Selective Excellence in Doctoral Education:

Targeting Analytics to Best Use

Dr. Cathy J. Lebo
Johns Hopkins University
March 2011

Abstract

The NRC Assessment of Research-Doctorate Programs is described as "an unprecedented collection of data on over 5,000 doctoral programs in 62 fields at 212 universities in the United States." These data can be used to compare programs across campus and across universities. The data can be used to compare programs on a single scale or on unique scales, customized to each program. In turn each of our doctoral programs tells its own story, providing historical, political, and financial context for their NRC ranking. There is no shortage of analytical opportunities here, but how can we target our efforts to identify the truly distinctive capabilities of our programs and understand how they are differentiated from competitors? How can we gain insights from the NRC study to shape key decisions?

This paper introduces an analytical approach for using the NRC results to assess doctoral education. First, we need to think about the big picture and define key performance factors for each program. What factors drive the performance of each program? Is the program sufficiently distinct from other programs? What are the implicit assumptions behind each program, and which of these may be obsolete or reversible? What problems are resistant to solution? Second, consider how you make decisions about graduate education. Identify areas that would benefit the most from better analytics. Can analytics be used to reduce time to degree or improve completion rates? Can programs effectively share scarce or expensive resources? Can we identify interdependencies and places where programs could work together? Do we need better forecasts for graduate enrollment? What level of investment and effort is warranted in the potential benefits from these studies?

The NRC Assessment of Research-Doctorate Programs can be used to identify to define key performance factors and judge where analytics would have the greatest impact. Doctoral programs tend to be primarily guided by internal pressures and may stray from their original purpose. The rankings force departments to take a hard look and reevaluate objectives for their programs. The analytical approach introduced in this study encourages a purposeful response, targeting analysis at strategic differences and key processes.

Higher education is an institution that evolves slowly. American universities are linked to institutions in England, France, and Germany dating back to the 11th century. We have an excellent illustration of these historic traditions in graduate education, especially in the nature of training for the doctor of philosophy degree. Today we face new challenges – demands for more consumer information and external accountability, the rising cost of instruction and laboratory costs for the sciences, changes in federal support for research, and limited returns from endowments. We have many reasons to think carefully as we make decisions about academic programs. Present circumstances also provide an opportunity to consider how we make decisions in higher education. Typically we do not have the kind of information needed to manage and evaluate doctoral programs. Essential data have not been available to all decision makers – presidents, provosts, graduate deans, and directors of doctoral programs. The NRC Assessment of Research-Doctorate Programs, released in September 2010, provides a wealth of information about more than 5,000 doctoral programs in 62 fields at 212 universities. How should the results of this study be used to inform academic decision making? We can work our way through the results, providing reports and specialized studies useful to programs, departments, and senior executives. We have prepared studies to explain the results and illustrate the competitive position of programs and universities. There will be numerous specific examples of how the NRC results are being used. However, it would be more effective if we started down this path with a plan for making analytical choices.

What are the distinctive qualities of our doctoral programs and what are the most important decisions that have to be made about those programs? Just a little more than 1% of the U.S. population age 25 or older has earned a doctorate (US Census, Current Population Survey 2009, Table 2 – Educational Attainment of the Population 25 Years and Over by Selected Characteristics).

Most people who start a doctoral program have the ability to do the work, yet many Ph.D. programs are plagued by high attrition and long time to degree. More than twice as many men as women have earned a doctorate. Minorities are sparsely represented in many fields. How can we target our efforts at analysis, tapping the full potential of the information released by the NRC? Finally, we also need to consider the limits of the data from the current NRC assessment of doctoral programs. What work remains to realize the full potential of this type of analysis?

As more information becomes available about graduate education, we have to learn how to put these facts into context and to best use. The objective is to inform decisions -- bringing facts to bear on key questions, interpreting and attaching meaning to those facts by viewing information through the lens of experience. The facts don't speak entirely for themselves. Each doctoral program can provide historical, political, and financial context for its NRC ranking. In many cases the NRC data by itself will not provide a clear direction. However, our goal is to find insights in a mountain of information, taking both facts and context into account. Thomas Davenport, lead author of Analytics at Work: Smarter Decisions, Better Results (2010), offers a framework for making decisions. His approach is also outlined in a recent article in Harvard Business Review ("Make Better Decisions", 2009). Davenport proposes a matrix of analytical questions that can be used to define choices within business domains. Can this approach be adapted to higher education? While much is written about the value of business analytics, even its proponents concede that many companies fail to use information in a meaningful way. Decision-making in higher education, as in business, is still largely individually based. We can take opportunities to apply the NRC data as we find them or we can be more purposeful about our analytical efforts and first define the big picture questions. We need to lead with strategic questions, not with illustrations of a hundred data elements.

