

Are High-Quality PhD Programs at Universities Associated with More Undergraduate Students Pursuing PhD Study?

Todd R. Jones and Ronald G. Ehrenberg*

April 19, 2018

Abstract

This paper investigates which attributes of a Carnegie PhD-level institution predict the share of its undergraduate BA recipients that will earn a PhD. Four broad PhD fields are studied: humanities, physical sciences, natural sciences, and social sciences. We use restricted-access, individual-level *Survey of Earned Doctorates* data to determine both where and when PhD recipients received their BA, and truncation-correction methodology to account for PhD receipt after the data end. Across fields, we find that PhD production is positively related to the number of highly-ranked PhD programs an institution has, suggesting such departments may play a dual role in both producing PhDs as well as encouraging undergraduates to earn PhDs themselves. We also find that PhD production is negatively related to the total number of students and the share of total BAs that are received in the field, and is positively related to student test scores.

*Todd R. Jones is a Ph.D. Candidate in Economics at Cornell University, 330 Ives East, Ithaca, NY 14853. Email: trj25@cornell.edu. Ronald G. Ehrenberg is the Irving M. Ives Professor of Industrial and Labor Relations and Economics and is the Director of the Cornell Higher Education Research Institute (CHERI), 271 Ives Hall, Ithaca, NY 14853. Email: rge2@cornell.edu. CHERI received financial support from the Andrew W. Mellon Foundation during part of the time this paper was prepared, however the conclusions reported in this paper are strictly our own. The use of NSF data does not imply NSF endorsement of the research, research methods, or conclusions contained in this paper. The authors would like to thank two anonymous referees for helpful comments.

1 Introduction

This study examines the characteristics of an institution that correlate with the share of its BAs that ultimately receive a PhD, or its PhD production rate. We examine four fields separately, the humanities, life sciences, physical sciences, and social sciences. Despite long reported poor job market prospects for new humanities PhDs, graduate programs in the humanities have continued, sometimes with smaller cohort sizes, to churn out PhD students in the humanities. On the other hand, STEM fields, the life sciences and physical sciences, strongly contribute to the economy, foster innovation, and lead PhDs in these fields to receive high salaries. We study a number of factors of institutions, such as expenditure measures, the number of total students, and other student body characteristics that might be associated with predicting the PhD production rate. In addition to producing PhDs, departments also educate undergraduate students, and one focus of this paper is to see whether the scope and quality of an institution's PhD programs are related the likelihood that its undergraduate students go on to receive PhDs. Put another way, we are interested in the dual role that departments at universities play in producing new PhDs and in generating undergraduate student interest in going on for doctoral degrees, typically at other institutions.

We focus our attention on doctoral institutions.¹ Using data from the National Science Foundation's *Survey of Earned Doctorates* (SED) linked to the IPEDS *Completions Survey*, we calculate that for 1980-2001 BA recipients, the fraction of BAs in a given field that receive PhDs in the same field is higher at *baccalaureate institutions*² than at *doctoral institutions*.³ However, because of the differences in the sizes of these two types of

¹We define define doctoral institutions as 2015 Carnegie Categories 15: Doctoral Universities: Highest Research Activity; 16: Doctoral Universities: Higher Research Activity; and 17: Doctoral Universities: Moderate Research Activity.

²We define baccalaureate as 2015 Carnegie Categories 21: Baccalaureate Colleges: Arts and Science Focus; 22: Baccalaureate Colleges: Diverse Fields; and 23: Baccalaureate/Associate's Colleges: Mixed Baccalaureate/Associate's. We do not consider category 14: Baccalaureate/Associate's Colleges: Associate's Dominant.

³This calculation is for BA-years 1980-2001 and consider only those with a time-to-degree of 15 years or less, which contrasts from the approach later in the paper, where the maximum time-to-degree varies by field. The total number of PhDs in a given field is divided by the total number of BA recipients in the same field. Some of the PhD recipients received BAs in other fields. We include only observations for which there

institutions, doctoral institutions produce far more total PhDs than baccalaureate colleges. Doctoral institutions produce 5.8 times as many PhDs than do baccalaureate colleges for the physical sciences; they also produce 2.7 times as many for the humanities, 4.7 times as many for the life sciences, and 3.7 times as many for the social sciences.⁴ With the exception of life sciences, the ratios are even higher for Master’s colleges.⁵ Therefore, knowledge of the characteristics of the Carnegie doctoral institutions that lead to more undergraduates going on to receive PhDs is of importance in understanding the determinants of PhD production.

Lemke (2009) is among the latest studies to examine the determinants of undergraduate institutions in “generating” (producing) PhD recipients.⁶ This study computed a PhD production rate by dividing publicly-available data on the number of PhDs in all fields earned by alumni of selective liberal arts colleges by the number of BAs earned by the students at these institutions. Because of this study’s use of public-use data, while it was known *where* a PhD recipient received her BA, it was not known *when* he or she received it. Lemke (2009) assumed a 6-year time-to-degree (TTD) window from BA to PhD, which often is too short of a window. We use restricted-access, individual-level SED data to observe the actual matches of where and when each individual PhD graduated from his or her BA institution. For each field, we compute one observation per institution-BA year on the percentage of BAs who received PhDs. This is in contrast to the one observation per institution in the Lemke (2009) study, where both the numerator and denominator of the dependent variable are averages over different 10-year periods.

When using individual-level data to compute the PhD production rate, there is an

is a match with the Completions Survey data. The institution must have produced at least 5 PhDs over the SED data, without regard to TTD. Other sample restrictions are as in Section A.

⁴This calculation makes use of sample restrictions as in Section A, including that the institution must have produced at least 5 PhDs, but here consider all TTDs, not just up to 15 years. The BA years are 1980-2005.

⁵We define Master’s institutions as 2015 Carnegie Categories 18: Master’s Colleges and Universities: Larger Programs; 19: Master’s Colleges and Universities: Medium Programs; and 20: Master’s Colleges and Universities: Small Programs.

⁶Many other studies have addressed the determinants of PhD students’ times to degree and completion rates. See, for example, Knapp and Goodrich (1951), Tidball (1986), Schapiro et al. (1991), Ehrenberg et al. (2007), Groen et al. (2008), Ehrenberg et al. (2009), Bowen and Bok (2016), and Groen (2016).

inherent truncation problem. Not all PhDs from a given BA cohort are observed because some individuals complete their PhD after the conclusion of the data. The truncation problem increases in severity for later cohorts. We use truncation correction methodology, where early cohorts' patterns of time-to-degree are used to predict the number of PhDs that later cohorts will eventually produce. After correcting for truncation, we examine which features of the institution predict the share of graduates that ultimately receive a PhD.

In our main specification using the full sample, we find the number of an institution's departments that are in the top 10th percentile of the National Research Council (NRC) 1995 doctoral rankings in each field is one of the three variables that is correlated with the PhD production rate across each field. Put simply, the number of high-quality doctoral programs is an important predictor of the rate at which undergraduates go on for PhD study. The other two variables that are statistically significant across all fields are the incoming student 75th percentile test scores, which is positively related with the PhD production rate, and the total number of students, which is negatively related. Other variables, such as institutional expenditures per student and the percentage of total BAs in the field, are significant for some, but not all, fields. We also restrict the sample to Carnegie category 15 institutions (Highest Research Activity), which have the highest level of research activity. Finally, we conduct an analysis along the lines of that done in Lemke (2009). For both, we find generally similar results to the baseline specification, although the NRC top 10th percentile variable is significant in only two of the four fields, life sciences and social sciences.

2 PhD Production

Following Lemke (2009), we assume that characteristics of the undergraduate institution generate, or produce, PhDs. We first list the explanatory variables considered in the analysis. We break down these characteristics into four categories: institutional, expenditure, student body, and graduate programs. Appendix A provides details on the data.

Institutional Characteristics

First, we include the student-faculty ratio. The smaller the class size, the greater the potential for students to interact with the professor, which may lead to the students being more interested in pursuing graduate studies. Next, we include a public institution indicator. If the quality or amount of interaction between faculty and students differs between public and private institutions, this may lead to differing propensities to pursue a PhD. The mission of the institution may be generally different between the two types and may also influence the propensity.

Expenditures

We include instructional expenditures per student. Greater instructional expenditures—of which faculty salaries are a large part—may attract faculty who are more research productive. Students may then want to emulate the career paths of these faculty by pursuing PhDs themselves. We also include research expenditures per full-time faculty. A higher level of research spending likely leads to higher quality research. This may influence students, making a career in research via a PhD more attractive.

Student Body Characteristics

As a proxy for the institution's academic ability, we consider first-year incoming student test scores, specifically a combination of 75th percentile SAT and ACT scores, weighted by the percentage of scores from each. We emphasize that the scores are for first-year, not seniors. We additionally include several measures of the size and composition of the student body: the total number of students (including graduate and professional students), the percentage of the students that are undergraduates, the percentage of undergraduates that are female, and the percentage of undergraduates that come from underrepresented minority groups. Because the importance on campus of, say, the humanities could play a role in generating interest in humanities PhD study, in the humanities regressions, we include the percentage of BAs that were received in the humanities in our humanities equations. We

similarly include the percent of degrees received in the other three fields in their relevant equations.

Graduate Programs

Finally, we are interested if having highly-ranked doctoral programs is associated with having a higher PhD production rate. More highly-ranked doctoral programs is presumably correlated with having more faculty at the cutting edge of their field, and, to the extent that these faculty interact with undergraduate students, this may influence students' decisions to obtain a PhD in the humanities. Thus, we include a set of variables indicating the number of humanities' doctoral programs an institution has in several percentile bins: 0-10, 11-25, 25-50, and 51-100.

