*** PRELIMINARY, DO NOT QUOTE OR DISTRIBUTE ***

The (Un)Productivity of American Colleges: From "Cost Disease" to Cost-Effectiveness

Douglas N. Harris
Sara Goldrick-Rab

University of Wisconsin at Madison

March 19, 2010

Presented at the 2010 Annual Meeting of the American Education Finance Association, Richmond, VA

Abstract: We show that the productivity of American colleges and universities in terms of degree production is plummeting. This is partly due to rising costs that economists typically attribute to Baumol's well known "cost disease," but we demonstrate two other basic problems contributing to productivity declines. First, using cost-effectiveness analysis, we find that several popular programs and strategies in higher education are not cost-effective. Second, there is little rigorous evidence with which to judge the cost-effectiveness of most postsecondary programs, perhaps reflecting a larger absence of incentives to improve productivity. Rather than simply a "cost disease," we argue that the problem is more a "system disease"—one that is partly curable.

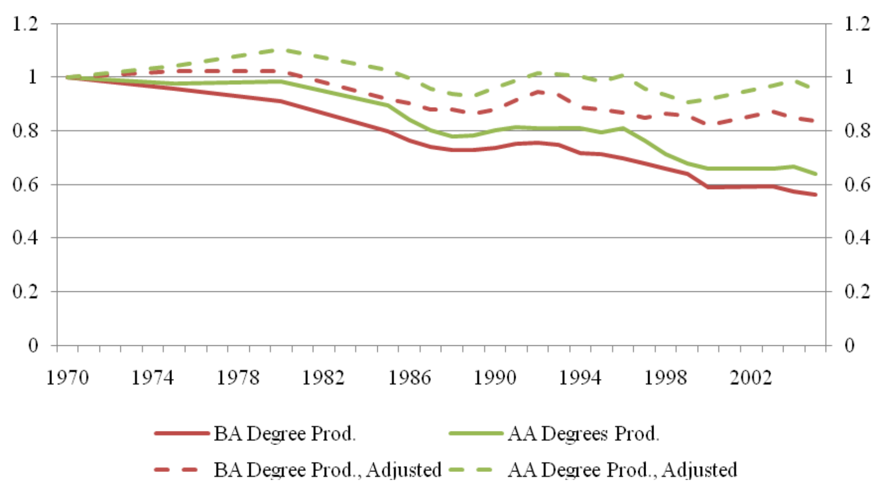
Acknowledgements: We gratefully acknowledge financial support of the Lumina Foundation. For their useful advice, we wish to thank Kristin Conklin, Noel Radomski, and John Wiley. For research assistance, we thank Gregory Kienzl, Regina Brown, Alan Nathan, Byoung-Ik Jeoung, and So Jung Park. All remaining errors are our own.

Educational Attainment and College Productivity

The productivity of American colleges and universities in terms of academic degrees is plummeting. 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). Despite some fluctuations in the last decade due to surging enrollments, total per-student spending is the highest in the world (Hauptman & Kim, 2009).

The number of degrees has also risen, but much more slowly. 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 shrunk. The class of 1972 and class of 1992 high school cohorts shows that eight-year college completion rates declined by 4.6 percentage points (from 50.5% to 45.9%) (Bound, Lovenheim, & Turner, 2009). The decline in productivity is shown more explicitly in Figure 1 with a direct comparison of degrees and the resources colleges use to produce them.

Figure 1: The Productivity Decline, 1970-2006
(Ratio of Degrees-to-Expenditures in Public Colleges)[1]



[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 constant 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.

Degrees granted divided by total sector expenditures is shown for each year, normalized with a base year of 1970. The (unadjusted) decline has been especially large in the four-year sector where today productivity is less than half what it was 40 years ago. 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. We also report productivity trends adjusted for the growth in overall labor costs in the economy (dashed lines), since colleges compete for labor with other industries. While this accounts for a substantial share of the productivity decline, it does not completely, especially looking at the trends since 1978 (more on this below). No matter how we look at the data, productivity has declined.

While some might argue that declines in productivity are the result of increases in the *quality* of higher education, in fact there is little evidence to support that argument (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 today's college students are gaining in terms of higher-order thinking skills (Arum, Roksa, & Velez, 2008).

Declining productivity is problematic for several reasons. First, while there is disagreement about to what degree we need to increase the number of degrees[2], President Obama and other national leaders have made increased degree attainment a national priority. Any attempt to reach this goal will be made more expensive by declining productivity. This is partly why the Spellings Commission report concluded that "we can take pride in our Nobel Prizes, our scientific breakthroughs, our Rhodes Scholars. But we must not be blind to the less inspiring realities of postsecondary education in our country" (2006, p.ix). The report not

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

only mentions “efficiency,” but laments the “lack of clear, reliable information about the cost and quality of postsecondary institutions” (p.x), the lack of accountability, and the out-dated nature of instructional and managerial practice. But the report provides little evidence to suggest exactly how colleges should move to enhance productivity.

What, then, accounts for the negative trend in productivity? A number of explanations have been given for this trend (Bowen, 1980; Getz & Siefried, 1991). The most common argument is that “cost disease” makes it inherently difficult to raise productivity in education and other service sectors (Baumol & Bowen, 1966). As Figure 1 shows, the increased cost of labor inputs has no doubt gone up, but there still appear to be declines. But the extension of this argument, that productivity should decline because colleges cannot use inputs more efficiently, is one we show below to be probably incorrect.

Others argue that increased access to higher education and corresponding increases in the enrollment of less qualified and less motivated students may reduce graduation rates.[3] 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 positive (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). These last two explanations are perhaps less convincing than the first

[3] Bound et al. (2009) find little support for the conclusion that declining preparation is responsible for the stagnant graduation rates, though new entrants might be less motivated. 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 explain declining productivity.

two simply because these problems have always existed and it is unclear why they would have worsened in recent years.

