the cost multipliers are lower in the two-year sector because two-year degrees take fewer years; this means that student participation in programs involves fewer years, and fewer total resources. The fact that some programs target disadvantaged students is important because these students may benefit more from certain types of programs; and policymakers, for purposes of equity, might be willing to pay more to increase college graduation for disadvantaged populations.

Some Innovative Programs Have Considerable Potential

We now begin to incorporate evidence about the impacts of some of the programs where at least some convincing evidence is available. We have considered two programs that show promise for various reasons and that are in some sense innovative. Online learning can be costly, yet this approach represents one of the few potentially transformative ways to improve instruction. Call centers, in contrast, are very inexpensive and the question is whether they generate measureable impacts. For those few programs where data are available, we combine the above evidence about costs with evidence on impacts.

Online and distance learning. A recent meta-analysis published by the U.S. Department of Education (Means, Toyama, Murphy, Bakia, & Jones, 2009) finds that online education yields greater achievement compared with face-to-face instruction with an effect size of 0.24-0.35

standard deviations.¹⁹ This conclusion is based primarily on online learning in higher education but also mix some K-12 studies. While we are not aware of a single study of the effects of online learning on persistence or graduation, these achievement effects would suggest that there is real potential in online learning, with regard to degree quality.

At least two more recent studies have questioned the USDOE conclusions, however. Jaggars and Bailey (2010) provide a more nuanced interpretation, showing that the positive achievement effects do not hold for fully online courses and may not hold for more disadvantaged groups. If the goal is to increase graduation rates, this is a legitimate concern since students who are now dropping out of college are also more disadvantaged. Figlio, Rush, and Lin (2010) reinforce this argument in one of the very few randomized trials; they find that live instruction is better than online delivery for lower-achieving students. They also go further in their critique of the USDOE review, concluding that “the [prior] evidence base on the relative benefits of live versus online education is therefore tenuous at best” (p.4). Overall, it remains difficult to generalize about online instruction and there remains essentially no evidence about persistence and graduation.

Call centers. The Des Moines Area Community College (DMACC, 2009) created a call center and kept track of the persistence rates of students they reached. The college conducted something akin to a quasi-experiment, keeping track of who they could not reach by phone, who they only left messages for, and who they talked with in person, as well as who ended up staying in college. As they point out in their report, comparing those students they could not reach by

¹⁹ A second review by Canadian researchers finds somewhat smaller but still positive achievement impacts (Abrami et al., 2006). The USDOE report cited in the text also found that courses blending online and face-to-face instruction were even more effective. Another review by Cavanaugh et al. (2005) focuses on studies from 1999 to 2004 and finds no statistically significant impact. However, a large proportion of the rigorous studies of higher education have been conducted since 2004 and these tend to show more positive results, perhaps because online technology has improved, as has the ability of students and faculty to operate in online environments.

phone to students they did reach would not be a reasonable comparison because it is likely that some of those who could not be reached at all (not even a voice mail) had moved, had their phones disconnected, or had some other issue that would also be associated with their likelihood of college persistence. Not surprisingly, the rates of persistence are much higher for the students having voice calls and voice mails versus those students not reached at all.

A more reasonable comparison is between the groups who received a voice mail and those spoken to by phone. Those comparisons are not perfect either because students reached only by voice mail might have a job or family responsibilities that make it less likely they would answer the phone and simultaneously less likely they would continue in college. Nevertheless, it is plausible that a substantial share of the variation in the ability of college representatives to reach the students was due to random chance in the timing of the call in relation to students' other responsibilities, and it is therefore worth comparing the outcomes of students who had in-person conversations with those who received only voice mails. In these cases, student persistence was 2-15 percentage points higher depending on exactly what triggered the call (not registering, etc.).²⁰ We therefore take the lowest number in this range—2 percentage points—as our estimated impact. Even this apparently “small” impact yields a very large ECR because the program is so inexpensive. While this evidence no doubt falls into the low range by our standards of rigor, it is worth pointing out the anecdotal evidence that the use of call centers is more widespread in for-profit colleges who have more to lose when a student drops out. So, online learning and call centers are both unproven, but they do show potential.

Common Programs Fail Simple Cost-Effectiveness Tests

²⁰ The estimates at the high end of this range are statistically significant, but the estimates at the low end are not.

While the above programs show promise, we show below that most typical programs in higher education do not. After discussing each program individually, we summarize the effectiveness-cost ratios in Table 2. Two-year college programs with ECRs higher than 0.0400 and four-year college programs with ECRs higher than 0.0170 are those that would increase college productivity (see above explanation).

College access. Using a randomized trial to study Upward Bound, there was no detectable effect on overall postsecondary enrollment or the type or selectivity of postsecondary institution attended for the average eligible applicant (Seftor, Mamun, & Allen, 2008). While not statistically significant, they estimated an impact of less than two percentage points in the overall rate of enrollment (across college types, etc.). For enrollment at four-year colleges and universities, the estimated impact is one percentage point.

Upward Bound increased the likelihood of earning a postsecondary certificate or license from a vocational school (these results were statistically significant). While about four percent of control group members received a vocational certificate or license, nearly nine percent of treatment group members did, implying an impact of five percentage points. Because of the focus of current initiatives on two- and four-year degrees, we use two percentage points (see above) as the baseline impact for average low-income students.²¹ Based on the range of costs from Albee (2005), the adjusted ECR for average students is 0.0008, which is far below current overall degree productivity, suggesting questionable cost-effectiveness. (Note that the ECRs in Table 2 should be compared with the current cost per degree reported in the upper rows of Table

²¹ Upward Bound increased postsecondary enrollment or completion rates for the 20% of eligible students who had lower educational expectations (no expectation of earning a Bachelor's) at baseline. However, because being eligible for Upward requires a considerable degree of disadvantage to begin with, those students who also have low educational expectations are extremely disadvantaged. Therefore, for comparability with other program ECRs in this study, we use the two percentage point impact as the baseline. The focus on overall impacts is also preferable because Upward Bound does not limit access based on college expectations nor is likely to do so in the future.