3 Estimating Equation

In order to obtain estimates, we employ weighted least squares regression with the PhD production rate, defined in Section 4.3, as the dependent variable. We run the regression separately by PhD field. Specifically:

$$PhdProdRate_{i,f,y} = \alpha_0 + \alpha_1 IndepVars + Carnegie_i + Year_y + \epsilon_{i,f,y},$$

where, for institution i , field f in BA-year y , $PhdProdRate_{i,f,y}$ is the PhD production rate defined in Section 4.3, $Year_y$ are year fixed effects, and $\epsilon_{i,f,y}$ is the error term. $Carnegie_i$ is the Carnegie category of the institution, with category 15 (Highest Research Activity) being the omitted category. We include these because institutions vary substantially across Carnegie categories (see Section 5). $IndepVars_{i,f,y}$ is a vector containing the explanatory variable (and missing variable indicators) as described in Section 2: *Stud – FT Fac Ratio*, *Public*, *Instr. Exp./Student*, *Rsch. Exp./FT Fclty*, *75 Percentile Score*, *Total Students*, *% UG*, *% Female*, *% Minority*, *%(Field) BAs of Total*, *#NRC 0 – 10 in (Field)*, *#NRC 11 – 25 in (Field)*, *#NRC 26 – 50 in (Field)*, *#NRC 51 – 100 in (Field)*. (*Field*)

refers to field of study (Humanities, Life Science, Physical Science, and Social Science), with the field matching the field currently in consideration.

If an observation is missing or has 0 BAs—the denominator of the PhD production rate—then it is not included in the regression. We weight the institution-field-BA year observation by the number of field-specific BAs.⁷ Robust standard errors are clustered at the institution-level.

In addition to the main regression above, we also present results where we restrict the sample to Carnegie category 15 institutions (Research Activity), which have a higher PhD production rate as well as a smaller amount of missing data as compared to the other two categories.

Finally, we perform an analysis in the spirit of Lemke (2009), although there are some differences. As opposed to constructing the PhD production rate as described in Section 4, we instead divide the total number of PhDs produced over the 1994-2003 time period by the total number of BAs produced over the 1989-1998 time period. Truncation correction is explicitly not used in this analysis because TTD is assumed and a PhDs are not matched to their actual BA-year. We use the same set of institutions as appear in the main analysis, and there is now only one observation per institution. If an observation produced PhDs but had no BAs, it is excluded from the regression. For the explanatory variables, we take the average value over the same 1989-1998 period among non-missing observations. If all observations for a variable are missing, we assign that variable a value of 0 and a corresponding missing variable a value of 1. We use robust standard errors, and weight by total BAs over the 1989-1998 period.⁸

⁷It can happen that an institution is predicted to have a very high (even greater than 1) PhD share. In such cases, it is likely that the institution awards a very small number of BAs in this field, but produced a disproportionately high number of PhDs. This can be because a relatively large share of the BAs in the field went on to earn a PhD or BAs in other fields went on to earn a PhD. Because there are very few BAs, these observations will receive a small weight, and not be very influential in the estimates.

⁸Results differ somewhat if we do not weight by BAs. An advantage of using weighting is that there are some observations that have very few BAs in a given field, but produce many PhDs, including those with BAs from a different field. This causes the PhD production rate to be very large, even over 100. Weighting greatly reduces the influence that such observations have in the regression.

4 Constructing the PhD Production Rate and Correcting for Truncation

We define the PhD production rate as the fraction of BAs for a given institution-BA year that obtain a PhD. We return to the details of its construction in Section 4.3 after first introducing the concept of maximum TTD in Section 4.1 and, in Section 4.2, how we employ truncation correction to address not being able to observe PhDs granted after the SED sample ends.

4.1 Maximum TTD

We measure the PhD production rate for students who earn a PhD within a given time, which we denote as the the maximum TTD. Earning a PhD often takes many years. To illustrate, Figure 1 plots the distribution of TTD by field for BA years 1980 and 1981, the two earliest years in the sample. Because the final year of the SED sample is 2016, the longest observable TTD for BA-year 1980 is 36 years and for BA-year 1981 is 35 years, and all PhDs earned after this are truncated. This figure shows that nearly all PhDs are earned by 30 years for the life and physical sciences, with a larger tail for the humanities and social sciences. The peaks of these latter two fields are also longer, suggesting a longer average TTD.

This truncation problem is the most pronounced for BA-year 2005, the final BA-year in the sample. For this year, the longest observable TTDs is only 11 years. We use the truncation-correction technique presented below in Section 4.2 to estimate how many PhDs will be produced by the maximum TTD.

The higher the maximum TTD, the higher the number of PhDs are observed. But there is a tradeoff: The higher the maximum TTD, the fewer the number of BA-years there are for which we observe all PhDs for the maximum TTD, and the more truncation correction is necessary. In this paper, we use field-specific maximum TTDs based on Table 1. This

table displays separately by field the cumulative percentage of PhDs that were received by each TTD for BA-years 1980 and 1981. For each field, we select the smallest TTD for which at least 90% of PhDs had been earned.⁹ Thus maximum TTDs of 23, 20, 16, and 22 years are used for humanities, life sciences, physical sciences, and social sciences, respectively. Choosing the maximum TTD this way puts all of the fields on an equal playing field, taking into account the different TTD distributions. This means, however, that the extent to which we must use truncation correction and thus artificially generate the outcome variable differs by field, being the most in the humanities and the least in the physical sciences. For the humanities, we observe the maximum TTD for BA-years 1980-1993 (because we observe PhDs until 2016), and use truncation correction for remaining BA-years, 1994-2005.

4.2 Truncation Correction

We employ the simple truncation correction technique used in Prenovitz et al. (2016) to estimate how many PhDs would be produced if we had observed PhDs up to the maximum TTD. Using the number of PhDs produced by 2016 as a baseline, we predict how many more would be produced had we observed the entire maximum TTD. The training sample that goes into the prediction uses the BA-years for which we observe the maximum TTD; as an example, for the humanities this is 1980-1993. Using these years only, we compute the percentage of PhDs were received by each TTD, up to to maximum TTD. These cumulative percentages are then applied to the remaining years (1994-2005 for humanities) at the institution-BA year level to estimate the number of PhDs.

An example will illustrate. Suppose that in the humanities, a given institution-field with BA-year 2003 produced 6 PhDs by 2016. Because the maximum observed TTD is 13, the cumulative percentage for a TTD of 13 will be used. Suppose that this is calculated to be x . The predicted number of PhDs that will be produced by 2026 (23 years after BA-year) is computed to be $(\frac{1}{x})6$.

⁹In this table, TTDs of 36 years are combined with TTDs of 35 years; only the 1980 cohort has an observed TTD of 35 years.

Here, we formalize the above idea. The prediction is based on the years for which the full TTDs exist (1980-1993 for the humanities). Restricting the sample to these years only, we compute the cumulative percentage of degrees that occur after TTDs of 4,...,MaxTTD, where MaxTTD is the maximum TTD (4,...,23 years for the humanities).

Specifically, we first compute the cumulative number of degrees by each TTD for each institution, separately by field. Define the number of degrees granted by institution i , field f , BA-year y , at TTD t , to be $PhD_{i,f,y}^t$, where $t \in \{4, MaxTTD\}$.¹⁰ The cumulative number of degrees for institution i by TTD t is then:

$$CumPhD_{i,f,y}^t = \sum_{t=4}^t PhD_{i,f,y}^t.$$

These are aggregated across institutions to the TTD-level to produce $CumPhD_f^t$:

$$CumPhD^t = \sum_i CumPhD_{i,f,y}^t.$$

Sum across TTDs to produce the total number of PhDs at the maximum TTD:

$$TotalPhD_f = \sum_{t=4}^{MaxTTD} CumPhD_f^t.$$

Now, obtain the cumulative percentage of degrees granted for each TTD:¹¹

$$CumPct_f^t = \frac{CumPhD_f^t}{TotalPhD_f}.$$

We next apply the cumulative percentages to each institution-BA year for years for which we do not observe the max TTD (1994-2005 for humanities) to obtain the estimated total number of degrees for each institution-BA year. This is done by multiplying the number of degrees produced by the maximum observed TTD by the inverse of the corresponding

¹⁰We reassign TTDs of less than 4 years to be 4 years.

¹¹The cumulative percentage is actually a fraction, not a percentage.

cumulative percentage:

$$PredictedPhD_{i,f,y} = \left(\frac{1}{CumPct_f^{MaxObsTTD}} \right) CumPhD_{i,f,y}^{MaxObsTTD},$$

where $MaxObsTTD$ is the maximum TTD that it was possible for the BA-year to have; this is equivalent to $2016 - y$. For example, $MaxObsTTD$ for BA-year 2003 is 13. Note that the $CumPct$ value takes on the value corresponding to $MaxObsTTD$. If an observation produced 0 PhDs by 2016, we assign it a predicted value of 0. This method uses information from past years to predict PhD rates for future years. It assumes that the production patterns of PhDs does not vary by year, or that the pattern from the earlier estimating sample will occur in the later prediction sample. For each field, we use the entire sample of institutions in the training sample, regardless of which Carnegie category the institution is in.