Selective Excellence in Doctoral Education

What are the big picture questions in doctoral education? Can the results of the NRC study be used to address those questions? Fundamentally we all grapple with the same issue – selective excellence in doctoral programs. Program strengths vary within a division, within a university, and over time. No university has immunity on this point. Even the very best schools have higher and lower ranked programs. They have some programs with a narrow range on their ranking and other programs where the ranking is less clear. They have new doctoral programs and established programs. Some programs are in transition, shifting to new areas of research or changing with faculty retirements. The comprehensive Survey and Regression models used to rank doctoral programs in the NRC study confirmed several key points. There are multiple paths to success. Even the very best programs are not excellent at every aspect of the ranking. In some cases, the range calculated for the ranking is quite broad because programs performed well on some key indicators but poorly on other indicators.

Selective excellence may also mean that we have limited coverage across subfields of a discipline. The NRC study does not address coverage within a field, an element of some importance to prospective graduate students. One strategy for success is to limit coverage in smaller programs. Some Anthropology programs cover all four subfields, while others focus exclusively on social-cultural anthropology. We all work within the limits of available resources, insufficient to make every program excellent at every aspect of doctoral education. The result is that all of us, even the very best institutions in the country, have to make decisions about which programs and which aspects of programs deserve attention.

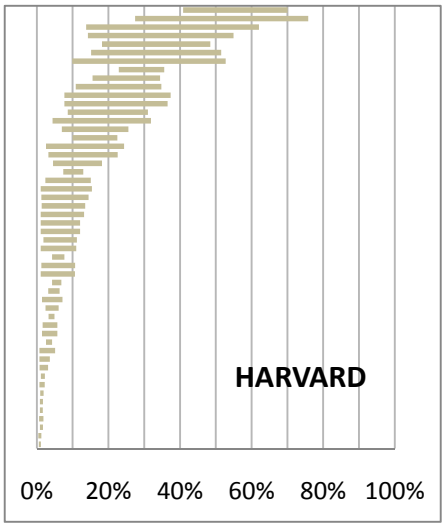
Figure 1 shows the S ranking (from 5th to 95th percentile) for all of the doctoral programs at Harvard and MIT. The rank for each program is divided by the number of programs in the field, to

provide a comparable measure across disciplines. The highest ranked programs are closest to the left side of the chart. The ranking for each program is sorted by the middle of the range.

These profile charts give a quick overall view of the array of doctoral programs at each university. There are clear differences even between two of the best universities in the country. MIT has fewer doctoral programs and very tight ranges on their S ranking. Harvard has more doctoral programs and more programs with broad ranges on their S ranking. The profiles for two universities can be compared by the degree of divergence from the y axis. How quickly do the rankings drift right? There is a face validity to these profiles which suggests that in spite of unresolved issues for specific programs (e.g. should books have been counted for this field, should this program have been placed in a different field in the taxonomy) there is a fundamental difference between doctoral programs at MIT and Harvard and doctoral education at many other institutions.

Figure 2 and Figure 3 show the number of doctoral programs and total number of doctoral students at a group of selective private universities. These figures reflect the number of programs submitted to the NRC ranking and doctoral enrollment reported for fall 2005. Clearly the academic array is different at MIT and they have limited the number of doctoral programs offered. MIT has 28 doctoral programs ranked in the NRC study, compared to 52 programs at Harvard and Johns Hopkins. Within this group of very similar private universities, the number of doctoral programs ranges from 25 (Cal Tech) to 63 (Cornell). The total doctoral enrollment ranges from 1188 (Cal Tech) to 3812 (Harvard). MIT and Cal Tech have a similar number of programs (Cal Tech 25, MIT 28), but the doctoral enrollment at MIT is nearly three times the enrollment at Cal Tech (Cal Tech 1188, MIT 3539). Even where institutional missions are similar, strategies for doctoral education vary.

Fig. 1 – S Ranking of Doctoral Programs at Harvard and MIT,
adjusted by number of programs in each ranking field.