2, *not* with the break-even impacts in Table 1, which are provided as a basis of comparison with future research on program impacts.)

Other studies of similar programs have used less rigorous propensity score matching (PSM) methods and found much larger effects. Constantine et al. (2006) find Talent Search improves college enrollment by 6-18 percentage points. Likewise, Domina (2009) finds that these types of college “outreach” programs have improve college enrollment by about six percentage points (though essentially no impacts on high school educational performance). Because of both the methodology, and the much smaller effects in the Upward Bound experiment, it appears likely that these estimates using PSM are inflated. We do not report ECRs for Talent Search in Table 2 for this reason.

Financial aid. We briefly summarize results in Table 3 from a large number of quasi-experimental studies as well as two are randomized trials on merit aid (indicated with an asterisk). See also Deming and Dynarski (2009) for a review of this literature. Except where indicated, our interpretation and review of the evidence is essentially the same as theirs. Their conclusions are more general than our own and we instead focus on specific impact parameter estimates. There is much less evidence on the impact of loans, though the two quasi-experimental studies we are aware both find positive impacts (Reyes, 1995; Dynarski, 2005). The adjusted ECRs in Table 2 are 0.0087 for loans and 0.0056 for grants (both are below current productivity in both sectors). The first figure means that an additional \$1,000 spent on loans for a group of 100 students would yield less than 1/10 of one new graduate. The ECR estimate for grants suggest that simply cutting costs and reducing tuition (which is what grants do) would not improve productivity in the long run. The ECRs are larger for performance-based aid, only slightly so in the case of the Georgia and Arkansas programs, but more so in Canada STAR.

Financial aid combined with services. For the Canada STAR study, Angrist et al. (2006) report a point estimate for the effect of the funding-only treatment on first-to-second year persistence of three percentage points. This figure was not statistically significant, but a much larger impact (more than six percentage points) was significant for males. This is important partly because the program ended after the first-year, so any effect on second-year enrollment would have been based on residual benefits from the first year rather than the expectation of continued funding. This implies a three percentage point impact of aid only and an adjusted ECR of 0.0065, slightly below current productivity in the four-year sector. The ECR for the combined financial aid and service is roughly half that size because the services almost doubled the costs and the impact size remained roughly the same.

Like Canada STAR, MDRC's Opening Doors also included an experiment that combined services with performance-based aid. MDRC has released a series of reports suggesting that performance-based financial aid increases credit accumulation and enrollment in classes between first and second semester (Richburg-Hayes et al., 2009). More recently, they summarize new findings that the program increased persistence from roughly 31 percent to 37.5 percent over four semesters, for an effect of 6.5 percentage points. This suggests the adjusted ECR is 0.0132, which is below current overall productivity.

Student-faculty ratios. We found several studies of the effects of class size on achievement and these tend to suggest that smaller classes do yield more learning.²² Because achievement effects cannot be readily translated into graduation rates, we rely on the recent work of Bound, Lovenheim, and Turner (2009). They find that reducing the student-faculty ratio by one increases degree completion by 1.11 (0.03) percentage points. (The large differences

²² Kokkelenberg, Dillon, and Christy (2008) find that larger classes lead to lower grade-point averages in one U.S. public university. De Paola and Scoppa (2009) study class size using data from Italy.

between 4-year and 2-year results here are noteworthy.) While these results are based on fairly simple regression analyses, we do report the results because this is such an important component of college costs. The adjusted ECR is 0.0116 (0.0016), just below current productivity levels in the four-year sector, and far below it in the two-year sector.

Full-time faculty and adjuncts. Ehrenberg and Zhang (2004) estimate the effects of full-time faculty by comparing graduation rates and percent full-time faculty across time within colleges (an “interrupted time series” quasi-experiment). Using this approach, their results imply that reducing the percent part-time by one percentage point would reduce the graduation rate by 0.14 percentage points. Multiplying this by 33 percent (the actual percent part-time) implies that eliminating part-time faculty would increase the graduation rate by 4.6 percentage points. They do not report results for two-year colleges, but Jacoby (2006) reports an almost identical finding for two-year sector: increasing the full-time faculty by one percentage point increases the graduation rate by 0.15 percentage points. To move from 33% to 100% full-time would therefore increase graduation rates by 10 percentage points.²³ The adjusted ECRs are 0.0405 (0.2280). The figure is much higher in the two-year sector because the cost of switching to full-time faculty is much lower compared with the four-year sector, as reported earlier. However, both are above current productivity rates.

Remediation. While some studies have identified positive short-term impacts on early persistence (Attewell et al. 2006; Calcagno & Long, 2008), two rigorous studies find no impact on degree completion (Calcagno & Long, 2008; McFarlin and Martorell, 2007). However, a quasi-experiment by Bettinger and Long (2005) find that remediation increases the probability of

²³ Not all the evidence on part-time faculty is so positive; Bettinger and Long (forthcoming), using a quasi-experimental technique, find that adjuncts increase interest in subsequent course enrollment, relative to full-time faculty.

receiving a degree by 10 percentage points.²⁴ This leads to a bit of a conundrum. If we accept the McFarlin and Martorell results, then the ECR is zero, but if we accept Bettinger and Long, the ECR is 0.0588 and above overall productivity.²⁵ This is one of the few program areas where we have multiple rigorous studies.

