

College Ratings: What Lessons Can We Learn from Other Sectors?

March 2014

Nicholas Hillman

Last summer, President Obama directed the U.S. Department of Education to develop a federal college ratings system. A key goal behind this system is to provide students and families with information for selecting “schools that provide the best value.”^[i] In early 2014, this plan started to take shape by convening a technical symposium for developing the new “Postsecondary Institution Ratings System” (PIRS). By the 2015 academic year, the White House hopes to roll out this ratings system and eventually (after Congressional approval) link federal student aid to these ratings.

Proponents believe this is a long-overdue reform and one that will ultimately improve student success. With ratings, students ostensibly will be able to make better educational choices since they will have information about how accessible and affordable a college might be, and whether it provides a quality education. This market logic works fairly well when purchasing products like refrigerators, food, or cars. Energy Star labels help customers decide which refrigerator will be the most energy efficient. Organic labels help customers determine which tomato is healthier for them. Consumer Reports helps shoppers determine which car will offer the best value for their money. Following this logic, PIRS will achieve the federal government’s goal of helping students make better choices about where to go to college.

But what makes for a “quality” education is difficult to measure in any meaningful way. Even the most well-intentioned and well-designed ratings system will have difficulty reducing a college’s “quality” into a standardized measure. Nevertheless, proponents have offered the following measures: graduation rates and time-to-degree, job placement rates and future earnings, along with debt-to-income ratios and student loan default rates.^[ii] Some have even proposed publishing a list of the names and qualifications of instructors, accreditation self-studies, and the ratio of tuition revenue relative to instructional spending as part of this reform.^[iii] Whether these indicators truly represent a college’s “quality” is open for debate. Regardless of what measures are ultimately included in the ratings system, there are fundamental theoretical and empirical reasons why ratings will have little value in the higher education marketplace.

These reasons are explored below, focusing on the unintended consequences of tying aid to ratings. From this vantage point, we can see several ways ratings can create *greater* confusion and stratification in an already highly unequal higher education marketplace. Despite these shortcomings, ratings proponents have rushed into a discussion about “what to measure” and “where to find” performance data before explaining “how” ratings might affect students’ behaviors. Their explanation should answer two basic questions.

Viewpoint

www.learningoutcomesassessment.org

First, is there any evidence that consumers decide what to purchase *because of* ratings systems? Second, do higher education markets operate according to the same principles that drive consumer product markets? Proponents of college ratings seem to think so, while the economics of education literature and a mounting body of evidence suggest otherwise.

Taking the second question first, three defining features separate higher education markets from markets for consumer products:

1. Students do not know the “value” of their purchase until after they have completed their education.
2. Students are both consumers and producers of their own educational experience.
3. Colleges maximize reputation and prestige to stay competitive in student marketplaces, and there is no empirical evidence that either are good surrogates for educational quality and student learning outcomes.

These features are not unique to higher education; they are common among most other service industries. For example, a hospital patient does not purchase a treatment (or know its value) until *after* they receive it. They arrive at the hospital as an *input to*, but also a *consumer of*, their own treatment. And the hospital is expected to maximize its *reputation and prestige* to ensure a steady demand of patients. To complicate this even further, every patient has a unique definition of what makes for a “high quality” or “valuable” experience.

The concept of “quality” is very subjective because each individual evaluates, according to their own preferences, whether something is “good or bad.” Education, like health care, is highly individualized where there is no universal definition of what makes for a “high quality” experience. Furthermore, “quality” may be better measured *qualitatively*. When one tries to quantify it, it immediately suffers from measurement error, aggregation bias, and raises construct validity concerns. A quick example will illustrate this point.

A patient may judge the quality of a hospital in terms of their doctor’s bedside manner, helpfulness of staff, or how quickly they are able to receive their treatment. If these factors are not included in a rating system, then a patient could derive a “high-quality” experience even from a poorly-rated hospital. A different patient receiving the same treatment at the same hospital may judge “quality” according to the pain they felt after surgery or whether their condition was cured. If these factors are not reflected in the rating score, then this patient would not find the hospital’s ratings very useful.

When considering “quality” in higher education, a student may judge a college according to the way their faculty creates engaging teaching and learning experiences inside the classroom, or how safe they feel on campus. Another may place a high priority on convenience and flexibility, so they would judge a college’s quality according to these preferences. These examples illustrate the extent to which “quality” varies according to each individual person’s goals, needs, expectations, constraints, preferences, and host of other contextual and personal factors.

