

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

Todd R. Jones and Ronald G. Ehrenberg\*

May 2019

(Forthcoming in *Education Economics*)

\*Todd R. Jones is a Postdoctoral Research Associate in the Economics Department at Georgia State University. During much of writing, he was a Ph.D. Student in Economics at Cornell University. Email: [tjones250@gsu.edu](mailto:tjones250@gsu.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](mailto: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 anonymous referees and Cassandra Benson for making this paper better.

## **Abstract**

Using restricted-access *Survey of Earned Doctorates* data, we investigate which attributes of a doctoral institution predict the share of its undergraduate BAs that will earn a PhD. We use truncation-correction to account for PhD receipt post-sample. Across the fields of humanities, physical, natural, and social sciences, PhD production is positively related to the number of an institution's highly-ranked PhD programs, suggesting such departments may play a role in both producing PhDs and encouraging undergraduates to earn PhDs themselves. It is negatively related to total students and the share of BAs received in the field, and positively related to test scores.

**Keywords:** Higher Education, PhD Production, Doctoral Education

# 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 to 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.<sup>1</sup> 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*<sup>2</sup> than at *doctoral institutions*.<sup>3</sup> However,

---

<sup>1</sup> We 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.

<sup>2</sup> We 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.

<sup>3</sup> This 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

because of the differences in the sizes of these two types of 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.<sup>4</sup> With the exception of life sciences, the ratios are even higher for Master's colleges.<sup>5</sup> 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.<sup>6</sup> 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

---

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 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.

<sup>4</sup> This calculation makes use of the sample restrictions described later, including that the institution must have produced at least 5 PhDs, except that here, all TTDs are considered. The BA years are 1980-2005.

<sup>5</sup> We 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.

<sup>6</sup> Many 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).

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 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 (with Carnegie category 15–17 institutions), 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 four 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 three other variables that are statistically significant across all fields are the incoming student 75th percentile test scores—which is positively related to the PhD production rate—and the total number of students as well as the percentage of total BAs that are in the field, which are both negatively related. Other variables, such as institutional expenditures per student, are significant for some, but not all, fields.

In a robustness check, we restrict the sample to Carnegie category 15 institutions (Highest Research Activity), which have the highest level of research activity. We also 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 and the percentage of total BAs that are in the field are each significant in only two fields in the Carnegie category 15

regressions.

## 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. Having 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, say the, 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.<sup>7</sup> We calculate these variables as follows: For a given subject (e.g., comparative literature) in the field, we compute percentiles for all programs that appear in the NRC data for that subject, including programs not in our paper's sample. After repeating this for all subjects in the field, we then construct an institution's percentile bin counts by summing up the number of programs an institution has—across all subjects in the field—in each percentile bin. Not all institutions have doctoral programs for a given subject; it also happens that even though an institution does have a doctoral program, that the NRC also does not consider it in its rankings. We include a “missing”

---

<sup>7</sup> The smaller the percentile, the higher ranked the department.

NRC variable for institutions for which we do not observe *any* program in any NRC subject ranking (within the field).<sup>8</sup>

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.<sup>9</sup> For example, many more departments are ranked in English Language and Literature than in French Language and Literature, and a department being ranked, e.g., in the 11-25 range (the original ranking, not percentile), may mean something very different for the former subject than the latter.<sup>10</sup>

We illustrate the National Research Council (NRC) variables using the counts of several universities; all examples are specific to the humanities. Each university is presented with four numbers, corresponding to the numbers of departments ranked in the following percentile ranges: 0-10, 11-25, 26-50, and 51-100. The subjects that make up the counts for each school are found in the corresponding footnote. UC-Berkeley: [0-10]: 7, [11-25]: 3, [26-50]: 0, [51-100]: 0. UCLA: 3, 0, 6, 1. UC-Riverside: 0, 0, 1, 3. UC-Irvine: 0, 3, 4, 0. Stanford: 3, 4, 3, 1. Wisconsin: 1, 3, 3, 3. Illinois: 0, 3, 2, 5. Georgia: 0, 0, 2, 4.<sup>11</sup> It is very rare for a university to have more than one ranked department in a given subject. These examples highlight the fact that universities with more

---

<sup>8</sup> We compute percentiles over all institutions in the NRC data, including those not in the sample. However, some institutions do not appear at all in any subject in the NRC data (within a field); these institutions take on 0 values for each percentile range, and take on 1 for the NRC missing variable. If an institution has at least one program ranked in a subject in the field, it takes on 0 for the NRC missing variable. We note that due to issues such as these, our definition of percentile is specific to this paper.

<sup>9</sup> A robustness check uses the raw counts, in this case of the top-ranked program in the field, as opposed to percentiles.

<sup>10</sup> For example, if subject A has roughly 100 ranked programs in the NRC data, there will be roughly 10 subjects in the 0-10 percentile bin (and roughly 50 in the 51-100 bin). If subject B has roughly 300 ranked programs in the NRC data, there will be roughly 30 subjects in the 0-10 percentile bin (and roughly 150 in the 51-100 bin). However, both subject A and subject B will have the same number in the, say, 11-25 range (original ranking, not percentile).

<sup>11</sup> The baseline subjects are: Art History, Classics, Comparative Literature, English Language and Literature, French Language and Literature, German Language and Literature, History, Music, Philosophy, Religion, Spanish and Portuguese Language and Literature. Stanford includes all subjects. UC-Berkeley, UCLA, Wisconsin, and Illinois include all but Religion. UC-Irvine includes all but Art History, Classics, Music, and Religion. UC-Riverside includes Comparative Literature, English Language and Literature, and Philosophy. Georgia includes Art History, English Language and Literature, History, Music, Philosophy, and Spanish and Portuguese Language and Literature. None of these examples have more than one ranked department in a subject.

programs (which are often bigger and/or richer) will have higher counts.<sup>12</sup> In a robustness check, we use only the university's highest ranked department to construct mutually-exclusive ranking categories.

One shortcoming of this paper is that our NRC measures are constant over all years for a given university because we use only the 1995 NRC data. Results should be interpreted accordingly.

### 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 estimate the regression separately by PhD field. Specifically:

$$PhDProdRate_{i,f,y} = \alpha_0 + \alpha_1 IndepVars_{i,f,y} + Carnegie_i + Year_y + \epsilon_{i,f,y},$$

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*, *% UG Female*, *% UG 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

---

<sup>12</sup> A university seldom has more than one ranked department within a subject. The reason a university has higher counts in these variables is mostly due to having more ranked subjects.