Student services. Webber and Ehrenberg (2009) point out that spending on non-instructional student services such as student organizations, intramurals, student health services (including psychological counselling) supplemental instruction (e.g., tutoring), and admissions and registrar offices have grown more rapidly in recent years than instructional expenditures. They use institution-level data, from the Integrated Postsecondary Education Data System (IPEDS), to study the potential impacts of different categories of student services as well as other typical categories college spending.²⁶ Using interrupted time series techniques, they find that spending on student services tends to increase student persistence, especially at colleges where students have low college entrance exam scores and lower family incomes. Instructional spending is also positively associated with graduation. Specifically, they find that a \$500 per student increase in student services spending would increase the college graduation rate by 0.7 percentage points (compared with 0.3 percentage points for instructional spending). This yields an adjusted ECR of 0.0038, well below current productivity.

Student counseling. Other studies focus on more specific student services programs. The MDRC Opening Doors initiative was studied with a randomized trial. Impacts were statistically

²⁴ Bettinger and Long (2005) specifically use an instrumental variables (IV) that takes advantage of the fact that: (a) different colleges in Ohio have different remediation policies; and (b) different students are located in closer proximity to, and are therefore more likely to attend, colleges with policies that affect whether they are placed in remediation.

²⁵ The point estimates in McFarlin and Martorell (2007) are a fraction the size of Bettinger and Long (2005), so even if we ignored statistical significance the ECR based on the McFarlin and Martorell estimate would be close to zero.

²⁶ We did not discuss this study in the cost sections because, like the financial aid studies, the costs are already expressed in dollar terms. The IPEDS categories are: instructional, academic support (libraries, museums, academic computing), research, and student services.

significant during the one year of the services were provided, though most of the initial effects diminished over time. The treatment increased persistence by seven percentage points in the first semester by the end of the first year after which the program was stopped. Follow-up analysis suggests that the post-program impact was cut in half (to 3.7) the first full semester after the program stopped and declined further thereafter. It is unclear what would have happened had the program continued. The impacts might have diminished even if the program had continued. More plausibly, the total impact of the program might be reflected in the impact measured at the time the program ended: increasing the graduation by 3.7 percentage points. As an upper bound, consider that the impacts could have continued to accumulate if the program continued—i.e., the program might have impacted the persistence rate and the initial benefits might have compounded. We take 3.7 percentage points as a middle ground estimate of the impact on graduation.

The adjusted ECR is 0.0218, which is six times larger than that suggested by Webber and Ehrenberg. While general spending on student services reflects somewhat different types of services than in this experiment, and notwithstanding Webber and Ehrenberg's careful analysis, the large differences in ECRs may suggest that the quasi-experimental studies are biased downwards. Also, note that Webber and Ehrenberg focused on four-year students while Opening Doors focused on two-year students.

Discussion of cost-effectiveness results. Table 2 reports the results of the cost-effectiveness analysis for those programs where we have convincing evidence of impacts. The second column reflects the costs from Table 1 (multiplying the unadjusted costs by the cost multiplier). The third column indicates the outcome variable used in the study, which is necessary for identifying the appropriate multiplier (for the studies of enrollment and

persistence), as well as for providing an overall picture of the strength of evidence. The fourth column reflects the impact reported on that outcome in the study and this is followed by the appropriate multiplier to obtain estimated graduation impacts. Finally, the last column indicates ratio of the multiplier-adjusted impact divided by the multiplier-adjusted cost—the ECRs discussed throughout this section.

The top two most cost-effective programs include both the widely used and the nearly unknown. Call centers appear most cost-effective. The reported ECR of 0.6288 is larger than all the others and implies that we could create more than one-half of one additional graduate for only \$1,000 per student. The ECR for full-time faculty in the two-year sector comes next at about one-third that size, or 0.2280.

The remediation ECR is the next largest at 0.0588, but this estimate is somewhat uncertain because another rigorous study finds no statistically significant effects. The variation in measured effects could be related to differences in the specific policy designs. For example, students are placed in remediation based on a cut-off on a test score and the impact is likely to depend on where that cut-off is placed. The importance of program design and implementation may also help explain why no strategy stands out—cost-effectiveness may depend as much on execution of specific programs as the general strategy into which the programs fall.

Alternatively, it may be that there is simply not enough good evidence yet with which to identify a strategic pattern. In any event, except for call centers, full-time faculty, and remediation, it does not appear that any of the remaining programs and strategies increase productivity over current levels.

There are no clear patterns in the types of strategies likely to deserve the most attention. Two of the three most cost-effective strategies—remediation and full-time faculty—both fall into

the instruction category, though call centers, as the most cost-effective option in the list, falls under the student services umbrella. Financial aid, even when combined with student services, does not seem likely to improve overall productivity.

The majority of the ECRs are related to four-year colleges and bachelor's degrees, but the distinction between that and two-year degrees should not be forgotten. Two of the three most cost-effective strategies were identified with data from the two-year sector (call centers and full-time faculty). In the case of call centers, this might be an area where one sector can learn from the other, once additional corroborating evidence is available. On the other hand, we also see evidence that programs, such as student services, that appear to be effective in one sector (MDRC Opening Doors study of two-year colleges) seem ineffective in another sector (the Canada STAR study of a four-year college).

It may not be a coincidence that the options appearing to be most cost-effective are also the ones with where our confidence in the evidence is lowest. We included call centers because the idea is unusual, it is used more widely in the for-profit sector, and it provides a good example of a very inexpensive option that could have real impacts. The evidence on full-time faculty is based on one interrupted times series study and one simple regression analysis.

Some programs that are not cost-effective overall might still be worthwhile because they target disadvantaged groups and therefore increase equity. Upward Bound appears much less cost-effective than counseling and performance-based aid that are also targeted to this disadvantaged group. Remediation could be added to the list because, while family background is not an explicit consideration in assignment to remediation, disadvantaged students are more likely to have the lower level of academic skills and test scores that trigger remediation

participation. In this respect, a well-executed remediation may be the most cost-effective option for this group.

The results of this cost-effectiveness analysis are, however, different when we switch to a cost-benefit framework. Dynarski (2008) finds that financial aid passes a simple cost-benefit test, without comparison to other programs. The evidence summarized in Table 2 suggests that other programs would pass a cost-benefit analysis more easily than financial aid and such programs should be adopted before programs that pass the cost-benefit test less easily. The analysis here suggests that there are likely more cost-effective uses of those resources than financial aid.