4.3 PhD Production Rate

After implementing the truncation correction, we compute the PhD production rate for each institution BA year, and do so separately by field. For BA-years for which we observe the max TTD (1980-1993 for humanities), this is:

$$PhdProdRate_{i,f,y} = 100 \cdot \left(\frac{CumPhD_{i,f,y}^{MaxTTD}}{TotalBA_{i,f,y}} \right),$$

and for BA-years for which we do not observe the max TTD (1994-2005 for humanities):

$$PhdProdRate_{i,f,y} = 100 \cdot \left(\frac{PredictedPhD_{i,f,y}^{MaxTTD}}{TotalBA_{i,f,y}} \right),$$

where $CumPhD_{i,f,y}^{MaxTTD}$ and $PredictedPhD_{i,f,y}^{MaxTTD}$ are defined the same as previously, and $TotalBA_{i,f,y}$ is the total number of BA degrees for institution i in field f in year y .¹² Because we multiply by 100, the PhD production rate is a percentage.

¹²Observations with 0 BAs produced in the field in the BA-year are not included in the regressions.

Note that the numerator includes all individuals who earned a PhD in the field, whether or not they also earned a BA in the same field. To illustrate, Table 2 shows the BA fields of study for PhDs in a given field. Over 93% of Physical Science PhDs also received a BA in Physical Sciences, while only 64% of Social Science PhDs received a BA in the Social Sciences. The denominator of the PhD production rate includes only the major that is listed as the first major. If the individual dual-majored in a field other than their first-listed major, then this will not be included in the other field.

5 Descriptive Statistics

We first describe several plots relating to the PhD production rate, and then proceed to present summary statistics for the independent variables considered in the analysis.

PhD Production Rate

Figure 2 plots the total number of PhDs produced for each BA-year from 1980 to 2005. This is broken up by PhD field. There were fewer PhDs granted in the humanities and social sciences than in the life and physical sciences. These latter two fields both experienced a large amount of growth beginning in the mid-1990s; the former fields also experienced growth, but to a lesser degree.

Panel B of Figure 2 also plots the total number of PhDs produced over time, but breaks it up by the Carnegie category of the university. Relatively few PhDs came from institutions in category 16, and even fewer from category 17. Both categories remained fairly steady over the time horizon, with a small increase towards the end. Institutions in category 15 (Highest Research Activity), meanwhile, produced many more PhDs, and steadily increased this number beginning in the late-1990s.

The plot in Figure 3 Panel A displays the PhD production rate, in contrast to the total number of PhDs. Here, at the aggregate level, this is calculated by taking the overall number of PhDs produced divided by the overall number of BAs.¹³ The PhD production rates in

¹³This is as opposed to taking the average of individually-computed PhD production rates.

the life and physical sciences largely track each other and are very similar in magnitude for much of the time period. These rates are also generally at least two times as large as the rates for the humanities and social sciences, which do not experience the same growth at the end of the sample.

Finally, Figure 3 Panel B shows that the trends in the PhD production rate among the various Carnegie categories are similar, but the levels differ substantially. The rate is highest for category 15 (Highest Research Activity), while lower rates for category 16 (Higher Research Activity), and again for category 17 (Moderate Research Activity).

Independent Variables

Table 3 displays summary statistics for the independent variables, specifically means, standard deviations, and the percent of observations that are still missing after filling in and interpolating as described in Section A. Means and standard deviations are conditional on the observation not being missing. The unit of observation is an institution-BA-year; thus, each institution contributes 26 observations.¹⁴ Observations are not weighted. Recalling that the set of institutions included in the samples of the different field varies somewhat due to the requirement that institutions produced at least five PhDs over the time frame, the sample used in this table is the life sciences sample, which contains the greatest number of institutions. For brevity, we do not present the summary statistics for the three other fields.

The 1st–3rd columns display statistics for the full sample, while the 4th–6th, 7th–9th, and 10th–12th columns display statistics for institutions with Carnegie category 15 (Highest Research Activity), 16 (Higher Research Activity), and 17 (Moderate Research Activity), respectively. After filling in and interpolating, most variables have a very small percentage of observations missing, exceptions being the 75th percentile score, and research expenditures per full-time faculty, both of which have a larger percentage missing across the Carnegie category gradient. Within field, institutions that do not have any department ranked in the NRC data are considered missing. There is a greater percentage of institutions without

¹⁴Observations that do not contribute to regression results because e.g., they did not have any humanities BAs, are still included in the summary statistics.

ranked NRC departments in the humanities and social sciences than in the life and physical sciences. Across all four fields, the percentage of category 15 (Highest Research Activity) institutions that do not have any departments ranked is small, while it is very large for category 17 institutions (Moderate Research Activity).

As a whole, the mean values across Carnegie category show a clear pattern. Compared to category 16 (Higher Research Activity) and even more so to category 17 (Highest Research Activity) institutions, category 15 (Highest Research Activity) institutions have smaller student-faculty ratios, are more or equally often public institutions, and have higher institutional expenditures per student and research expenditures per full time faculty. They also have a smaller percentage of undergraduate students—meaning a higher percentage of graduate and professional students—and a smaller percentage of female and minority students. Carnegie category 15 (Highest Research Activity) institutions also have a higher average share of undergraduate majors in the humanities, life sciences, and physical sciences. They have a much smaller share in other majors, which include subjects such as business. Finally, they have many more departments ranked in the 1995 NRC doctoral rankings.

6 Results

Baseline Regression

Table 4 presents the baseline weighted least squares results, where each column is from a separate regression and corresponds to one of the PhD fields: humanities, life sciences, physical sciences, and social sciences. The mean value of the dependent variable, the PhD production rate, is displayed directly below the regression coefficients. Recall that this variable is a percentage, not a share. In line with Figure 3 Panel A, this value is much larger for the life and physical sciences. The explanatory power for these models is fairly high—around 0.5—for all but the social sciences, where it is lower.

There are three variables that are consistently statistically significant across the fields:

the 75th percentile test score, the number of total students at the institution, the percentage of total BAs that are in the given field, and the number of doctoral programs ranked in the top 10th percentile. The direction of each is also consistent across fields. The higher the test scores, the higher the PhD production rate. The coefficient is largest for the life sciences. Students with higher college entrance exams are more likely to excel in school and have lower costs of continuing to acquire education. That the total students variable has a negative coefficient means that smaller institutions produce more PhDs, everything else equal: having 1,000 fewer students is associated with a drop in the PhD production rate of between .04 for social sciences and .11 for physical sciences. The percentage of the total BAs that are in a given field is negatively related to the PhD production rate. Interestingly, this negative relationship goes against the hypothesis that a field having a larger presence is positively related to producing PhDs. One possible explanation is that those in smaller programs are less likely to get lost in the shuffle and are more likely to gain more individual attention from mentors such as faculty members. Having more highly-ranked doctoral programs, especially in the top 10th percentile, is strongly related to producing PhDs. For the humanities, an increase of one program in the top 10th percentile is associated with a higher PhD production rate of 0.44, which is 17% of the (weighted) mean. The coefficient is even higher for the life sciences (0.65; 13%) and the physical sciences (0.54; 12%). For the social sciences, it is 0.29 (13%). Top-ranked departments tend to have stellar faculty, which may influence students in several ways, including attracting them to the institution in the first place.

Several more variables are statistically significant for at least two of the four fields: instructional expenditures per student, the percentage of undergraduates who are female, the percentage of total BAs that majored in a particular field, and the institution being in Carnegie category 17 (Moderate Research Activity) as opposed to category 15 (Highest Research Activity). Instructional expenditure per student, measured in \$1,000s, is positively related for the humanities and physical sciences, but the magnitude is rather small. It may be the case that having higher quality faculty inspires students. Research expenditures per

full time faculty is not related to the PhD production rate. For life and social sciences, the more men there are, the more likely students are to get a PhD. One factor that may be involved is the lower rate in general that women pursued PhDs, especially at the beginning of our sample timeframe. The percentage of total students who are undergraduates is never significant. The percentage of undergraduates who are a minority is positively related to the PhD production rate, but only for social sciences, and the coefficient is small.

Compared to Carnegie category 15 institutions, those in category 17 have a lower PhD production rate for both the physical and social sciences. This is after controlling for all other variables and is thus not a raw difference as in Figure 3 Panel B. Neither the student-faculty ratio nor being a public (versus private) institution relate to the PhD production rate.

In tables not shown, results are quantitatively unchanged if we use a maximum TTD based on the first year in which at least 80% or 85% of individuals receive their PhD.

Restricting to Carnegie 15 Institutions

In Table 5 everything is the same as above except that we now restrict the sample to Carnegie category 15 institutions (Highest Research Activity). The PhD production rates and explanatory of these regressions are both higher. The 75th percentile test score and total students variables are again statistically significant across all fields, with coefficients larger in magnitude compared to the baseline. The coefficients across fields for the number of doctoral programs ranked in the top 10th percentile are roughly of the same magnitude as those found in Table 4, although only humanities and physical sciences continue to be statistically significant.

With some exceptions, the same general findings of significance and sign hold for the other variables even though the standard errors have increases as the sample size has decreased.

Lemke-Style Regression

Table 6 shows the results from performing the analysis along the lines of Lemke (2009). Compared to the baseline regression, the findings are again largely the same. The test

score and total student variables exhibit very similar patterns. The doctoral 0-10 percentile rankings variables are statistically significant for the same fields as when we restrict the sample to Carnegie category 15 institutions. Compared to the baseline regression, most of the same coefficients are significant statistically and have similar magnitudes.

7 Conclusion

By using the restricted access *Survey of Earned Doctorate* individual-level data, we are able to determine for each field each PhD recipient's time-to-degree and, given the heterogeneity in time-to-degree, to compute better estimates of the share of bachelor's graduates at an institution in a year that subsequently received PhDs in the field than did the early method used by Lemke (2009). Our method required us to address the truncation problem that occurs when using the individual level SED data, that we will not observe PhDs earned after the data end.