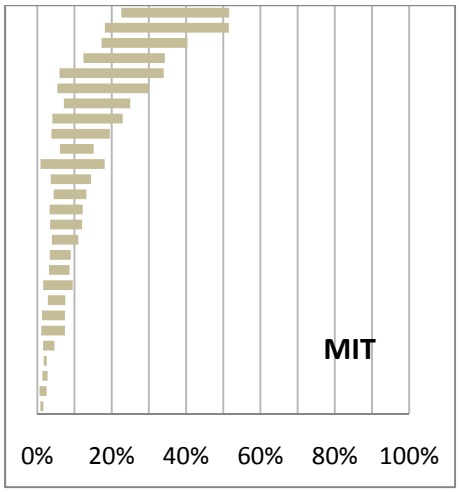


Fig 2 – Number of Doctoral Programs at selected private universities, fall 2005.

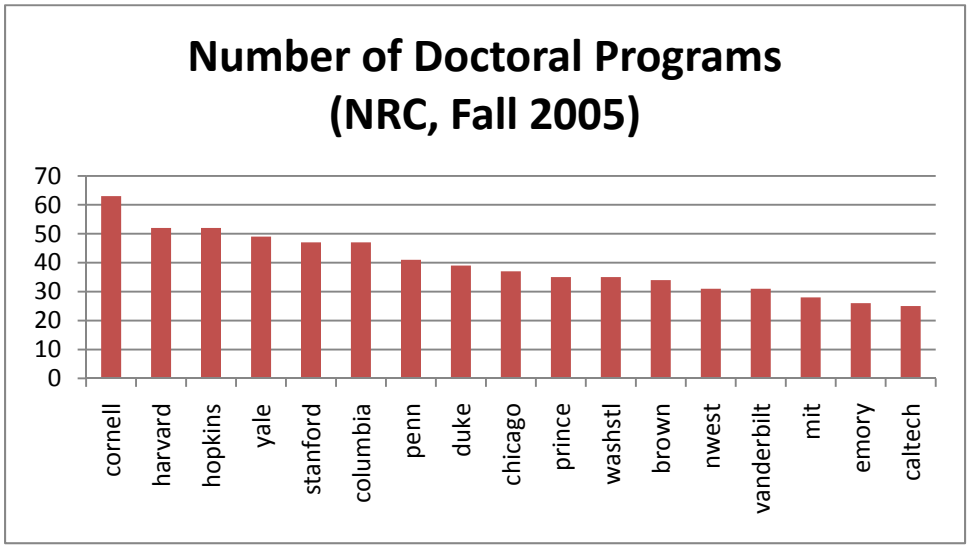
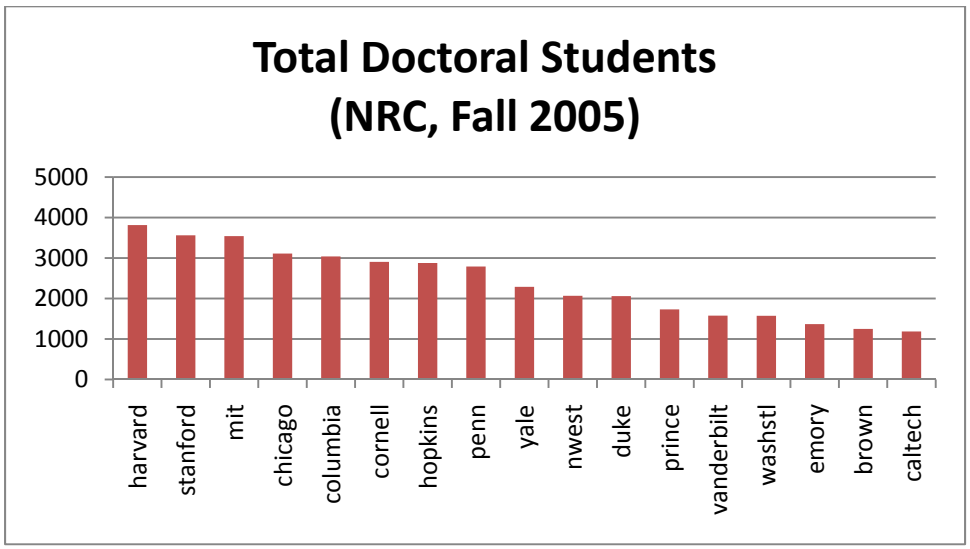


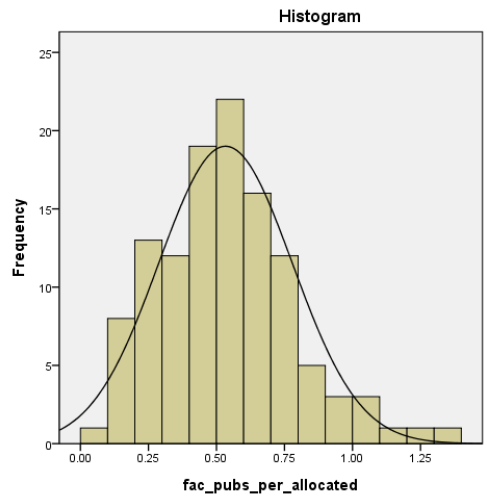
Fig 3 – Total Doctoral Enrollment at selected private universities, fall 2005.



