

THE SUPPLY SIDE OF HIGHER EDUCATION:
HIGHER EDUCATION FINANCE AND THE POTENTIAL OF USING INSTITUTIONAL
INCENTIVES TO SUPPORT STUDENT SUCCESS

Bridget Terry Long
Harvard Graduate School of Education
National Bureau of Economic Research (NBER)
National Center for Postsecondary Research (NCPSR)
longbr@gse.harvard.edu

prepared for:

“Reform and Innovation in the Changing Ecology of U.S. Higher Education”

inaugural strategy session, 2-3 December 2010

School of Education, Stanford University

WORKING PAPER PREPARED FOR DISCUSSION
DO NOT CITE OR QUOTE WITHOUT AUTHOR’S PERMISSION

“... remarkably little is known about how access institutions strategize to maintain and enhance their various revenue streams, or the extent to which current government funding systems encourage access institutions to do the things that matter most: enable students to *persist and complete* college, not merely enroll.”

Reform and Innovation in the New Ecology of U.S. Higher Education
Proposal to the Bill & Melinda Gates Foundation
Michael W. Kirst and Mitchell L. Stevens

I. INTRODUCTION: THE SUPPLY SIDE OF HIGHER EDUCATION

There have been many contributions from the Economics of Education literature to build understanding about the demand side of higher education. Numerous studies have examined how background and various inputs, policies and other factors affect student behavior (Long, 2007). However, little is known about the supply side of the equation: the thousands of postsecondary institutions. Related to this is the paucity of information on how to produce better student outcomes and/or alter the incentives colleges face with the goal of improving student persistence and completion. Such answers are necessary as less than 60 percent of students at four-year colleges graduate within six years. Completion rates are especially alarming for low-income and minority students. While insufficient academic preparation is part of the problem, it does not fully explain differences in graduation rates by background. Among students who were identified as being college-qualified, only 36 percent of low-income students completed a bachelor's degree within eight years while 81 percent of high-income students did so (Adelman, 2006). Colleges and universities have some role to play in improving these numbers as even colleges with similar student bodies and resources can have very different graduation rates.¹

Studying the supply side of higher education is not a straightforward matter as colleges and universities do not fit the traditional economic model in which a firm maximizes profits and uses a combination inputs in a production process to make its outputs. Given most institutions are non-profit, it is difficult to specify exactly what their objective function is. Bowen (1980) suggests that the dominant goals of institutions are educational excellence, prestige, and influence, but this is debatable and does not easily lead to simple models of postsecondary

¹ See the Education Trust website: College Results Online.

finance. The lack of a simple profit motive makes it difficult to understand how to influence college behavior.

Another aspect of schools that is unique involves the inputs they use to produce “education.” In a typical model, the inputs are entirely separate entities from the outputs, but in education, students complicate this basic framework as they are not only the outputs of the process (i.e., an educated college graduate) but also important inputs. Peers have been shown to influence each other’s academic and social outcomes (Winston and Zimmerman, 2004; Kremer and Levy, 2003). In fact, as noted by Winston (1998), the important role of students in the production of education helps to explain many of the anomalies present in higher education. For example, unlike a regular firm, colleges will turn away a majority of potential customers who are willing and able to buy their product, and require elaborate application procedures before allowing a purchase. They also do not expand to meet demand and will use extensive price discrimination (charging people different prices) in an effort to redistribute income.

The important role of students as both inputs and outputs complicates understanding of the resources and costs a school faces. Unfortunately, there is no good, systematic way to measure the value added by one student to another. Institutions implicitly compensate some students for their positive peer effects by giving them merit-based financial aid, but this is an imperfect measure of the role of students as inputs given that there are multiple motivations for giving merit aid (e.g., attracting students with high test scores with the hope of rising in the college rankings), and measures of institutional financial aid are also flawed.² The costs of building a cohort of students with reinforcing, positive peer effects on each other should also include the costs of many enrollment management activities such as advertising, outreach, admissions staff, and so on. To summarize, there is no convention on the best way to incorporate the costs or benefits of students into finance models.

Another interesting feature of the production function of postsecondary institutions is the role of joint products. This issues stems from the fact that institutions have multiple missions (e.g., education, research, public service), and these missions overlap and may even be reinforcing. For example, a professor’s research may contribute to his or her teaching. To

² Most data only gives institutional aggregates in terms of financial aid. Even when aid figures are divided into types (e.g., need-based versus merit-based aid), it is clear that aid is often awarded due to a combination of need and merit criteria, and so the categorizations may not be that helpful. Only in rare instances is data available on student-level aid awards.

address each of these multiple missions, there are fixed costs, such as administration and operations, which are shared. While one would often like to isolate the costs of one particular function, such as the production of educational services, this can be difficult because of the joint nature of the missions, hence joint products. The goals are not easily separable from each other, and therefore, something like parsing out which costs go with each mission can be impossible. For instance, an analyst is often not able to separate out what proportion of total instructional expenditures was spent for one purpose versus another (Cunningham, et al., 2001). This is partly a problem of aggregation, but more importantly, given the overlap between those missions, it would be difficult to determine the relative cost of each purpose even with a more detailed data collection. In the case of the professor whose research is used in his or her teaching, how should the costs between the research function and the teaching function be divided?

The rest of this paper details some of the challenges to understanding the supply side of higher education, with particular attention to the role of postsecondary finance. Unfortunately, good data are not available at the level of specificity necessary to answer many pertinent questions. There are also other considerations that make interpreting institutional resources difficult. When considering the possible role of incentives to encourage certain institutional behaviors, there are also multiple factors to take into account. Finally, while there are many challenges involved when attempting to compel colleges to improve outcomes, many states have tried to do so, and the paper reviews these efforts and the lessons to be learned from them.

II. THE CHALLENGES OF MEASURING AND INTERPRETING RESOURCES

Before considering how one might judge or influence the behavior of institutions, it is important to accurately measure how an institution uses its resources. College and university spending generally falls into one of five major expenditure categories. The first is education and related services, including instruction, student services, and some share of administration, academic support, and operations and maintenance related to educational services. The second relates to research and includes all research activities as well as another share of administration, academic support, and operations and maintenance. Public service and related activities are the third category. The last two categories are financial aid (scholarships and fellowships) and

auxiliary enterprises and hospitals. In this section, I discuss how these are measured, particularly expenditures related to educational services such as instruction.