Some Programs May Reduce Costs While Maintaining Quality

While we have chosen degree attainment as the primary outcome of interest, we can still learn something important by considering effects on quality measures like achievement. If programs save resources and do not reduce quality, then this suggests there may be no reduction in persistence and degrees—in fact, if the saved resources were used to reduce tuition or expand other cost-effective programs, they could improve other outcomes. We specifically focus on programs and systems that represent a break from the tradition of relying on “market accountability” from students and parents (based on U.S. News and other information) and requiring a fixed amount of seat time.

Competency-based learning is a prime example of breaking the higher education mold. Underlying the cost disease explanation is an assumption that education is like a symphony—an educational “concert” played in front of an audience. The competency-based approach sets aside this assumption and allows students and faculty to focus their attention on what each student needs to learn rather than (perhaps inefficiently) requiring a fixed amount of seat time and giving

all students a standardized education. This approach has become a significant part of the Bologna Process in Europe, as a potential way to rethink higher education (Adelman, 2008).

We discussed earlier the Program in Course Redesign (PCR) as one competency-based example. The Western Governors' University (WGU), a private online university built on the competency-based model through an alliance of state governors in western U.S. states, has had students graduate in as little as six months paying just \$2,790 for tuition (Witkowsky, 2006). This is a fraction of the cost of traditional credit-oriented degrees. The lower cost can be attributed to two main factors: fully online programs require very little physical infrastructure and competency-based programs demand less time of students and faculty than traditional colleges.²⁷

Yet, it is difficult to directly compare this model, with a goal of competency, with traditional higher education and its more varied goals. Part of the value of higher education is as a signal to employers—not just of academic competencies but persistence and social skills that competency-based models do not clearly reflect. Also, the “impact” of the competency-based approach on degree attainment is likely to derive as much from lower demands as anything else. With competency-based learning, there is very likely a trade-off between quality and degrees, but the size of the trade-off remains unclear.

“Accelerated” degree programs share with competency-based learning the shortened time to degree, though most maintain an orientation toward seat time and credit hours. For accelerated programs that only compress credit hours into a shorter span, there is likely to be

²⁷ Recall that the cost/degree estimates used in the above break-even analysis do not incorporate students' opportunity costs because those programs do not alter student time, but the situation is different with competency-based learning. Suppose that average BA degree requires 120 credits, or 40 15-week courses, and six hours per week per course in total time. This means 3,600 student hours. At \$7 per hour, this amounts to \$25,200 per degree. If the time to degree is cut in half with the competency-based model, then this saves resources worth \$12,600 in student time alone. Another way to see this is to note our earlier figure that the cost of student time is similar to the cost of faculty, which is one of the largest budgetary costs of college.

little cost savings. The opportunity costs of student and faculty time—the main costs of higher education—are unchanged, though this approach might improve retention for students who are in hurry to finish.

The Lumina Foundation's effort to increase productivity has called for similar approaches that reward non-traditional colleges like WGU. The foundation also recommends providing rewards to colleges and students for the completion of courses, degrees, and certificates. As in K-12 education, there is also growing interest in “accountability” policies that would hold colleges responsible for outcomes ranging from achievement scores to employment outcomes (Goldrick-Rab, Harris, Mazzeo, & Kienzl, 2009; Spellings Report, 2009).

Having more students start in community colleges instead of four-year colleges is another commonly considered idea (see above differences in costs per degree). Currently, students who start off at two-year colleges are less likely to receive degrees, even when comparing students with similar aspirations and background characteristics (Rouse, 1995; Leigh & Gill, 2003; Gonzales, Hilmer, & Sandy, 2006). While it remains unclear how these less positive outcomes balance out against the lower costs, one recent study finds that there might be no cost savings at all from having more students start in two-year colleges, partly because two-year colleges have much smaller class sizes (Romano & Djajalaksana, 2010).

Even more so than the programs for which we can estimate ECRs, systemic reforms, like accountability, competency-based learning, and community college transfers have low costs and what remains is considerable uncertainty about impacts, both on degrees and the quality of those degrees. The fact that the rationale for these systemic reforms is focused on cost reduction does raise red flags about the potential reduction in quality.

Conclusion

The combination of growing demand for college credentials and declining degree productivity is a serious problem. It will be very difficult to reach the lofty education goals that policymakers are setting without improving productivity.

We find, in contrast to the “cost disease” and the larger debate on higher education, that some productivity improvements are possible. Some programs are extremely expensive with little evidence to justify those high costs. Other programs such as call centers and online learning appear to have more credibility than is often assumed. While we are cautious about making comparisons between studies of graduation and studies where evidence is available about graduation, the differences in ECRs are so large in many cases that it seems unlikely the results are driven by the need to convert the effects to a common metric.

We do not claim that colleges could get back to productivity rates from decades past. The cost disease and other pressures driving costs up and degrees down are real, powerful, and to a certain extent unavoidable. But that should not excuse the dearth of rigorous evidence or the failure to include costs as part of the conversation. The absence of the type of information that would be needed to improve productivity—a hole that this study helps to fill—is perhaps the strongest evidence that we are falling short of our productivity potential.

The limitations of research fall partly on researchers. Data to study higher education are increasingly available from research purposes through, for example, the National Student Clearinghouse. Moreover, adding analysis of costs has long been possible but rarely carried out. We have provided a template on which future cost-effectiveness research can be based, but this should not be left to just the occasional research synthesis. Every study of impacts should at least briefly discuss program costs, or else they tell only half the story (Harris, 2009).

The larger point is that colleges are not completely helpless in addressing productivity as some appear to assume. Our results suggest a need to break out of this mindset, to actively search for new and better ways to help students, and to study program costs and effects more carefully so that policymakers and college leaders can make more informed decisions.