We conduct an weighted least squares analysis to examine which factors are associated with the PhD production rate. We point out that because we are not exploiting exogenous variation, our results should not be interpreted causally, as there are likely variables that we did not control for. With that said, our major empirical finding is the strong positive association between the number of highly-ranked PhD programs in a field at an institution in a year and the number of its undergraduate students that go on for PhDs in that field. Strong doctoral programs in a field thus seem to contribute to the supply of PhDs both through the number of PhDs they directly generate and through their impact on the number of their undergraduate students that go on for PhDs. We also find that higher entering test scores for undergraduate students and a lower percentage of BAs that are received in the field are both associated with more undergraduates at the institution going on for PhDs.

References

- Bowen, William G and Derek Bok**, *The Shape of the River: Long-Term Consequences of Considering Race in College and University Admissions*, Princeton University Press, 2016.
- Ehrenberg, Ronald G, George H Jakubson, Jeffrey A Groen, Eric So, and Joseph Price**, “Inside the Black Box of Doctoral Education: What Program Characteristics Influence Doctoral Students? Attrition and Graduation Probabilities?,” *Educational Evaluation and Policy Analysis*, 2007, 29 (2), 134–150.
- , **Harriet Zuckerman, Jeffrey A Groen, and Sharon M Brucker**, *Educating Scholars: Doctoral Education in the Humanities*, Princeton University Press, 2009.
- Goldberger, Marvin L, Brendan A. Maher, and Pamela E Flattau**, *Research Doctorate Programs in the United States: Continuity and Change*, National Academies Press, 1995.
- Groen, Jeffrey A**, “The Impact of Labor Demand on Time to the Doctorate,” *Education Finance and Policy*, 2016.
- , **George H Jakubson, Ronald G Ehrenberg, Scott Condie, and Albert Y Liu**, “Program Design and Student Outcomes in Graduate Education,” *Economics of Education Review*, 2008, 27 (2), 111–124.
- IPEDS**, “2017-18 Survey Materials, Calculation of FTE students (using fall student headcounts),” <https://surveys.nces.ed.gov/ipeds/VisGlossaryPopup.aspx?idlink=854> 2018. Accessed: 3-19-18.
- Jones, Lyle V, Gardner Lindzey, and Porter E Coggeshall**, *An Assessment of Research-Doctorate Programs in the United States: Humanities*, National Academies Press, 1982.

- Knapp, Robert H and Hubert B Goodrich**, “The Origins of American Scientists,” *Science*, 1951, *113* (2941), 543–545.
- Lemke, Robert J**, “Generating Doctoral Degree Candidates at Liberal Arts Colleges,” in Ronald G Ehrenberg and C V Kuh, eds., *Doctoral Education and Faculty of the Future*, Ithaca, NY: Cornell University Press, 2009, pp. 93–108.
- Lenihan, C**, “IPEDS Analytics: Delta Cost Project Database 1987-2010 (NCES 2012-823),” *Washington, DC: National Center for Education Statistics. Retrieved from <http://nces.ed.gov/pubsearch>. America: Its History, Mission and Management*, 2012, pp. 111–122.
- National Science Foundation**, “Survey of Earned Doctorates Survey Description,” <https://www.nsf.gov/statistics/2018/nsf18304/survey-description.cfm> 2017. Accessed: 3-19-18.
- Ostriker, Jeremiah P, Charlotte V Kuh, and James A Voytuk**, *A Data-Based Assessment of Research-Doctorate Programs in the United States*, National Academies Press Washington, DC, 2011.
- Prenovitz, Sarah J, Gary R Cohen, Ronald G Ehrenberg, and George H Jakubson**, “An Evaluation of the Mellon Mays Undergraduate Fellowship’s Effect on PhD production at Non-UNCF Institutions,” *Economics of Education Review*, 2016, *53*, 284–295.
- Schapiro, Morton Owen, Michael P O’Malley, and Larry H Litten**, “Progression to Graduate School from the ?Elite? Colleges and Universities,” *Economics of Education Review*, 1991, *10* (3), 227–244.
- Tidball, M Elizabeth**, “Baccalaureate Origins of Recent Natural Science Doctorates,” *The Journal of Higher Education*, 1986, *57* (6), 606–620.

Table 1: Cumulative Percentage of PhDs Received by Time-to-Degree, by Field, BA-Years 1980 and 1981

TTD	Humanities	Life Science	Physical Science	Social Science	Total
4	0.8	1.3	3.6	1.2	2.1
5	2.0	7.6	14.5	4.4	9.1
6	5.1	20.2	30.7	10.9	20.9
7	10.8	32.9	45.2	17.8	32.6
8	17.5	44.4	56.0	27.3	42.8
9	25.4	53.0	64.6	35.5	51.2
10	34.4	60.5	70.6	44.0	58.5
11	42.7	66.9	75.8	52.9	65.0
12	50.7	71.8	80.0	59.1	70.2
13	57.6	75.6	83.4	65.0	74.6
14	63.2	78.9	86.3	69.9	78.3
15	68.6	81.5	88.5	74.0	81.4
16	72.8	83.6	90.7	76.8	83.9
17	76.7	85.7	92.2	80.1	86.2
18	79.9	87.3	93.5	82.8	88.1
19	82.7	88.9	94.3	85.5	89.7
20	85.2	90.2	94.9	86.8	90.9
21	87.2	91.4	95.5	88.6	92.0
22	89.0	92.4	96.0	90.2	93.0
23	90.6	93.3	96.4	91.3	93.9
24	91.8	94.1	96.9	92.4	94.7
25	93.1	94.9	97.3	93.5	95.4
26	94.2	95.7	97.6	94.3	96.0
27	95.2	96.4	97.9	95.2	96.6
28	95.8	96.8	98.1	95.7	97.1
29	96.5	97.3	98.5	96.3	97.6
30	97.2	97.9	98.7	97.0	98.0
31	98.1	98.4	99.0	97.6	98.5
32	98.6	98.8	99.3	98.4	98.9
33	99.1	99.3	99.5	99.0	99.3
34	99.5	99.6	99.7	99.6	99.6
35	100.0	100.0	100.0	100.0	100.0

Notes: This table shows the cumulative percentage of PhDs received by the time-to-degree, TTD. This is displayed separately by field as well as aggregated together. BA-years of 1980 and 1981 are used. TTDs of less than 4 are combined with TTD of 4. TTD of 36 is combined with TTD of 35. The entry corresponding to the first TTD for which at least 90 percent of PhDs were earned is indicated in bold.

Table 2: BA Fields of Study of PhD Recipients, by PhD Field

BA Field	PhD Field			
	Humanities	Life Science	Physical Science	Social Science
Humanities	82.4	2.2	0.9	13.2
Life Science	1.3	78.0	3.6	4.5
Physical Science	2.3	11.7	93.6	7.0
Social Science	6.6	2.7	0.7	64.0
Other	7.5	5.4	1.2	11.3
<i>N</i>	45521	89488	104591	34509

Notes: This table displays the percentage of PhD recipients in a given field (Humanities, Life Science, Physical Science, Social Science) who obtained their BA in each of these four fields as well as in subjects outside of these; this category is referred to as “Other.” PhD fields are shown in the columns; BA fields are shown in the rows. Observations with no BA field listed are excluded. Observations are not restricted by TTD, and only PhDs actually observed are shown (meaning no truncation correction was used).

Table 3: Summary Statistics, Full Sample and by Carnegie Classification

	All			Carnegie 15			Carnegie 16			Carnegie 17		
	Mean	SD	Miss	Mean	SD	Miss	Mean	SD	Miss	Mean	SD	Miss
Stud-FT Fac Ratio	18.2	14.4		15.5	20		18.9	9.6		21.2	7	
Public	.7	.5		.7	.5		.7	.4		.5	.5	
Instr Exp/Student	7.5	6.8		10.5	8.8		6.3	4.7		4.6	2.3	
Rsch Exp/FT Fac	48.4	75.6		76.1	93.2		37.3	58.6		10.1	13.6	
75 Percentile Score	26.6	3.1		28.3	3		26	2.7		24.6	2.3	
Total Students	13.1	8.9		19.7	9.4		10.7	5.3		6.7	4.5	
% UG	77.3	13.1		72.5	12.8		79.8	12.6		81.1	12.1	
% Female	50.7	9.7		48.5	6.8		49.5	10.8		55.5	9.9	
% Minority	16.4	18.2		12.8	8.7		16.7	20.8		21.2	23.1	
% Hum	12	7.7		14.4	7.4		10.8	8.2		10	6.4	
% LifSci	11.7	7.3		12.2	5.2		11.1	6.6		11.5	10.2	
% PhysSci	15.6	16.3		17.5	13.1		18.7	21.5		9	9.4	
% SocSci	10.1	8.1		14.1	9		8	6.7		7.1	5.5	
% Other	50.6	18.2		41.7	16.8		51.5	18.1		62.4	12.3	
Hum NRC 0-10	.5	1.6		.7	1.9		0	0		0	0	
Hum NRC 11-25	.7	1.2		1	1.3		0	0		0	0	
Hum NRC 26-50	1.1	1.4		1.5	1.4		.1	.3		0	0	
Hum NRC 51-100	2.2	1.7		2.4	1.8		1.8	1.5		1	0	
LifSci NRC 0-10	.6	1.5		.9	1.7		0	0		0	0	
LifSci NRC 11-25	.8	1.3		1.2	1.5		0	.1		0	0	
LifSci NRC 26-50	1.3	1.7		1.8	1.8		.3	.8		0	0	
LifSci NRC 51-100	2.3	2		1.9	2		3.1	1.7		2	1.4	
PhySci NRC 0-10	.8	2.4		1.4	3		0	0		0	0	
PhySci NRC 11-25	1	1.9		1.7	2.2		.1	.4		0	0	
PhySci NRC 26-50	1.8	2.4		2.8	2.6		.4	1.2		.1	.3	
PhySci NRC 51-100	3.4	2.7		3.5	3		3.4	2.3		1.8	1.6	
SocSci NRC 0-10	.4	1		.5	1.2		0	0		0	0	
SocSci NRC 11-25	.5	.9		.6	1		0	0		.3	.5	
SocSci NRC 26-50	.8	1.1		1	1.1		.1	.3		0	0	
SocSci NRC 51-100	1.7	1.5		1.7	1.6		1.8	1		1	0	
N	7592			2964			2600			2028		