There are many potential uses for the results of the NRC assessment of doctoral programs. We can use this information to rank the competitive position of programs and universities. Who is better -- Harvard or MIT? Should we be asking this question if the mission of these two universities is fundamentally different? Harvard and MIT made different choices about fields of study, the number and size of doctoral programs. The NRC study provides consumer information useful to prospective graduate students, including information on time to degree, completion rates, and financial support. We can also use the results of the NRC study to manage academic programs, by looking at individual data elements that shed light on components of doctoral education. These are three distinct purposes (rankings, consumer information, and program management) that should not be conflated.

The most important potential use of the NRC results is to manage academic programs. Questions about rankings can be a distraction from this purpose, ignoring differences in mission and role across universities and across doctoral programs. Another reason to curb our natural enthusiasm for rankings is that many of the key data elements in the NRC study don't provide the kind of information needed to support linear rankings. There is very little spread on some key variables, resulting in large clusters of programs with the same value. Significant differences in rank are then based on truly insignificant differences in the underlying value. Figure 4 shows the distribution of publications per allocated faculty for all Economic programs in the NRC study. More than a hundred programs are distributed between a value of 0 and 1.36. The difference between the average score (.53) and the best score (1.36) is based on a difference of a fraction of a publication per person. Yes, we can make a decision about the relative importance of publications in determining the quality of faculty. However, not much is gained by plugging a data element with so little range into the regression equation. We need better measures of publications and citations.

Fig. 4 – Publications per Allocated Faculty
in Economics Doctoral Programs.



ECONOMICS	
programs	117
mean	.53
max	1.36

When the results of the NRC Assessment of Research-Doctorate Programs were released in September 2010, our immediate problem was to explain, translate, and unpack the results of the study for graduate deans and program directors. The Wikipedia definition of "information overload" should say "see NRC study". The NRC study provides so much information on doctoral education that it is difficult to make decisions. The people who lead academic programs want to know two things. They want to know where to start in the midst of so much information. They also want to know how they can improve specific aspects of programs. They need actionable information. We were asked to reduce the complexity of the results to a clearer picture. In many ways, the comprehensive S and R rankings obscure what you need to know to improve programs. It is easy to be trapped into discussions about how to win the rankings game, rather than affirming local standards for quality.

Davenport's method for targeted analytics is one way to answer their questions, making the most effective use of analytical resources, without getting pulled astray onto a thousand possible questions. He suggests that we begin by defining key performance factors for each program. What are the prime drivers for each program and how is program A different from program B? The results of the NRC study can be used to identify key characteristics by comparing programs within and across fields. For example, some programs have higher rates of admissions coupled with lower completion rates. A program that operates this way is a different kind of program from one that admits very few students, provides full funding to each, and has high completion rates. The way a particular program is functioning may be rooted in the discipline or be purely local. It could also be an atypical portrayal of the program at one moment in time during the NRC study. Programs have been created and terminated since the NRC study began. Some faculty rosters have changed in significant ways since 2006, especially in small programs. Nobel prizes have been awarded and faculty inducted into the national academies since 2006. Universities continue to invest in and redefine doctoral programs.

Keeping these caveats in mind about the limits of the NRC data, the NRC assessment of doctoral education can be used to define key performance factors for each program. This exercise alone will help decision-makers see the strategies being used across campus, place these strategies in perspective compared to peer programs in the same field, ask questions about the value of these strategies, and see opportunities to address impediments to progress. What factors drive the performance of each program? What are the implicit assumptions underlying each doctoral program and which of these may be obsolete or reversible? Are the key performance factors for a given program typical or unusual for the discipline?