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.<sup>13</sup> Robust standard errors are clustered at the institution-level.

In addition to the main regression above, in a robustness check, 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.

We next 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 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 per field. 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.<sup>14</sup>

Finally, we perform several additional robustness checks. First, there is a concern that our

---

<sup>13</sup> It 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.

<sup>14</sup> Results 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.

main sample may rely too heavily on the truncation correction, which might lead to an upward bias in the number of degrees produced. This would be the case if, for example, TTD has decreased over time.<sup>15</sup> In order to lessen the reliance on the truncation correction, we use a shorter maximum TTD (defined below). Second, our main NRC variable—the number of departments in a given percentile range that a university has—gives preference to universities that have more departments, perhaps because they are larger or richer. These variables are also difficult to interpret. In this robustness check, we therefore define new NRC variables: mutually-exclusive indicator variables for the highest-ranked department within the field for two different measures: a) percentiles (0-10, 11-25, 26-50, 51-100, and no departments in the rankings), and b) the raw count ranking (1-10, 11-25, 26-50, 51 and up, and no departments in the rankings). Third, our main sample classifies PhDs to field of study based on the PhD field of study, but does not take into account the BA field of study. In this robustness check, we restrict the sample only to PhDs who also received their BA in the same field of study as the PhD. Finally, our fields as currently classified are quite broad, and use information on subjects within the same field that may not be particularly relevant to students who majored in other subjects. Here, we consider several well-defined subjects separately: economics, history, mathematics,<sup>16</sup> physics, and chemistry. These also have the advantage of having NRC fields that have rankings that are easily put into the mutually-exclusive indicator variables, as in the second robustness check.

## **4 Constructing the PhD Production Rate and Correcting for Truncation**

---

<sup>15</sup> We do see evidence that TTD has fallen over time. See, for example, <https://www.nsf.gov/statistics/2018/nsf18304/report/what-influences-the-path-to-the-doctorate/time-to-degree.cfm>.

<sup>16</sup> We group mathematics and statistics together.

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.

First, though, we clarify our matching process. We classify an individual to a given field (humanities, life science, physical science, or social science) *based solely on their PhD field*. Their BA field of study is not used (except in a robustness check in which we restrict the sample to be those with the same BA field as PhD field). The individual is associated with their BA institution (the institution from which they graduated with a BA) and BA year (the year they graduated with their BA). Then, for each field and for each BA institution-BA year combination, the number of individuals who ultimately earn a PhD and who also graduate from the given BA institution in the given BA year are tabulated. This produces the numerator of the PhD production rate, as described in the sections below. The denominator of the PhD production rate is taken from the IPEDS Completions Survey, and is the total number of individuals who graduated with a BA from the BA institution-BA year; this does not depend on the BA major of PhDs in the SED.

## 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 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.

[Figure 1 near here.]

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.<sup>17</sup> 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.

[Table 1 near here.]

## 4.2 Truncation Correction

---

<sup>17</sup> In this table, TTDs of 36 years are combined with TTDs of 35 years; only the 1980 cohort has an observed TTD of 35 years.

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 the 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  $\left(\frac{1}{x}\right) 6$ .

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 PhDs that occur after TTDs of  $4, \dots, MaxTTD$ , where  $MaxTTD$  is the maximum TTD ( $4, \dots, 23$  years for the humanities).

Specifically, we first compute the cumulative number of PhDs produced by each TTD for each institution, separately by field.<sup>18</sup> Define the number of PhDs produced by institution  $i$ , field  $f$ , BA-year  $y$ , at TTD  $t$ , to be  $PhD_{i,f,y}^t$ , where  $t \in \{4, MaxTTD\}$ .<sup>19</sup> The cumulative number of PhDs for institution  $i$  by TTD  $t$  is then:

---

<sup>18</sup> Recall that a BA institution produces a PhD when one of its BA graduates earns a PhD, regardless of where they earned the PhD.

<sup>19</sup> We reassign TTDs of less than 4 years to be 4 years.

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

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

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

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 PhDs granted for each TTD:<sup>20</sup>

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

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 PhDs for each institution-BA year. This is done by multiplying the number of PhDs produced by the maximum observed TTD by the inverse of the corresponding 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

---

<sup>20</sup> The cumulative percentage is actually a fraction, not a percentage.

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  $y$ .<sup>21</sup> Because we multiply by 100, the PhD production rate is a percentage.

[Table 2 near here.]

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. This is because the Completions Survey data on BAs includes only this major, with the

---

<sup>21</sup> Observations with 0 BAs produced in the field in the BA-year are not included in the regressions.

exception of the final years of our BA data. In order to be consistent across years, we do not include dual majors in the analysis.<sup>22</sup>

## 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.

[Figure 2 near here.]

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

---

<sup>22</sup> Dual majors are not included in the analysis because the IPEDS Completions Survey only has dual majors for the final years of our BA data. In order to be consistent across years, we do not include them. The SED includes information on only one BA. Only this BA is used in the robustness check where we restrict to BAs in the same field as the PhD.

PhDs produced divided by the overall number of BAs.<sup>23</sup> The PhD production rates in 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.

[Figure 3 near here.]

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.<sup>24</sup> Observations are not weighted. This table uses the set of institutions in Life Science sample for all variables except the PhD Production Rate variables, which use only the institutions in their respective field's sample.<sup>25</sup>

[Table 3 near here.]

The 1st–3rd columns display statistics for Carnegie categories 15-17 combined together,

---

<sup>23</sup> This is as opposed to taking the average of individually-computed PhD production rates.

<sup>24</sup> Observations that do not contribute to regression results because e.g., they did not have any humanities BAs, are still included in the summary statistics.

<sup>25</sup> Summary statistics for the PhD Production Rate variables in Table 3 are computed using only the institutions of their respective field's sample. All other variables are computed using the institutions that are in Life Sciences sample. We use the Life Science sample because it is the sample with the most institutions—recall that that the set of institutions included in the samples of the different fields varies somewhat due to the sample requirement that institutions produced at least five PhDs over the time frame. Table 4 shows summary statistics for all variables computed separately by the field's sample; but only for all Carnegie Classifications together (analogous to Columns 1-3 of Table 3; Columns 4-6 of Table 4 are equivalent to Columns 1-3 of Table 3 for variables appearing in both tables).

while the 4th–6th, 7th–9th, and 10th–12th columns separately 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 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.