Notes: This table displays means, standard deviations, and percent of observations that are missing. Observations are unweighted and are displayed separately by Carnegie category of the institution. The sample shown is the Life Science sample, as this was the field with the most institutions. Observations are at the institution-BA year, and each institution contributes 26 observations, one for each year between 1980 and 2005. Means and standard deviations do not include missing observations in the calculations. Total students, instructional expenditures per student, and research expenditures per student are in units of 1,000. The missing column displays the fraction of observations that are missing after interpolating and filling in as described in the text. The missing column for the NRC variables is the fraction that were not observed with at least one department for the respective field (Humanities, Life Science, Physical Science, Social Science).

Table 4: Baseline Regression

	Humanities	Life Sci.	Physical Sci.	Social Sci.
Stud-FT Fac Ratio	0.005 (0.008)	-0.026 (0.017)	0.006 (0.008)	-0.004 (0.007)
Public	-0.251 (0.403)	-0.317 (0.597)	-0.208 (0.570)	-0.092 (0.273)
Instr Exp/Student	0.061*** (0.023)	0.041 (0.046)	0.098* (0.051)	0.023 (0.017)
Rsch Exp/FT Fac	-0.001 (0.002)	0.008 (0.006)	-0.003 (0.002)	0.001 (0.002)
75 Percentile Score	0.218*** (0.043)	0.757*** (0.087)	0.376*** (0.072)	0.225*** (0.034)
Total Students	-0.046*** (0.014)	-0.112*** (0.023)	-0.107*** (0.034)	-0.044*** (0.012)
% UG	0.004 (0.013)	0.002 (0.025)	-0.009 (0.024)	-0.002 (0.012)
% Female	-0.013 (0.015)	-0.077** (0.035)	-0.034 (0.037)	-0.042*** (0.013)
% Minority	0.005 (0.005)	0.006 (0.007)	-0.003 (0.007)	0.008** (0.004)
% of Total BAs in (Field)	-0.041* (0.021)	-0.145*** (0.028)	-0.048*** (0.018)	-0.064*** (0.016)
#NRC 0-10 in (Field)	0.444*** (0.135)	0.648** (0.254)	0.539*** (0.090)	0.288** (0.136)
#NRC 11-25 in (Field)	0.238** (0.101)	0.168 (0.163)	-0.028 (0.123)	0.104 (0.121)
#NRC 26-50 in (Field)	-0.070 (0.110)	0.110 (0.116)	0.128* (0.075)	-0.023 (0.080)
#NRC 51-100 in (Field)	0.004 (0.050)	-0.109 (0.097)	-0.126* (0.068)	-0.003 (0.045)
Carnegie 16	-0.220 (0.217)	-0.541 (0.395)	-0.959** (0.406)	-0.034 (0.209)
Carnegie 17	-0.543 (0.337)	-0.146 (0.625)	-1.802*** (0.523)	-0.635** (0.284)
PhD Prod. Rate Mean	2.61	5.11	4.65	2.21
Missing Dummies	Yes	Yes	Yes	Yes
N	7237	7494	7368	7122
Adjusted R-squared	0.48	0.55	0.58	0.25

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table displays regression results. The dependent variable is the PhD production rate, which is a percentage, not a share. Maximum TTD and truncation correction are as described in Section 4. Each column is a separate regression and corresponds to a field. Observations are weighted by the number of field-specific BAs from the BA-year. Observations are at the institution-BA year, and each institution in the field's sample contributes 26 observations, one for each year between 1980 and 2005. Total students, instructional expenditures per student, and research expenditures per student are in units of 1,000. Percentage undergraduate, female, minority, and total majors in (Field) are percentages, out of 100. BA-year dummies are included, as are missing variable dummies. Robust standard errors are clustered at the institution level.

Table 5: Regression, Carnegie Category 15 Institutions Only

	Humanities	Life Sci.	Physical Sci.	Social Sci.
Stud-FT Fac Ratio	0.001 (0.012)	-0.052** (0.022)	0.012 (0.013)	-0.010 (0.010)
Public	0.036 (0.618)	0.023 (1.119)	0.546 (0.714)	0.321 (0.408)
Instr Exp/Student	0.067** (0.028)	0.031 (0.054)	0.083 (0.063)	0.013 (0.019)
Rsch Exp/FT Fac	0.000 (0.003)	0.016** (0.007)	-0.004 (0.004)	0.003 (0.003)
75 Percentile Score	0.239*** (0.065)	1.003*** (0.127)	0.436*** (0.105)	0.287*** (0.051)
Total Students	-0.056*** (0.020)	-0.133*** (0.031)	-0.112*** (0.037)	-0.052*** (0.015)
% UG	0.004 (0.021)	0.017 (0.041)	-0.049 (0.036)	-0.013 (0.016)
% Female	-0.038 (0.023)	-0.131*** (0.050)	-0.075 (0.067)	-0.047*** (0.017)
% Minority	0.012 (0.015)	0.013 (0.018)	-0.035 (0.024)	0.013 (0.011)
% of Total BAs in (Field)	-0.035 (0.043)	-0.229*** (0.069)	-0.062 (0.042)	-0.066*** (0.022)
#NRC 0-10 in (Field)	0.406** (0.161)	0.427 (0.301)	0.479*** (0.117)	0.215 (0.145)
#NRC 11-25 in (Field)	0.227** (0.109)	0.074 (0.165)	-0.088 (0.152)	0.100 (0.127)
#NRC 26-50 in (Field)	-0.078 (0.149)	0.082 (0.152)	0.074 (0.108)	-0.014 (0.094)
#NRC 51-100 in (Field)	0.040 (0.059)	-0.116 (0.146)	-0.224** (0.108)	0.034 (0.052)
PhD Prod. Rate Mean	3.05	6.44	5.56	2.42
Missing Dummies	Yes	Yes	Yes	Yes
N	2923	2959	2959	2943
Adjusted R-squared	0.52	0.58	0.61	0.32

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table displays regression results. The sample is restricted to Carnegie category 15 institutions. The dependent variable is the PhD production rate, which is a percentage, not a share. Maximum TTD and truncation correction are as described in Section 4. Each column is a separate regression and corresponds to a field. Observations are weighted by the number of field-specific BAs from the BA-year. Observations are at the institution-BA year, and each institution in the field's sample contributes 26 observations, one for each year between 1980 and 2005. A maximum TTD of 15 years is used, and BA-years 2002 and after are truncation-corrected. Total students, instructional expenditures per student, and research expenditures per student are in units of 1,000. Percentage undergraduate, female, minority, and total majors in (Field) are percentages, out of 100. BA-year dummies are included, as are missing variable dummies. Robust standard errors are clustered at the institution level.