To apply an aggregate measure of “quality” to a hospital, college, or any other service enterprise, there is bound to be a high degree of variation in terms of how useful any sort of “quality rating” is for consumers. This bears out in the literature, where a growing body of evidence suggests ratings are not very useful tools for changing consumer preferences or behaviors, let alone outcomes. This body of research can teach us some valuable lessons

The concept of “quality” is very subjective because each individual evaluates, according to their own preferences, whether something is “good or bad.”

since hospitals, nursing homes, and other service sectors are subject to similar “awkward economics” as higher education.[iv]

The Journal of the American Medical Association has published several reviews concluding that ratings, report cards, and other public “quality information” documents rarely and non-systematically affect consumer behaviors.[v] Despite the mounting evidence, the growing sense of “naïve optimism” overstates the role information plays in shaping people’s behaviors.[vi]

The evidence is simply too weak to conclude that ratings or other forms of “quality information” positively affect consumer’s choices or outcomes. In fact, ratings have modest behavioral effects and can even have the unintended consequence of *worsening* outcomes, as briefly summarized by these important studies:

- Cardiac surgeons turned away the sickest and most severely ill patients after adopting performance-based health report cards.[vii]
- Health disparities widened among White, Black, and Hispanic patients after introducing physician report cards.[viii]
- The total cost of care increased after implementing medical performance ratings.[ix]
- Physicians more frequently dismissed patient preferences to meet “target rates” for interventions, even when these were not in the best interest of the patient.[x]
- Information related to service fees and health costs did not change how much patients spent on healthcare or how often their visits occurred.[xi]
- There is scant evidence that ratings changed patients’ choices of medical care providers. [xii]

It is not difficult to see why these outcomes might occur. Consider a nursing home that has poor health outcomes, underpaid staff, out-of-date equipment, etc. This place would likely receive a low rating and conventional economics would say it is irrational for anybody to choose to attend this nursing home. But what if there were few alternative nursing homes nearby? What if this poorly-rated facility actually left some residents better-off than if they had never received the care? Many people are place-bound, culturally and socially entwined in their communities, so it is very rational for them to stay close to home. Plus, the outcomes are impossible to know in advance, so it may be worth the risk despite the poor rating. Several factors beyond “better information” drive people’s behaviors, and a growing body of research shows how ratings, report cards, and rankings have negligible (and sometimes harmful) impacts on consumer behaviors.

The implications of what is known about ratings and consumer behavior raise serious concerns about the utility of PIRS and the future of higher education equity and access:

- Not all students are mobile or have the luxury of “shopping around” for colleges;
- The concept of “quality” means something different for various types of students;
- Students are both the input and output of their education;
- It is impossible to know the outcomes of an education until after;
- Colleges may have perverse incentives to reduce access in exchange for higher ratings.

Ratings proponents believe PIRS is a long-overdue reform and that tying ratings to federal financial aid will improve students’ success. But they do not consider the unintended consequences outlined here. Instead of answering questions about “*bon*” ratings might

The evidence is simply too weak to conclude that ratings or other forms of “quality information” positively affect consumer’s choices or outcomes. In fact, ratings have modest behavioral effects and can even have the unintended consequence of worsening outcomes, as briefly summarized by these important studies:

affect students' behaviors, proponents are jumping to decisions about “*what to measure*” and “*where to find*” performance data. To justify their hasty actions, they characterize ratings critics as defending the status quo.

Rather than defending the status quo, critics are calling for public policy informed by evidence and grounded in theory. This is not too much to ask, particularly when trying to solve educational equity and access problems. There is a naïve optimism surrounding the ratings debate, where the provision of “better” consumer information will somehow transform the higher education marketplace. I have yet to hear an evidence-based and theoretically sound answer to the fundamental question: *how will ratings affect student behaviors?*

The onus is on the proponents to answer this question. But judging from the evidence outlined here, and considering the shortcomings of applying market-based solutions to education, it is unlikely even a well-designed college rating system will help *all* students make better educational choices. In fact, there is a very good chance that ratings will negatively affect students who are already under-served and under-represented in higher education. Rationing student aid according to poorly conceived measures of “quality” is worrisome enough, but pursuing this course of action without an evidence-based or theoretically supported rationale is just bad public policy. Moving forward, proponents should be attentive to these concerns and, at a minimum, offer a convincing justification for why ratings are the best solution to the problems outlined here.

References

[i] The White House. (2013, August 22). Fact Sheet on the President’s Plan to Make College More Affordable: A Better Bargain for the Middle Class. Retrieved February 21, 2014, from <http://www.whitehouse.gov/the-press-office/2013/08/22/fact-sheet-president-s-plan-make-college-more-affordable-better-bargain->, Also see Federal Register (2013). Request for Information to Gather Technical Expertise Pertaining to Data Elements, Metrics, Data Collection, Weighting, Scoring, and Presentation of a Postsecondary Institutions Ratings System. 78(242), pp. 76289-76291, <http://www.gpo.gov/fdsys/pkg/FR-2013-12-17/pdf/2013-30011.pdf>