[Table 4 near here.]

For brevity, we do not present the summary statistics split by Carnegie category for the

three other fields, but in Table 4 we do show the Carnegie categories grouped together (this is analogous to the first three columns of Table 3). Because the sample institutions are quite similar across fields, most of the means are very similar across fields. The exception (besides the PhD production rate) is the NRC variables, which vary much more across fields. For example, the values for physical science are roughly twice those of social science.

## 6 Results

### *Baseline Regression*

Table 5 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.

[Table 5 near here.]

There are four 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. 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 PhDs at a greater rate,

everything else equal: having 1,000 more 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 coefficients and percentages for the other fields are as follows: life sciences (0.65; 13%); physical sciences (0.54; 12%); and social sciences (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, 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.

### ***Robustness Checks***

#### *Restricting to Carnegie 15 Institutions*

[Table 6 near here.]

In Table 6 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 power 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 5, although only humanities and physical sciences continue to be statistically significant. Similarly, the coefficients for the percentage of total BAs that are in the given field are of similar magnitude as in Table 5, but only life sciences and social sciences remain 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 7 shows the results from performing the analysis along the lines of Lemke (2009). Compared to Table 5, the sample size is much smaller because each institution includes only 1

observation in this table, while each institution contributes 26 observations in Table 5;<sup>26</sup> this smaller sample size could influence statistical significance. 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.

[Table 7 near here.]

#### *Additional Robustness Checks*

First, for each field, we define the maximum TTD to be the smallest TTD for which at least 70% of PhDs had been earned. This is as opposed to 90%, which is used in the main results. While we observe fewer PhDs, the number of years that rely on truncation correction is fewer. The maximum TTDs for humanities, life sciences, physical sciences, and social sciences are 16, 12, 10, and 15 years; there is no truncation correction for physical sciences, and only one year of truncation correction for life sciences. Online Appendix Table A.1 shows that while the PhD production mean is smaller, results are largely the same as the baseline, Table 5.

Next, we address the concern that the NRC variables as originally defined favor universities that have more PhD programs, we redefine these variables using mutually exclusive indicator variables based only on the ranking of the university's highest-ranked program within the field as opposed to the number of programs within each category as before. First, we use percentile rankings, and create the following categories: 0-10, 11-25, 26-50, 51-100, and not ranked. Compared to the number of programs within each category, these categories are more easily interpreted: the coefficient for each group is the increase in the PhD production rate as

---

<sup>26</sup> As noted previously, observations are excluded if there are no field-specific BAs produced in the relevant time frame, so actual numbers are sometimes less.

compared to the omitted category, not ranked. Second, we use the raw count rankings, meaning that they are not converted to percentiles. The categories are as follows: 1-10, 11-25, 26-50, 51 and up, and no departments in the rankings.<sup>27</sup> Appendix Table 1 shows results for the percentile indicators. Focusing on the NRC variables, across fields, all 0-10 variables are significantly different than the reference category of not being ranked. For humanities, the coefficients decrease in magnitude across the percentile groups. This pattern holds for life sciences and physical sciences, although the 51-100 group is not different from the reference group. The only category that is significant for social sciences is 0-10. Online Appendix Table A.2, which is for the raw count indicators, follows roughly the same pattern. In other results in the Online Appendix,<sup>28</sup> when we restrict to be Carnegie Category 15 institutions only, results for both percentile and raw count indicators are similar for humanities and life sciences, but no categories are significant for physical sciences and social sciences.

In Appendix Table 2 we include only those whose BA is in the same field of study as their PhD. Noting that coefficients are expected to be of a smaller magnitude due to the smaller PhD production rate mean, results are fairly similar to Table 5.

Appendix Table 3 shows our final robustness check, which is to consider several subjects individually: economics, history, mathematics, physics, and chemistry. Here, we show the version using NRC percentile indicators. The PhD production rate mean varies greatly across subjects, ranging from 1.6 in economics to 12.3 in chemistry to 21.5 in physics. To a large degree, results are substantively similar to the main results. Across all subjects, student test scores are positive and significant. Both total students and the percentage of total BAs in the subject are negative and significant across all subjects but physics. Percentage female is negative and significant for all but

---

<sup>27</sup> We note that the number of departments ranked varies substantially across field; some fields have fewer than 50 departments ranked.

<sup>28</sup> These are Online Appendix Tables A.3 and A.5.

history. The coefficients on the NRC variables are mostly monotonically decreasing, with most of the 0-10, 11-25, and 26-50 categories being statistically significant. The coefficients on the 0-10 category are quite large for mathematics, chemistry, and physics; relative to the PhD production mean, the coefficient for mathematics is particularly large. This may mean that these programs are influential to the students; it could also mean that students who are likely to pursue PhDs are attracted to the types of universities ranked highly in their PhD programs.

The Online Appendix shows a number of additional robustness checks, which are different versions and combinations of the above. Results are, for the most part, substantively similar.

## **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 a 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 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.

## **8 Declaration of Interest Statement**

The authors declare no conflicts of interest.

## 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 head-counts),” <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.

## Tables and Figures

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	<b>90.7</b>	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	<b>90.2</b>	94.9	86.8	90.9
21	87.2	91.4	95.5	88.6	92.0
22	89.0	92.4	96.0	<b>90.2</b>	93.0
23	<b>90.6</b>	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, by Carnegie Classification, Life Science Sample