Table 6: Lemke-Like Regression

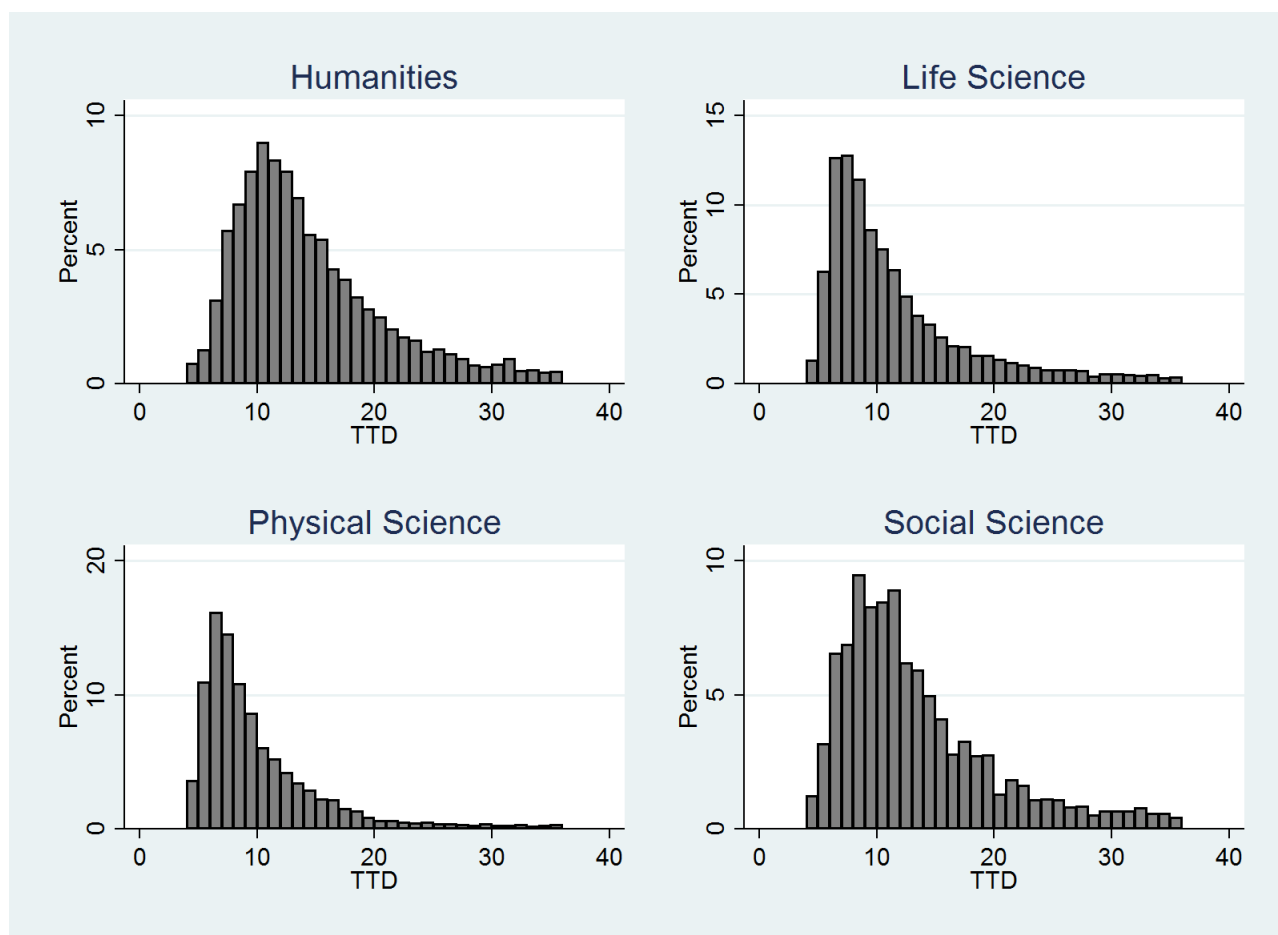
	Humanities	Life Sci.	Physical Sci.	Social Sci.
Stud-FT Fac Ratio	0.021 (0.016)	-0.032 (0.034)	0.011 (0.027)	0.010 (0.013)
Public	0.135 (0.427)	0.281 (0.587)	0.369 (0.596)	0.199 (0.243)
Instr Exp/Student	0.102* (0.053)	-0.006 (0.066)	0.157* (0.091)	0.030 (0.029)
Rsch Exp/FT Fac	-0.001 (0.004)	0.015* (0.008)	-0.004 (0.006)	0.003 (0.003)
75 Percentile Score	0.327*** (0.066)	0.942*** (0.127)	0.623*** (0.139)	0.315*** (0.052)
Total Students	-0.047*** (0.015)	-0.082*** (0.025)	-0.104*** (0.039)	-0.035*** (0.012)
% UG	0.010 (0.013)	-0.000 (0.027)	0.010 (0.025)	0.001 (0.011)
% Female	0.008 (0.018)	-0.056 (0.040)	0.000 (0.043)	-0.021 (0.014)
% Minority	0.005 (0.007)	0.008 (0.010)	-0.000 (0.010)	0.008* (0.005)
% of Total BAs in (Field)	-0.050** (0.021)	-0.121*** (0.026)	-0.045** (0.021)	-0.057*** (0.015)
#NRC 0-10 in (Field)	0.423*** (0.129)	0.321 (0.252)	0.529*** (0.118)	0.208 (0.132)
#NRC 11-25 in (Field)	0.197* (0.104)	-0.063 (0.166)	-0.047 (0.134)	0.007 (0.112)
#NRC 26-50 in (Field)	-0.157 (0.136)	-0.047 (0.118)	0.135 (0.094)	-0.021 (0.067)
#NRC 51-100 in (Field)	0.050 (0.054)	-0.135 (0.100)	-0.100 (0.072)	0.003 (0.043)
Carnegie 16	-0.191 (0.214)	-0.192 (0.442)	-0.514 (0.454)	0.167 (0.197)
Carnegie 17	-0.258 (0.402)	0.708 (0.648)	-1.187* (0.646)	-0.033 (0.292)
PhD Prod. Rate Mean	2.32	4.53	4.97	1.83
Missing Dummies	Yes	Yes	Yes	Yes
N	281	291	284	276
Adjusted R-squared	0.48	0.65	0.67	0.45

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

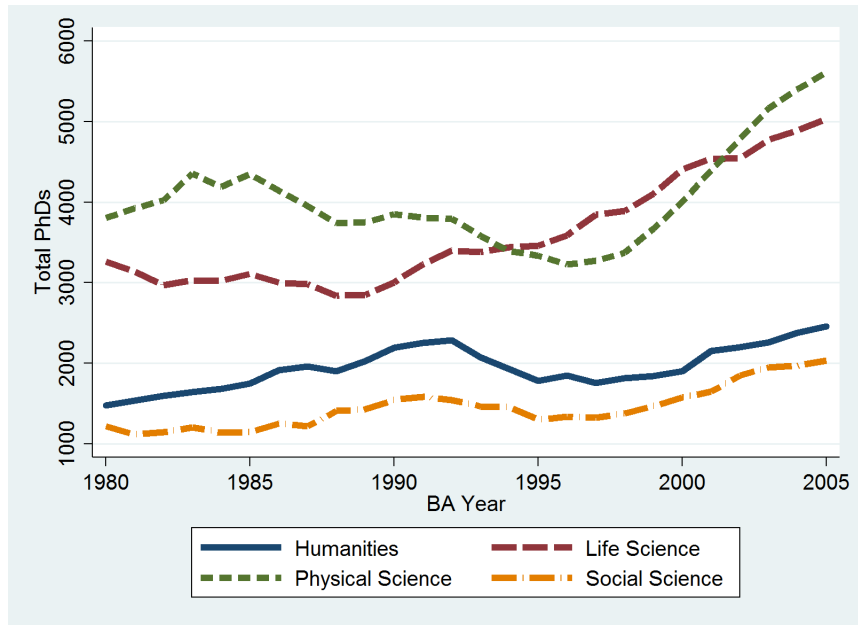
Notes: This table displays regression results. The dependent variable is the number of PhDs produced between 1994-2003 divided by the number of BAs earned from 1989-1998; this is a percentage, not a share. Each column is a separate regression and corresponds to a field. Observations are weighted by the number of field-specific BAs from 1989-1998. Each institution in the field's sample contributes one observation. The independent variables are averages over the 1989-1998 period. Averages are computed among non-missing years; if all years are missing, the missing indicator is one. Total students, instructional expenditures per student, and research expenditures per student are in units of 1,000. Percentage undergraduate, female, minority, and total majors in (Field) are percentages, out of 100. Missing dummies are included. Robust standard errors are used.

Figure 1: Distribution of Time-to-Degree, BA-Years 1980 and 1981

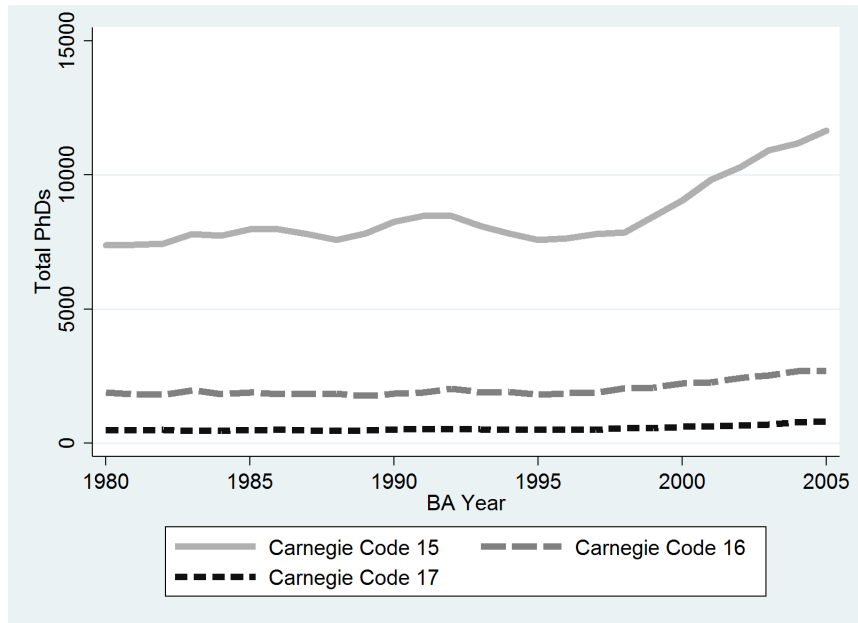


Notes: This figure plots the number of PhDs received by TTD, the number of years after the BA-year that the PhD was received. Only BA-years of 1980 and 1981 are considered. TTDs of less than 4 are grouped with TTD of 4; TTD of 36, which is only observed for 1980, is grouped with TTD of 35.