Problems in Measuring Higher Education Costs

Expenditure studies are only as good as the underlying data. As summarized by the researchers leading the Delaware Study of Instructional Expenditures, “full cost analyses start with accounting data and rely on adjustments to, and allocations of these financial data to arrive at answers, making the analyses captives of the purposes, conventions and limitations of such data” (Jones, 2000). Before even starting the process, one must first determine how to define certain costs using the available data. In studies of college finances, researchers have developed a variety of cost models and approximations of expenditures on educational activities, but the analysis is never straightforward due to the assumptions that must be made and the availability of data.

One issue that must be resolved when measuring institutional expenditures concerns how one should deal with direct versus indirect costs. Early cost models suffered from the fact that the calculations of one institution might include a battery of indirect cost formulas while another institution might use an entirely different set of formulas. One standard formula was developed by Halstead and published annually in *Higher Education Revenues and Expenditures*. However, as part of the 1998 National Commission on the Cost of Higher Education, Winston (2001) posited an alternate definition. He treated some expenditures as directly related to instruction while other were assumed to be only partially related, but like Halstead, he had to make assumptions about what would be counted and to what degree (i.e., 100 percent of the cost or only a proportion). To these he added a percentage of capital costs, which Halstead did not include at all.

The more recent Delta Cost Project on Postsecondary Costs, Productivity, and Accountability focuses on what it calls education and related spending per full-time equivalent (FTE) student rather than the previous convention to highlight “instructional costs.” This is calculated by dividing “spending on instruction, student services, and the educational share of academic and institutional support (including administration as well as computing services and libraries) by the number of full-time equivalent students enrolled for a particular year.” This definition includes the cost of departmental research but excludes contracted research, public

service, and auxiliary enterprises. However, because of inconsistencies in how capital spending is accounted for across different types of institutions, the Delta Cost Project (2009) does not include capital outlays such as building and renovating facilities, which Winston chose to do. All three of the above studies used the same data source, and their decisions to use slightly different definitions underscores how there is no one way to define instructional or educational costs.

The issue of how to define particular costs is further complicated by differences in accounting standards across institutions. In IPEDS, for instance, private institutions use standards established by the Financial Accounting Standards Board (FASB) while public institutions use standards established by the Governmental Accounting Standards Board (GASB). This difference makes some comparisons difficult across the sectors. Past changes in FASB and GASB standards have improved the IPEDS data collection, but problems in comparability remain.³

Another complication in cost accounting involves not a definitional issue but instead how measurements are reported due to the limits of data. One cannot often distinguish spending on different kinds of students. This is especially a concern for accountability systems in which it may be important to understand whether and how resources are directed towards students who might need more support to complete college. Students with strong academic preparation and social capital are likely to succeed in most places, but the true value-added of an institution could be in what supports it offers to students who might be at some disadvantage, perhaps due to academic deficiencies, low income, or lack of information about how to navigate the system. Within an institution, one can rarely do the accounting necessary at that level of detail. Yet there are some clues that students with higher entrance exam scores tend to receive more resources. For instance, institutions with honors programs or colleges often advertise that students in those programs get better facilities, special access to top faculty, and smaller class size. Instead of being able to detect these kinds of differences, college expenditures are often reported per student or credit hour to standardize across institutions of different sizes. These estimates, in essence, give the average cost within the system, and so they mask differences in the resources one student might receive versus another.

³ As Cunningham, et al. (2001) point out, one must also consider differences across institutions when considering the issue of land. Technically, the land and buildings of public colleges and universities are owned by either the state or local government.

The Shortcomings of National Measures of College Expenditures

The primary source of information on college finances nationally is the Integrated Postsecondary Education Data System (IPEDS). Institutions that participate in any federal student aid program are required to complete the IPEDS survey annually, and the finance data collection includes institutional revenues by source, expenditures by category, and assets and liabilities. The information is reported in aggregate terms, and so it is difficult to isolate the costs of particular activities to get a true sense of costs, spending efficiency, and funding streams. Still, because the coverage of IPEDS is extensive, it has been the foundation of multiple large-scale efforts to understand college expenditures. For example, as part of the Higher Education Amendments of 1998, Congress required that the National Center for Education Statistics (NCES) conduct studies on of expenditures in higher education (Cunningham, et al., 2001).⁴

More recently, the Delta Project (2009) examined college spending for nearly 2,000 public and private colleges from 2002 to 2006. As noted above, their primary indicator for their studies was education and related spending per FTE student. It includes all spending for instruction and student services, plus a portion of spending on academic and institutional support and for operations and maintenance of buildings, and so it is sometimes termed as the “full cost of education.” Their definition is average spending from all revenue sources across all kinds of students and all types of courses of instruction.

In their 2009 report, *Trends in College Spending*, the Delta Cost Project researchers found substantial variation in the resources available to institutions. Most students attend colleges that have very limited budgets while the richest institutions appear to be getting richer. Second, much of the new money to higher education outside of tuition increases is restricted, meaning that it can only be used for special functions. Meanwhile tuition increases are only partially making up for reductions in state appropriations at public institutions. Finally, efforts to investigate changes in productivity were hampered by a lack of outcomes and quality measures. However, they did find that the relationship between spending and the number of certificates and degrees produced has changed little in recent year. For several types of institutions, there is some evidence of a lower cost per credential, but this analysis is not definitive.

⁴ It was highly influenced by National Commission on the Cost of Higher Education. (1998) *Straight Talk College Costs and Prices*. Phoenix, AZ: Oryx Press.

There are limits to cost accounting using the national IPEDS data. Even the researchers of the Delta Cost Project recognize that “aggregate data are not a substitute for the more granular analysis that institutions and states need to perform regularly to examine their own spending patterns” (p. 6). Some of IPEDS’ limitations include that fact that it does not distinguish between expenditures by discipline or level (remedial versus undergraduate versus graduate education). Institutional financial aid and tuition discounting are also not reported as spending in IPEDS. Instead, IPEDS uses a measure of “scholarships and fellowships,” which is only a fraction of institutional aid. Still, as one of the only sources for national-level studies of higher education spending, IPEDS is the foundation of much of what is known about college expenditures.