	All			Carnegie 15			Carnegie 16			Carnegie 17		
	Mean	SD	Missing	Mean	SD	Missing	Mean	SD	Missing	Mean	SD	Missing
Hum PhD Prod Rate	2.9	6.7	.01	3.8	7.6	.01	2.5	7.6	.01	1.9	2.5	0
LifSci PhD Prod Rate	6	8	.01	8.9	9.5	0	4.9	7.1	.03	3.1	4.3	0
PhySci PhD Prod Rate	4.7	7.7	0	6.4	5.7	0	3.5	2.9	0	3.7	12.9	.01
SocSci PhD Prod Rate	3.1	10.1	.02	3.4	9	.01	3.2	11.1	.02	2.6	10.4	.03
Stud-FT Fac Ratio	18.2	14.4	0	15.5	20	0	18.9	9.6	0	21.2	7	0
Public	.7	.5	0	.7	.5	0	.7	.4	0	.5	.5	0
Instr Exp/Student	7.5	6.8	.06	10.5	8.8	.04	6.3	4.7	.09	4.6	2.3	.03
Rsch Exp/FT Fac	48.4	75.6	.14	76.1	93.2	.04	37.3	58.6	.09	10.1	13.6	.32
75 Percentile Score	26.6	3.1	.13	28.3	3	.1	26	2.7	.11	24.6	2.3	.2
Total Students	13.1	8.9	0	19.7	9.4	0	10.7	5.3	0	6.7	4.5	0
% UG	77.3	13.1	0	72.5	12.8	0	79.8	12.6	.01	81.1	12.1	0
% UG Female	50.7	9.7	0	48.5	6.8	0	49.5	10.8	.01	55.5	9.9	0
% UG Minority	16.4	18.2	0	12.8	8.7	0	16.7	20.8	.01	21.2	23.1	0
% Hum	12	7.7	0	14.4	7.4	0	10.8	8.2	0	10	6.4	0
% LifSci	11.7	7.3	0	12.2	5.2	0	11.1	6.6	0	11.5	10.2	0
% PhysSci	15.6	16.3	0	17.5	13.1	0	18.7	21.5	0	9	9.4	0
% SocSci	10.1	8.1	0	14.1	9	0	8	6.7	0	7.1	5.5	0
% Other	50.6	18.2	0	41.7	16.8	0	51.5	18.1	0	62.4	12.3	0
Hum NRC Perc 0-10	.5	1.6	.52	.7	1.9	.17	0	0	.61	0	0	.91
Hum NRC Perc 11-25	.7	1.2	.52	1	1.3	.17	0	0	.61	0	0	.91
Hum NRC Perc 26-50	1.1	1.4	.52	1.5	1.4	.17	.1	.3	.61	0	0	.91
Hum NRC Perc 51-100	2.2	1.7	.52	2.4	1.8	.17	1.8	1.5	.61	1	0	.91
LifSci NRC Perc 0-10	.6	1.5	.43	.9	1.7	.04	0	0	.5	0	0	.92
LifSci NRC Perc 11-25	.8	1.3	.43	1.2	1.5	.04	0	.1	.5	0	0	.92
LifSci NRC Perc 26-50	1.3	1.7	.43	1.8	1.8	.04	.3	.8	.5	0	0	.92
LifSci NRC Perc 51-100	2.3	2	.43	1.9	2	.04	3.1	1.7	.5	2	1.4	.92
PhySci NRC Perc 0-10	.8	2.4	.35	1.4	3	.03	0	0	.3	0	0	.9
PhySci NRC Perc 11-25	1	1.9	.35	1.7	2.2	.03	.1	.4	.3	0	0	.9
PhySci NRC Perc 26-50	1.8	2.4	.35	2.8	2.6	.03	.4	1.2	.3	.1	.3	.9
PhySci NRC Perc 51-100	3.4	2.7	.35	3.5	3	.03	3.4	2.3	.3	1.8	1.6	.9
SocSci NRC Perc 0-10	.4	1	.55	.5	1.2	.12	0	0	.73	0	0	.96
SocSci NRC Perc 11-25	.5	.9	.55	.6	1	.12	0	0	.73	.3	.5	.96
SocSci NRC Perc 26-50	.8	1.1	.55	1	1.1	.12	.1	.3	.73	0	0	.96
SocSci NRC Perc 51-100	1.7	1.5	.55	1.7	1.6	.12	1.8	1	.73	1	0	.96
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, with the first set of columns grouping categories 15, 16, and 17 together, and the remaining columns displaying each category separately. The PhD Production Rate variables are computed using the institutions in the samples of their respective fields. The remaining variables are computed from the institutions in the Life Science sample. 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: Summary Statistics, Carnegie Classification 15-17 Combined, Each Field's Sample Displayed Separately

	Humanities			Life Science			Physical Science			Social Science		
	Mean	SD	Missing	Mean	SD	Missing	Mean	SD	Missing	Mean	SD	Missing
PhD Prod Rate	2.9	6.7	.01	6	8	.01	4.7	7.7	0	3.1	10.1	.02
Stud-FT Fac Ratio	18.1	14.4	0	18.2	14.4	0	17.9	14.3	0	18.1	14.5	0
Public	.7	.5	0	.7	.5	0	.7	.5	0	.7	.5	0
Instr Exp/Student	7.5	6.9	.06	7.5	6.8	.06	7.6	6.8	.06	7.7	6.9	.05
Rsch Exp/FT Fac	48.7	76.9	.14	48.4	75.6	.14	48.6	75.8	.12	48.3	76.1	.11
75 Percentile Score	26.6	3.1	.13	26.6	3.1	.13	26.6	3.1	.13	26.6	3.1	.13
Total Students	13.5	8.9	0	13.1	8.9	0	13.4	8.8	0	13.6	8.8	0
% UG	77.7	12.7	0	77.3	13.1	0	77.2	13.1	0	77.1	13.2	0
% UG Female	51.4	8.5	0	50.7	9.7	0	50.3	9.3	0	51.1	8.9	0
% UG Minority	16.2	18	0	16.4	18.2	0	16.3	18	0	16.7	18.5	0
% Hum	12.5	7.5	0	12	7.7	0	11.8	7.1	0	12.1	7.5	0
% LifSci	11.6	6.3	0	11.7	7.3	0	11.7	7.3	0	11.8	7.2	0
% PhysSci	14.2	13.1	0	15.6	16.3	0	16	16.4	0	14.8	13.8	0
% SocSci	10.5	8	0	10.1	8.1	0	10.2	8.1	0	10.4	8.1	0
% Other	51.2	17.3	0	50.6	18.2	0	50.3	18.1	0	51	17.7	0
NRC 0-10 (in Field)	.5	1.6	.5	.6	1.5	.43	.8	2.4	.33	.4	1	.54
NRC 11-25 (in Field)	.7	1.2	.5	.8	1.3	.43	1	1.9	.33	.5	.9	.54
NRC 26-50 (in Field)	1.1	1.4	.5	1.3	1.7	.43	1.8	2.4	.33	.8	1.1	.54
NRC 51-100 (in Field)	2.2	1.7	.5	2.3	2	.43	3.4	2.7	.33	1.7	1.5	.54
N	7306			7592			7384			7254		