Figure 2: Total PhDs Produced over BA-Year, by Field and by Carnegie Classification



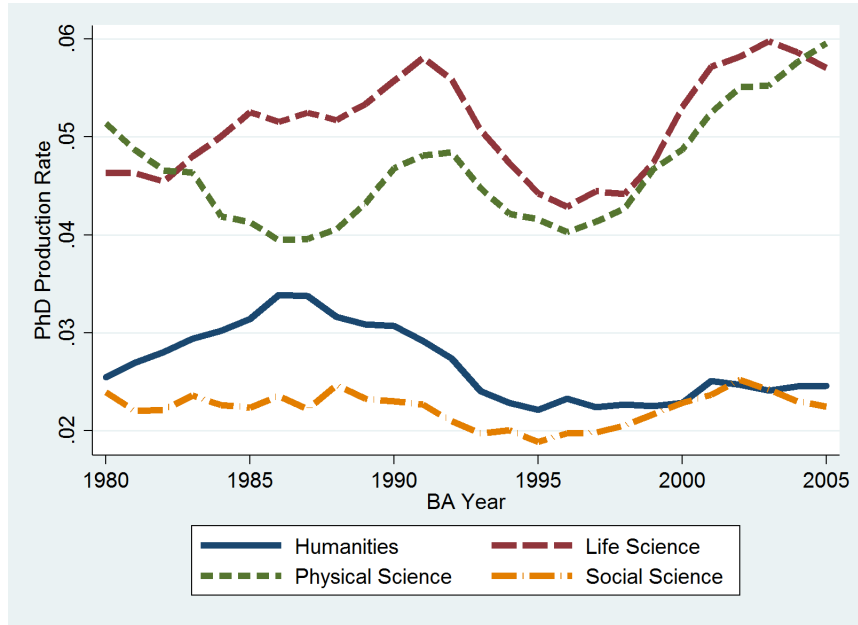
(A)



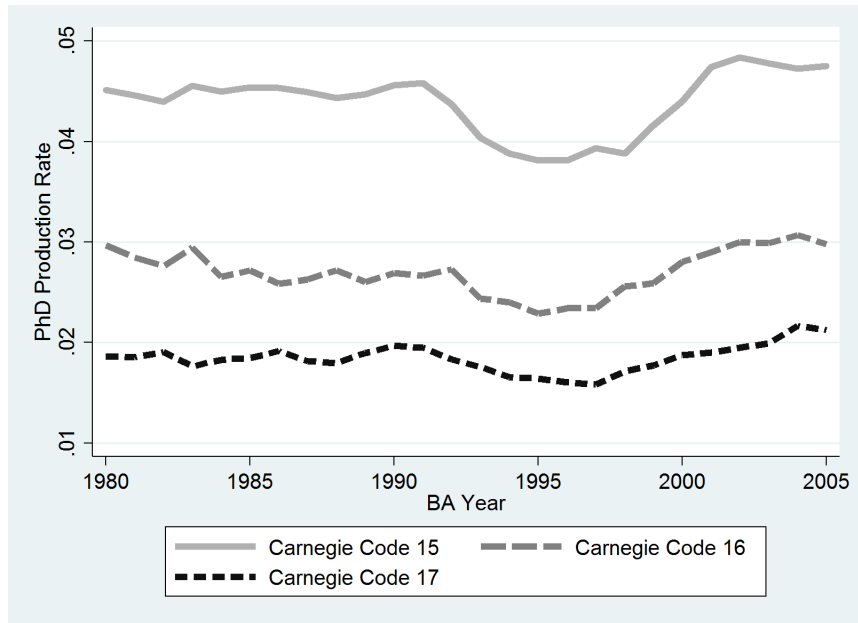
(B)

Notes: Panel A plots the total number of PhDs produced each BA-year, separately by PhD field (Humanities, Life Science, Physical Science, Social Science). Panel B plots the same, but splits by the Carnegie category of the institution. Carnegie categories are: 15—Doctoral Universities: Highest Research Activity, 16—Doctoral Universities: Higher Research Activity, and 17—Doctoral Universities: Moderate Research Activity. Maximum TTDs for fields are as described in Section 4.1.

Figure 3: PhD Production Rate over BA-Year, by Field and by Carnegie Classification



(A)



(B)

Notes: Panel A plots the PhD production rate each BA-year, separately by PhD field (Humanities, Life Science, Physical Science, Social Science). Panel B plots the same, but splits by the Carnegie category of the institution. The PhD production rate in Figure A is computed by dividing the total number of PhDs of a given field (for a given BA-year) by the total number of BAs from the corresponding field and BA-year. The PhD production rate in Panel B is computed by dividing the total number of PhDs (for a given BA-year and across all four fields) of a Carnegie Classification by the total number of BAs across all four fields. Carnegie categories are: 15—Doctoral Universities: Highest Research Activity, 16—Doctoral Universities: Higher Research Activity, and 17—Doctoral Universities: Moderate Research Activity. Maximum TTD and truncation correction are as described in Section 4.

Appendix A Data

We make use of several datasets in our analysis. The primary dataset is the SED, a survey administered to new PhD recipients that has a very high response rate.¹⁵ These individual-level data contain information on those who received their PhD by 2016. We consider the PhD recipients who received their BA between 1980 and 2005, where 1980 refers to the academic year 1980-1981.¹⁶ We drop observations with missing year of BA or PhD receipt as well as those missing BA institution. We construct TTD, used in correcting for truncation (Section 4.2), by computing the length of time between BA and PhD receipt. Our measure of TTD thus includes any time pursuing other degrees, working, etc., before completing the doctoral degree.

As our focus is separately on the humanities, life sciences, physical sciences, and social sciences, we consider the number of PhDs earned in each of these fields; Appendix B.1 describes how these fields are classified. We focus our analysis on doctoral institutions as defined in the Carnegie 2015 classifications, those with Carnegie categories 15 (Doctoral Universities: Highest Research Activity), 16 (Doctoral Universities: Higher Research Activity), and 17 (Doctoral Universities: Moderate Research Activity). We exclude for-profit institutions, institutions that are only graduate schools, and institutions not in the 50 states and Washington DC. We include only institutions that produced at least five PhDs over the sample timeframe. This restriction is made without taking into account TTD and only includes observed PhDs, not those filled in by truncation correction. This is done separately by field, meaning that one institution may be represented in one field and not another. For this reason, the life and physical sciences have a somewhat larger sample than the humanities and social sciences.

¹⁵For example, the response rate in 2016 was 91.8%. See National Science Foundation (2017).

¹⁶The SED data indicates the month and year of the completion of both the BA and the PhD. For BAs, we consider completions occurring between September-December of year $x-1$ and January-August of year x as corresponding to school year x . Because of this, our data includes individuals who graduated in the latter half of 1979. If the month of completion is missing, we use the year variable as the year of completion. For PhDs, we do not attempt to determine the school year, and only use the year of completion.

We use the IPEDS *Completions Survey* to obtain the number of BA degrees granted by each undergraduate institution each year, by field. See Appendix B.2 for classification details. We consider only first reported BA major, which is problematic if the student received two (or more) majors in different fields. The number of BA degrees is combined with the SED data on PhDs produced to compute the humanities PhD production rate, described more fully in Section 4.3. We also calculate the percentage of total BAs earned in each field from this dataset. The Completions Survey data includes a FICE code (from the older HEGIS survey) for all years and a UnitID code (from the IPEDS survey) for 1987 and after. There are a small number of instances where a FICE code maps to more than one UnitID; this may happen, for example, when there are satellite campuses. If the first year that we observe an institution having more than one UnitID is 1987, we aggregate all UnitIDs together for this institution for all years. Because 1987 is the first year that we observe UnitID, it is likely that the multiple campuses were included in the FICE in 1986 and before. This results in an inflated number of BAs for some institutions. If the first year we observe a FICE mapping to multiple UnitIDs is after 1987, we exclude the UnitIDs associated with institution names that do not appear to correspond to the main campus. In general, we will observe too many BAs for observations where the satellite UnitIDs were not broken out into their own UnitIDs.¹⁷

We obtain additional explanatory variables, which are matched at the undergraduate institution-BA year level, from a number of sources. The public institution indicator comes from the SED. We use the IPEDS *Fall Enrollment* dataset to obtain the total number of students, including graduate and professional students. Specifically, we sum full-time students, having a weight of 1, with part-time students, having a weight of 0.4, regardless of institution type or level of student¹⁸ This dataset is also used to compute the share of

¹⁷In one case, the FICE code, which appears to sometimes include satellite campuses and sometimes not, began in 1987; we used one of the several other FICE codes for prior years which correspond to the main campus and which stops in 1986.

¹⁸The weight of 0.4 for part-time students is very similar to the weight used for undergraduate in the IPEDS “Calculation of FTE Students” (IPEDS, 2018). The weight listed for Public 4-year institutions is .403543 and for private 4-year institutions is .392857. The weights are somewhat further from 0.4 for part-time

students who are undergraduates, the share of undergraduates who are female, and the share of undergraduates who are from underrepresented minority groups.¹⁹ Each of these is measured as percentage (out of 100). These variables also use the same weights. We divide total students by full-time faculty counts from the IPEDS Fall Staff Survey to compute the student-faculty ratio.²⁰

Data from the *Delta Cost Project* (Lenihan, 2012) allow us to compute instructional expenditures per student and research expenditures per full-time faculty. The former is obtained by dividing instructional expenditures by full-time students; the latter comes from dividing research expenditures by full-time faculty.²¹

We construct the 75th percentile score variable from the Annual Survey of Colleges *Standard Research Compilation*. We combine the incoming freshmen 75th percentile ACT scores and 75th percentile SAT scores together; we do not use just one of the tests because institutions vary widely in the propensity of their students to take these tests.²² This dataset includes the percentage of students with reported SAT scores and reported ACT scores. We weight the ACT score by the percentage of the total that reported ACT scores, and do similarly for SAT scores. For example, if 70% reported a SAT score and 35% reported an

graduate students: the weights at public 4-year and private 4-year institutions are .361702 and .382059.