While these explanations have merit, the focus on them contributes to a perception among college leaders and scholars that college productivity is impossible to control. As Archibald and Feldman put it, "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). Also, 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. I've worked in my share of states, now six altogether. There are no more efficiencies to be had" (Immerwahr, Johnson & Gasbarra, 2008)[4], The clear implication is that any improvement in productivity necessitates sacrificing some other outcome.

In this paper, we provide evidence that productivity improvement is possible, and perhaps without any decline in quality. Specifically, we use an approach outlined by Harris (2009a) to calculate cost-effectiveness ratios for higher education programs that have been evaluated using rigorous methods. For most programs, this involves using program impacts reported in existing studies and then imputing costs based on publicly available information about program design. We find that many common hallmarks of quality are quite costly and that these costs are not matched by comparable impacts. This is especially true of programs aimed at disadvantaged students. In short, these programs appear to reduce productivity both on the average and for disadvantaged groups specifically.

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

Beyond our conclusions about specific programs, we lay out a methodology for analyzing productivity and identifying cost-effective programs. Costs are rarely considered in education research and even less often in higher education. We outline the issues involved and various ways to address them and argue for a new productivity agenda among higher education scholars.

A limitation of this research, and impediment to the future research agenda, is that there is little rigorous evidence about either the impacts of higher education programs. But the absence of evidence reinforces the conclusion that higher education has room for improvement in productivity. Colleges cannot raise productivity if they do not know which of their programs are least productive.

Cost-Effectiveness Analysis: Methodology

As noted in the last section, there are two significant limitations that make assessments of cost-effectiveness in higher education difficult. First, as we explain in more detail below, most analyses of programs and policies in this sector cannot be plausibly interpreted as causal impacts because they fail to address the standard selection bias problem. Second, even when causal impact estimates are available, they are disconnected from costs (Levin, 1991; Rice, 2002). These problems are particularly severe in higher education research because, for example, data systems are less centralized than K-12 and key outcomes such as labor market success require tracking students after they leave college.

The data problem is starting to diminish with the advent of the National Student Clearinghouse and federal incentives to create K-20 and employment data systems, but there is no guarantee that this will improve research that might inform decisions and improve productivity. Consideration of cost is especially rare in education compared with other areas of public policy (Monk and King, 1993; Rice, 2002). One reason is that economists generally have shown little interest in applying these economics-based techniques to education. Another is that the norms of practice in education research do not encourage consideration of costs. As a result, there are few studies that consider costs and therefore little basis of comparison for researchers who might wish to consider them. Harris (2009a) 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). This problem too is worse in higher education because, even when costs are considered, the focus is usually on the cost of enrolling more students rather than helping them graduate (Hossler, 2004). The perception that productivity is uncontrollable (because of the "cost disease") only adds to the problem.

Opportunity and Budgetary Costs

We use the standard economics-based approach to cost-effectiveness rooted in opportunity costs and operationalized through the "ingredients method" (Levin & McEwan, 2001). Opportunity costs are typically larger than budgetary costs. For example, the costs of students' own time can be quite large in relation to what colleges spend. 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 30x24=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 student opportunity cost per faculty member is $75,000, about as large as the opportunity costs of the faculty itself. Because budgetary costs will tend to under-state total economic costs, we focus more on the latter in order to capture all costs to society (Levin & McEwan, 2001).[5]

[5] A more ambiguous example of budgetary cost is financial aid. Cash transfers are budgetary costs that are considered opportunity costs only if they involve economic rents; that is, if the payments are in exchange for behavioral change. Otherwise, money is simply being transferred from one group to another. This means that if financial aid affects what students do then it is both a budgetary and an opportunity cost. More broadly, this example highlights a methodological challenge in estimating costs pertaining to persistence and graduation: the costs depend on the effects, since students who drop out do not receive services. However, because program

Economists also distinguish between marginal costs, which vary by the number of students, and fixed costs, which are unrelated to the number of the number of students. Average costs can differ greatly from the marginal costs, especially for very small changes in policy. However, since the underlying policy debate is about larger changes (e.g., President Obama's goal of being first in the world in college attainment), we take average cost calculations as reasonable estimates of marginal cost.

Since few studies estimate costs per student, we often have to impute costs in our analysis by making informed assumptions about each ingredient or type of economic resource. All costs in this study are expressed in 2007 dollars, unless otherwise noted. Costs are discounted at the standard rate of three percent, based on analysis by Barnett (1996), Lipscomb, Weinstein, & Torrance (1996), Moore et al. (2004), and Muennig (2002).[6] Also, the taxation necessary to garner resources for government programs incurs a loss in social welfare (Browning, 1987), called the "marginal cost of taxation" (MCT). We use a value of 25 percent, applied only to budgetary costs.

Calculating and Standardizing Cost-Effectiveness Ratios

To measure program productivity, we use the cost-effectiveness ratio (CER), or ratio of effects (E)-to-costs (C). Larger CERs imply greater productivity. To standardize the measures, we report the costs of implementing a program for an entire cohort of 100 students who enter college in a given year. For targeted programs (e.g., need-based financial aid), the cohort involves 100 students in the targeted population. Some programs are applied only

effects are generally small relative to typical (control group) persistence and graduation rates, ignoring the cost-effect linkage in this case introduces little bias in cost estimates. The authors wish to thank David Weimer for his advice on this issue.

[6] All of the studies cited in the text recommend a baseline discount rate of exactly three percent with the slight exception of Moore et al. (2004) who present different discount rates that depend, for example, on the likelihood that the intervention "crowds out" other investments. In the conditions of the present analysis, the Moore et al. (2004) discussion suggests a discount rate of 3.5 percent, quite close to the 3.0 percent figure recommended in other texts.

while students are in college at four-year colleges and are applied equally in all years of college. For pre-college programs, the standardization technique is analogous and we consider a cohort of 100 high school freshmen and follow them through to (potential) college enrollment and graduation. While taking a different strategy, the CERs of pre-college programs and be directly compared with CERs of college programs. It might turn out, after considerable additional research, that pre-college programs *in general* are more cost-effective than college programs (or vice versa). As Harris (2009a) points out, choices among strategies such as this are as important as choices among specific policies.