The Delaware Study of Instructional Costs and Productivity was also developed in response to the 1998 mandate by Congress to study college costs and prices. The report focuses on direct instructional expenditures at four-year colleges and universities (Middaugh, Graham, and Shahid, 2003). However, rather than using IPEDS, the source of the data is from multiple cycles of the Delaware Study of Instructional Costs and Productivity, which was begun in 1992 by the Office of Institutional Research and Planning at the University of Delaware. The data contain information on teaching loads by faculty category, instructional costs, and externally funded scholarly activity, all at the level of the academic discipline, for over 300 four-year institutions. Therefore, compared to IPEDS, much more detailed analysis is possible taking into account teaching and differences across disciplines. However, because participation was voluntary, the data do not give a national picture of college instructional expenditures. This also raises the issues of non-response bias, and the authors of the study acknowledge that institutions which participated in the study were more likely to have at least 5,000 students and be organizationally complex.

The focal measure in the Delaware study is the direct instructional cost per student credit hour taught. This was defined as total direct instructional expenditures divided by total student credit hours taught for 1998 to 2001. In many ways, this definition is much simpler than those discussed above, but this study used credit hours to standardize the measures rather than full-time equivalent student. While the Delaware study provides an example of how costs can be measured, the true aim of the study was to explain why there are cost differences across institutions. The researchers conclude that multiple factors are related to cost differences, and

importantly, relative to the accountability debate, is the fact that they identify factors that have nothing to do with the effective or efficient use of resources.

The authors conclude that most of the variance in cost is due to differences in the mix of disciplines across institutions. For example, the Delaware study found that the difference in direct expense per student credit hour taught between English and mechanical engineering was \$239 (with engineering being more expensive) and \$140 between sociology and chemistry (with chemistry being more expensive) in 1998-2001. Carnegie classification, which captures some of the differences in institutional missions, was another important factor in differences in costs. The authors surmise that the Carnegie classifications are associated with different faculty responsibilities; for example, faculty at research universities are expected to teach fewer student credit hours so that they can be more engaged in research activities.

The researchers identified other factors related to cost, many of which are similar to Brinkman's work. Brinkman (1990) suggested that the important determinants of college costs are likely size (or the quantity of activity), the scope of services offered, the level of instruction, and the particular discipline, and these assertions are backed by the Delaware study. First, cost per student is inversely related to volume, defined as the total number of student credit hours taught. Second, the larger the department in terms of the number of faculty, the higher the cost per student. The proportion of the faculty with tenure is also related to higher costs. As expected, graduate instruction is more expensive than undergraduate teaching, but this factors appears to be less important than volume, department size, and the tenure rate.

Documenting the important role of disciplinary mix, Carnegie classification, and other factors suggests the need for nuance in comparisons across colleges and universities. Differences in mission, student body characteristics and environment are important, but so too is the academic mix of departments, the number of credit hours taught, faculty characteristics, as well as the role of research. It is also worth noting that researchers found “no apparent relationship between the level of instructional expenditures at an institution and the tuition rate charged by that institution.” (p. xi) This is counter to the issues that were originally raised by Congress to motivate the need for such a cost study. They caution not to use price (i.e., tuition) and cost (i.e., institutional expenditures) as interchangeable constructs as price appears to be much more related to revenues than expenditures. For example, the Delaware study points out, “it is not practical for an institution to charge engineering majors a tuition rate three times that charged to

sociology majors” (p. xi) just because of the differences in instructional costs. However, some schools do charge discipline-specific fees (e.g., equipment or lab fees), though those do not ordinarily amount to substantial differences in total price.

How Should One Judge College Spending?

The National Commission on the Cost of Higher Education called for more detailed data to enable better cost measurement analysis, and better data would certainly help. However, such measures still require some interpretation and judgment about what are good versus bad uses for resources. As Shulman (2007) writes: “Counting without narrative is meaningless.”⁵ Unfortunately, cost accounting in higher education does not naturally lead to easy answers about what the narrative should be. As noted by Jones (2001), accounting efforts naturally result in additional questions:

“Even if the total level of resources is the same, the way institutions choose to utilize these resources will vary for reasons of both choice and circumstance. Some institutions will solve their developmental education problems through the instruction program, while others will address the same problem through student services activities (tutoring and advising rather than through classroom instruction). The faculty at one institution may be relatively young, at another relatively old. One may be the victim of locally high energy prices while another may have small utility bills as a function of either location or energy providers—the list goes on. Once started down this path, one is inevitably drawn to the next set of “whys.” (pp. 50-51)

The examples provided by Jones are only the tip of the iceberg in terms of the multiple ways resources could be used due to the situation and preferences of the institution. However, what is not clear is if the way in which the funds are allocated in one scenario versus another is better or worse. Is it better to deal with developmental education problems through instructional investments or by bolstering student support services? Is it better to invest in a faculty that skews old or young? How should one take into consideration local prices and providers? The questions go on and on.

The true difficulty in higher education finance accountability is judging what is an “effective” (in terms of bringing about positive student outcomes) and “efficient” use of resources, as the focus of many of the calls for accountability have been in response to feeling that colleges are wasteful and/or fail to focus on producing good results. Because there is no clear standard for what these two key criteria mean in absolute terms, they can only be measured

⁵ He also acknowledges the important of counting: “Narrative without counting is suspicious.”

relative to other institutions or changes within an institution over time. For example, in comparison to other schools, if two institutions have the same outcomes but one spends more, all else equal, then one might conclude that the school with higher expenditures is not using resources as efficiently or effectively. Likewise, overtime one would expect to see better outcomes as institutions spend more money. However, these comparisons across institutions and over time are still unsatisfying as measures of efficiency and effectiveness. For instance, these types of comparisons do not give one a sense of how close institutions are to using resources in the *most* effective way possible.