Bibliography

- Abrami, Phillip C. Robert Bernard, Anne Wade, Richard F. Schmid, Eugene Borokhovski, Rana Tamin, Michael Surkes, Gretchen Lowerison, Dai Zhang, Iolie Nicolaidou, Sherry Newman, Lori Wozney, Anna Peretiatkowicz (2006). A Review of e-Learning in Canada: A Rough Sketch of the Evidence, Gaps and Promising Directions. *Canadian Journal of Learning and Technology*, 32(3).
- Adelman, C. (2008). *Learning accountability from Bologna: A higher education policy primer*. Institute for Higher Education Policy: Washington, DC.
- Albee, Amy (2005). A Cost Analysis of Upward Bound And GEAR UP. Unpublished manuscript. Tallahassee, FL: Florida State University.
- Angrist, J., Lang, D., Oreopoulos, P. (2006). Lead them to the water and pay them to drink: An Experiment with services and incentives for college achievement. Working Paper 12790. Cambridge, MA; National Bureau of Economic Research.
- Angrist, J., Lang, D., Oreopoulos, P. (2009). Incentives and Services for College Achievement: Evidence from a Randomized Trial, *American Economic Journal: Applied Economics* 1(1) 136-163.
- Archibald, R.B. & Feldman, David H. (2008a). Explaining Increases in Higher Education Costs. *The Journal of Higher Education*, 79 (3), 268-295.
- Archibald, R.B. & Feldman, David H. (2008b). Why do higher education costs rise more rapidly than prices in general? *Change*, May/June, 25-31.
- Arum, R., Roksa, J. & Velez, M. (2008). *Learning to reason and communicate in college: Initial report of findings from the CLA longitudinal study*. Social Science Research Council: NY.
- Attewell, P. A., Lavin, D., Domina, T., & Levey, T. (2006). New Evidence on College Remediation. *Journal of Higher Education*. 77.5 886-924
- Baumol, W. J., & Bowen, W. G. (1966). *Performing arts: The economic dilemma*. New York: Twentieth Century Fund.
- Baumol, W. J., & Blackman, S. A. (1995). How to think about rising college costs. *Planning for Higher Education*, 23(Summer), 1-7.
- Bettinger, Eric P. & Bridget Terry Long (2004). Shape up or ship out: The effects of remediation on students at four-year colleges. NBER Working Paper 10369. Cambridge, MA: National Bureau of Economic Research.
- Bettinger, Eric P. & Bridget Terry Long (2005). Addressing the needs of under-prepared students in higher education: Does college remediation work? Working Paper 11325

Cambridge, MA: National Bureau of Economic Research.

- Bettinger, Eric P. & Bridget Terry Long (2006) Remedial and developmental courses. In Stacy Dickert-Conlin and Ross Rubenstein, *Economic Inequality and Higher Education: Access, Persistence, and Success* (pp.69-100). New York: Russell Sage.
- Bettinger, Eric P. & Bridget Terry Long (forthcoming). Does Cheaper Mean Better? The Impact of Using Adjunct Instructors on Student Outcomes. *Review of Economics and Statistics*.
- Black, Dan and Jeffrey Smith. 2006. "Evaluating the Returns to College Quality with Multiple Proxies for Quality." *Journal of Labor Economics* 24(3): 701-728.
- Bound, J., Lovenheim, M. & Turner, S. (2009). *Why Have College Completion Rates Declined?* NBER Working Paper No. 15566.
- Bowen, H. R. (1980). *The costs of higher education: How much do colleges and universities spend per student and how much should they spend?* San Francisco: Jossey-Bass.
- Bowen, W. G., Chingos, M.M., and McPherson, M. (2009). *Crossing the Finish Line: Completing College at America's Public Universities*. Princeton University Press.
- Breneman, D. W. (2001). An essay on college cost. In *Study of college costs and prices, 1988-89 to 1997-98, Volume 2: Commissioned papers* (NCES 2002-158, December 13-20, pp. 13 – 20). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.
- Brown, C. & Belfield, C. (2002). How cost-effective are lectures? A review of the experimental literature, *Cost-Effectiveness and Educational Policy*, Henry M. Levin & Patrick J. McEwan eds., Eye on Education: Larchmont, NY.
- Calcagno, J.C. & Long, B.T. (2008). The Impact of Postsecondary Remediation Using a Regression Discontinuity Approach: Addressing Endogenous Sorting and Noncompliance. NBER Working Paper 14194 (July 2008).
- Cameron, S. V., & Taber, C. (2004). Estimation of educational borrowing constraints using returns to schooling. *Journal of Political Economy*, 112(1), 132–182.
- Cavanaugh (2005). *The Effects of Distance Education on K-12 Student Outcomes: A Meta-Analysis*.
- Clotfelter, Charles T. (1996). *Buying the Best: Cost Escalation in Elite Higher Education*. Princeton: Princeton University Press.
- Constantine, J. M., Seftor, N., Martin, E. S., Silva, T., & Myers, D. (2006). A study on the effect of the Talent Search program on secondary and postsecondary outcomes in Florida,

Indiana, and Texas. Washington, DC: U.S. Department of Education, Office of Planning, Evaluation, and Policy Development.

Cornwell, C., David B. Mustard & Deepa J. Sridhar (2006). The Enrollment Effects of Merit-based Financial Aid: Evidence from Georgia's HOPE Program. *Journal of Labor Economics*, 24(4), 761-786.

Delta Cost Project (2009). *Trends in College Spending*. Washington, DC.

De Paola, Maria and Vincenzo Scoppa (2009) Effects of Class Size on Achievement of College Students. Munich Personal RePEc Archive Working Paper #16945. Downloaded September 1, 2009 from: <http://mpa.ub.uni-muenchen.de/16945/>.