We focus on degree completion as the main outcome of interest for several reasons. It is clearly of great interest to our national leaders from both political parties and graduation is what most students are interested in when they start college. One limit of this approach, however, is that focusing on any single outcome is that education has many goals and it is possible that programs vary in their relative productivity in addressing each (Harris, 2009a). This would not be problematic if all the studies measured impacts on all the various outcomes, but this is far from the case. Achievement, persistence, and graduation are the most common outcomes but few studies measure more than one of them. By focusing on graduation only, we may give the false appearance that spending time in college without obtaining a degree is not useful, apparently contradicting evidence that there are positive returns to education separate from degrees.

We address the multiple outcomes problem partly by translating persistence effects into graduation effects about the relationship between these two measures. Graduation and persistence are closely linked—the former cannot happen without the latter (the converse is not true). Therefore, based on data from nationally representative data about the persistence-graduation link, we calculate multipliers (see appendix) to convert the impacts to graduation when necessary. The impact multipliers are less than one because of the compounding effect of persistence rates (e.g., a one-year persistence rate of 0.7 becomes a two-year persistence rate of 0.49). This might sound questionable, but it is reinforced by the one case in which we have estimates of the same program on both persistence and graduation; in Table 3 we show estimates from Cornwell et al. and from Dynarski for the Georgia (and Arkansas) HOPE programs. The estimated impact on persistence is roughly twice the estimated impact on graduation, which is very close to what our multipliers imply.

Achievement is a conceptually different outcome and we therefore do not attempt such translations in that case and instead we view achievement impacts as indicators of potential impacts on degree quality.[7]

The appendix also explains and reports multipliers for costs. In the discussion of specific programs below, we report the cohort cost for a single year. The cost multipliers then make the same assumptions about student participation as are made in calculating the impact multipliers. 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 years, then the cost multiplier would be five (ignoring discounting). Because some students will drop out before graduation, the actual cost multiplier is lower than five.

Once costs and effects (and multipliers) are calculated and applied, the cost-effectiveness ratio (CER) is mainly a matter of arithmetic. For example, CER=1 means that it costs a total of $1,000 for each new graduate. As we will see below, the more realistic CERs are much smaller than this—it costs much more than $1,000 to produce an additional college graduate, even for the most cost-effective programs.

The CERs are based on a variety of assumptions. In particular, they implicitly assume that the costs and impacts of programs are independent of one another (Harris, 2009a). This is true in many cases, but not in others. Since we know relatively little about even the "main

[7] We report impacts on achievement with the "effect size" which is the effect expressed in raw test scale points divided by the population or sample standard deviation (at the student-level, not aggregated to a school for example). This approach is standard in education and psychology research involving test scores.

effects" of these programs, consideration of potential interactions must necessarily be set aside as a secondary issue for future consideration.

Rigorous Research and "Break-Even" Impacts

Most of the uncertainty about the CERs is arguably due to uncertainty about impacts. In identifying promising practices and programs, we searched databases for studies of college programs or policies that used rigorous research methods and/or programs presenting cost data, and consulted with colleagues and experts in the field. We also identified programs that are widely utilized (e.g., Upward Bound and financial aid) and input factors that have a significant bearing on costs (e.g., student-faculty ratio). We make no claim that the programs or studies we have identified represent a comprehensive set of studies that meet our criteria.

Next, we placed programs into categories based on the rigor of available research, using the standards established by the U.S. Department of Education, Institute for Education Sciences, What Works Clearinghouse (2010). Levels of evidence are "strong" with randomized control trials that exhibit high internal and external validity, "moderate" with quasi-experimental methods that have equivalence between treatment and comparison groups, and "low" when relying on other strong findings and/or theories based on expert opinion. The vast majority of research is unfortunately in the low category. Consistent with these standards, we do not assume that all randomized trials exhibit high internal (or external) validity. The key is whether the authors can rule out alternative interpretations that would call into question the causal impacts (Shadish, Cook, & Campbell, 2002). Studies from outside the U.S. (except one from Canada) are omitted as are many studies from before 1980.

In summary, there are several sources of uncertainty in these CERs: (1) uncertainty about the true costs; (2) uncertainty about the true effects on the measured outcome; and (3) uncertainty about the conversion of the effect on the measured outcome to graduation (through multipliers). But the differences in measured cost-effectiveness reported below across a range of programs are so large that these patterns almost certainly reflect real differences in cost-effectiveness. Because 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 equal the current average spending per degree.

Results: Costs and Cost-Effectiveness

We organize the discussion below according to five themes, focusing first on the cost estimates and then those used to make CER calculations. Programs are distinguished based on whether they are pre-college or during college and whether they pertain to instruction or other student services, broadly defined to include counseling and financial aid. Details underlying the cost and CER calculations are provided in a detailed technical appendix (available upon request). Cost and impact multipliers are reported only in Table 1.

Typical Hallmarks of College Quality are Costly

Since the goal of this analysis is to inform college decisions about resources, there is perhaps no better place to start than with two resource decisions that have the greatest effect on budgets and are widely seen as key indicators of quality—student-faculty ratios and adjuncts.[8] These two factors also comprise 13 percent of *U.S. News and World Report* undergraduate rankings. Below, we consider the costs of each and briefly summarize available evidence on impacts.