¹⁹Due to problems with subcategories not adding up correctly in early years, when computing certain of these variables, we did not use certain subcategories such as FT or PT UG Unclassified. When computing race, the sum of the race categories were often less than expected; we assume that the race categories do not include the non residence variable.

²⁰Full-time faculty counts come from the sum of the “Faculty, full-time men” and “Faculty, full-time women” variables.

²¹Instructional expenditures come from the instruction01 variable—“Expenditures for instruction—current year total.” Research expenditures come from the research01 variable—“Expenditures for research—current year total”.

²²The score data include two identifies, ID and FICE. However, there is not always a 1:1 mapping between the two. After cleaning the test score data, in cases where the FICE code maps to multiple IDs, we select one ID to use based first on the one that seems to match the institution name the best, second based on the one that has non-missing test scores, and finally based on the one with the highest test scores. If ID maps to multiple FICE, we use the FICE that is in the main dataset for matching. Finally, there are cases where there is no match between datasets via FICE, but the institution name is associated with a different FICE code in the score data. We manually make these matches. Finally, we manually match a number of cases by institution name if the FICE codes do not match between datasets. We dropped the data from 1983 as it is unreliable. In one case, the campus the ID referred to apparently changed over time; we used all years. In addition, there were no reported percentage taking ACT or SAT for 1984-1987, so we backfilled the 1984 values with the earliest observation before 1993. We then interpolated the percentage taking ACT and SAT.

ACT score (a student can take both tests), then the SAT will get twice the weight as the ACT. We do this after converting the SAT to be on the ACT scale.²³ We note that after conversion and among observations with both scores, the mean ACT is 1.5 points lower than the mean SAT; the standard deviation for both tests is approximately 3.

We obtain doctoral program rankings from the “scholarly quality of program faculty” ranking from the 1995 National Research Council (NRC) doctoral program rankings (Goldberger et al., 1995). The newer, 2000s rankings do not provide point estimates of program rankings and are thus not used (Ostriker et al., 2011). We also do not use the 1980s rankings, which contain a smaller set of institutions (Jones et al., 1982). These rankings have a high correlation with the 1995 rankings.²⁴ Because we only use the 1995 data, the ranking variables are constant within institution-field. We first classify the different subjects into the four broad fields; classification details appear in Appendix B.3. For a given subject, such as comparative literature (in the humanities), we construct percentile rankings over all departments, including departments at institutions not in our sample. We then count the number of departments a given institution has across all subjects in the field that are in each of the following percentile ranges: 0-10, 11-25, 26-50, and 51-100.²⁵ Importantly, the

²³The composite SAT score, the sum of an institution’s verbal and math 75th percentile scores, was converted to its ACT equivalent using the conversion found at <https://web.archive.org/web/20160130223549/http://act.org/solutions/college-career-readiness/compare-act-sat/>—accessed April 27, 2016. We rounded up to the ACT score if the SAT score was outside one of the listed ranges. For example, if the range was 1290-1320 for a 29 and 1330-1350 for a 30, we converted a score of 1325 to 30. We made note of the re-centering of the SAT scores in 1995 and have converted all scores prior to 1995 to their current day equivalent (separately for math and verbal) using the SAT I Individual Score Equivalents conversion table from the CollegeBoard (see <http://research.collegeboard.org/programs/sat/data/equivalence/sat-individual>—accessed April 27, 2016). The chart presented single conversion numbers. We rounded to the higher number. For example, a verbal score of 690 is re-centered to 750 and 700 is re-centered to 760. We re-centered 695 to 760. For cases where there was a test score for SAT, but not ACT, we used the SAT score entirely and vice versa. As there was a great deal of missing data on earlier years of percentage taking each test, we backfill years missing this data with the earliest reported percentages. We also fill in some of the missing score data using linear interpolation.

²⁴The 1980s ranking include fewer institutions than the 1995 rankings, but for departments that appeared in both rankings (and after re-ranking them to reflect only the departments in both rankings), the correlations of the number of departments at an institution in each rankings interval are: 1 to 10 (.942), 11 to 25 (.888), 26 to 50 (.867), and 51+ (.805). Note that these are not percentiles, but the number in the top 10, etc. If one does not restrict to departments in both, somewhat smaller, but still high correlations are observed: 1 to 10 (.917), 11 to 25 (.842), 26 to 50 (.797), and 51+ (.593).

²⁵Using the set of institutions in the life sciences, correlations within percentile ranges and across fields

smaller the percentile, the higher ranked the department. We use percentiles as opposed to the number in the top 10, etc., because there is a large amount of heterogeneity in the number of departments ranked in each subject. For example, many more departments are ranked in English Language and Literature than in French Language and Literature, and a ranking being ranked, e.g., in the 11-25 range may mean something very different for the former subject than the latter.

We face the challenge of missing data. For many of the control variables, we use a linear extrapolation to fill in data for years that are bookended by two other years with non-missing data. When variables do not contain bookended data due to the survey starting after 1980, we backfill using the first possible year of data (for example, we fill in 1980-1983 with the 1984 value).²⁶ After this process, if an observation is still missing, we code it as 0 and create a missing data dummy that takes on value 1 for missing and 0 otherwise. Universities not included in the 1995 NRC are coded as having zero programs in each of our quality intervals and we do not include a missing dummy variable for such institutions.

Appendix B Classification of Subjects to Fields

B.1 SED

Our classification of degrees from the SED comes from the subjects listed in Appendix E of the 2011 Paper Questionnaire.²⁷ The SED uses fine-grained codes to classify subjects into broader categories. We generally follow these with some exceptions. We list the more general categories, but generally not list the more-specific categories.

range from 0.33 to 0.75, with progressively higher correlations for the 0-10 range than lower ranges in general. These calculations include 0's.

²⁶More specifically, we fill in full-time faculty of years 1980-1986 with 1987; test scores of 1980-1983 with 1984; student-faculty ratios of 1980-1986 with 1987; and instructional and research expenditures of 1980-1986 with 1987. We note that this is not ideal as it does not take into account trends, only takes care of cases that have a non-missing value for the first year of data, and results in a lack of variation over these several years. We compute student-faculty ration, instructional expenditures per student, and research expenditures per faculty after backfilling.

²⁷See pages 6 and 7 of https://www.nsf.gov/statistics/srvydoctorates/surveys/srvydoctorates_2011.pdf.

Humanities

Codes 700–708, 718–770, 775–799. History; Foreign Languages & Literature; Letters; and Other Humanities.

Life Sciences

Codes 0–299. Agricultural Sciences/Natural Resources; Biological/Biomedical Sciences; Health Sciences.

Physical Sciences

Codes 300–599. Engineering; Computer & Information Sciences; Mathematics; Astronomy; Atmospheric Science & Meteorology; Chemistry; Geological & Earth Sciences; Ocean/Marine Sciences; Physics.

Social Sciences

Codes 650–699, 710, 773. Social Sciences. 710 is History, Science & Technology & Society. 773 is Archaeology.

Other

Codes 600–649, 800–989. Psychology; Research & Administration (Education); Teacher Education; Teaching Fields; Other Education; Business Management/Administration; Communication; and fields not classified elsewhere: Architecture/Environmental Design; Family/Consumer Science/Human Science; Law; Library Science; Parks/Sports/Rec./Leisure/Fitness; Public Administration; Social Work; Theology/Religious Education; Other Fields, NEC.

B.2 IPEDS Completions

We classify the subjects, which are at the “Academic Discipline, Detailed”-level, into fields as follows:

Humanities

History; English and Literature; Foreign Language; Other Humanities; Religion and Theol-

ogy; Arts and Music.

Life Sciences

Agricultural Sciences; Biological Sciences; Medical Sciences; Other Life Sciences.

Physical Sciences

Aerospace Engineering; Chemical Engineering; Civil Engineering; Electrical Engineering; Mechanical Engineering; Materials Engineering; Industrial Engineering; Other Engineering; Astronomy; Chemistry; Physics; Other Physical Sciences; Atmospheric Sciences

Social Sciences

Economics; Political Science and Public Administration; Sociology; Anthropology; Linguistics; History of Science; Area and Ethnic Studies; Other Social Sciences.

Other

Psychology; Science Technologies; Engineering Technologies; Health Technologies, Other Science and Engineering Technologies; Interdisciplinary or Other Studies; Architecture and Environmental Design; Science Education; Mathematics Education; Social Science Education; Other Science/Technical Education; Non-Science Education; Business and Management; Communication and Librarianship; Law; Social Science Professions; Vocational Studies and Home Economics; Other Non-Science or Unknown Disciplines.

B.3 NRC

We classify the subjects into fields as follows:

Humanities

Art History; Classics; Comparative Literature; English Language and Literature; French Language and Literature; German Language and Literature; Linguistics; Music; Philosophy; Religion; Spanish and Portuguese Language and Literature; and History.

Life Sciences

Biochemistry and Molecular Biology; Cell and Developmental Biology; Ecology, Evolution,

and Behavior; Molecular and General Genetics; Neurosciences; Pharmacology; and Physiology.

Physical Sciences

Aerospace Engineering; Biomedical Engineering; Chemical Engineering; Civil Engineering; Electrical Engineering; Industrial Engineering; Materials Science; Mechanical Engineering; Astrophysics and Astronomy; Chemistry; Computer Sciences; Geosciences; Mathematics; Oceanography; Physics; Statistics and Biostatistics.

Social Sciences

Anthropology; Economics; Geography; Political Science; Psychology; and Sociology.

Other subjects are not listed here. We do not include institutions that are listed as a medical school in assigning the number of departments an institution has in the percentiles of the fields.