The approach used for targeted analytics is fairly straight forward. Identify key processes in doctoral education -- like Admissions, Student Persistence, and Diversity. Select appropriate indicators

and define boundary conditions for each of these components of doctoral education. Identify peer universities for each discipline. Calculate national and peer standards for each component. Compare the results for the local program to these national and peer standards. Is the local program operating like all programs in the field, like the best peer programs, or in an atypical fashion? Finally, use this process to define the key analytical questions. What questions should we be asking about each aspect of our doctoral programs?

Performance Domain – Admissions

First, let's consider Admissions as a performance domain. The NRC study includes a weighted average GRE score (math for most fields, verbal for fields in the humanities). Completion rates could be considered part of a success metric for Admissions. You hope to admit highly qualified students who complete the program. Figure 5 compares national and peer data on GRE scores and completion rates for doctoral programs in Economics. The national data includes all participating programs in the NRC study. The peer universities are highly selective, private universities including Ivy League and non-Ivy institutions. The proportion of doctoral programs with high GRE scores (equal to or greater than 750) and high completion rates (equal to or greater than 40% of students who start the program complete the doctorate) varies across each discipline. The shaded cells report the count of the top programs, with high GRE scores and high completion rates.

The peer universities have a clear advantage in attracting students with the highest GRE scores and most of the peer programs have completion rates above 40%. A significant number of doctoral programs in Economics at institutions that aren't typically considered peer universities also meet or exceed these standards. These tables raise several basic questions. Why are the Economics programs with the highest GRE scores equally split between high and low completion rates (on the table for all

universities)? What is different about the four Economics programs at peer universities that have low completion rates? Are there more programs outside the small group of peer universities we should be considering as competitors for the top students?

The Ph.D. Completions Project (Council of Graduate Schools 2009) examined factors that affect completion rates and found little to no difference in ability between people who finish the Ph.D. and those who don't complete the degree. Individual ability was measured by GRE scores and undergraduate grade point average. Factors that were found to influence completion rates included mentoring, climate, financial support, and processes/policies. With the exception of financial support for the first year of study, the factors that affect completion rates were not measured by the NRC study.

If your doctoral program was one of the four Economics programs at peer universities that had completion rates below 40%, what questions should you be asking about the peer programs with higher completion rates? What policies and processes do you have in common with the three other programs at peer universities that had completion rates below 40%?

Fig. 5 – Admissions,
at National and Peer Programs in Economics

ALL UNIVERSITIES					
	GRE SCORES				
COMPLETION RATE	< 650	650 TO 699	700 TO 749	>= 750	Total
EQUAL OR GREATER 40%	0	4	8	41	53
LESS THAN 40%	3	6	15	40	64
Total	3	10	23	81	117
percentage of programs with high completion rates					
GREATER THAN 40%	0%	40%	35%	51%	45%

PEER UNIVERSITIES					
	GRE SCORES				
COMPLETION RATE	< 650	650 TO 699	700 TO 749	>= 750	Total
EQUAL OR GREATER 40%	0	0	0	16	16
LESS THAN 40%	0	0	0	4	4
Total	0	0	0	20	20
percentage of programs with high completion rates					
GREATER THAN 40%				80%	80%

shaded cells show count of top performing programs
with high GRE scores and high completion rates

Performance Domain – Student Persistence

Completion rates could also be coupled with time to degree as part of a Student Persistence domain. The desired objective is a high completion rate on students who start the program and a limit on the number of students who take extra years to complete their doctorate. Programs can be categorized as having low or high completion rates and time to degree. Figure 6 compares national and peer data on completion rates and time to degree for doctoral programs in Economics. Roughly a third of all programs and more than half of peer programs have high completion rates (>40%) and low time to degree (students finish in less than 6 years).

Economic programs at peer universities are performing better than the national rate (55% compared to 37%) on this joint measure. No peer program in Economics fell into the worst category that combines low completion rates and long time to degree. It is worth noting that more than half of the doctoral programs in Economics in the U.S. have completion rates under 40%.

The Graduate Education Initiative (GEI) funded by the Andrew Mellon Foundation studied ways to reduce attrition and time to degree in doctoral programs in the Humanities and Social Sciences. Seven peer universities participated in the GEI study (Cornell, Columbia, Chicago, Harvard, Princeton, Stanford, and Yale). This research suggests that student persistence in doctoral programs is not entirely about funding. Financial support and multiyear commitments matter, but good advising, clear program requirements, and the scope of the dissertation are also important (Ehrenberg et al 2009, Educating Scholars).