Student-faculty ratio and class size. Some theorize that smaller student-faculty ratios translate into greater student engagement in the learning process, which in turn increase

[8] According to U.S. News (2010), the "faculty resources" component of the criteria is weighted 65 percent to three indicators of student-faculty ratio and class size, plus two additional measures of adjuncts (percent full-time faculty and percent of faculty with highest terminal degree in field).

achievement and persistence. Student-faculty ratio is related to class size, but also relates to the general accessibility of faculty inside and outside the classroom.

Jacoby (2006) reports full-time faculty salaries of $46,000 and $60,000 for full-time faculty in 1999, or $57,040 and $74,400, respectively in 2007. 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). Reducing the student-faculty ratio from 15 to 13 at four-years (from 19 to 16 at two-years) would therefore cost $73,656 ($56,470) per year for 100 students, excluding capital costs.

Adjuncts. One way colleges have attempted to reduce costs in recent decades is through hiring adjuncts or part-time faculty. While adjuncts are certainly less costly, as we show below, the question is whether and to what degree they reduce student persistence and graduation. Jacoby (2006) also presents evidence on the cost savings of adjuncts. Based on these data, the shift from all part-time faculty to all full-time faculty would cost $213,551 at four-year colleges and $29,470 at two-year colleges. The reason for the large differences in costs between two- and four-year colleges is that full-time faculty in four-year colleges earn much larger salaries and teach many fewer hours (because of their research and service responsibilities). The full-time faculty in four-year colleges in two-year public colleges teach five times as many class hours as four-year full-time faculty (Jacoby, 2006).

Four-year (two-year) colleges currently already have 2/3 (1/3) of their courses taught by full-time faculty. Therefore, the more realistic comparison requires reducing the above estimates accordingly so that the unadjusted cohort cost of switching from current practice to all full-time faculty would be $70,993 ($19,646).

Summary. These cost estimates show that it is costly for 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 Especially Costly

Policymakers have focused for making 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 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 may 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 counselling, and parent education about access to higher education, as well as college scholarships.[9] GEAR UP chooses entire middle schools to participate and requires provision of services to an entire grade cohort according to their needs, but individual student participation is voluntary. GEAR UP services must begin no later than the seventh grade. The funds are intended to support students through high school, through college-supported activities. 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. Like GEAR UP, The stated goals of the Upward Bound pre-college program is to increase both the number and percentage of college-going and college-

[9] It is important to note that a substantial share of the costs of GEAR UP come in the form of volunteer time. This is a good example of a cost that would not be included in an accounting cost framework, but which is counted here in this opportunity cost approach.

completing students from low-income, non-college completing families. Program components include tutoring, SAT and ACT test preparation, summer and after school sessions aimed at improving language arts and math skills as well as campus visits. Data from Albee (2005) suggests an average cohort cost per year of $516,000- $677,000.

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 situated language acquisition.

Jenkins (2002) reports costs of: Vocational ESL $684 per enrollee for 140 hours, Workplace Basics $684 per enrollee for 140 hours, and Technology Career, $4,791 per enrollee for 320 hours.

Summary. GEAR UP and Upward Bound are widely regarded pillars of college access, but they are also costly. At a half-million dollars, the Upward Bound costs are equivalent to hiring seven full-time faculty at a four-year college. 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.

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. 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. Again, while it might be tempting to think such divergent strategies cannot be compared, college leaders and policymakers make such broad strategic decisions on a regular basis (Harris, 2009a).

Student counseling. Other studies focus on more specific programs that fall into the broader “student services” category. 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 $150 per semester when they used those services in two Ohio community colleges. 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. The time of counselors and the student stipends are the main costs in the Opening Doors study and these add to an unadjusted cohort cost of $54,898.

Call centers. An alternative student service approach, the call center, involves literally making phones 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 what issues the student might have and direct them to other services or assistance. It could be that the information itself was important to students or that the effort made by the college representatives gave students a closer connection to the campus and motivated them to return. The report provided to us did not include cost information, but we estimated cohort costs of $200-$500 based on information from call centers in other sectors and our own back-of-the-envelope calculations (see technical appendix available from authors upon request).

Financial aid (no services). Tuition is the heavily subsidized price of college paid by students. The U.S. federal government, and to a lesser degree state governments, have had programs for college financial aid to further reduce the price to students (grants) and/or to address short-term credit constraints (loans). These programs date back to the GI Bill of the post-World War II era and these expanded significantly in the 1960s and beyond. Since loans have to be paid back, grants would likely have a larger impact than loans, albeit with much greater costs to the government. This is especially true given evidence that availability of credit is not a significant problem (e.g., Stinebrickner and Stinebrickner, 2003; Cameron and Taber, 2004).

In contrast to tradition forms of financial aid which are made available based on financial need, many states and philanthropies have begun experimenting with merit-based programs that require students to meet specific academic standards in order to receive the funding. Some of these programs are set up so that middle and high school students are promised a certain amount of money for college if they do sufficiently well academically in high school. Others impose the academic requirements only on students once they have reached college and still others combine these elements. Since the cost of financial aid is essentially just the amount of money given to students (or the interest on loans), we do not discuss costs further here.

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 merit-based aid. Like the Canada experiment, this too incorporated service along with financial aid. The most obvious costs of Opening Doors in New Orleans are the costs for merit-based scholarship, as well as advising and counseling students. From Richburg-Hayes et al. (2009), in the Delgado Community College and Louisiana Technical College-West Jefferson, average total scholarship payment per student over 2 semesters is $1,246. We estimate the costs of the counselors to be $340 so that the total average cohort cost is $1,246+$340=$1,473 (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.[10] In addition to the control group, there 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, respectively for these options are reported as $302, $366, and $739 (Angrist et al., 2009).

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 merit-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 sign up and to 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. Because the vast majority of low-income students going on to college attend less selective—and therefore less costly—colleges, we assume actual average awards of half this amount, or $2,586 per student per year.

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 the university: instruction.