[ii] For a summary of the PIRS symposium, see: Field, K. (2014, February 6). Skepticism Abounds at Education Dept.’s College-Ratings Symposium. *The Chronicle of Higher Education Blogs: The Ticker*. Retrieved from <http://chronicle.com/blogs/ticker/live-blog-the-education-dept-s-technical-symposium-on-college-ratings/72377>

[iii] Robert Shireman (2014). Federal college ratings: three modest steps. College Guide Blog. Washington Monthly. http://www.washingtonmonthly.com/college_guide/blog/federal_college_ratings_three.php

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

[v] Krumholz, H., Rathore, S., Chen, J., Wang, Y., & Radford, M. (2002). Evaluation of a consumer-oriented internet health care report card: The risk of quality ratings based on mortality data. *Journal of the American Medical Association*, 287(10), 1277–1287. doi:10.1001/jama.287.10.1277 And Marshall, M., Shekelle, P., Leatherman, S., & Brook, R. (2000). The public release of performance data: What do we expect to gain? a review of the evidence.

I have yet to hear an evidence-based and theoretically sound answer to the fundamental question: *how will ratings affect student behaviors?*

Journal of the American Medical Association, 283(14), 1866–1874. doi:10.1001/jama.283.14.1866; Schneider, E., & Epstein, A. (1998). Use of public performance reports: A survey of patients undergoing cardiac surgery. *Journal of the American Medical Association*, 279(20), 1638–1642. doi:10.1001/jama.279.20.1638 And Walter, L., Davidowitz, N., Heineken, P., & Covinsky, K. (2004). Pitfalls of converting practice guidelines into quality measures: Lessons learned from a performance measure. *Journal of the American Medical Association*, 291(20), 2466–2470. doi:10.1001/jama.291.20.2466; Werner, R. M., Norton, E. C., Konetzka, R. T., & Polsky, D. (2012). Do consumers respond to publicly reported quality information? Evidence from nursing homes. *Journal of Health Economics*, 31(1), 50–61.

[vi]Folland, S. T. (1985). The effects of health care advertising. *Journal of Health Politics, Policy and Law*, 10(2), 329–345. doi:10.1215/03616878-10-2-329

[vii]Omoigui, N. A., Miller, D. P., Brown, K. J., Annan, K., Cosgrove, D., Lytle, B., ... Topol, E. J. (1996). Outmigration for coronary bypass surgery in an era of public dissemination of clinical outcomes. *Circulation*, 93(1), 27–33. And Burack, J. H., Impellizzeri, P., Homel, P., & Cunningham Jr, J. N. (1999). Public reporting of surgical mortality: a survey of New York State cardiothoracic surgeons. *The Annals of Thoracic Surgery*, 68(4), 1195–1200.

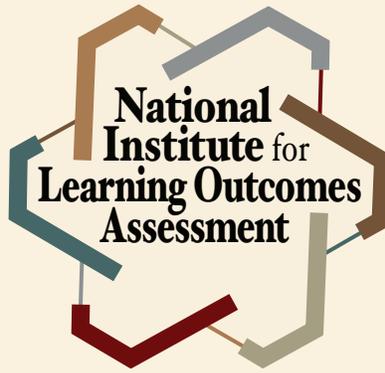
[viii]Werner, R. M., Asch, D. A., & Polsky, D. (2005). Racial profiling: The unintended consequences of coronary artery bypass graft report cards. *Circulation*, 111(10), 1257–1263.

[ix]Dranove, D., Kessler, D., McClellan, M., & Satterthwaite, M. (2003). Is more information better? The effects of “Report Cards” on health Care providers. *Journal of Political Economy*, 111(3), 555–588. doi:10.1086/374180

[x]Werner, R. M., & Asch, D. A. (2005). The unintended consequences of publicly reporting quality information. *Journal of the American Medical Association*, 293(10), 1239–1244.

[xi]Hibbard, J. H., & Weeks, E. C. (1989). Does the Dissemination of Comparative Data on Physician Fees Affect Consumer Use of Services? *Medical Care*, 27(12), 1167–1174.

[xii]Fung, C. H., Lim, Y.-W., Mattke, S., Damberg, C., & Shekelle, P. G. (2008). The evidence that publishing patient care performance data improves Quality of care. *Annals of Internal Medicine*, 148(2), 111–123. doi:10.7326/0003-4819-148-2-200801150-00006 And Werner, R. M., Norton, E. C., Konetzka, R. T., & Polsky, D. (2012). Do consumers respond to publicly reported quality information? Evidence from nursing homes. *Journal of Health Economics*, 31(1), 50–61.



Please Cite As:

Hillman, N. (2014, March). *College ratings: What lessons can we learn from other sectors?* Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

Follow us on social media:



@NILOA_web



@LearningOutcomesAssessment

Sign up to receive our monthly **NILOA Newsletter** and stay up to date with our research and publications.



Viewpoint

www.learningoutcomesassessment.org