There are two interesting analytical questions for the peer programs in Economics. There is a small group of programs (n=5) where completion rates are high, but students take longer to finish. Why

does it take longer for students in these programs to finish and what could be done to reduce the number of students who take longer than 6 years to finish? There is a second group of programs (n=4) where students finish in less than 6 years, but the percentage of students who finish is low. Why is the attrition rate higher for this group of programs and when do students leave? Is this problem related to selectivity in admissions? How is it related to graduate funding? Do these programs admit more students at lower levels of funding? The pattern of completion rates and time to degree exhibited by these four programs is uncommon among peer universities, but more typical among all Economics programs in the nation.

This approach to analysis helps to frame the key questions. We could set the break points for the categories at different levels, depending on local context and objectives. For example, if students are fully funded for four years, we might want to look at the percentage of students who finish within four years. We can determine whether student persistence for our doctoral program is typical for the discipline or among programs at peer universities. We can begin to see the truly distinctive capabilities of our doctoral programs and understand how they are differentiated from competitors. More importantly, this analysis can foster a conversation about what we are trying to achieve with a given program. If we were setting a deliberate course of action, then which pattern of student persistence is best?

Fig. 6 – Student Persistence,
at National and Peer Programs in Economics

ALL UNIVERSITIES			
	TIME TO DEGREE		
COMPLETION RATE	< 6 YRS	>= 6 YRS	Total
	Count	Count	Count
GREATER THAN 40%	43	10	53
LESS THAN 40%	42	22	64
Total	85	32	117
HIGH COMPLETION RATE, LOW TTD		43	37%
LOW COMPLETION, HIGH TTD		22	19%

PEER UNIVERSITIES			
	TIME TO DEGREE		
COMPLETION RATE	< 6 YRS	>= 6 YRS	Total
	Count	Count	Count
GREATER THAN 40%	11	5	16
LESS THAN 40%	4	0	4
Total	15	5	20
HIGH COMPLETION RATE, LOW TTD		11	55%
LOW COMPLETION, HIGH TTD		0	0%

shaded cells show count of top performing programs
with high completion rates and low time to degree

In English and Physics, the peer universities are not more efficient at producing doctoral degree recipients. Fewer peer programs have high completion rates and low time to degree. You are just as likely to take more than 6 years to complete a Ph.D. in English or in Physics at one of these select private universities. Among all universities, Physics programs are evenly split between programs where the time to degree is under or over six years. The majority of all English programs take more than 6 years to finish a degree. In this situation is there any competitive advantage in reducing time to degree? What is gained (e.g. graduate stipends) and what is lost (e.g. number of teaching assistants) if time to degree is reduced? What are the lost opportunity costs if many programs are operating with high attrition and longer time to degree? We tend to make and view programmatic decisions in higher education in isolation. Department chairs may not know how their doctoral program compares to peers in the discipline or how it compares to other programs in the same division. The results of the NRC Assessment can provide important lessons by showing alternative strategies. There is more value in asking these questions than in assuming that all programs can be judged by a single standard.

In some cases universities operate more than one program in a field. There are a few disciplines where this is more common, including Math, Statistics, Physics, and Psychology. The second program may focus on a specialized area within the field (e.g. theoretical and applied math) or be part of a different organizational unit (e.g. connected to a Medical School). Location may dictate the need for a second program at schools where the campus is spread across an urban region. The results of the NRC study can be used to understand the advantages and disadvantages of these dual programs. Is one program ranked higher than the other? Do the two programs have similar rates of student persistence? Could the two programs share resources or combine faculty?

Performance Domain – Diversity

Diversity could also be analyzed as a performance domain. The NRC Assessment of Research-Doctorate Programs provides benchmark data on the percentage of female and the percentage of under-represented minority students and faculty. Figures 7 and 8 provide national and peer data on the student and faculty diversity of doctoral programs in Economics. We should consider whether diversity is a prime driver for these programs. Are initiatives underway to increase student or faculty diversity? Is either student or faculty diversity a defining characteristic of the program? Longitudinal data are critical for looking at slow moving trends like increasing the number of under-represented minority faculty and students. A single snapshot like the NRC study reminds us that there is much room for improvement.

Figure 7 provides the percentage of under-represented minority (URM) students and faculty in doctoral programs in Economics. There is less diversity in programs at peer universities. Doctoral students are more diverse than doctoral faculty. There is roughly a 20% gap between student diversity and faculty diversity. The percentage of under-represented minorities is at least 5% for both students and faculty in only 1 in 4 programs in Economics at peer universities. Analysis should focus on student retention and pipeline issues. Student diversity has improved, but it has not translated into noticeable improvement in faculty diversity. Are minority students completing the degree and seeking academic jobs?