[10] The Wisconsin Scholars Longitudinal Study (WSLS) is a randomized trial of need-based aid taking place at all public two-year and four-year colleges in Wisconsin. The authors of this article are co-directors of the WSLS.

We therefore conclude this section by discussing evidence about online learning, other modes of instruction, and remediation.

Online and distance learning. 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 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 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 upfront costs, 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 are, to our knowledge, 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 seem to require more faculty time. So, while online education no doubt expands access to courses and appears to increase quality, it seems most likely more expensive.

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. Material 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. While users report reduced costs from PCR, it is unclear how these estimates were made.

The difficulties of measuring cost of online learning 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.

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).

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 approximately $1,000 per student.

Some have expressed concern, however, that the quality of the typical remediation problems 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 costs of these programs vary widely, from as low as $500 for call centers to 1,000 times as much for access programs. From a productivity standpoint, call centers could produce tiny impacts and still be worthwhile.

Since we do not have convincing estimates of causal impacts on graduation for the majority of the programs considered, 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 CER equals 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 cost for a bachelor's degree is $30,000-40,000. His full cost attribution method which, as in our analysis, includes costs of students who do not graduate, is arguably closest to the economic concept of opportunity cost and to the method of calculation used below; this represents the upper end of the range—$40,000. We believe these are underestimates for several reasons: (1) Florida spends less than average on higher education compared with other states; and (2) These estimates include only 73% of total spending on undergraduate education. Also, graduate education in some ways subsidizes undergraduate education 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 only slightly 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 cost-effectiveness ratio in thousands), we obtain baseline four-year (two-year) CERs of 1/60=0.017 (1/25=0.04). Below, we estimate what impact for each program would be necessary to generate these same CERs and thus break-even compared with current practices.

The calculations could also be carried out using comparisons other than current cost per degree, but the main point here is to show that 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 explore the effects further.

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

Program	*Unadjusted Cohort Cost ($)*	*Cost Multiplier (Appendix)*	*Break-Even Graduation Impact*
College Instruction			
Stud/Fac Ratio 15-to-13 (4y)	73,656	3.00	3.68
Stud/Fac Ratio 19-to-16 (2y)	56,470	1.74	3.93
Adjuncts (4y)	70,993	3.00	3.55
Adjuncts (2y)	19,646	1.74	1.37
Online Education	80,000	3.00	4.00
Other Modes of Instruction			
Independent Study	-48,204	3.00	-2.41
Discussion/Inquiry	-45,396	3.00	-2.27
Personalized Instruction	286,650	3.00	14.33
Other	131,976	3.00	6.60
Remediation	170,000	1.00	2.83
College Non-Instruction			
Student Services (Webber/Ehr)	50,000	3.00	2.50
Student Counseling	54,898	3.00	2.74
Call Centers (2y)	500	3.00	0.03
Pre-College			
GEAR UP	264,000	4.50	19.80
Upward Bound	677,000	3.80	42.88
Bridges to Opportunity			0.00
Vocational ESL	58,500	1.00	0.98
Workplace Basics	58,500	1.00	0.98
Technology Career Bridge	409,500	1.00	6.83
Financial Aid Only			
Tuition and grants	100,000	3.80	6.33
Loans	10,000	3.80	0.63
Merit Aid (GA/AR)	100,000	3.80	6.33
Merit Aid (Canada STAR)	36,600	3.80	2.32
Dreamkeepers	39,300	1.74	1.14
Angel Funding	26,600	1.74	0.77
Financial Aid with Services			
Canada STAR	73,000	3.80	4.62
Opening Doors (Louisiana)	147,300	1.74	4.27
Indiana 21st Century	258,600	3.80	16.38

Notice 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 not coincidentally also the one with the smallest break-even impact (0.05). This means that spending $1,000 on call centers need only produce 1/20 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 43 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.

Some Innovative Programs Have Considerable (But Unproven) Potential

We have considered two programs that show promise for various reasons. Online learning are 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. To test this potential, we combine the above evidence about costs with evidence on impacts.

Online and distance learning. A recent meta-analysis conducted by the U.S. Department of Education (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.[11] We also conducted our own review and corroborated these findings.[12] These are primarily based 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 suggest that there is real potential in online learning.

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. While the program was not studied with a randomized trial, what the college did do was something akin to a quasi-experiment. They kept 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

[11] 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.

[12] 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 with the general spread of internet.

college. As they point out in their report, comparing those they could not reach to those they did reach by phone to those 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 not reached at all.

A somewhat more reasonable comparison is between the groups who received a voice mail and those talked 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 continue in college. Nevertheless, it is plausible that most 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 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.). 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 CER because the program is so inexpensive. While this evidence no doubt falls into the low-medium range by our standards of rigor, it is worth pointing out the anecdotal evidence that the use of call centers is, anecdotally, more widespread in for-profit colleges who have more to lose when a student drops out.

Common Programs Fail Simple Cost-Effectiveness Tests

While the above programs show promise, we show below that typical programs in higher education do not. Those programs mentioned above but not in this section did not have evidence of impacts with sufficient rigor to be included.

Student-faculty ratios. We found only one unpublished study that met our criteria for rigor and focused on the U.S. context and even this focused only on achievement rather than persistence and graduation (Kokkelenberg, Dillon, and Christy, 2008).[13] Because achievement effects cannot be readily translated into graduation rates, we are forced to rely on simpler regression estimates from Jacoby (2006); reducing the student-faculty ratio from 15 to 13 at four-years (or from 19 to 16 at two-years) would raise the graduation rate by one percentage point. We do report CERs because the evidence meets only the lowest standard. But it is worth pointing out that the Jacoby estimate of 1.0 is far lower than the break-even impact of 3.68 (3.93) in Table 1.