Notes: This table displays means, standard deviations, and percent of observations that are missing. Summary statistics are computed separately using each field's sample. In all cases, Carnegie categories 15, 16, and 17 together. Observations are unweighted. 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 5: 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)
% UG Female	-0.013 (0.015)	-0.077** (0.035)	-0.034 (0.037)	-0.042*** (0.013)
% UG 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 Perc (in Field)	0.444*** (0.135)	0.648** (0.254)	0.539*** (0.090)	0.288** (0.136)
#NRC 11-25 Perc (in Field)	0.238** (0.101)	0.168 (0.163)	-0.028 (0.123)	0.104 (0.121)
#NRC 26-50 Perc (in Field)	-0.070 (0.110)	0.110 (0.116)	0.128* (0.075)	-0.023 (0.080)
#NRC 51-100 Perc (in Field)	0.004 (0.050)	-0.109 (0.097)	-0.126* (0.068)	-0.003 (0.045)
NRC Missing	-0.518** (0.207)	-1.201** (0.511)	-0.988*** (0.360)	-0.137 (0.203)
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

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 (excluding years with 0 BAs). Total students, instructional expenditures per student, and research expenditures per student are in units of 1,000. Percentage undergraduate, undergraduates that are female, undergraduates that are minority, and total BAs (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. Standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 6: 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)
% UG Female	-0.038 (0.023)	-0.131*** (0.050)	-0.075 (0.067)	-0.047*** (0.017)
% UG 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 Perc (in Field)	0.406** (0.161)	0.427 (0.301)	0.479*** (0.117)	0.215 (0.145)
#NRC 11-25 Perc (in Field)	0.227** (0.109)	0.074 (0.165)	-0.088 (0.152)	0.100 (0.127)
#NRC 26-50 Perc (in Field)	-0.078 (0.149)	0.082 (0.152)	0.074 (0.108)	-0.014 (0.094)
#NRC 51-100 Perc (in Field)	0.040 (0.059)	-0.116 (0.146)	-0.224** (0.108)	0.034 (0.052)
NRC Missing	-0.608** (0.297)	-1.987*** (0.738)	-0.339 (1.478)	-0.015 (0.353)
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

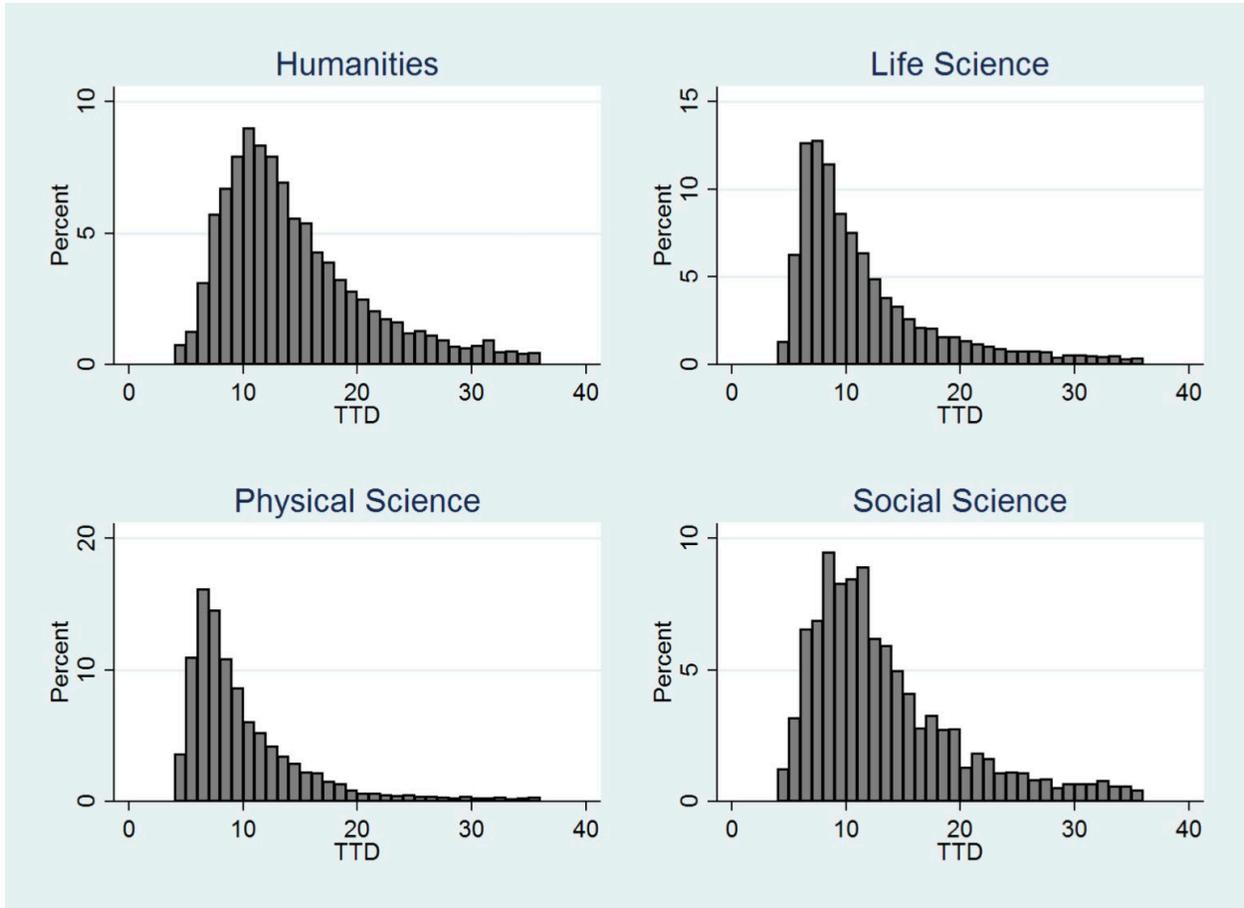
Notes: This table displays regression results. The sample is restricted to Carnegie category 15 institutions. The dependent variable is the PhD production rate, which is a percentage, not a share. Maximum TTD and truncation correction are as described in Section 4. Each column is a separate regression 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 (excluding years with 0 BAs). Total students, instructional expenditures per student, and research expenditures per student are in units of 1,000. Percentage undergraduate, undergraduates that are female, undergraduates that are minority, and total BAs (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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 7: 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)
% UG Female	0.008 (0.018)	-0.056 (0.040)	0.000 (0.043)	-0.021 (0.014)
% UG 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 Perc (in Field)	0.423*** (0.129)	0.321 (0.252)	0.529*** (0.118)	0.208 (0.132)
#NRC 11-25 Perc (in Field)	0.197* (0.104)	-0.063 (0.166)	-0.047 (0.134)	0.007 (0.112)
#NRC 26-50 Perc (in Field)	-0.157 (0.136)	-0.047 (0.118)	0.135 (0.094)	-0.021 (0.067)
#NRC 51-100 Perc (in Field)	0.050 (0.054)	-0.135 (0.100)	-0.100 (0.072)	0.003 (0.043)
NRC Missing	-0.414** (0.188)	-0.995* (0.538)	-0.809* (0.415)	-0.152 (0.187)
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