The allocation of resources is not only a question; there are also questions about the amount of expenditures. How much is too much to spend on a particular goal? There is no clear answer. As pointed out by Bowen (1980), “In quest of excellence, prestige, and influence, there is virtually no limit to the amount of money an institution can spend for seemingly fruitful educational ends.” In other words, nearly all spending could be justified, especially because learning and development happen in more than just college classrooms. Bowen’s point also highlights that it can be difficult to distinguish between essential activities and those that are luxuries. To sort through these issues, policymakers must have some agreement on the principal aims of an institution, but as noted above, they will also likely need to compare one institution to another.

The above challenges would apply if one were trying to discern the finances and performance of a set of near-identical institutions. However, in reality, accountability systems have to deal with the added difficulty of being applied to institutions with differing missions, student bodies, and goals. This further clouds our understanding of spending patterns and standards for what might be effective or efficient. Lederman (2006) summarizes the view that a one-size-fits-all set of criteria for judging institutions should not be applied to higher education. Writing about the effects of applying a common set of standards or measurements of what students learn, he quotes the former president of the American Council of Learned Societies: “Either there won’t be agreement, and it will be overly controversial, or it will be reduced to an elastic, lowest common denominator, as in No Child Left Behind, in which case it will become trivial.” And so the act of evaluating colleges needs the additional nuance of taking into account institutional mission and aims.

III. USING INCENTIVES TO CHANGE INSTITUTIONAL BEHAVIOR

Although most colleges and universities are non-profit, they still have a documented history of responding to incentives. For example, research has found that colleges react to the criteria used by the *US News and World Report* ranking system to inflate their scores and attract additional applicants (Ehrenberg, 2000). In another example, Long (2004) found that colleges strategically raised their room and board prices when provisions from the Georgia Hope Scholarship curtailed their ability to raise list tuition price. This suggests that incentives could be used to encourage colleges and universities to have certain types of behavior. This section considers how those incentives might be designed.

Sticks versus Carrots: Which Type of Incentives should be Used?

Incentive systems are based on some combination of sticks (i.e., penalties) and carrots (i.e., rewards). Central to this balance is the question of why there is a problem in the first place. In other words, *why is there a need for some kind of accountability?* The answer has implications for what types of incentives might be most appropriate.

One possible response is that incentives are needed because colleges and universities are not using their funds effectively due to laziness, intentionally wasteful behavior, and/or the lack of consequences for spending resources in an irresponsible way. For example, in a 2004 article published in the *Los Angeles Times*, Vedder writes:

“Colleges could increase teaching loads and use more online instruction; they could cut back on administrative staff, subsidies for intercollegiate athletics and high-cost, low-enrollment graduate programs; they could abolish tenure and contract out food and lodging operations. The only thing missing so far is *a will to change* [emphasis added]... American universities have made our nation a better place. But their inefficiency and indifference to costs could in the end bring them down.”

If true, the implication is that an accountability system should focus on closely monitoring college finances and creating a set of penalties that punish colleges for being wasteful. The key to this view is that colleges and universities are capable of doing a better job, but they fail to act due to a lack of urgency or negative consequences.

Another possibility, which lies at the other end of the spectrum, is that the lack of strong performance is due to colleges not having sufficient funds to meet the standards demanded by stakeholders. There is clearly a great deal of variation in the expenditure patterns and amount of

resources available to different kinds of institutions. In particular, colleges that focus on serving students with the most financial and academic needs have much less to spend relative to their more selective counterparts. Therefore, there is a case to be made that for some parts of higher education, the problem is that institutions need additional resources. The implication is that rewards might be the best type of incentive to use in an accountability system: if an institution demonstrates that they are using their limited funds in effective and efficient ways, then it could be given additional resources to help meet its needs.

Yet a third viewpoint on the key problem is that colleges and universities do not perform better because they lack a clear sense of best practices in terms of spending and so the failure of schools to work more efficiently is due to ignorance. Certainly, research on this issue is scant and there is little information about the educational process of colleges. In general, the production of higher education is largely considered a “black box,” in which a number of inputs hopefully mix together to produce high-quality outputs. As Lake (2009) points out, much more needs to be understood about the “science of higher learning,” meaning how students learn, which teaching tools are the most effective, and how institutions can help even those with lower academic performance succeed. Without such information, it is difficult to know how to improve teaching and student outcomes in higher education even with an influx of additional resources.

If one believes the lack of professional guidance is the true problem, then standard forms of accountability may not be the best solution. Instead, one might consider creating incentives and opportunities for the constant evaluation of funding practices linked to outcomes and then compare these across institutions to establish a set of best practices. The emphasis would be on making sure institutions share information and learn from each other. Another option would be to provide grants to help institutions develop and evaluate new and innovative practices. However, this still has complications in that successful practices might differ by institutional mission, student body, and setting as well as vary over time. In other words, the lack of a clear research foundation in this area is at least partly due to the complexity of the issue.

In all likelihood, the true problem in higher education is a combination of all three scenarios. There are examples of institutional wastefulness, cases in which an institution had too few resources, and challenges that lack a clear set of funding solutions practices. When considering the design of any accountability system, one must consider which of the views is best supported by the available data and seems to apply to the target set of institutions.

Unfortunately, within any system, it is unlikely that one story will fit all situations, and thus the challenge of an accountability system is balance allowing for flexibility while imposing constant standards.

The Size of the Incentive: How large is large enough?

While the type of incentive is important, neither sticks nor carrots will work if the incentive is too small. The key question is how large is large enough to spur real change or performance. For many public institutions, the main source of financial control from the state is in the form of operational appropriations. Awarded by the state legislature every year, these are sometimes determined by an enrollment-based or similar formula. Targeting these funds could provide clear incentives to institutions if a large enough proportion is at stake. Alternatively, many policies have focused on other sources of state funds, such as special grants or new money. Because these resources are not essential to the institution, putting them under an accountability system is less likely to have an impact even if the incentive is large.

A serious concern about the size of the incentive, regardless of source, is trends over time. The business cycle can greatly influence state revenues generally, which in turn can affect the amount of state appropriations to colleges and universities and other funds directed towards higher education. When times are bad, the financial incentives of accountability programs are often at risk as the more straightforward appropriations become a critical focus of institutions. On the other hand, a carefully-designed accountability program might take advantage of the importance of small amounts of money during recessions. If an incentive was structured to counter-balance the loss of other resources, it might be highly influential in terms of college behavior.