Adjuncts. Bettinger and Long (forthcoming), using a quasi-experimental technique, find that adjuncts increase interest in subsequent course enrollment. But, we were unable to find any rigorous studies of the effects on graduation, or effects that could be reasonably translated to graduation. Jacoby (2006) reports that going from an entirely part-time faculty to an entirely full-time faculty is associated with a graduation rate increase of 15 percentage points at two-year colleges, but it is far from clear whether these represent causal impacts. The fact that Bettinger and Long find a *positive* impact of adjuncts on subsequent course enrolment within a major suggests that the Jacoby results might suffer from selection bias, but we cannot be sure. We do not report CERs on student-faculty ratios and adjuncts in Table 3 for this reason, though it is worth comparing these impacts to the break-even values in Table 1. Again, we do not report the CERs for this study, but in this case the estimated

[13] Badiera, Larcinese and Rasul (2008) study class size using UK data. De Paola and Scoppa (2009) do the same using data from Italy.

impact is much larger than the break-even, so it is at least plausible that hiring more full-time faculty is cost-effective.

Upward Bound. Using a randomized trial, 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 2 percentage points in the overall rate of enrollment (across college types, etc.). For enrollment at four-year colleges and universities, the estimated impact is 1 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 4 percent of control group members received a vocational certificate or license, nearly 9 percent of treatment group members did, implying an impact of 5 percentage points. Because of the focus of current initiatives on two- and four-year degrees, we use 2 percentage points as the baseline impact for average low-income students.[14] Based on the range of costs from Albee (2005), the unadjusted CER for average students is: 2/677=0.003.

Remediation. The most rigorous evidence on remediation suggests that there is no impact on degree completion. 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). The CER is therefore zero.

Student services. Webber and Ehrenberg (2009) point out that spending on non-instructional student services such as student organizations, intramurals, student health

[14] 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. Upward Bound increased the rate of postsecondary enrollment and the likelihood of receiving a degree, license, or certificate by 6 and 12 percentage points, respectively. 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 CERs in this study, we use the 2 percentage point impact as the baseline.

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.[15] 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 unadjusted cohort CER of 0.7/50=0.014.

Student counseling. Other studies focus on more specific programs that fall into the broader "student services" category. The MDRC Opening Doors initiative was studied with a randomized trial. Impacts were statistically significant during the one year of the services were provided, though most of the initial effects diminished over time. The treatment increased persistence by 7 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

[15] 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

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 unadjusted CER is therefore 3.7/55=0.0673, which is five 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 CERs suggest most likely that the quasi-experimental studies are biased downwards, or that the specific types of student services were different.

Financial aid. We briefly summarize results in Table 2 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 costs of loans are quite different from grants, however. Because loans have to be paid back by the borrower, and because they are paid back at very high rates, the true cost is much lower than the face value of the loans. For this same reason, we expect that the impact of loans is much smaller. We estimate that the true cost to the government of a $1,000 loan is only about 20 percent of the face value. We do not report CERs for the Indiana 21st Century Scholars program because of the absence of rigorous evidence, but the correlational results suggest that the impacts might have been as large as 10 percentage points. Even this large figure is smaller than the break-even value in Table 1.

Table 2: Impacts of Traditional Financial Aid

Study	*Program Type*	*Dependent Variable*	*Impacts* *(Perc. Points/$1,000)*
Need-Based Aid			
Kane (1995)	Tuition	Attendance	two-year: 3 four-year: 1
Kane (2007)	DC Tuition	Attendance	3-4
Dynarski (2003)	Grant (Need)	Attendance	3.6
Reyes (1995)	Loans (Need)	Attendance	1.5
Dynarski (2005)	Loans (Need)	Attendance	0-5.1
Merit-Based Aid			
Cornwell et al. (2006)	Georgia HOPE	Attendance	4-6
Kane (2003)[16]	CalGrant	Attendance	1.8
Angrist et al. (2006) *	Canada STAR	Persistence	8.2
Dynarski (2008)	GA/AR HOPE	Completion	0.9
MDRC *	Open Doors		2.2

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

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 3 percentage points. This figure was not statistically significant, but a much larger impact (more than 6 percentage points) was significant for males.[17] 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 3 percentage point impact of aid only and an unadjusted CER of 3/0.366/100=0.082. The CER for the combined financial aid and service is roughly half that size because the services almost doubled the costs and the impact

[16] Results for the Kane (2003) study are obtained as follows: He concludes that the impact on enrolment 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.

[17] In their summary of results, Angrist et al. report only impacts on females, which conflicts with this finding regarding males. The apparent reason is their main focus is on outcomes such as grades where there are no impacts on males.

size remained roughly the same. The point estimates were similar for the combined funding-services treatment, but none of these estimates were statistically significant.

MDRC's Opening Doors also included an experiment that combined services with merit-based aid. Like the Canada experiment, this too incorporated service along with financial aid. MDRC has released a series of reports suggesting that merit-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.

Based on the above evidence, the unadjusted CER is 6.5/1.473/100=0.011. While the impact here is larger than in Canada STAR, the costs are five times larger, which explains the smaller CER.

Summary of cost-effectiveness results. Table 3 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 multiplier). The third column indicates the outcome variable used in the study, which is necessary for identifying the appropriate multiplier in the appendix, 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 from the Appendix to obtain estimated graduation impacts. Finally, the last column indicates ratio of the multiplier-adjusted impact divided by the multiplier-adjusted cost—the CER.