Deming, D. & Dynarski, S. (2009). Into college, out of poverty? Policies to increase the postsecondary attainment of the poor. Working Paper 15387. Cambridge, MA: National Bureau of Economic Research.

Des Moines Area Community College (2009). *DMACC Call Center Cost Benefit Analysis-FY07-08*. Unpublished manuscript.

Domina, T. (2009). What Works in College Outreach: Assessing Targeted and Schoolwide: Interventions for Disadvantaged Students, *Educational Evaluation and Policy Analysis*, 31(2), 127-152.

Dowd, Alicia C. & Laura M. Ventimiglia (2008). A Cost Estimate of Standards-Based Remediation in a Community College Developmental Education Program. Unpublished manuscript.

Dynarski, S. (2003a) Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion, *American Economic Review*, 279-288.

Dynarski, Susan (2003b). Loans, Liquidity, and Schooling Decisions. Unpublished working paper. University of Michigan. Downloaded June 3, 2010 from: http://www.nber.org/dynarski/Dynarski_loans.pdf.

Dynarski, S. (2008). Building the stock of college-educated workers. *Journal of Human Resources*, 43(3), 576-610.

Ehrenberg, R. G. (2000). Tuition rising: why college costs so much. Cambridge, MA: Harvard University Press.

Ehrenberg, R.G. & Zhang, L. (2004). Do Tenured and Tenure-Track Faculty Matter? *NBER Working Paper 10695*. Cambridge, MA; National Bureau of Economic Research.

Garms, Walter I. (1971) "A Benefit-Cost Analysis of the Upward Bound Program." *The Journal of Human Resources* v.6 n.2. 206-20.

- Geckeler, C., Beach, C., Pih, M., & Yan, L. (2005). *Helping Community College Students Cope with Financial Emergencies: Lessons from the Dreamkeepers and Angel Fund Emergency Financial Aid Programs*. Oakland: Manpower Demonstration Research Corporation.
- Getz, M., & Siegfried, J. J. (1991). Cost and productivity in American colleges and universities. In C. Clotfelter, R. Ehrenberg, M. Getz, & J. J. Siegfried (Eds.), *Economic challenges in higher education* (261–392). Chicago: University of Chicago Press.
- Glater, J. & Finder, A. 2006. “In tuition game, popularity rises with price.” *The New York Times*. December 12.
- Goldin, C. & Katz, L. (2008). *The Race Between Education and Technology*. Cambridge: Harvard University Press.
- Goldrick-Rab, S., Harris, D., Maseo, C., Kienzl, G. (2009). *Transforming America’s Community Colleges: A Proposal to Expand Opportunity and Promote Economic Prosperity*. Washington, DC: The Brookings Institution.
- Gonzalez, A., Hilmer, M.J., & Sandy, J. (2006). “Alternative Paths to College Completion: The Effect of Attending a Two-Year School on the Probability of Completing a Four-Year Degree,” *Economics of Education Review* 25(4): 463-471.
- Gordon, S., He, W., & Abdous, M. (2009). Using a Web-Based System to Estimate the Cost of Online Course Production. *The Online Journal of Distance Learning Administration*, 12(3).
- Harris, D. (2007). “Class size and school size: Taking the trade-offs seriously.” In F.M. Hess and T. Loveless (Eds.), *Brookings Papers on Education Policy 2006-2007* (pp.137-161). Washington, DC: Brookings Institution.
- Harris, D. (2009). Toward policy-relevant benchmarks for interpreting effect sizes: Combining effects with costs. *Educational Evaluation and Policy Analysis*. 31(1), 3-29.
- Hauptman, Arthur & Young Kim (2009). *Cost, Commitment, and Attainment in Higher Education: An International Comparison*. Lumina Foundation and Jobs for the Future: Boston, MA
- Immerwahr, J., Johnson, J. and Gasbarra, P. (2008). *The iron triangle: College presidents talk about costs, access, and quality*. The National Center for Public Policy and Higher Education and The Public Agenda.
- Jacoby, D. (2006). Effects of Part-Time Faculty Employment on Community College Graduation Rates, *The Journal of Higher Education*, Vol. 77, No. 6, pp. 1081-1103.
- Jenkins, D. (2002). The potential of community colleges as bridges to opportunity for the disadvantaged: Can it be achieved on a large scale? Unpublished manuscript. Great Cities

Institute, University of Illinois at Chicago.

- Johnson, Nate (2009). *What does a college degree cost?* Washington, DC: Delta Cost Project.
- Jones, D. & Wellman, J. (2009) "Bucking Conventional Wisdom on College Costs." *Inside Higher Ed*. July 20. <http://www.insidehighered.com/views/2009/07/20/wellmanjones>
- Kane, T. (1995). Rising public college tuition and entry: How well do public subsidies promote access to college? NBER Working Paper 5164. Cambridge, MA: National Bureau of Economic Research.
- Kane, T. (2003). A Quasi-Experimental Estimate of the Impacts of Financial Aid on College-Going. Working Paper 9703. Cambridge, MA: National Bureau of Economic Research.
- Kane, T. (2007). Evaluating the Impact of the D.C. Tuition Assistance Grant Program, *Journal of Human Resources*, 42(3), 555-582.
- Kokkelenberg, C., Dillon, M. and Christy S. (2008), "The Effects of Class Size on Student Grades at a Public University", *Economics of Education Review*.....
- Leigh, D.E. & Gill, A.M. (2003). "Do Community Colleges Really Divert Students from Earning Bachelor's Degrees?" *Economics of Education Review*, 22(1): 23-30.
- Lipscomb, J., Weinstein, M.C., & Torrance, G.W. (1996). Time preference. In M.R. Gold, L.B. Russell, J.E. Siegel, & M.C. Weinstein (Eds), *Cost-Effectiveness in Health and Medicine* (pp.214-246). New York: Oxford University Press.
- Massy, W. F. (1996). Productivity issues in higher education. In W. F. Massy (Ed.), *Resource issues in higher education* (pp. 49–86). Ann Arbor: University of Michigan Press.
- McPherson, Michael, and Schapiro, Morton Owen, and Gordon C. Winston. (1993). *Paying the Piper: Productivity, Incentives, and Financing in U.S. Higher Education*. Ann Arbor: The University of Michigan Press.
- Means, Barbara, Yukie Toyama, Robert Murphy, Marianne Bakia, and Karla Jones (2009). *Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies*. U.S. Department of Education Office of Planning, Evaluation, and Policy Development Policy and Program Studies Service. Washington, DC.
- Moore, M.A., Boardman, A.E., Vining, A.R., Weimer, D.L. & Greenberg, D.H. (2004). Just give me a number! Practical values for the social discount rate. *Journal of Policy Analysis and Management*, 23, 789-812.
- Morgan, Brian M. (No Date) *Is Distance Learning Worth It? Helping to Determine the Costs of Online Courses*. Unpublish manuscript, Marshall University.