IV. STATE EFFORTS TO LINK INSTITUTIONAL FINANCES WITH OUTCOMES

As noted above, there are limits to the usefulness of national comparison of expenditures, whether using the IPEDS data or the more detailed accounting of the Delaware study. In contrast, states have more specific budget information to understand the finances of their particular public institutions, and with a better understanding of the particular context, they may be better able to measure and interpret cost estimates. States have experimented with various forms of finance accountability for many years. The examples below showcase the range of

decisions states have made regarding their policies, from the criteria used, to how they have evaluated college performance, to the type, size, and timing of incentives. The examples highlight the diverse actions and experiences of systems across the country.

Which Criteria and Outcomes should be used?

Performance funding or other types of finance accountability are usually not the first attempt by a state to engage with postsecondary institutions about their activities. When choosing indicators with which to judge the performance of colleges and universities, some states have based the criteria of their accountability systems on previous goals and priorities. For example, Missouri had long had discussions about the importance of assessment, and the 1991 creation of the Missouri Assessment Consortium served as a precursor to the state's later approach to performance funding. The criteria used in the accountability system were marked by their direct links previous planning priorities, and as noted by Burke (2002), using these familiar measures helped the state to avoid extreme reactions from colleges and universities.

Even if the idea of assessment is not new, when it is time to link performance indicators to finances, states must make concrete decisions about exactly what will be evaluated. Beyond the types of criteria, they must decide how these measures will be applied. Some have chosen to apply similar criteria to all college and universities, regardless of level or mission. For example, from 1994 to 1997, Arkansas judged colleges using six major set of criteria. Retention measures received the most weight (39 percent), followed by quality indicators (e.g., exam passage rates), program efficiencies, workforce development, the diversity of the faculty and staff, and graduation rates. Because the indicators were applied to two- and four-year colleges alike, they were widely criticized (Burke & Associates, 2002).

Other states have instead developed criteria that differ by institutional level. In Florida, for instance, two- and four-year colleges were judged using a different set of indicators. The community colleges were evaluated based on degree awards, graduates from particular backgrounds (e.g., required remediation, economically disadvantages, or disabled), the time to degree completion, and numbers on placements and transfers. In contrast, the four-year colleges were judged on their graduation rates (six-year rate for first-time students and four-year rate for transfer students), the percentage who graduated with credits close to the degree requirement, the

percentage who went on to graduate school in Florida, and the ratio of external research funds to state research funds.

While the criteria chosen by Florida acknowledge differences between community colleges and four-year colleges or universities, other states have allowed the criteria to vary at a finer level. There are several examples in which states have used a combination of common indicators with other criteria chosen by the institutions or their local stakeholders. In Illinois, for instance, the community colleges all had to address five statewide goals related to student satisfaction, educational advancement, success in employment or graduate school, the proportion of the population serves, and the success of academically disadvantaged students. In addition, each community college was also subject to a goal that could be related to their local district. Each institution had to select one of the following areas on which to focus: workforce preparation, technology; or responsiveness to a local need. Virginia allowed even greater institutional autonomy. Although they require public institutions to gauge and report their own performance in a range of areas, they left it up to the individual institutions to decide which measures to use.

Allowing institutions to choose their criteria can sometimes backfire. For example, in Kentucky, the Higher Education Review Commission chose 26 criteria that all campuses had to have, but the campuses were allowed to select the weights applied to each indicator. Some institutions set such low standards that their targets were below then-current levels of performance. What resulted were several years of negotiation between the commission and university presidents, but by the time there was some agreement, politicians no longer believed the policy would be successful in bringing about meaningful change (Burke, 2002).

Kansas is an example of a state that has asked colleges and universities to think not only of how to showcase their past performance but also set goals for the future. The state instructed institutions to draft goals that were then linked to new funding. “Each institution proposes its own performance contract, complete with proposed goals, proposed performance measures and proposed performance targets. The Board requires that the goals included in the proposed agreements be ‘stretch’ goals that truly challenge the institutions to step up from business as

usual” (Bogue and Hall, 2003). Institutions earn increases in year-to-year funding only if they meet certain percentages of the goals.⁶

After choosing criteria for an accountability system, states have also had to decide how to apply those indicators and make judgments about the performance of institutions. Some have done this by comparing how an institution has done in a given year relative to its own prior performance. This appears to be the preferred method by institutions as other designs pit one institution versus another. For example, in Arkansas, because the funds not claimed by low-performing schools went instead to high-performing institutions, the campuses felt it created an unhealthy atmosphere of competition (Bogue and Hall, 2003). The community colleges in Florida also criticized a plan that measured a college’s performance improvement against that of other colleges (Dougherty and Natow, 2009).

The Role of Incentives in Finance Accountability Policies

In an accountability system focused on higher education finance, the role of incentives is particularly important. The type, size, and timing of the incentives created by the policy are major factors in the determination of whether the system can spur better performance by institutions. States must determine whether the policy will incorporate rewards for meeting standards, just maintain funding levels for doing so, or enact penalties for failing to perform adequately. Then, the timing of when the reward or penalty is executed can be important to how institutions respond. Finally, the size of the incentives must be enough to encourage the intended behavior among colleges and universities.