Finally, the percentage of female students and faculty is shown in Figure 8. In this case, the key indicator is the number of programs with at least 25% female students or faculty. Again many of the same patterns are evident. Peer universities are less diverse. No Economics faculty at peer universities is more than 25% female. There are 22 doctoral programs in Economics where the

percentage of female faculty is at least 25%, but not one of those programs is at a peer institution. If you were trying to increase the percentage of female faculty in Economics, there are several obvious questions for analysis. There are a significant number of programs with female students, but not female faculty. What happens to female doctoral students? Do they complete the program and do they seek academic jobs? What is different about the recruitment and role of female faculty at peer universities? Do fewer female faculty hold leadership positions, like department chair, at peer universities?

Fig. 7 – Under-Represented Minority Diversity,
at National and Peer Programs in Economics

ALL UNIVERSITIES				
	FACULTY URM			
STUDENT URM	< 5%	5 TO 10%	>10%	Total
> 10%	18	13	12	43
5 TO 10%	15	12	4	31
< 5%	33	11	0	44
Total	66	36	16	118
number and percentage of programs with				
STUDENT URM >= 5%			74	63%
FACULTY URM >= 5%			52	44%
FACULTY AND STUDENT URM >= 5%			41	35%

PEER UNIVERSITIES				
	FACULTY URM			
STUDENT URM	< 5%	5 TO 10%	>10%	Total
> 10%	2	2	1	5
5 TO 10%	4	2	0	6
< 5%	7	2	0	9
Total	13	6	1	20
number and percentage of programs with				
STUDENT URM >= 5%			11	55%
FACULTY URM >= 5%			7	35%
FACULTY AND STUDENT URM >= 5%			5	25%

Fig. 8 – Gender Diversity,
at National and Peer Programs in Economics

ALL UNIVERSITIES				
	FACULTY FEMALE			
STUDENT FEMALE	< 25%	25 TO 50%	>50%	Total
> 50%	5	5	0	10
25 TO 50%	77	15	1	93
< 25%	14	1	0	15
Total	96	21	1	118
number and percentage of programs with				
STUDENT FEMALE >= 25%			103	87%
FACULTY FEMALE >= 25%			22	19%
FACULTY AND STUDENT FEMALE >= 25%			21	18%

PEER UNIVERSITIES				
	FACULTY FEMALE			
STUDENT FEMALE	< 25%	25 TO 50%	>50%	Total
> 50%	0	0	0	0
25 TO 50%	14	0	0	14
< 25%	6	0	0	6
Total	20	0	0	20
number and percentage of programs with				
STUDENT FEMALE >= 25%			14	70%
FACULTY FEMALE >= 25%			0	0%
FACULTY AND STUDENT FEMALE >= 25%			0	0%

shaded cells show count of top performing programs
with high student and faculty diversity

Moving Beyond Analytics – Making Better Decisions

Davenport's method for making better decisions begins by defining the key drivers and distinctive capabilities of each program within specific performance domains. We've used the results of the NRC Assessment of Research-Doctorate Programs to target specific areas for analysis in three domains – Admissions, Student Persistence, and Diversity. The data provided in Figures 5 – 8 frame the questions that need to be addressed and provide direction for subsequent analysis. They also identify areas where administrators need to make decisions. The second step of the process in making better decisions is to consider how we make decisions about graduate education. Who identifies and prioritizes the decisions that need to be made? When and how often do these decisions need to be made? What information is available to support the decision and does it reach the person(s) making the decision? Is expert help available from Institutional Research to support this process -- gathering institutional and peer data for analysis, building and refining analytical models, helping managers interpret information, and pointing out when additional analysis would improve the decision? Are resources and time allocated so these activities can occur? Was everyone involved in the decision working from the same set of assumptions and conclusions? In the end, did we improve the process for making decisions? Once we have begun to answer these questions, we have moved beyond data and beyond analytics to institutionalize a process for making decisions.

Efforts to improve student persistence and campus diversity illustrate why we need to pay attention to how we make decisions. Who sets and monitors goals for completion rates and time to degree? Who establishes goals for minority or female faculty? What secondary decisions need to be made (e.g. how to identify a diverse pool of applicants) to be successful in hiring and retaining minority or female faculty? There are typically many more people involved in hiring faculty than participate in

campus discussions about diversity. How does the right information get to the right people at the right time to support efforts to increase student persistence or diversity? Open channels of communication aid this process, but it ultimately depends on a clear, shared understanding of the key decisions that need to be made.