- Muennig, P. (2002). *Designing and conducting cost-effectiveness analyses in medicine and health care*. San Francisco: Jossey-Bass.
- Murray, C. (2008). "For Most People, College is a Waste of Time." *The Wall Street Journal*. August 13. P. A17.
- Myers, D., Olsen, R., Seftor, N., Young, J., & Tuttle, C. (2004). The impacts of regular Upward Bound: Results from the third follow-up data collection. Washington, DC: Mathematica Policy Research.
- Provasnik, S., and Planty, M. (2008). Community Colleges: Special Supplement to The Condition of Education 2008.
- Romano, R. & Djajalaksana, Y.M. (2010). Using the Community College to Control College Costs: How Much Cheaper Is It? Unpublished working paper.
- Rouse, C.E. (1995). Democratization or Diversion? The Effect of Community Colleges on Educational Attainment." *Journal of Business and Economic Statistics*, 13(2): 217-224.
- Shadish, W.R., Cook, T.D., and Campbell, D.T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. New York: Houghton Mifflin.
- Seftor, Neil S. Mamun, Arif and Schirm, Allen (2008) The Impacts of Regular Upward Bound on Postsecondary Outcomes 7-9 Years After Scheduled High School Graduation. Final Report. Washington, DC: Mathematica Policy Research, Inc.,. November 2008
- Smith, J. 2008. "Heterogeneity and Higher Education." In Michael S. McPherson and Morton Owen Schapiro, eds. *Succeeding in College: What it Means and How to Make it Happen*. (New York: College Board, 2008).
- Stinebrickner, T., & Stinebrickner, R. (2003). Understanding educational outcomes of students from low-income families: Evidence from a liberal arts college with a full tuition subsidy program. *Journal of Human Resources*, 38(3), 591–617.
- Terry, Brooke D. (2007). *The Costs of Remedial Education*. Austin, Texas: Texas Public Policy Foundation.
- The National Center for College Costs (2010). *Indiana College Costs 2009-2010*. Downloaded June 3, 2010 from:
<http://www.collegecosts.com/HotTopics/IndianaCollegeCosts/tabid/102/Default.aspx>.
- Twigg, C. A. (2003). Improving quality and reducing costs: Designs for effective learning. *Change*, 35(July-August), 22–29.
- U.S. News and World Report (2010). *Methodology: Undergraduate Ranking Criteria and Weights*. Downloaded March 8, 2010 from:

<http://www.usnews.com/articles/education/best-colleges/2009/08/19/methodology-undergraduate-ranking-criteria-and-weights.html>

U.S. Department of Education (2004), National Center for Education Statistics, 2004 National Study of Postsecondary Faculty.

U.S. Department of Education (2010a), National Center for Education Statistics. 2003-04 through 2006-07 Integrated Postsecondary Education Data System (IPEDS). Table 362: Expenditures of public degree-granting institutions, by purpose of expenditure and type of institution: 2003-04 through 2006-07.

U.S. Department of Education NCES (2010b), National Center for Education Statistics. Higher Education General Information Survey (HEGIS), "Fall Enrollment in Colleges and Universities" surveys, 1969 through 1985; and 1986 through 2008 Integrated Postsecondary Education Data System, "Fall Enrollment Survey" (IPEDS-EF:86-99).

U.S. Department of Education (2010c), Institute for Education Sciences, What Works Clearinghouse. Downloaded March 1, 2010 from: www.whatworks.ed.gov.

Vedder, R. (2004). *Going Broke by Degree: Why College Costs Too Much*. Washington, DC: AEI Press.

Vedder, R. (2007). *Over-invested and over-priced: American higher education today*. Center for College Affordability and Productivity: Washington DC

Webber, Douglas A. and Ronald G. Ehrenberg (2009) Do expenditures other than instructional expenditures affect graduation and persistence rates in American higher education. National Bureau of Economic Research. Cambridge, MA.

Weimer, D.L. (2009). "Making Education Research More Policy-Analytic," in Gary Sykes, Barbara Schneider, and David N. Plank, eds. *Handbook of Education Policy Research*. New York: Routledge, 93–100.

Weimer, D.L. & Vining, A.R. (2005) *Policy Analysis: Concepts and Practice*. Upper Saddle River, NJ: Pearson.

Weisbrod, B. Ballou, & Asch, (2008). *Mission and Money: Understanding the University*. New York: Cambridge University Press.

Winston, G. C. (1999). Subsidies, hierarchy and peers: The Awkward economics of higher education. *Journal of Economic Perspectives*, 13(Winter), 13–36.

Zhang, L. (2005). Do measures of college quality matter? The effect of college quality on graduates' earnings. *Review of Higher Education*, 28(4), 571–596.