Table 3: Cost-Effectiveness Ratios for Higher Education Programs

Program	*Adjusted Cohort Cost ($)*	*Outcome Variable*	*Unadjusted Effect*	*Effect Multiplier*	*CER*
Average Costs per Degree					
2-year degree	--	--	--	--	*0.0400*
4-year degree	--	--	--	--	*0.0170*
College Instruction					
Remediation	170,000	Graduation	0.00	1.00	0.0000
College Non-Instruction					
Student Services (Webber/Ehr)	150,000	Graduation	0.70	1.00	0.0047
Student Counseling	164,694	Persistence	3.70	0.56	0.0126
Call Centers (2y)	1,500	Persistence	2.00	0.30	0.4000
Pre-College					
Upward Bound	2,572,600	Enrollment	2.00	0.43	0.0003
Financial Aid					
Tuition and grants	380,000	Attendance	3.00	0.64	0.0051
Loans	76,000	Attendance	2.00	0.64	0.0118
Merit Aid (GA/AR)	380,000	Attendance	4.00	0.64	0.0067
Merit Aid (Canada STAR)	139,080	Enrollment	3.00	0.51	0.0110
Merit Aid with Services					
Canada STAR	277,400	Persistence	3.00	0.51	0.0055
Opening Doors (Louisiana)	256,302	Persistence	6.50	0.56	0.0142

The top two most cost-effective programs include both the widely used and the nearly unknown. Call centers appear most cost-effective. The reported CER of 0.4 is much larger than all the others. This number means that almost one-half of one new graduate could be produced for $1,000, well above current rates of productivity as reflected in the break-even analysis above.

Loans also appear relatively cost-effective with a CER around 0.03. Again, student loans provide the cash students need to pay their bills, but they cost little because students have to eventually pay the money back. On the other hand, some have argued that loans also drive up costs and tuition, as discussed above, and this is not captured in the above estimate. So, the net impact and cost-effectiveness may be somewhat smaller than indicated. Merit-based aid, along with some forms of student counselling appear next in line, just above 0.01.

The majority of the CERs are related to four-year colleges and bachelor's degrees, but the distinction between that and two-year degrees should not be forgotten. As we pointed out above, two-year degrees require less time and therefore lower costs. Note that the only evidence on call centers comes from a community college. This means that, in the four-year sector, loans appear to be the most cost-effective option. As we showed in our introductory remarks, however, some argue that this partial equilibrium result masks the fact that, in general equilibrium, loans may drive up costs and tuitions by making students less price sensitive. While we do not question the general presence of this phenomenon, it is not obvious that these secondary effects would be on the same scale as the direct effects.

We are also interested in finding the most cost-effective ways to improve graduation chances for disadvantaged student populations. Considering the results for all the programs we study that are targeted to those populations, we find that a combination of loans and merit aid coupled with student counselling services would be most cost-effective. The potential role of student services is suggested both by both the services-only intervention and the Opening Doors merit aid combined with services. The fact that these services did not seem to have an impact in the Canada STAR study suggests that this approach might only work in two-year colleges and perhaps openly for disadvantaged student populations.

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 efficiency.

We find, in contrast to the message of 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 warrant greater attention than they have received.

We do not claim that colleges could get back to productive 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 is perhaps the strongest evidence that we are falling short of our productivity potential.

The limitations of research partly fall on researchers. Data to study higher education are increasingly available. 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, 2009a).

This research also justifies additional consideration of systemic reforms such as accountability and competency-based systems such as that recommended in Europe as part of the Bologna Process. 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 students need to know. Like call centers, systemic reforms have low costs and what remains is uncertainty about impacts.

The larger point, however, has to do with the perceived helplessness among some scholars and practitioners who see productivity as something uncontrollable. 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 those programs more carefully so that we can make more informed decisions.

Appendix: Translating Costs and Impacts for Analysis of College Graduation

This appendix details the data and assumptions necessary to translate costs and impacts pertaining to high school graduation, college entrance, and college persistence into impacts on college graduation. Table A1 below provides descriptive information about the relationship between persistence and graduation, much of it taken directly from U.S. Department of Education data sources (see footnotes for details).

In some cases, programs will have baseline persistence rates that are below or above the average and this would changed the multipliers, but these baseline assumptions are arguably better for the purposes here, focused on making generalizations about the effects (and costs) in "typical" settings.

Table A1:
Assumptions for Typical Persistence and Completion Rates for College Entrants

	Disadvantaged Students			Average Students		
	H.S.	two-year	four-year	H.S.	two-year	four-year
% 12th graders who enter college	0.60	0.32	0.25	0.77	0.28	0.46
1-2nd year		0.74	0.87		0.72	0.90
2-3rd year		0.74	0.87		0.72	0.90
3-4th year		0.74	0.87		0.72	0.90
4-5th year		0.75	0.68		0.79	0.61
5-6th year		0.75	0.68		0.79	0.61
6-7th year		0.75	0.68		0.79	0.61
Actual graduation rate		0.43	0.64		0.41	0.70

Tables A2 and A3 below use the information from Table A1 to produce the appropriate cost and impact multipliers. Recall from the main text that the cost multipliers are intended to take the annual cohort costs and turn them into total program costs. For programs that are applied to students in all years, the main issue is how many students drop out along the way and therefore reduce the costs.

For the impact multipliers, the issue is quite different: that most studies use some outcome other than college graduation. The key assumption underlying this is that an impact on these intermediate outcomes yields the same impact on graduation. That might not be the case—the same impact on persistence might mean different impacts on graduation—but this is the best that can be done given the information available. In future analyses, we will also apply a discounting rate of 3.5 percent to these multipliers.

Now, consider the adjustments made for programs that occur pre-college. The cost aspect of these adjustments is the same. Because persistence and graduation rates are usually much higher in high school, the multipliers are much closer to simply summing across the number of years. For example, the cost multipliers for programs that last three years are just slightly lower than three whereas in college programs they are closer to two.

Table A2: Cost Multipliers

Program Starts in This Year	Program Lasts This Many Years:				
	1	2	3	4	5
9th grade	1.00	1.98	2.94	3.80	
10th grade	1.00	1.98	2.86		
11th grade	1.00	1.90			
12th grade	1.00				
1st year of college	1.00	1.74	2.29	2.69	3.00
2nd year	1.00	1.74	2.29	2.70	
3rd year	1.00	1.74	2.30		
4th year	1.00	1.75			
5th year	1.00				

Table A3: Impact Multipliers

Outcome Measure	Last Measure of Outcome Variable is This Many Years from the Start of College:		
	1	2	3
Two-year coll entrance	0.43		
Two-year coll persist	0.41	0.56	0.75
Four-year coll entrance	0.64		
Four-year coll persist	0.51	0.59	0.68

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.