Limits of Available Data

It has been fifteen years since the release of the last NRC study of doctoral education in 1995. A vastly expanded data collection in the 2010 study makes it possible to consider the type of targeted analytics proposed by Davenport. We now have a massive amount of information on doctoral education, but it comes with some clear limitations. Anyone using the results should be aware of three sets of issues – the taxonomy used to classify academic programs, missing information, and factors contributing to noise in the data.

Analytics ultimately depend on being able to describe and compare like entities. However, it is nearly impossible to avoid lumping and splitting issues in any taxonomy. The NRC classification of academic programs used in the 2010 study produces different levels of granularity across the ranking fields. A good example is Public Health, where generalized programs in Public Health and specific programs in subfields are ranked in the same category. Institutions made choices about where to rank doctoral programs, in some cases choosing between competing options (e.g. Biostatistics could be ranked in Public Health or Statistics and Probability). Known variants in a discipline are disadvantaged in some cases (e.g. social-cultural programs in Anthropology and theoretical programs in Political Science, because books were not counted for faculty in the Social Sciences). As a result of both the structure of the taxonomy of fields and the limited range of variation on some data elements, programs

are clustered within ranks. Often a small and insignificant difference in the program value on a particular element contributes to a significant increase in rank.

Some doctoral programs and some key pieces of information are missing from the study. Doctoral programs with fewer than 5 graduates in the 5 years prior to 2005-06 were excluded. Some areas of study were not ranked (e.g. Language, Societies, and Cultures; Engineering Science and Materials; Computer Engineering). Doctoral programs have been created since the NRC study began. Certain types of publications, including articles in edited volumes and foreign publications are not included. Some fields of study (e.g. Anatomy, Classics) were disadvantaged by this approach for counting publications. Independent reviews suggest that publication counts in the NRC study under-report scholarly activity.

The NRC study does not include information on the amount of research funding by discipline. This key piece of scholarly productivity is missing. The NRC faculty survey records only whether each faculty member had a grant, of any size and source of funding, and is limited by the response rate on the survey. Johns Hopkins University has led the nation in federal research funding each year since 1979. Research universities are dedicated to advancing knowledge and human conditions through scholarly inquiry and discovery. A picture of Johns Hopkins and other leading research universities, excluding information on competitive funding for research is at best incomplete. Asking what research doctorate programs look like without research funding is like asking what does the Pacific Ocean look like without water. You lose all sense of the scope, energy, and sustainability of these programs.

Realistically the best effort to collect information on a national scale cannot eliminate all sources of noise in the data. However, steps could be taken to reduce these problems. Data standards need to be clearly defined before collection begins. At times, universities followed different definitions and the data are not comparable. To be able to rank some programs, the NRC substituted the average value for

all programs in the ranking field for some missing elements. The 2010 study included more per capita measures, but there are still areas in which small programs are disadvantaged. Depending on who responded to the faculty survey, some reported funding rates (the percentage of faculty with grants) are lower than the actual rates for the program.

Next Steps

In the coming months we will work together to refine the data standards proposed by the NRC Assessment of Research-Doctorate Programs. We hope to reach consensus on definitions and changes to the collection process that would improve the comparability of information across programs. Gaps in information on research funding and publications can be addressed with new sources of information. Further discussion is needed on issues like whether missing data should be replaced or perturbed. We need to establish a cycle for exchanging information on doctoral education. We must dramatically reduce the effort required to produce this information. With these steps, we hope to realize the full potential of the NRC assessment of doctoral programs with targeted analytics.

Selected References

Council of Graduate Schools
2009 Ph.D. Completion and Attrition: Findings from Exit Surveys of Ph.D. Completers.

Davenport, Thomas H.
2009 Make Better Decisions.
Harvard Business Review, November :2-7.

Davenport, Thomas H., Jeanne G. Harris, and Robert Morison
2010 Analytics At Work: Smarter Decisions, Better Results.
Boston: Harvard Business Press.

Ehrenberg, Ronald G., Harriet Zuckerman, Jeffrey A. Groen, and Sharon M. Brucker
2009 Educating Scholars: Doctoral Education in the Humanities. Princeton, New Jersey: Princeton University Press.

U.S. Census Bureau
2009 Current Population Survey, Annual Social and Economic Supplement.
Internet release: April 2010.