Tennessee is an example of state that uses rewards as incentives in their accountability program. Institutions in the state can earn up to 5.45 percent of their state operating appropriations. Quoting the state’s performance funding website, "This program is a rare incentive opportunity for institutions to earn resources above and beyond formula-based appropriations." Instead of introducing new resources, Missouri designed its accountability system to reward institutions with an inflationary increase in their funding. Put another way,

⁶ According to the Performance Agreements Guidelines and Procedures: “Beginning July 1, 2005, "each postsecondary educational institutions' receipt of new state funds shall be contingent on achieving compliance with its performance agreement...The state board shall determine the amount of new state funds to be received by each postsecondary institution, taking into account the postsecondary educational institution's level of compliance with its performance agreement and the funds available for distribution." (retrieved July 30, 2009 from http://www.kansasregents.org/download/aca_affairs/initiatives/perfagree/guidelines.pdf)

institutions that met standards had their funding maintained with an adjustment for inflation. The policy was meant to be a change in philosophy—inflationary increases were no longer automatic (Burke, 2002, p. 118). Starting in 2005 in Kansas, colleges and universities that did not meet performance goals lost money from the pot of new state funds. In 2006, all but three schools received full funding. One lost 2/3rds of the funding while two institutions lost all their funding. In making these decisions, the state board takes into account the school's level of compliance with its performance agreement and the funds available for distribution.⁷

Regardless of whether the incentives are in the form of rewards or penalties, the timing of the incentive also matters. The experience of Florida emphasizes the importance of providing the incentives in a timely fashion. In that state, the community colleges criticized one accountability program, the Workforce Development Education Fund (WDEF) because of the way it left school uncertain about their funding (Dougherty and Natow, 2009). Given the importance of planning in the administration of colleges, uncertainty could undermine the incentive created by a policy.

If incentives are not large enough to elicit a response, the policy will fail. There are many examples. In Arkansas, the reward for performance was only a modest share of total state appropriations. In Florida, Performance Based Budgeting (PBB) encompassed only about one percent of state appropriations to community colleges, or \$8.3 million in 2000.⁸ Likewise in Illinois, the accountability system only put at stake 0.4 percent of state appropriations to community colleges in 2000-01. These funds were in addition to the funding schools received based on an enrollment-based formula. Minnesota serves as a fourth example: schools that met their performance indicators and standards could only get up to a one percent increase in their non-instructional budgets. One reason for the lack of strong incentives has been that most systems have avoided putting base funding at risk. Instead, funding in accountability efforts like performance funding has most often been confined to new sources of money (Burke, 2002).

There are also state models with large incentives. As noted above, Florida had WDEF for several years. Its incentive ranged up to 5.6 percent of state appropriations, and the state could withhold up to 15 percent of the prior year's workforce appropriations. In Missouri, over

⁷ Source: Kansas Performance Agreements: Guidelines and Procedures, retrieved July 30, 2009 from http://www.kansasregents.org/download/aca_affairs/initiatives/perfagree/guidelines.pdf

⁸ For some time, Florida also had WDEF, which was much larger.

time, the FFR program resulted in an increase of \$66 million in funding to the core budgets of postsecondary institutions (Burke, 2002).

Over time, the size of the incentive may grow in importance. If institutions continually do better and better, they may expect that their reward will also grow over time. Funding for the accountability system in Florida did not grow, thereby drawing criticism from the community colleges who wanted additional rewards for their improvements (Dougherty and Natow, 2009).

The Sustainability of State Accountability Efforts

Although there have been many state experiments with accountability linked to higher education finance, programs are often cut after several years, and few are around longer than a decade. There are a number of reasons for this. In Ohio during the mid 1990s, the state legislature adopted performance funding for the community colleges. However, it ended due to a myriad of problems. As noted by Burke (2002): “It suffered from too little funding, vague objectives, and uneven implementation.” For other states, there is a key problem that caused the termination of an accountability policy.

Foremost, budget cuts have been the blame for the dissolution of many state accountability systems. When there are budget cuts, colleges often prefer to cut incentive funding rather than core, formula-based funding. Such was the case in Florida. Illinois and Missouri also cut their programs during the recession of the early millennium. While fiscal crises explain the demolition of several accountability policies, economic booms can also be a culprit. In Minnesota, when the economy improved and state appropriations to higher education increased, there was less interest in performance funding, and the incentives were dwarfed relative to the main pot of money. Performance funding was then replaced by performance reporting (Burke, 2002).

Declining political support has also been the reason why some finance accountability policies have been eliminated. In Florida, after a few years, the legislators who had originally championed WDEF were no longer around, and so support for the program disappeared. Likewise, in Illinois, the key champions of the accountability effort on the state community college board were no longer there after a while. Because the new governor was not interested in performance accountability, the policy ended. Other key constituents, such as the legislature and business, also had little interest in the topic (Dougherty and Natow, 2009). Instead of a lack of

political support, sometimes the relative power of college presidents can derail accountability efforts. In Kentucky, for instance, campus presidents and local boards of trustees were able to garner greater influence on education policy after a new law limited governors to one term. As noted by Burke, this shift in power helped to kill the accountability program (Burke, 2002, p. 230).

Impatience can also have a negative effect on the sustainability of a policy. It is not clear how quickly colleges can and should respond to incentives with improvements in performance, but the political time frame is short. As discussed in a previous section, the Higher Education Review Commission and university presidents in Kentucky spent several years negotiating appropriate criteria for the performance plan. In the meantime, however, the governor created the Task Force on Postsecondary Education and resolved that "efforts to implement a meaningful system of performance funding have been ineffective, and efforts to improve the formula for funding higher education have not resulted in meaningful change."⁹

Lessons Learned from State Accountability Efforts

While research has not shown definitively that finance accountability can have positive effects, and the low rate of success among states remains disconcerting, there are, nonetheless, lessons that can be learned. First, the size of the incentive matters a great deal. If it is not large enough, it will not have an effect. For example, if the size of the incentive is dwarfed by other sources of state funding, then the accountability program will not have much of an effect. Second, to ensure sustainability, the funding for accountability systems must be maintained and from a source not susceptible to easy reductions. There are several examples of states that cut their programs during fiscal crises. Sustainability is also threatened by changes in political power. Over time, policies often lose their original champions or become the victim of growing power among the colleges and universities. The above examples also highlight criticisms about how colleges are evaluated, whether they all face the same criteria or are pitted against each other. Uncertainty about funding also can wreak havoc on the reactions of postsecondary institutions.