Aghion, Philippe, Matthias Dewatripont, Caroline Hoxby, Andreu Mas-Colell, and Andre Sapir (forthcoming). The Governance and Performance of Research Universities: Evidence from Europe and the U.S. *Economic Policy*.

Albee, Amy (2005). A Cost Analysis of Upward Bound And GEAR UP. Unpublished manuscript. Tallahassee, FL: Florida State University.

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 eduation 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

Bandiera, O., Larcinese, V., Rasul, I. (2008), Heterogeneous Class Size Effects New Evidence from a Panel of University Students, Department of Economics, London School of Economics and Political Science, mimeo

Baum, Sandy & Kathleen Payea (2004). *Education Pays: The Benefits of Higher Education for Individuals and Society,* College Board.

Baumol, W. J. (1967). Macroeconomics of unbalanced growth: The anatomy of urban crisis. The American Economic Review, 57, 415–426.

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 (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.

Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas Accountability System, American Educational Research Journal, 42(2), 231-268.

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.

Breneman, David and William Haarlow. (1997). "Remedial Education: Costs and Consequences." Remediation in Higher Education. Washington, D.C.: Thomas B. Fordham Foundation.

Brown, C. & C. Belfield (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.

Browning, E.K. (1987) On the Marginal Welfare Cost of Taxation, *American Economic Review*, 77(1), 11-23.

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.

Card, David (2001). "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems" *Econometrica*.

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.

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.

Dannenberg, M. (2009). *10 New Higher Education Ideas for a New Congress*. The New America Foundation: Washington DC.

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://mpra.ub.uni-muenchen.de/16945/.

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

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

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

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

Gates, Bill and Melinda. 2009. "A Bold Vision for Stimulus, Education Reform." *Roll Call.* February 10, online edition.

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.

Gordon, S., He, W., & Abdous, M. (2009). Using a Web-Based System to Estimate the Cost of Onine Course Production. *The Online Journal of Distance Learning Administration*, 12(3).

Harris, D. (2009a). Toward policy-relevant benchmarks for interpreting effect sizes: Combining effects with costs. Educational Evaluation and Policy Analysis. 31(1), 3-29.

Harris, D. (2009b). *First Things First: Replacing 'Status' Models with Value-Added.* Washington, DC: American Enterprise Institute.

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.

Jenkins, D. & Thomas Bailey (2009). Downloaded October 10, 2009 from: http://www.makingopportunityaffordable.org/sites/default/files/resources/The%20Dreaded%20P%20Word.pdf.

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

Kadden, J. (2009). Education Secretary Criticizes Steep Rise in College Costs. *The New York Times*. August 20. http://thechoice.blogs.nytimes.com/2009/08/20/education-secretary-criticizes-steep-rise-in-college-costs/

Kokkelenberg, C., Dillon, M. and Christy S. (2008), "The Effects of Class Size on Student Grades at a Public University", Economics of Education Review…..

Lieber, Ron (2009). Why College Costs Rise, Even in a Recession. *The New York Times*. Sept 5. B1.

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.

Leslie, Larry and Paul Brinkman (1988). *The Economic Value of Higher Education*. New York: Macmillan.

Lewis, D.R. & L.A. Kallsen (1995) Multiattribute evaluations: An aid in reallocation decisions in higher education, The Review of Higher Education, 18(4), 437-465.

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.

Massy,W. F. (2003). Honoring the trust: Quality and cost containment in higher education. Bolton, MA: Anker.

McPherson, Michael, and Schapiro, Morton Owen. (1991). Keeping College Affordable: Government and Educational Opportunity. Washington, DC: The Brookings Institution.

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.

National Commission on the Cost of Higher Education. (1998). Straight talk about college costs and prices. Phoenix, AZ: The Oryx Press.

Provasnik, S., and Planty, M. (2008). Community Colleges: Special Supplement to The Condition of Education 2008 (NCES 2008-033). Table SA-10. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Provasnik, S., and Planty, M. (2008). Community Colleges: Special Supplement to The Condition of Education 2008.

Pusser, B. (2008). "The State, the Market, and the Institutional Estate: Revisiting Contemporary Authority Relations in Higher Education." In J. Smart (Ed), *Higher Education: Handbook of Theory and Research.v23* P. 105./

Rothstein, R. (2008). *Grading Education: Getting Accountability Right*. New York: Teachers College Press.

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

Seftor, Neil and Sarah Turner (2002). "Back to School: Federal Student Aid Policy and Adult College Enrollment." *Journal of Human Resources*.

Smith, Burck (2009). College by Subscription. Washington, DC: American Enterprise Institute. Downloaded October 14, 2009 from http://www.aei.org/outlook/100071.

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).

Soares, Louis (2009). A Postsecondary Degree or Credential in Every Pot. Center for American Progress: Washington DC.

Standing, K., Judkins, D., Keller, B., & Shimshak, A. (2008). Early outcomes of the GEAR UP program. Final report US Department of Education

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.

Trostel, P. 2003. The Long-Term Economic Effects of Declining State Support for Higher Education Are States Shooting Themselves in the Foot? Wisconsin Center for the Advancement of Postsecondary Education.

Twigg, C. A. (2003). Improving quality and reducing costs: Designs for effective learning. Change, 35(July-August), 22–29.

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

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

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.

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

Wolfe, Barbara & Robert Haveman (2003). Social and Nonmarket Benefits from Education in an Advanced Economy. in Yolanda Kodrzycki, *Education in the 21st Century: Meeting the Challenges of a Changing World*.

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.