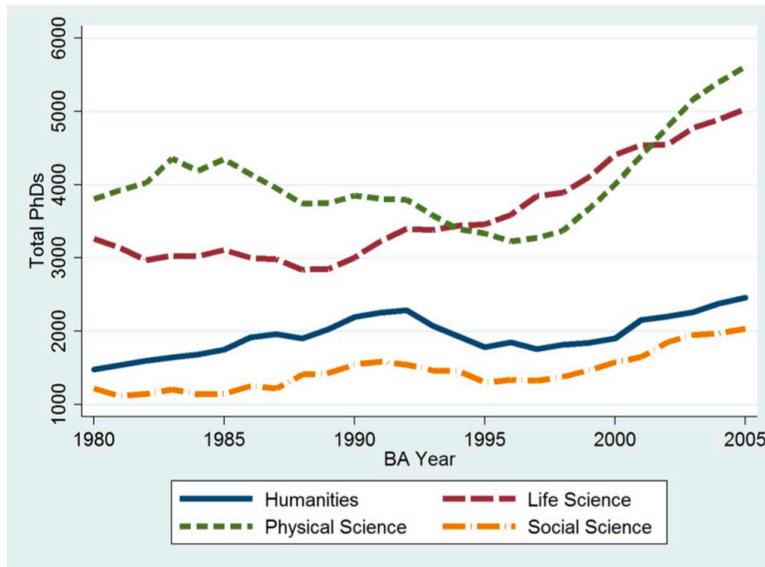
Notes: This table displays regression results. The dependent variable is the number of PhDs produced between 1994-2003 divided by the number of BAs earned from 1989-1998; this is a percentage, not a share. Each column is a separate regression and corresponds to a field. Observations are weighted by the 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, undergraduates that are female, undergraduates that are minority, and total BAs (in Field) are percentages, out of 100. Missing dummies are included. Robust standard errors are used. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

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

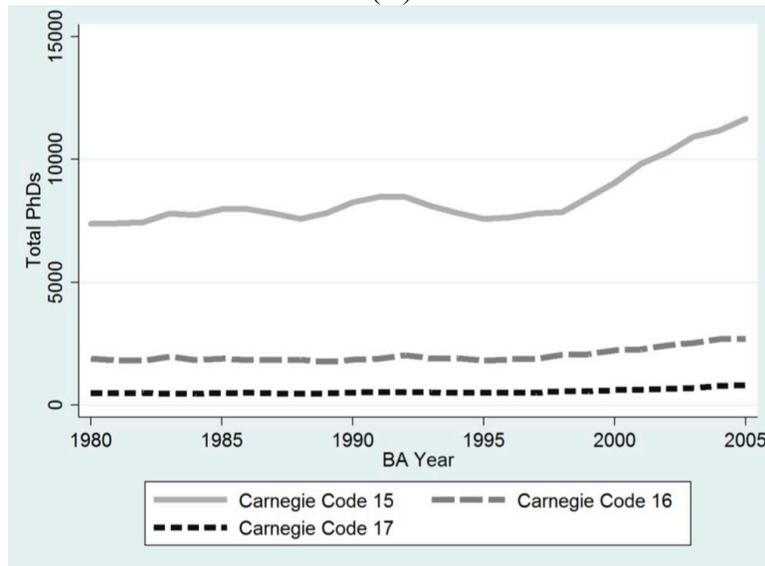


Notes: This figure plots the number of PhDs received by TTD, the number of years after the BA-year that the PhD was received. Only BA-years of 1980 and 1981 are considered. TTDs of less than 4 are grouped with TTD of 4; TTD of 36, which is only observed for 1980, is grouped with TTD of 35.

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



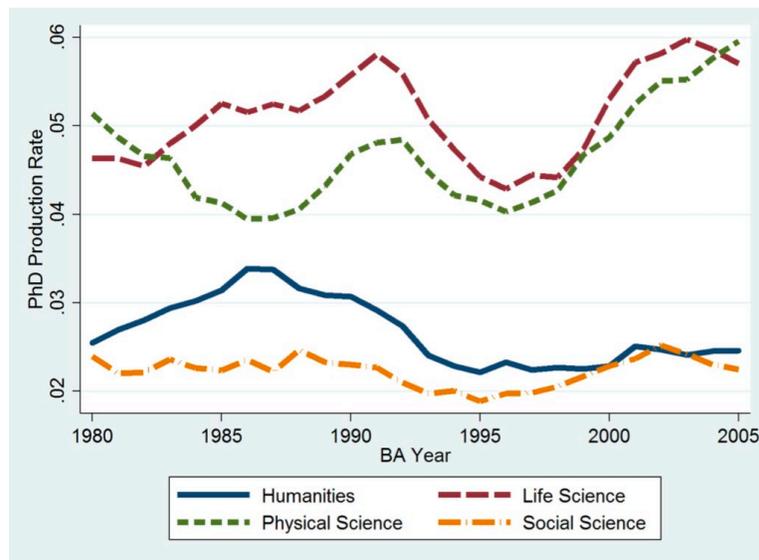
(A)



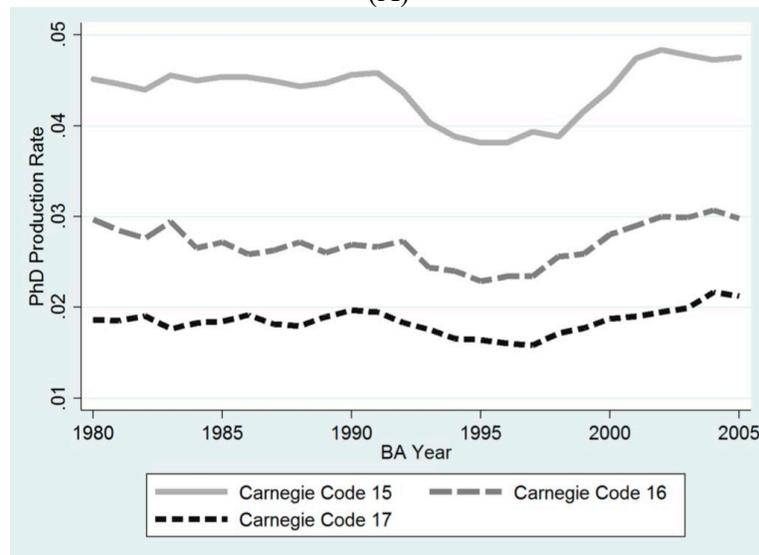
(B)

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 Panel 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.<sup>29</sup> 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.<sup>30</sup> 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; Online Appendix A.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. Within each field, we include only institutions that produced at least five PhDs over the sample timeframe, irrespective of time-to-degree. 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.