The literature highlights other lessons from the experiences of states. The first focuses on a major problem many systems have faced: the lack of good information. Without informative

⁹ General Assembly (1996), quoted in Burke (2002), p. 235

basic indicators and a system that helps to interpret that information, it is difficult to believe that an accountability initiative would have much success. The measures chosen can also be problematic. Ultimately, states hope that their investments into higher education yield public and private benefits such as the production of degrees (i.e., human capital) along with the new information that might be beneficial to society and the local economy (i.e., innovation). However, in their approaches, states have tended to focus on aggregated measures, such as the total number of degree awarded or the average credits taught by faculty. As emphasized by Carey (2007), very little attention has been paid to what one hopes underscores these measures: student learning (Carey, 2007). On the other hand, the positive side of the accountability movement of the 1990s is that fact that nearly every state now has institutions publicly report information on some aspects of its activities. However, more information has not necessarily translated into greater understanding of institutional performance or how that ties to higher education finance.

V. CONSIDERATIONS FOR THE FUTURE

Given current higher education trends, it is imperative that we gain a better understanding of higher education finance, productivity, and efficiency. That is the first step towards building a system that gives institutions incentives to support student access, persistence, and completion. Recent research like the Delta Cost Project underscores the holes in current finance data systems. A finer level of detail is necessary to better gauge how money is being spent and to what end. Aggregating instructional expenditures to the school level tells us little; instead, one might want to learn about the multiple forms and types of instruction, by whom, for whom, and in what subject. Standard methods of categorizing the educational enterprise are needed to start such a data collection with some understanding of the diversity of missions of postsecondary institutions. Because significant costs are also spent on other functions, such as administration, operations, and research, methods of categorizing these expenditures at a finer level is also necessary.

With more information, institutions must also do a better job explaining how these resources relate to specific goals, populations, and outcomes. Mississippi is a state that changed the conversation about budgeting from the typical set of requests to the more fundamental issue

of how all resources are being used. The Board of Trustees of State Institutions of Higher Learning discuss with institutions how spending in major functional areas does or does not match with priorities. If spending appears to be out of line with a stated goal in comparison to other Mississippi colleges or national peers, then there is a discussion about changes that should be made. The Board has also used information on spending to review expenditures for administrative and support functions, and as a result, they have made reductions in energy and purchasing.¹⁰ As shown in Mississippi, even the simple exercise of checking to make sure that institutional priorities are reflected in spending levels can be informative and helpful. Core to such discussion is examining what spending practices are based on old assumptions or decisions no longer applicable. It is no longer sufficient to justify spending patterns based on the adage that what was done last year is appropriate for this year.

The goal of improving higher education efficiency and effectiveness must also be held by multiple stakeholders, so support for the initiative does not decline over time with the changing of leaders or economic conditions. The experience of the University System of Maryland is an example of this principle. In 2000, the Chancellor began an “effectiveness and efficiency” initiative with the goal of attacking the perception that the campuses did not pay enough attention to fiscal stewardship. Multiple stakeholders worked together to develop a set of system-wide goals, including optimizing the use of resources. Wellman writes that the system has increased teaching loads and limited funding to 120 credits for baccalaureate programs. The system estimates that it has saved \$40 million during the last three years, and as a result, tuition increases have been mitigated, and political leaders have been willing to fund additional enrollments. According to Wellman (2008), “Maryland is an example of a place where transparency about spending has directly paid off in increased public credibility for higher education and a growth in state support when other states were reducing funding or raising tuitions.”

While there are certainly great challenges ahead, the accountability movement is not without examples of practices that might hold promise for future, better-designed accountability policies. The need and desire to find a way to improve postsecondary efficiency and effectiveness continues to be strong, and success is possible. While the answers are not clear, higher education is a business in which researchers work for years to uncover what once seemed

¹⁰ Wellman (2008) discusses several examples of promising practices.

inconceivable or impossible. Each year, each discussion, and each new data set refines what is understood about higher education finance and college performance.

REFERENCES

- Armstrong, Thomas, Bohl-Fabian, Lou, & Yazdi, Khalil. (2008) *The Pennsylvania State System of Higher Education's Performance Funding Program 2005-06*.
- Bogue, Ernest G. and Kimberley B. Hall (2003). *Quality and Accountability in Higher Education: Improving Policy, Enhancing Performance*. Westport, CT: Greenwood Publishing Group.
- Bowen, Howard R. (1980) *The Costs of Higher Education: How Much Do Colleges and Universities Spend per Student and How Much Should They Spend?* San Francisco, CA: Jossey-Bass Inc.
- Brinkman, Paul T. (1990) "Higher Education Cost Functions," in *The Economics of American Universities: Management, Operations and Fiscal Environment*, S.A Hoenack and E.L. Collins, Eds. Albany, NY: State University of New York Press.
- Burke, Joseph C. & Associates (2002). *Funding Public Colleges and Universities for Performance: Popularity, Problems, and Prospects*. Albany, NY: Rockefeller Institute Press.
- Burke, Joseph C. & Associates (2005). *Achieving Accountability in Higher Education: Balancing Public, Academic, and Market Demands*. San Francisco, CA: Jossey-Bass.
- Burke, Joseph C. and Henrik Minnassians. (2003). *Performance Reporting: "Real" Accountability or Accountability "Lite": Seventh Annual Survey 2003*. Albany, NY: Nelson A. Rockefeller Institute of Government.
- Burke, Joseph C. and Henrik Minnassians. (2002). *Performance Reporting: Preferred "No Cost" Accountability Program – The Sixth Annual Report 2002*. Albany, NY: Nelson A. Rockefeller Institute of Government.
- California Legislative Analyst's Office. *Analysis of the 2002-03 Budget Bill*. http://lao.ca.gov/analysis_2002/education/ed_30_6870_an102.htm#_Toc1448325
- Carey, Kevin. (2007) "Truth Without Action: The Myth of Higher-Education Accountability," *Change*, September-October.
- Cunningham, Alisa F., Jane V. Wellman, Melissa E. Clinedinst, Jamie P. Merisotis, and C. Dennis Carroll. (2001) *Study of College Costs and Prices, 1988–89 to 1997–98: Volume I*, National Center for Education Statistics Report 2002–157. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.
- Delta Cost Project. (2009) *Trends in College Spending: Where does the money come from? Where does it go? What does it buy?*, Supported by Making Opportunity Affordable, an initiative of Lumina Foundation for Education. Washington, DC: Delta Cost Project.
- Dougherty, Kevin J. and Rebecca S. Natow (2009). *The Demise of Higher Education Performance Funding Systems in Three States*. CCRC Working Paper No. 17. Accessed July 29, 2009 from <http://ccrc.tc.columbia.edu/Publication.asp?UID=693>.
- Dougherty, Kevin J. and Monica Reid. (2007). *Fifty States of Achieving the Dream: State Policies to Enhance Access to and Success in Community Colleges Across the United States*. <http://ccrc.tc.columbia.edu/Publication.asp?UID=504>.