Table 1: Costs and “Break-Even” Effects of Higher Education Programs

<i>Program</i>	<i>Unadjusted Cost (2007 \$)</i>	<i>Cost Multiplier</i>	<i>Break-Even Impact</i>
<i>College Access (disadv)</i>			
GEAR UP	264,000	3.00	13.20
Talent Search	39,200	1.50	0.98
Upward Bound	677,000	1.50	16.93
<i>Bridges to Opportunity</i>			
Vocational ESL	68,400	1.00	1.14
Workplace Basics	68,400	1.00	1.14
Technology Career Bridge	479,100	1.00	7.99
<i>Financial Aid</i>			
Grants	100,000	3.73	6.22
Loans	43,000	3.73	2.67
Merit Aid (GA/AR)	100,000	3.73	6.22
Merit Aid (Canada STAR)	36,600	3.73	2.28
Dreamkeepers (2y; disadv)	39,300	1.00	1.57
Angel Funding (2y; disadv)	26,600	1.00	1.06
<i>Financial Aid w/ Services</i>			
Canada STAR	73,900	3.73	4.59
Opening Doors (2y; disadv)	158,300	2.29	14.53
Indiana 21st Century	517,200	3.73	32.15
<i>Instruction</i>			
Stud/Fac Ratio (4y)	25,561	3.73	1.59
Stud/Fac Ratio (2y)	8,064	2.29	0.74
Full-time faculty (4y)	30,425	3.73	1.89
Full-time faculty (2y)	19,153	2.29	1.75
Online education (Marshall Univ.)	80,000	3.73	4.97
<i>Other Modes of Instruction</i>			
Independent Study	-48,204	3.73	-3.00
Discussion/Inquiry	-45,396	3.73	-2.82
Personalized Instruction	286,650	3.73	17.82
Other	131,976	3.73	8.20
Remediation	170,000	1.00	2.83
<i>Student Services</i>			
Student Services (Webber/Her)	50,000	3.73	3.11
Student Counseling (2y; disadv)	54,898	2.29	5.03
Call Centers (2y)	1,000	2.29	0.09

Notes: Unadjusted costs come from the authors’ analysis described in the text. Cost multipliers are discussed in the methods section and the Technical Appendix (available upon request). The break-even impact in the final column is calculated by solving for “Impact” in the following: $(Impact / (Unadjusted Cost * Cost Multiplier)) = (1 / (Current Average Cost per Degree))$.

Table 2: Effectiveness-Cost Ratios for Higher Education Programs

<i>Program</i>	<i>Adjusted Cohort Cost (2007 \$)</i>	<i>Outcome Variable</i>	<i>Unadjusted Impact</i>	<i>Impact Multiplier</i>	<i>ECR</i>
<i>Average Costs per Degree</i>					
<i>2-year degree</i>	--	--	--	--	<i>0.0400</i>
<i>4-year degree</i>	--	--	--	--	<i>0.0170</i>
<i>College Access (disadv)</i>					
Upward Bound	1,015,500	Enrollment	2.00	0.43	0.0008
<i>Financial Aid</i>					
Grants	373,000	Enrollment	3.00	0.70	0.0056
Loans	160,390	Enrollment	2.00	0.70	0.0087
Perf. Aid (GA/AR)	373,000	Enrollment	4.00	0.70	0.0075
Perf. Aid (Canada STAR)	136,518	Persistence	3.00	0.60	0.0132
<i>Financial Aid w/ Services</i>					
Canada STAR	275,647	Persistence	3.00	0.60	0.0065
Opening Doors (2y; disadv)	363,194	Persistence	6.50	0.74	0.0132
<i>Instruction</i>					
Stud/Fac Ratio (4y)	95,343	Graduation	1.11	1.00	0.0116
Stud/Fac Ratio (2y)	18,467	Graduation	0.03	1.00	0.0016
Full-time faculty (4y)	113,485	Graduation	4.60	1.00	0.0405
Full-time faculty (2y)	43,860	Graduation	10.00	1.00	0.2280
Remediation	170,000	Graduation	10.00	1.00	0.0588
<i>Student Services</i>					
Student Services (Webber/Her)	186,500	Graduation	0.70	1.00	0.0038
Student Counselling (2y; disadv)	125,716	Persistence	3.70	0.74	0.0218
Call Centers (2y)	2,290	Persistence	2.00	0.72	0.6288

Notes: The adjusted cohort cost comes from multiplying the unadjusted cohort costs in Table 1 by the cost multiplier. The unadjusted effect is what was reported by the authors for the “outcome variable” studied. The effect multiplier translates this into a ECR that is comparable across programs (subject to the assumptions discussed in the text). If the ECR is greater than the average cost per degree in the top rows for the particular sector, then the program would increase college productivity.

Table 3: Impacts of Traditional Financial Aid

<i>Authors</i>	<i>Program Type</i>	<i>Dependent Variable</i>	<i>Impacts (Perc. Points/\$1,000)</i>
<i>Need-Based Aid</i>			
Kane (1995)	Tuition	Enrollment	two-year: 3 four-year: 1
Kane (2007)	DC Tuition	Enrollment	3-4
Dynarski (2003a)	Grant (Need)	Enrollment	3.6
Reyes (1995)	Loans (Need)	Enrollment	1.5
Dynarski (2005)	Loans (Need)	Enrollment	0-5.1
<i>Performance-Based Aid</i>			
Cornwell et al. (2006)	Georgia HOPE	Enrollment	4-6
Kane (2003)†	CalGrant	Enrollment	1.8
Angrist et al. (2006) *	Canada STAR	Persistence	8.2
Dynarski (2008)	GA/AR HOPE	Completion	0.9
MDRC *	Opening Doors	Persistence	2.2

Notes: Studies using randomized trials are indicated by an asterisk (*). All others use quasi-experimental methods.

† Results for the Kane (2003) study are obtained as follows: He concludes that the impact on enrollment was 3.5 percentage points. Based on tuition rates in the UC and CSU systems, we estimate average grant amounts of \$2,000 per recipient. Dividing the impact by this figure yields an impact of 1.8 per \$1,000.