---

<sup>29</sup> For example, the response rate in 2016 was 91.8%. See National Science Foundation (2017).

<sup>30</sup> The 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.

For this reason, the life and physical sciences have a somewhat larger sample than the humanities and social sciences.

We use the IPEDS Completions Survey to obtain the number of BA degrees granted by each undergraduate institution each year, by field. See Online Appendix A.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.<sup>31</sup>

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,

---

<sup>31</sup> In 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.

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 student.<sup>32</sup> This dataset is also used to compute the share of students who are undergraduates, the share of undergraduates who are female, and the share of undergraduates who are from underrepresented minority groups.<sup>33</sup> 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.<sup>34</sup>

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.<sup>35</sup>

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.<sup>36</sup> This dataset includes the

---

<sup>32</sup> The 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 graduate students: the weights at public 4-year and private 4-year institutions are .361702 and .382059.

<sup>33</sup> Due 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 was often less than expected; we assume that the race categories do not include the non residence variable.

<sup>34</sup> Full-time faculty counts come from the sum of the “Faculty, full-time men” and “Faculty, full-time women” variables.

<sup>35</sup> Instructional 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”.

<sup>36</sup> The 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

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 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.<sup>37</sup> 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 and the 2000s rankings, with important exceptions, generally have a high correlation with the 1995 rankings.<sup>38</sup> Because we only

---

backfilled the 1984 values with the earliest observation before 1993. We then interpolated the percentage taking ACT and SAT.

<sup>37</sup> The 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/compareact-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.

<sup>38</sup> The 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 and using only the highest ranked department within a subject), the correlations (considering all fields at once) of the number of departments at an institution in each percentile rankings interval are: 0 to 10 (.94), 11 to 25 (.84), 26 to 50 (.78), and 51-100 (.90). For the percentile indicator bins (using only the highest-ranked subject in the field) are: 0 to 10 (0.82), 11 to 25 (0.49), 26 to 50 (0.40), 51 to 100 (0.75). The correlations of the 2000s and 1995 rankings follow a similar pattern, but are somewhat lower. For the number of departments ranked in each percentile, the rankings are: 0 to 10 (.90), 11 to 25 (.77), 26 to 50 (.72), and 51-100 (.86). For the percentile indicator bins, they are: 0 to 10 (.76), 11 to 25 (.43), 26 to 50 (.38), and 51-100 (.72).

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 Online Appendix A.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.<sup>39</sup>

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).<sup>40</sup> 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.

---

<sup>39</sup> Using the set of institutions in the life sciences, correlations within percentile ranges and across fields 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.

<sup>40</sup> More 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 ratio, instructional expenditures per student, and research expenditures per faculty after backfilling.

## Appendix B Tables

Appendix Table 1: Baseline Regression, NRC Percentile Indicators

	Humanities	Life Sci.	Physical Sci.	Social Sci.
Stud-FT Fac Ratio	0.007 (0.008)	-0.028* (0.017)	-0.008 (0.010)	-0.003 (0.007)
Public	-0.073 (0.433)	-0.367 (0.588)	0.093 (0.632)	0.040 (0.269)
Instr Exp/Student	0.064** (0.026)	0.033 (0.049)	0.085 (0.056)	0.020 (0.017)
Rsch Exp/FT Fac	-0.001 (0.002)	0.009* (0.005)	0.001 (0.003)	0.001 (0.002)
75 Percentile Score	0.282*** (0.047)	0.757*** (0.088)	0.518*** (0.100)	0.234*** (0.035)
Total Students	-0.054*** (0.014)	-0.108*** (0.025)	-0.112*** (0.027)	-0.040*** (0.011)
% UG	-0.002 (0.016)	-0.005 (0.027)	-0.025 (0.024)	-0.008 (0.012)
% UG Female	-0.031 (0.021)	-0.083** (0.037)	-0.024 (0.038)	-0.048*** (0.013)
% UG Minority	0.010** (0.005)	0.006 (0.007)	0.008 (0.009)	0.011*** (0.004)
% of Total BAs (in Field)	-0.044** (0.019)	-0.143*** (0.027)	-0.056** (0.022)	-0.054*** (0.014)
NRC 0-10 Perc I (in Field)	2.296*** (0.558)	4.125*** (0.793)	3.697*** (0.692)	0.775** (0.341)
NRC 11-25 Perc I (in Field)	0.954*** (0.235)	2.345*** (0.559)	1.905*** (0.553)	0.419 (0.321)
NRC 26-50 Perc I (in Field)	0.623** (0.246)	1.591*** (0.453)	1.056** (0.499)	-0.270 (0.231)
NRC 51-100 Perc I (in Field)	0.442*** (0.164)	0.610 (0.403)	0.210 (0.327)	0.128 (0.162)
Carnegie 16	-0.024 (0.237)	0.000 (0.385)	-0.132 (0.432)	0.013 (0.185)
Carnegie 17	-0.299 (0.348)	0.484 (0.602)	-0.803 (0.531)	-0.562** (0.254)
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	0.44	0.55	0.52	0.25

Notes: This table displays regression results. This table is similar to Table 5, except that the NRC variables are now mutually-exclusive indicator variables based on the percentile ranking of the top-ranked program (within the field) at the university. 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 (excluding years with 0 BAs). Total students, instructional expenditures per student, and research expenditures per student are in units of 1,000. Percentage undergraduate, undergraduates that are female, undergraduates that are minority, and total BAs (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. Standard errors in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

Appendix Table 2: Baseline Regression, Restrict Sample to BAs in Same Field as PhD