- Dowd, Alicia C. and Vincent P. Tong. (2007). Accountability, assessment, and the scholarship of "Best Practice". *J.C. Smart (ed.), Higher Education: Handbook of Theory and Research, Vol. XXII*, 57–119.
- Education Sector. (2008). Using Information: Funding. Accessed on July 29, 2009 from http://www.educationsector.org/usr_doc/funding.pdf
- Education Sector. (2008). Kansas Score Card. Accessed on July 30, 2009 from http://www.educationsector.org/usr_doc/Kansas.pdf.
- Ehrenberg, Ronald G. (2000) *Tuition Rising: Why Colleges cost so much*. Cambridge: Harvard University Press.
- Harbour, Clifford P. (2002). The Legislative Evolution of Performance Funding in the North Carolina Community College System. *Community College Review*, 29(4): 28-49.
- Hess, Frederick M., Mark Schneider, Kevin Carey, and Andrew P. Kelly. 2009. *Diplomas and Dropouts: Which Colleges Actually Graduate Their Students (and Which Don't)*. Washington, D.C.: American Enterprise Institute.
- Hossler, Don. (2004). Refinancing Public Universities: Student Enrollments, Incentive-Based Budgeting, and Incremental Revenue. *Edward P. St. John & Michael D. Parsons (eds.), Public funding of higher education: Changing contexts and new rationales*. Baltimore: Johns Hopkins University Press.
- Jones, Dennis P. (2000) "An Alternative Look at the Cost Question," In *Higher Education Cost Measurement: Public Policy Issues, Options, and Strategies*. Washington, DC: The Kane, Thomas and Peter Orszag. 2003. "Higher Education Spending: The Role of Medicaid and the Business Cycle," *The Brookings Institution Policy Brief #124*. Washington, DC: The Brookings Institution.
- Jones, Dennis P. (2001) "Cost Analysis and the Formulation of Public Policy," *Study of College Costs and Prices, 1988–89 to 1997–98: Volume 2*, National Center for Education Statistics Report 2002–158. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.
- Kremer, Michael and Dan M. Levy. 2003. "Peer Effects and Alcohol Use among College Students," National Bureau of Economic Research Working Paper No. 9876.
- Lake, Peter F. (2009) "Will Your College Be Sued for Educational Malpractice?" *The Chronicle of Higher Education, Commentary*, August 11.
- Lederman, Doug. (2006) "No College Left Behind?" *Inside Higher Ed*, February 15.
- Long, Bridget Terry (2004) "How do Financial Aid Policies affect Colleges? The Institutional Impact of the Georgia HOPE Scholarship." *Journal of Human Resources*, vol. 39, no. 3.
- Long, Bridget Terry. (2007) "The Contributions of Economics to the Study of College Access and Success." *Teachers College Record*.
- Long, Bridget Terry. (in press) "Higher Education Finance and Accountability." In Kevin Carey and Mark Schneider, Eds. *Best in the World? Accountability and Unaccountability in American Higher Education*. Washington, D.C.: American Enterprise Institute.
- Long, Bridget Terry and Dana Ansel. (2007) As Student Debt Increases, Colleges Owe more in

- Performance. *Connection: The Journal of the New England Board of Higher Education*, Winter: 23-24.
- Middaugh, Michael F, Rosalinda Graham, and Abdus Shahid. (2003) *A Study of Higher Education Instructional Expenditures: The Delaware Study of Instructional Costs and Productivity*, National Center for Education Statistics Report 2003–161. Washington, DC: U.S. Department of Education, Institute of Education Sciences.
- Miller, Charles. (2006) “Memo from the Chairman,” *Inside Higher Ed*, January 24.
- Missouri Board of Higher Education. (2008). Recommendations for Performance Funding Appropriations. <http://www.dhe.mo.gov/mdhe/boardbook2content.jsp?id=573>.
- Missouri Board of Higher Education. FY 2010 Higher Education Operating Budget Status. Downloaded July 29, 2009 from <http://www.dhe.mo.gov/files/budgetstatus.pdf>.
- Sacerdote, Bruce. (2001). “Peer Effects with Random Assignment: Results for Dartmouth Roommates,” *Quarterly Journal of Economics*, May.
- Shulman, Lee S. (2007). “Counting and Recounting: Assessment and the Quest for Accountability.” *Change*, January-February.
- Spence, David S. (2006). Accountability in the California State University System. *New Directions for Higher Education*, 135: 103-110.
- Tennessee Performance Funding: Frequently Asked Questions. Retrieved August 1, 2009 from http://tennessee.gov/thec/2004web/division_pages/academic_pages/performance_funding/performancefundingfaq.html.
- Vedder, Richard. (2004) “Colleges Have Little Incentive to Hold Down Costs,” *Los Angeles Times*, July 18.
- Washington State Board of Community and Technical Colleges. Student Achievement Initiative. http://www.sbctc.ctc.edu/college/e_studentachievement.aspx.
- Wellman, Jane V. 2008. “The Higher Education Funding Disconnect: Spending More, Getting Less,” *Change*, November-December.
- Winston, Gordon. (1998) “Economic research now shows that Higher Education is not just another Business.” *Chronicle of Higher Education*, March 27, page B6.
- Winston, Gordon. (2001) “Higher Education’s Costs, Prices, and Subsidies: Some Economic Facts and Fundamentals.” *Study of College Costs and Prices, 1988–89 to 1997–98: Volume 2*, National Center for Education Statistics Report 2002–158. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.
- Winston, Gordon and David Zimmerman. (2004). “Peer Effects in Higher Education.” In *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It*, ed. Caroline M. Hoxby. Chicago: University of Chicago Press: 394-421.