	Humanities	Life Sci.	Physical Sci.	Social Sci.
Stud-FT Fac Ratio	0.006 (0.006)	-0.019* (0.010)	0.005 (0.007)	-0.001 (0.004)
Public	-0.244 (0.341)	-0.342 (0.479)	-0.265 (0.527)	0.021 (0.180)
Instr Exp/Student	0.059*** (0.020)	0.008 (0.031)	0.091* (0.048)	0.016 (0.010)
Rsch Exp/FT Fac	-0.002 (0.002)	0.006* (0.003)	-0.003 (0.002)	-0.000 (0.001)
75 Percentile Score	0.164*** (0.036)	0.533*** (0.063)	0.337*** (0.068)	0.119*** (0.023)
Total Students	-0.038*** (0.012)	-0.070*** (0.018)	-0.097*** (0.031)	-0.024*** (0.008)
% UG	0.008 (0.010)	0.029 (0.018)	-0.005 (0.022)	-0.003 (0.008)
% UG Female	-0.019* (0.011)	-0.056** (0.025)	-0.027 (0.032)	-0.019** (0.009)
% UG Minority	0.003 (0.005)	0.002 (0.005)	-0.003 (0.007)	0.003 (0.002)
% of Total BAs (in Field)	-0.018 (0.018)	-0.076*** (0.019)	-0.038** (0.017)	-0.025*** (0.009)
#NRC 0-10 Perc (in Field)	0.377*** (0.106)	0.551*** (0.179)	0.511*** (0.087)	0.194** (0.090)
#NRC 11-25 Perc (in Field)	0.204*** (0.078)	0.147 (0.143)	-0.015 (0.114)	0.050 (0.077)
#NRC 26-50 Perc (in Field)	-0.046 (0.093)	0.137 (0.086)	0.126* (0.068)	-0.011 (0.050)
#NRC 51-100 Perc (in Field)	0.011 (0.042)	-0.039 (0.073)	-0.101 (0.062)	0.002 (0.028)
NRC Missing	-0.369** (0.173)	-0.799** (0.367)	-0.784** (0.323)	-0.062 (0.128)
Carnegie 16	-0.193 (0.182)	-0.492 (0.307)	-0.858** (0.379)	-0.021 (0.131)
Carnegie 17	-0.344 (0.261)	-0.242 (0.465)	-1.683*** (0.485)	-0.327* (0.176)
PhD Prod. Rate Mean	2.08	3.85	4.23	1.37
Missing Dummies	Yes	Yes	Yes	Yes
N	7012	7460	7344	6715
Adjusted R-squared	0.49	0.51	0.58	0.24

Notes: This table displays regression results. This table is similar to Table 5, except that only PhDs who received their BA in the same field as their PhD are considered. 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 (excluding years with 0 BAs). Total students, instructional expenditures per student, and research expenditures per student are in units of 1,000. Percentage undergraduate, undergraduates that are female, undergraduates that are minority, and total BAs (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. Standard errors in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

Appendix Table 3: Subject-specific Regressions, NRC Percentile Indicators

	History	Economics	Mathematics	Chemistry	Physics
Stud-FT Fac Ratio	0.013 (0.010)	-0.008 (0.008)	0.003 (0.013)	-0.038 (0.029)	0.011 (0.037)
Public	-0.145 (0.688)	0.074 (0.333)	-0.959 (0.666)	-0.999 (1.198)	-3.789 (2.530)
Instr Exp/Student	0.084*** (0.027)	0.001 (0.016)	0.109* (0.065)	0.003 (0.073)	0.096 (0.104)
Rsch Exp/FT Fac	-0.003 (0.003)	0.002 (0.003)	0.002 (0.003)	0.008 (0.009)	-0.008 (0.010)
75 Percentile Score	0.343*** (0.077)	0.211*** (0.043)	0.364*** (0.093)	0.520** (0.234)	0.829** (0.352)
Total Students	-0.084*** (0.020)	-0.043*** (0.014)	-0.164*** (0.034)	-0.126** (0.049)	0.065 (0.098)
% UG	0.009 (0.022)	-0.019 (0.015)	0.046 (0.033)	0.021 (0.039)	0.019 (0.098)
% UG Female	-0.032 (0.022)	-0.029* (0.015)	-0.102*** (0.028)	-0.184*** (0.048)	-0.252*** (0.088)
% UG Minority	0.029*** (0.010)	0.005 (0.008)	0.014 (0.014)	-0.034** (0.014)	0.012 (0.065)
% of Total BAs (in Subject)	-0.340*** (0.118)	-0.234*** (0.047)	-0.942*** (0.113)	-2.654*** (0.478)	-0.295 (0.706)
NRC 0-10 Perc I (in Subject)	2.858** (1.247)	0.907* (0.509)	6.196*** (1.350)	6.962*** (2.251)	10.104** (4.337)
NRC 11-25 Perc I (in Subject)	1.537*** (0.498)	-0.041 (0.281)	1.642** (0.775)	3.387* (1.762)	4.763* (2.637)
NRC 26-50 Perc I (in Subject)	0.912** (0.386)	-0.047 (0.248)	1.283** (0.539)	3.465** (1.434)	2.588 (2.270)
NRC 51-100 Perc I (in Subject)	0.405 (0.265)	-0.133 (0.189)	0.418 (0.393)	-0.397 (0.983)	1.528 (1.411)
Carnegie 16	-0.371 (0.391)	0.244 (0.229)	-0.934* (0.483)	0.884 (0.881)	-2.490 (2.215)
Carnegie 17	-0.831 (0.507)	-0.061 (0.395)	-1.899*** (0.695)	-0.359 (1.249)	-2.950 (2.733)
PhD Prod. Rate Mean	3.45	1.61	4.47	12.29	21.49
Missing Dummies	Yes	Yes	Yes	Yes	Yes
N	6244	4892	6141	7028	5792
Adjusted R-squared	0.18	0.13	0.29	0.18	0.19

Notes: This table displays regression results. Instead of fields, five subjects are used. This table is otherwise similar to Table 5, except that the NRC variables are now mutually-exclusive indicator variables based on the percentile ranking of the top-ranked program (within the subject) at the university. 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 subject. Observations are weighted by the number of subject-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 (excluding years with 0 BAs). Total students, instructional expenditures per student, and research expenditures per student are in units of 1,000. Percentage undergraduate, undergraduates that are female, undergraduates that are minority, and total BAs (in Subject) are percentages, out of 100. BA-year dummies are included, as are missing variable dummies. Robust standard errors are clustered at the institution level. Standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .