Volume Title: Productivity in Higher Education

Volume Authors/Editors: Caroline M. Hoxby and Kevin Stange, editors

Volume Publisher: University of Chicago Press

Volume ISBNs: 978-0-226-57458-5 (cloth); 978-0-226-57461-5 (electronic)

Volume URL:
https://www.nber.org/books-and-chapters/productivity-higher-education

Conference Date: May 31–June 1, 2016

Publication Date: November 2019

Chapter Title: Learning and Earning: An Approximation to College Value Added in Two Dimensions

Chapter Author(s): Evan Riehl, Juan E. Saavedra, Miguel Urquiola

Chapter URL:
https://www.nber.org/books-and-chapters/productivity-higher-education/learning-and-earning-approximation-college-value-added-two-dimensions

Chapter pages in book: (p. 105 – 132)

4

Learning and Earning
An Approximation to College Value Added in Two Dimensions

Evan Riehl, Juan E. Saavedra, and Miguel Urquiola

4.1 Introduction

Colleges produce outputs in various dimensions. Parents and students, for instance, care about colleges' ability to place graduates on good career trajectories. As a result, the United States and other countries now provide information on the labor market *earnings* of graduates from various colleges and majors.[1] A drawback of such measures is that they typically do not adjust for ability; some colleges might perform better, for instance, simply because they attract more able students.

The earnings dimension, however, is not the only one that parents, students, and especially policy makers care about. A second dimension of interest is *learning*—namely, the ability of colleges to enhance human capital and skills. System-wide measures of learning are uncommon, in part because most countries lack nationwide college graduation exams. Questions remain, therefore, about the extent to which these two dimensions of

Evan Riehl is assistant professor of economics at Cornell University.

Juan E. Saavedra is a research economist at the Dornsife Center for Economic and Social Research at the University of Southern California and a faculty research fellow of the National Bureau of Economic Research.

Miguel Urquiola is professor of economics and international affairs at Columbia University and a research associate of the National Bureau of Economic Research.

For useful comments, we thank Joseph Altonji, Caroline Hoxby, Kevin Stange, and participants of the NBER Conference on Productivity in Higher Education. All remaining errors are our own. For acknowledgments, sources of research support, and disclosure of the authors' material financial relationships, if any, please see http://www.nber.org/chapters/c13877.ack.

1. Other countries, such as Chile and Colombia, have similar initiatives. These are relevant in view of evidence that, at least in some cases, college identity can have a causal impact on graduates' earnings (e.g., Hoekstra 2009; Saavedra 2009; Dale and Krueger 2014; and MacLeod et al. 2015). This finding is not universal; see Stange (2012) for contrasting findings among community colleges.

college productivity relate to each other—whether colleges that improve student earning also improve their learning.

This is the first study to simultaneously analyze system-wide measures of the earning and learning productivity of colleges. We use data from the country of Colombia to arguably improve on the measures in the literature to date. Our detailed administrative records provide the earnings of nearly all graduates in the country upon labor market entry. With these data, we can control for a measure of ability—performance on a national standardized admission exam—and for characteristics related to students' socioeconomic backgrounds. Further, the Colombian setting allows us to propose and implement measures of college productivity in the learning dimension, as all graduates are required to take a national college exit exam. In measuring learning performance, we can similarly control for individual characteristics and precollege ability. In particular, some components of the college exit exam are also assessed in the entrance exam, enabling us to implement an approach akin to those commonly used in the teacher value-added literature.[2] In short, our earning and learning measures may not fully isolate college value added, but they have advantages relative to measures previously used in the context of measuring college productivity.

We then show how these measures of college productivity relate to each other and to characteristics of colleges' entering classes. This yields three findings. First, we find that measures of college productivity on earning and learning are far from perfectly correlated. This implies that college rankings based on earnings differ from those based on learning; in other words, the colleges that seem to add most to students' postgraduation earnings are not necessarily the ones that add most to their measured learning.[3] For instance, we find that on average the top private schools seem to do relatively better on earning, whereas the top public institutions perform better on learning.

Second, the measures of earnings productivity are significantly more correlated with students' socioeconomic status (SES) than the learning measures; not surprisingly, earnings are also more correlated with colleges' tuition levels. This leaves open the possibility that learning measures do a better job of isolating a college's contribution to students' human capital even when one focuses on early career earnings, as we do. For example, learning may be more easily influenced by factors that colleges can control

2. See, for instance, Chetty, Friedman, and Rockoff (2014). Our empirical approach is also closely related to the one in Saavedra and Saavedra (2011), discussed below.

3. With learning measures, a concern often arises regarding whether these capture anything that the market and therefore students actually value. In the Colombian setting, student performance on the field-specific component of the exit exam is predictive of student wages, even after controlling for students' performance on the admission exam, college reputation, and socioeconomic status.

directly, such as teaching, as opposed to factors such as parental connections and signaling. Consistent with this, we show that a college's measured performance can vary substantially depending on whether earnings are measured right after graduation or later in workers' careers. This illustrates that colleges have only partial control over the earnings paths of their graduates.

Our third finding is that a college's ranking under the earning and learning measures can differ depending on its mix of majors. We show that the earning measures tend to favor majors related to engineering, business, and law; more specialized majors, such as those in fine arts, education, and social/natural sciences, are relatively higher ranked under learning metrics. Thus if measures such as the ones we calculate became salient, they could lead colleges to make strategic choices on which majors they offer.

Taken together, our findings imply that the design of accountability systems may influence colleges' relative performance—and therefore applicants' school choices—as well as colleges' responses. Policy makers may wish to keep these implications in mind as they begin to release more college performance information to the public.

Our study relates to two strands of work on college productivity: those related to earning and learning. In terms of learning, a variety of standardized tests exist in the United States that could in principle be used to measure student-learning outcomes. These tests include the Measure of Academic Proficiency and Progress (MAPP), the Collegiate Assessment of Academic Proficiency (CAAP), the Collegiate Learning Assessment (CLA), the California Critical Thinking Skills Test (CCTST), the Watson-Glaser Critical Thinking Appraisal, and the Cornell Critical Thinking Tests (Pascarella and Terenzini 2005; Sullivan et al. 2012). However, these tests are not systematically used across the country.

Few studies investigate the extent to which variation in learning value added relates to institutional characteristics. In general, these studies find little systematic relationship between learning growth and institutional characteristics. Arum and Roksa (2011) use longitudinal CLA data from students at 23 US colleges and find no systematic relationship between critical thinking value added and institutional characteristics. The Council for Aid to Education (2013) uses cross-sectional CLA data from students at 158 US colleges to document how colleges exhibit similar growth of critical thinking skills regardless of ownership status, institution size, Carnegie Classifications, or selectivity. Hagedorn et al. (1999) use longitudinal data from students in 23 US colleges taking the CAAP test and find that peer composition modestly influences critical thinking in the first year of college but that its effect fades over an individual's college career. Saavedra and Saavedra (2011) use cross-sectional data from an administration of Australia's Graduate Skills Assessment (GSA) to estimate educational value added in a nationally representative sample of freshmen and seniors at 17 Colombian colleges.[4] After controlling for incoming student characteristics, Saavedra and Saavedra (2011) find that private ownership is related to value added but that measures of college quality—such as resources, selectivity, and reputation—are not.

Our work also relates to a long and growing literature measuring productivity in higher education (e.g., Cooke 1910; Sullivan et al. 2012). For instance, recent system-wide studies from Norway, the United States, and Chile that credibly address selection bias using administrative data find mixed evidence on the labor market payoffs to attending more-selective colleges (Hoxby and Bulman 2015; Hastings, Neilson, and Zimmerman 2013; Kirkeboen, Leuven, and Mogstad 2016). In chapter 2 of this volume, Hoxby uses administrative data to estimate the productivity of all postsecondary institutions in the United States. However, unlike prior studies that credibly address issues of selection bias, Hoxby is able to estimate both per-pupil lifetime earnings outcomes and per-pupil costs for each institution. She finds that more-selective colleges produce higher lifetime earnings but do so at a proportionally higher cost. As a result, among the 1,000 most-selective US colleges, there is little relationship between earnings value added per unit of input and institutional selectivity.

The remainder of the chapter is structured as follows. Section 4.2 presents background on the Colombian higher education sector, and section 4.3 describes our data and sample. Section 4.4 discusses the computation of our productivity measures, and section 4.5 presents results. Section 4.6 concludes with broader implications.

4.2 Background

This section provides background on Colombia's higher education system.

4.2.1 Access to College

In the past decades, Latin American countries have seen a marked expansion in access to secondary and tertiary education. Access to the latter has actually risen faster, although from a lower base. As figure 4.1 shows, the gap between secondary and tertiary enrollment in the region narrowed from 60 percentage points in 1996 to 50 percentage points by 2013. By 2013, about 43 percent of the population had enrolled in some type of tertiary education. The evolution in Colombia has generally mirrored that in the rest of the region, although the gap between both types of enrollment has remained stable at about 45 percentage points.[5]

4. The GSA, which is most similar to the CLA in the United States, measures four general skill domains: critical thinking, problem solving, writing, and interpersonal skills.

5. The salient difference between Colombia and the rest of the region is that secondary rose faster initially and then stagnated. Tertiary enrollment trends are essentially identical in Colombia and the region as a whole.

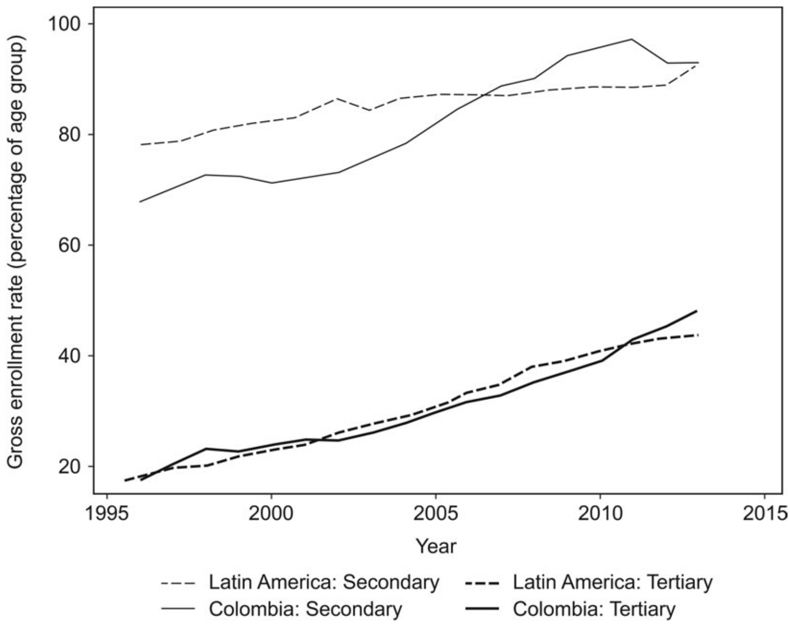

Fig. 4.1 Enrollment trends in Colombia and Latin America

Notes: The data come from the World Bank indicators (http://databank.worldbank.org, consulted on April 7, 2016). The figure plots gross secondary and tertiary enrollment rates for Colombia and the corresponding aggregate for Latin America as a whole. Gross secondary enrollment rate is the number of individuals enrolled in secondary school as a fraction of the total number of individuals 12 to 17 years of age. Gross tertiary enrollment rate is the number of individuals enrolled in tertiary education as a fraction of the total number of individuals 18 to 24 years of age.

Throughout the region, there are constraints for further tertiary expansion. In the case of Colombia, these partially reflect market structure. Private and public providers coexist, and while public colleges are significantly subsidized, their capacity is strained. Table 4.1 shows that public colleges account for 23 percent of institutions but 52 percent of total tertiary enrollments.[6]

There is little regulation on the entry of tuition-charging, unsubsidized private providers, and these generally offer few financial aid opportunities.[7] As a result, private colleges represent 77 percent of all institutions but only 48 percent of total enrollment.

6. Throughout this chapter, we use the term *colleges* to refer to both universities and technical institutions, as depicted in table 4.1.

7. Technically there are no for-profit colleges in Colombia. It is widely perceived, however, that many nonselective private colleges are de facto for-profit, as their owners are the residual claimants of excess revenue typically distributed through wages, rental charges, investments, and so on. In this sense, the situation resembles that which has existed during certain periods in other countries with large private college sectors, such as Chile.

Table 4.1 **Colombian higher education market structure**

	Institutions			Enrollment		
	Public	Private	Total	Public	Private	Total
Universities	47	142	189	495,855	799,673	1,295,528
	0.17	0.53	0.70	0.25	0.40	0.65
Technical schools	15	65	80	524,007	163,886	687,893
	0.06	0.24	0.30	0.27	0.08	0.35
Total	62	207	269	659,142	601,744	1,983,42
	0.23	0.77	1.00	0.52	0.48	1.00

Notes: Calculations based on the Colombian national higher education information system (SNIES) for 2013, the last year with data available. Enrollment data only include undergraduate students. The category "universities" combines universities and university institutes. "Technical schools" combines technical institutes, technological institutes, and the National Job Training Agency (SENA).

Colleges and universities are also geographically concentrated: 50 percent are in Colombia's three largest cities, which account for 26 percent of the population. Bogotá, the capital, is home to 31 percent of all colleges. About 75 percent of tertiary students attend a college in the city of their birth (Saavedra and Saavedra 2011). Furthermore, in our data, roughly 70 percent of graduates get their first formal sector job in the same municipality where they attended college. This suggests an important role for local labor markets in our analysis—part of the benefit of attending a college in an urban area, for example, is that it may increase access to high-wage jobs.[8]

4.2.2 College Entrance Exam

To apply to college, Colombian students must take a standardized entrance exam called the Icfes, which is administered by a government agency.[9] The Icfes is generally analogous to the SAT in the United States, but it is taken by the vast majority of high school seniors regardless of whether they intend to apply to college.[10] The Icfes also plays a larger role in admissions

8. We also find a positive relationship between college selectivity and the probability that an individual stays in the area upon graduation; a one-standard-deviation increase in a college's mean entrance exam score raises the likelihood that a graduate works in the municipality where she attended college by six percentage points.

9. Icfes stands for Institute for the Promotion of Higher Education, the former acronym for the agency that administers the exam. The Colombian Institute for Educational Evaluation, as it is now called, was created in 1968 and is a state agency under the authority of the national Ministry of Education. The Icfes exam is now known as Saber 11, reflecting the fact that students usually take it in the 11th grade. We use the name Icfes to match the designation during the period covered by our data.

10. Angrist, Bettinger, and Kremer (2006) and our personal communications with the Colombian Institute for Educational Evaluation suggest that more than 90 percent of high school seniors take the exam. The test-taking rate is high in part because the government uses Icfes exam results to evaluate high schools.

in Colombia than the SAT does in the United States. In addition to using it as an application requirement, many schools extend admission offers based solely on students' entrance exam performance. Others consider additional factors such as high school grades while heavily weighting the Icfes, and a handful administer their own exams. Applications and admissions are major specific; students apply to a college/major pair.

The Icfes tests multiple subject areas, including biology, chemistry, English, math, reading/language arts, social science, philosophy, and physics.

4.2.3 College Exit Exam

In 2004, the agency that administers the Icfes introduced, with considerable publicity, new field-specific college *graduation* exams. These exit exams are standardized and administered at every institution that offers a related program.[11] The exams are intended to assess senior students' competencies in fields ranging from relatively academic in orientation (e.g., economics and physics) to relatively professional (e.g., nursing and occupational therapy).

The creation of the exit exams was a major undertaking, as it required coordination among departments in multiple colleges. The stated intent of this effort was to improve quality, transparency, and accountability in the higher education sector. Consistent with this, school-level aggregate scores were made available and have been used by news outlets as part of college rankings.

Field-specific exams became available for most majors in 2004, with several majors receiving field exams in subsequent years. A few fields, such as political science, anthropology, history, and philosophy, never received a corresponding field-specific exam. In part because of this, for the first few years, taking the exit exam was optional, although the majority of students in tested fields took the exam. This changed in 2009, when the exit exam became a graduation requirement for all students. A generic test was introduced for majors that did not previously have a field-specific exam. In addition, from 2009 onward, the exam included several common components in subjects such as English and reading comprehension, which were taken by all students regardless of their field.

Increasingly, colleges and students use results on the college exit exam as a signal of ability. For example, students may report whether they obtained a top score nationally or their score in comparison to the university or the national average. Some universities use exit exam results in admissions to graduate programs, and the Colombian Student Loan Institute offers a postgraduate study credit line (of up to $16,000) exclusively to the best 10 nationwide scorers. In addition, every year the Colombian president and education minister publicly recognize the individuals with the top 10 scores

11. These tests were initially labeled Ecaes, which stands for *Exámenes de Calidad de Educación Superior*—that is, higher education quality exams. They are now called *Saber Pro*.

in each field. Anecdotally, the best scorers receive job offers based on public knowledge of their test scores, and MacLeod et al. (2015) provide evidence that the exit exams affect graduates' labor market earnings.

4.3 Data and Sample

This section describes our sources of data and the sample we use for our analysis.

4.3.1 Data

We use individual-level administrative data sets from three sources:

1. The Colombian Institute for Educational Evaluation, which administers the college entrance and exit exams, provided records for both tests. This includes scores for all high school seniors who took the entrance exam between 1998 and 2012 as well as college exit exam scores for all exam takers from 2004 to 2011.
2. The Ministry of Education provided enrollment and graduation records for students entering college between 1998 and 2012. These include each individual's college, program of study, and enrollment and graduation dates. These data cover roughly 90 percent of all college enrollees; the ministry omits a number of smaller colleges due to poor and inconsistent reporting.
3. The Ministry of Social Protection provided monthly earnings records for formal sector workers during 2008–12. These come from data on contributions to pension and health insurance funds.

We link these data sources using student names, birthdates, and national ID numbers. The resulting data set includes students from nearly all colleges in Colombia with information on their entrance exam scores and, if applicable, their exit exam performance and formal labor market earnings.

4.3.2 Sample

We select a sample that allows us to cleanly compare measures of college performance on earning and learning. Specifically, we restrict our sample to graduates who satisfy two important criteria. First, we include only students who took the college exit exam in 2009–11. As noted above, the exit exam was voluntary prior to 2009, so we exclude pre-2009 exam takers to limit selection into taking the exam. Second, we include only graduates for whom we observe initial labor market earnings. Since students typically take the exit exam one year before graduating, this means that we include only 2010–12 graduates with earnings observed in their graduation year.

This restriction sets aside other outcomes of interest to students and policy makers, such as graduation rates. In Colombia, as in the United States, the probability of graduating tends to increase with the selectivity of the

Table 4.2 **Sample and college types**

College type	No. of colleges	No. of grads	Admit rate	Annual tuition ($)	Mother went to college	Entrance exam percentile
Public (most selective)	12	15,642	0.20	369	0.42	0.82
Public (medium selective)	24	13,228	0.55	509	0.29	0.67
Public (least selective)	12	6,063	0.87	535	0.23	0.59
Top private	8	9,653	0.64	2,584	0.90	0.90
Other private (high cost)	51	19,229	0.82	1,696	0.59	0.72
Other private (low cost)	50	17,489	0.86	1,079	0.31	0.63
Total	157	81,304	0.65	1,134	0.46	0.72

Notes: Admission rate data are from Colombian national higher education information system (SNIES) and average over 2007–12. Tuition data are from the exit exam records, which report each exam taker's tuition in the previous year in six categories. We compute the average across all students using the midpoint of each category and convert to US dollars using 2012 exchange rates. Entrance exam percentiles are relative to all exam takers in each year, including those who did not attend college.

college a student attends. To the extent that graduation rates are highly correlated with college-level earnings, restricting the sample to graduates is unlikely to significantly change our findings.[12]

In addition to these restrictions, we drop individuals with missing values on any of the other variables we use, including entrance exam scores, high school of origin, mother's education, and previous year's tuition.[13] This ensures that all performance measures calculated below are based on the same set of individuals. Lastly, to obtain reasonable precision for each of our performance measures, we restrict our analysis to colleges that have at least 50 graduates satisfying the above criteria.

The resulting sample includes approximately 81,000 graduates from 157 colleges. This is much larger than samples available in previous studies that use longitudinal data to compute college performance measures (e.g., Klein, Steedle, and Kugelmas 2010). The last row in table 4.2 presents summary statistics on our sample.

4.3.3 College Categorization

Table 4.2 additionally categorizes colleges into six types with the aim of providing a useful portrayal of the college market in Colombia. The top

12. For example, we find that the correlation between mean college earnings in samples with and without college dropouts is 0.9.

13. The entrance exam underwent a major overhaul in 2000, so we also exclude the small number of students who graduated in 2010–12 but took the entrance exam prior to 2000. Since one of our learning outcomes below is a student's English exit exam score, we additionally drop the fewer than 1 percent of students who took the French or German entrance exams, which were offered until 2006, rather than the English exam.

three rows separate public colleges into three groups based on quartiles of their admission rates. We define the most-selective public colleges as those in the quartile with the lowest admission rates and the least-selective colleges as those in the highest admission rate quartile. Medium-selective colleges are those in the middle two quartiles.[14] Table 4.2 shows that the most-selective public colleges admit 20 percent of their applicants on average, while the least-selective are essentially open enrollment.[15]

Selectivity defined by admission rates has limited usefulness in categorizing private colleges in Colombia, as most private colleges admit nearly all of their applicants. Instead, sorting into private colleges is defined more strongly by the tuition rates they charge. We therefore define "top private" colleges as those few that are actually selective—that is, they reject some of their applicants—and in which average annual graduate tuition exceeds the equivalent of about $2,500.[16] This definition picks out eight colleges that represent the most elite private schools in the country. We divide the remaining private institutions—which we label "other private"—into two types based on the average tuition payments reported by their graduates. We define high-cost private colleges as those above the median tuition and low-cost colleges as those below.[17]

Average annual tuition varies significantly across private college types, with a mean of roughly $1,000 at low-cost private colleges. Average tuition is significantly lower at all public college types, as they offer substantial discounts to low-SES students.

The last two columns of table 4.2 summarize the socioeconomic and academic backgrounds of graduates from each college type. Graduates from private colleges are much more likely to have mothers with a college education; for instance, 90 percent of students at top private colleges do so. Academic preparation, as defined by each student's entrance exam percentile in the full distribution of test takers, also varies starkly across college types. Average entrance exam performance is at the 82nd percentile at the most-selective public colleges and the 90th percentile at top private schools. Graduates from the lowest college types, both public and private, have average entrance exam scores near the 60th percentile.

14. We use quartiles rather than terciles to define these three groups to provide more detail on colleges at the extremes of the distribution.

15. Note that nonselective colleges often have admission rates that are slightly less than one in table 4.2. This reflects that students may fail to follow all application procedures or may withdraw their applications before admission.

16. Specifically, we use a four million peso cutoff for top private colleges, and we define their selectivity using a 2002 report from the Colombian Institute for Educational Evaluation entitled *Estadísticas de la Educación Superior*. Selective private colleges are those for which the number of applicants exceeded the number of offered slots, according to this report.

17. We note that we do not use an institution's level of training (university or technical, as in table 4.1) to define these six college categories. We find that this distinction provides little additional information on average college characteristics conditional on the categories defined by financing, selectivity, and tuition.

We use the sample and college categorization in table 4.2 for our analysis of college performance measures below.

4.4 Measures

This section describes the outcome variables we use and the measures we employ to approximate college earning and learning productivity.

4.4.1 Earning and Learning Variables

Our earnings variable is log average daily formal labor market earnings, which we calculate by dividing base monthly earnings for pension contributions by the number of employment days in each month and averaging across the year. We use earnings in the year of each student's graduation (2010–12) and demean earnings in each year.

Our learning variables are based on students' scores on the college exit exam. During the exam years we analyze (2009–11), this test included a field-specific component related to a student's major (e.g., economics or mechanical engineering) as well as several components taken by all students. We focus on three of these: (1) the field-specific score, (2) a reading common component score, and (3) an English common component score.

These components have different strengths and weaknesses in measuring college productivity. The field exit score, because it typically reflects each student's college major, provides arguably the best measure of the material studied in college. However, in general, there is no direct analog on the entrance exam. The English component of the exit and entrance exams are very similar and thus well placed to measure progress, but English proficiency may be less directly related to college productivity. Since the exit and entrance exams include a similar but not identical reading / language arts component, the reading component lies arguably in the middle of the comparability and relevance spectrums.

Using these three exit exam scores, we calculate each student's percentile relative to all other students in our sample in the same exam field and cohort. We use exam score *percentiles* because the entrance and exit exams are not on a common scale and thus cannot measure growth in human capital. As a result, our learning measures will capture a college's relative rather than absolute performance. The same caveat applies to our earning measures, since we do not observe a precollege measure of earnings.

4.4.2 Calculation of Productivity Measures

We use four procedures to measure learning and earning performance. Some of these procedures are simple and require less-detailed information, and thus they correspond to measures that may be more commonly reported in the media or easier for policy makers to compute. Other procedures use comprehensive information on students' backgrounds and align more closely with "value-added" methods employed in other areas of economic research. These four procedures, which we describe in the following subsections, allow us to explore the sensitivity of our results to different data requirements and methodologies.

4.4.2.1 Raw Means

Our first performance measure is the average log earnings, or the average exit exam percentile, at each college:

$$\bar{\theta}_c = E\{y_{ic} \mid i \in c\}, \tag{1}$$

where y_{ic} is either outcome for individual i who graduated from college c. We label $\bar{\theta}_c$ the *raw means* measure, as it implements the simplest and least data intensive of our four procedures. Note that it does not adjust for differences across colleges in incoming student characteristics—that is, in the student "inputs" to college production.

4.4.2.2 Entrance Exam Residuals

Our second performance measure adjusts for differences in college inputs by controlling for students' entrance exam performance. We do this through an individual-level regression of the following form:

$$y_{ic} = \beta' t_i + \tilde{\theta}_c + \tilde{\varepsilon}_{ic}, \tag{2}$$

where t_i is a vector of student i's entrance exam percentiles on eight components, which include reading/language arts and English.[18] We decompose the residual from this regression into a school-specific term, $\tilde{\theta}_c$, and an idiosyncratic component, $\tilde{\varepsilon}_{ic}$. Our second college productivity measure, which we call *entrance exam residuals*, is the $\tilde{\theta}_c$ coefficient from equation (2).

4.4.2.3 Entrance Exam + SES Residuals

Our third performance measure is closely related to the second, but we include additional controls for students' socioeconomic background in regression (3):

$$y_{ic} = \beta' t_i + \gamma' x_i + \hat{\theta}_c + \hat{\varepsilon}_{ic}, \tag{3}$$

where x_i represents dummies for four categories of mother's education (primary, secondary, vocational, university), which are fully interacted with dummies for each of the approximately 6,000 high schools in our sample. The *entrance exam + SES residuals* measure for each college is the $\hat{\theta}_c$ coefficient from this regression. This coefficient is identified from variation in college attendance across students with the same high school and mother's educa-

18. The other components are biology, chemistry, math, social sciences, philosophy, and physics. As with the exit exam scores, we convert entrance exam scores into percentiles within each exit exam field and cohort.

tion combination. This measure is most analogous to benchmark "value-added" models in other work in economics, which control for a broad array of initial individual characteristics.

4.4.2.4 College-Level Residuals

Our fourth performance measure controls for college-level characteristics in addition to individual-level characteristics. This is motivated by research in Altonji and Mansfield (2014), which shows that in the estimation of group-level treatment effects, including group average characteristics can in some cases control for between-group sorting on unobservable individual traits.

We control for both individual and college characteristics using a two-step procedure. First, we estimate equation (3) and calculate residuals y_{ic}^* from the individual characteristics only. That is, we calculate $y_{ic}^* = y_{ic} - \hat{\beta}'t_i - \hat{\gamma}'x_i$, where $\hat{\beta}$ and $\hat{\gamma}$ are the estimated coefficients from regression (3).[19]

Second, we calculate the mean value of y_{ic}^* for each college, $y_c^* = E\{y_{ic}^* | i \in c\}$, and estimate the following college-level regression:

$$y_c^* = \beta' t_c + \gamma' x_c + \theta_c, \tag{4}$$

where t_c is the vector of college mean percentiles for each of the eight entrance exam components, and x_c is the fraction of students with a college-educated mother at college c.[20] The *college-level residuals* measure is the residual from regression (4), θ_c. As we discuss below, this measure has properties that differ from those of measures based on individual residuals because it is uncorrelated with college mean entrance scores by construction. Altonji and Mansfield (2014) note that under certain conditions, the variance in θ_c also serves as a lower bound to the true variance of college treatment effects, in part because these treatment effects are likely correlated with t_c and x_c.

4.4.3 Correlations of Productivity Measures with Inputs

For our earnings and each of our three learning variables, the above procedures yield four separate productivity measures—in short, 16 measures for each college in our sample. We normalize each of these to have mean zero and standard deviation one across the 157 colleges. This normalization is convenient because it makes the coefficient from a linear regression of one measure on another equal to their pairwise correlation coefficient.

To provide context on these measures, we show how they relate to a college characteristic that is, in principle, easily observable to many agents: colleges' mean entrance exam score. We begin with a graphical exposition using only

19. Note that this first-step regression also includes group-level (i.e., college) fixed effects, as is common in the teacher value added literature (Chetty, Friedman, and Rockoff 2014).

20. Observations in regression (4) are weighted by the number of graduates from each college. All college-level computations in this chapter use these same weights.

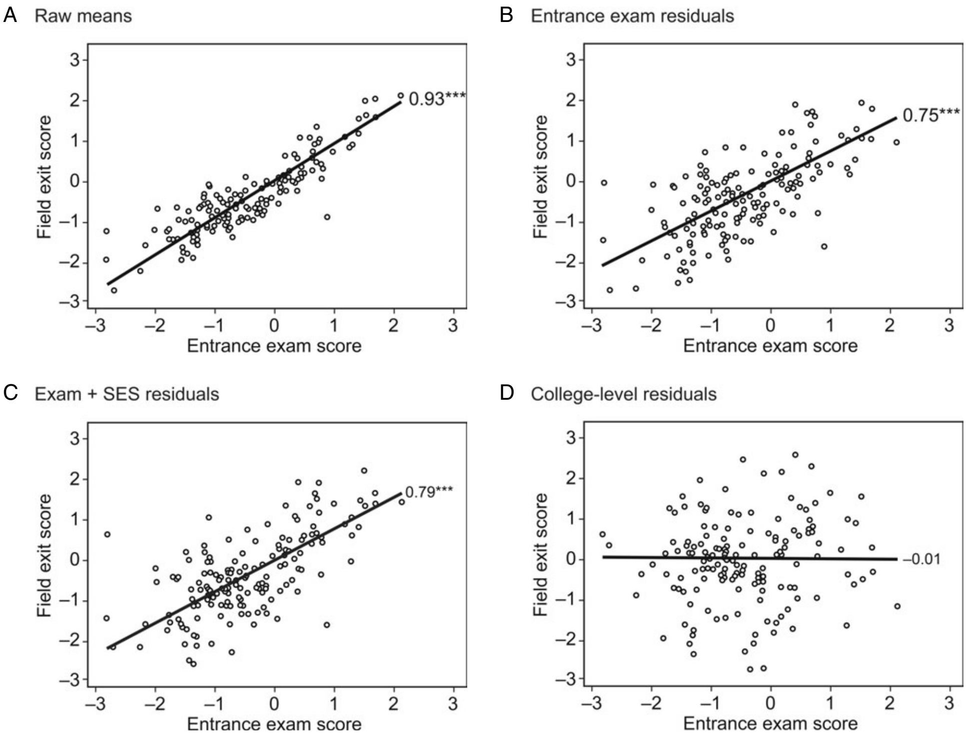

Fig. 4.2 Illustration of field-specific learning measures

Notes: Small circles represent the 157 colleges in our sample. The solid line depicts the linear relationship between the learning measures and college mean entrance scores, with colleges weighted by their number of graduates. Asterisks on the slope coefficients indicate statistical significance with robust standard errors.

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

one learning outcome: the field-specific exam score. The four panels of figure 4.2 depict our four measures for this outcome. The gray circles are the 157 colleges in our sample. The vertical axis in each panel represents the learning performance under each measure, while the horizontal axis depicts the raw mean entrance exam score at each college.[21] The solid line depicts the linear relationship between these two measures, with the slope indicated on the graph.

Panel A shows that the correlation between a college's raw mean field exit score ($\bar{\theta}_c$ from equation [1]) and its mean entrance exam score is 0.93. Panel B shows that controlling for individual entrance exam scores (using $\tilde{\theta}_c$ from equation [2]) reduces this correlation only slightly. Note that while $\tilde{\theta}_c$ ensures

21. Raw mean entrance score is the average percentile across the same eight components included in regressions (2)–(4), also normalized to mean zero and standard deviation one.

Table 4.3 **Correlations with college mean entrance scores**

	Raw means (A)	Entrance exam residuals (B)	Exam + SES residuals (C)	College-level residuals (D)
Field exit score	0.93***	0.75***	0.79***	–0.01
Reading exit score	0.90***	0.59***	0.65***	–0.03
English exit score	0.88***	0.73***	0.71***	–0.04
Log earnings	0.70***	0.63***	0.57***	0.06

Notes: This table displays coefficients from linear regressions of college mean entrance exam scores on each of our 16 learning and earning measures. All regressions have 157 observations with weights equal to each college's number of graduates. Asterisks indicate statistical significance with robust standard errors.

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

that *individual* exit residuals are uncorrelated with *individual* entrance exam scores, it allows *college-level* exit scores to be correlated with *college-level* entrance exam performance. This can arise if other individual characteristics that affect exit exam performance, such as socioeconomic background, also affect the colleges students choose to attend.

Panel C partially addresses this issue by using the entrance exam + SES residual measure ($\hat{\theta}_c$ from equation [3]), which controls for students' observable background. Panel C shows that these controls have little effect on the correlation of the exit field score with college mean entrance exam performance; in fact, the correlation coefficient increases slightly. This illustrates that our individual learning productivity measures may still be correlated with unobservable student characteristics that affect both college choice and exit exam performance.

Panel D illustrates that our last productivity measure, the college-level residual (θ_c from equation [4]), is uncorrelated with college mean entrance exam performance by construction.[22] This addresses the issue that individual characteristics may be correlated with college mean entrance scores (as well as college mean mother's education). However, the college residual measure, θ_c, rules out the possibility that colleges with high mean entrance scores systematically produce better learning outcomes than colleges with low average scores. Rather, this measure is better suited for comparing the performance of colleges with similar inputs as defined by mean entrance scores.

As stated, we have 16 outcome measures in total (log earnings plus three learning measures, each calculated using the procedures in equations [1]–[4]). Table 4.3 displays the correlations of each of these measures with college mean entrance scores. The top row refers to the field exit score and replicates

22. The correlation between the two measures in panel D is not strictly zero because the horizontal axis is the average of the eight entrance exam components, not any individual component from regression (4).

Table 4.4 **Correlations with earning measure**

	Raw means (A)	Entrance exam residuals (B)	Exam + SES residuals (C)	College-level residuals (D)
Field exit score	0.62***	0.45***	0.45***	0.07
Reading exit score	0.58***	0.29***	0.41***	0.16**
English exit score	0.71***	0.62***	0.51***	–0.09

Notes: This table displays coefficients from linear regressions of our earning measures on each of our learning measures. All regressions have 157 observations with weights equal to each college's number of graduates. Asterisks indicate statistical significance with robust standard errors.

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

the correlation coefficients depicted in figure 4.2. The remaining three rows cover the other measures. The manner in which the correlation measures change as one moves across columns is similar across all rows; in other words, the above discussion applies to all of our learning and earning measures. This provides an additional justification for using multiple methods to calculate productivity in examining our key findings below.

4.5 Results

This section presents empirical results related to three questions: (1) How are the earning and learning measures related to each other? (2) How are they related to other factors that influence students' choice of colleges? (3) How do these measures vary with the majors a college offers?

4.5.1 Comparing Learning and Earning Measures

Our first empirical task is to explore how the learning and earning measures relate to each other. Table 4.4 shows the correlation coefficients for each of our three learning measures with our earning measure, where each has been calculated according to the procedure listed in the column.

A simple but important result is that the learning measures are mostly positively related to our earning measure, but far from perfectly so, with correlations ranging from –0.09 to 0.71 across the learning outcomes and the four procedures. The raw mean learning and earning measures are more strongly correlated than those that control for individual characteristics. The college-level residual measures are mostly uncorrelated, with only one correlation coefficient that is statistically different from zero. It is also notable that the English learning measures are generally more correlated with earnings, which may reflect a stronger socioeconomic component to English education relative to the other subjects.

Figure 4.3 depicts the relation between the earning measures (vertical axis) and the field-specific learning measures (horizontal axis). The imperfect cor-

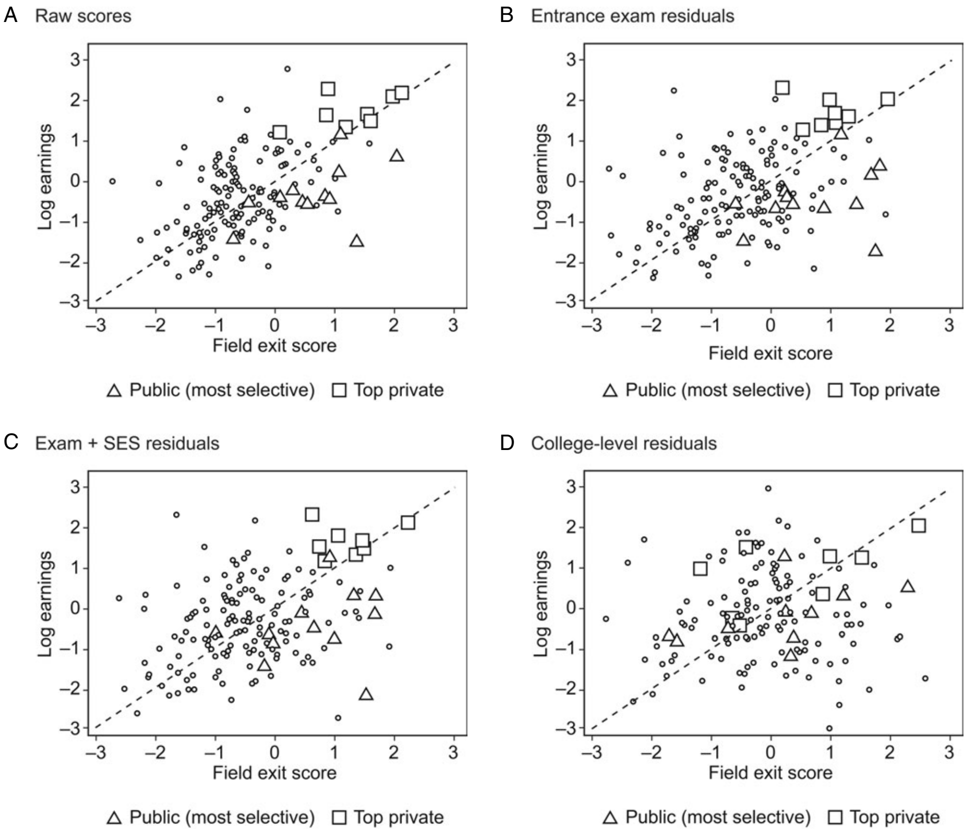

Fig. 4.3 Earning vs. field-specific learning

Notes: Triangles represent the most-selective public colleges as defined in table 4.2. Squares represent top private colleges, and small circles depict all other colleges.

relations from table 4.4 are evident here in the dispersion of the dots, which is most prevalent for the college-level residual method in panel D. Each panel also contains a 45-degree line that represents the boundary between whether colleges appear more productive on the learning or earning measures. In all four panels, the most-selective public colleges (indicated by the triangles) typically lie below the diagonal line—these colleges appear in a more favorable light when we define productivity by learning. Conversely, top private colleges (squares) mostly lie above the 45-degree line; this means that they appear in a more favorable light when performance is defined in terms of earnings. Note that these conclusions hold across all four procedures for calculating productivity despite the different properties discussed above. We also find that they hold when we measure earnings eight years after graduation rather than in the year of graduation. We note, however, that

Table 4.5 **Average institution rank by college type**

College type	Field exit score	Log earnings	Field exit score	Log earnings
	Panel A: Entrance exam residuals		*Panel B: College-level residuals*	
Public (most selective)	0.88	0.58	0.63	0.56
Public (medium selective)	0.54	0.44	0.47	0.57
Public (least selective)	0.26	0.20	0.45	0.48
Top private	0.89	0.95	0.44	0.68
Other private (high cost)	0.63	0.70	0.59	0.51
Other private (low cost)	0.36	0.49	0.42	0.58

Notes: This table displays percentile ranks of colleges using the measures listed in the column header. We sort all colleges according to each measure and then calculate average ranks within the college types depicted in table 4.2. Averages are weighted by each college's number of graduates.

this comparison requires that we calculate earning and learning measures using different samples, as we discuss in further detail below.

Table 4.5 elaborates on this point by presenting the average institution rank that arises from the use of learning or earning measures. Specifically, we sort colleges according to each measure and calculate their percentile rank among the 157 schools. We then compute the average rank in each of the six college types defined in table 4.2. We repeat this calculation for the field-specific learning measures and the earning measures from the entrance exam residual method (panel A) and the college-level residual procedure (panel B). For instance, using the field exit score and individual entrance exam residuals, the most-selective public colleges have an average rank at the 88th percentile, while the average rank of a top private college is in the 89th percentile.

The main conclusion from table 4.5 is that public colleges receive higher rankings from the learning measures than from the earning measures. Conversely, private colleges are relatively higher ranked using earnings. This finding holds for all college categories using the individual-level measures. It also holds for most categories under the college-level measures, though the result is flipped for middle-ranked public and private institutions.

The different measures can thus lead to starkly different conclusions about colleges' relative productivity. In panel A, for example, high-cost private colleges are ranked higher on average than the most-selective public colleges using earnings, but their average rank is 25 percentile points lower using the learning measure. As discussed above, comparisons of colleges with different mean entrance scores are more complicated under the college-level residual method of panel B. Nonetheless, a similar conclusion applies to the relative rankings of the most-selective public colleges and top private colleges, which have similar mean entrance scores (see table 4.2). Top private colleges receive higher ranks under the earning measure, while selective public colleges appear more favorably when one uses the learning measure.

The ranking differences between public and private institutions are consistent with the hypothesis that these colleges add value in different dimensions. For example, students at private colleges may benefit more from peer and alumni networks in the labor market. Conversely, public colleges typically offer a more diverse set of majors, which could allow for better sorting of students into their fields of comparative advantage. An alternative hypothesis is that these colleges vary in the types of students they attract and that these differences are correlated with students' earning and learning potential. We consider this possibility in the next section.

4.5.2 Correlations with Other College Characteristics

The fact that the learning and earning measures are not perfectly correlated suggests that they likely have different relationships with other student and college characteristics. In this section, we explore how learning and earning productivity are related to two other factors that influence students' college choice. We first consider socioeconomic status as defined by whether a student's mother attended college. We then consider a proxy for student demand: each graduate's annual tuition in the prior year.

For both the SES and tuition variables, we follow the same procedures described in section 4.4.2 to compute college averages. This yields measures of college mean SES and college mean tuition corresponding to the raw means, entrance exam residuals, entrance exam + SES residuals, and college-level residuals methods. Note that we do not present the SES measures from equation (3), as this method includes SES controls also defined by mother's education. Similarly, we exclude the SES variables (x_i and x_c) from equation (4) when we calculate the college-level residual measures for figure 4.4 and table 4.6 below; this allows us to compare their correlations with mother's education. As above, we normalize each measure to mean zero and standard deviation one across the sample of 157 colleges.

Figure 4.4 displays the correlations of SES with the field-specific learning measures and the earning measures. In all cases, the earning measures are more strongly correlated with SES than the learning measures, though the difference between the two is not statistically different from zero using raw means.[23]

Table 4.6 presents these correlations for all our learning and earning measures. The top panel displays the correlation of the measures with college mean SES, while the bottom panel displays the difference between each learning measure and the earning measure. In nearly all cases, the learning measures are less correlated with SES than the earning measures, and this difference is statistically significant using the two residual methods (columns B and C). The only exceptions arise with two of the English learning

23. The same patterns arise when we measure earnings eight years after graduation rather than in the year of graduation.

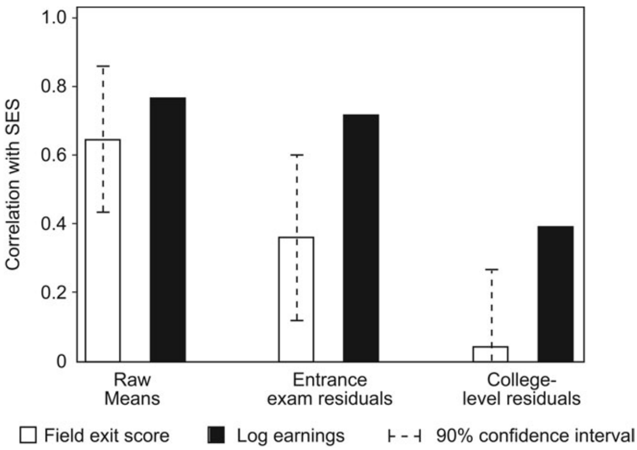

Fig. 4.4 Correlations with SES

Notes: White bars depict the correlations of our SES measures with our field-specific learning measures (the first row in table 4.6). Black bars show the correlation of our SES measures with our earning measures (the fourth row in table 4.6). Dashed lines are 90 percent confidence intervals using robust standard errors. We exclude the x_i and x_c variables in calculating the college-level residual measures for this figure (see equation [4]).

Table 4.6 Correlations with SES

		Raw means (A)	Entrance exam residuals (B)	College-level residuals (C)
Correlations	Field exit score	0.65***	0.36***	0.04
	Reading exit score	0.59***	0.16	0.08
	English exit score	0.83***	0.75***	0.20***
	Log earnings	0.77***	0.72***	0.39***
Differences from earnings	Field exit score	–0.12	–0.36**	–0.35**
	Reading exit score	–0.18	–0.56***	–0.31**
	English exit score	0.07	0.03	–0.19*

Notes: The top panel displays coefficients from linear regressions of SES (defined by mother's education) measures on each of our learning and earning measures. All regressions have 157 observations with weights equal to each college's number of graduates. The bottom panel shows the difference between each of the learning coefficients and the earnings coefficient. We exclude the x_i and x_c variables in calculating the college-level residual measures for this table (see equation [4]). Asterisks indicate statistical significance with robust standard errors.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

measures, which, as noted above, may be more influenced by socioeconomic background than the field and reading scores.

Table 4.7 is analogous to table 4.6, but it presents the correlations of learning and earning measures with tuition rather than with SES. The same pattern holds; the learning measures are in all cases substantially less correlated with graduates' average tuition than the earning measures.

Table 4.7 **Correlations with tuition**

		Raw means (A)	Entrance exam residuals (B)	Exam + SES residuals (C)	College-level residuals (D)
Correlations	Field exit score	0.32*	0.16	0.24	0.02
	Reading exit score	0.24	–0.05	0.10	0.02
	English exit score	0.59***	0.63***	0.54***	–0.03
	Log earnings	0.67***	0.67***	0.60***	0.27***
Differences from earnings	Field exit score	–0.36*	–0.52***	–0.36*	–0.26*
	Reading exit score	–0.44**	–0.72***	–0.50***	–0.25**
	English exit score	–0.08	–0.04	–0.06	–0.31**

Notes: The top panel displays coefficients from linear regressions of tuition (defined as in table 4.2) measures on each of our learning and earning measures. All regressions have 157 observations with weights equal to each college's number of graduates. The bottom panel shows the difference between each of the learning coefficients and the earnings coefficient. Asterisks indicate statistical significance with robust standard errors.

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

The results in tables 4.6 and 4.7 are consistent with a college's earning performance being a stronger driver of its demand than its learning performance. Though none of our measures may fully isolate college value added, these findings suggest that learning measures may be less related to other factors that affect student outcomes, which may not be observable in all contexts. This is particularly relevant if learning outcomes are ultimately under greater control on the part of colleges than earning results. In particular, earning measures, unlike those based on learning, have a natural dynamic component in the years after students enter the labor market. Throughout our analysis, we have used earnings measured in the year of each student's graduation, but there are both conceptual and data-related reasons why earnings might be measured later in a worker's career.

To explore the potential implications of the timing of earning measurement, we use a different sample than in the above analysis that allows us to measure earnings later in workers' careers. Specifically, we include 2003–12 graduates with earnings observed in 2008–12. With this sample, we can observe earnings between zero and eight years of potential experience, defined as earnings year minus graduation year.[24] Note that this analysis relies on cross-cohort earning comparisons, meaning that the sample differs across experience levels.

The earning measures analyzed above normalize measures to have a constant standard deviation. Before computing such measures, we display the raw data in figure 4.5. This figure shows average log earnings at the 128 col-

24. We can actually observe a ninth year of potential experience using 2012 earnings for 2003 graduates, but these ninth-year measures are noisy because they come from only a single cohort and year.

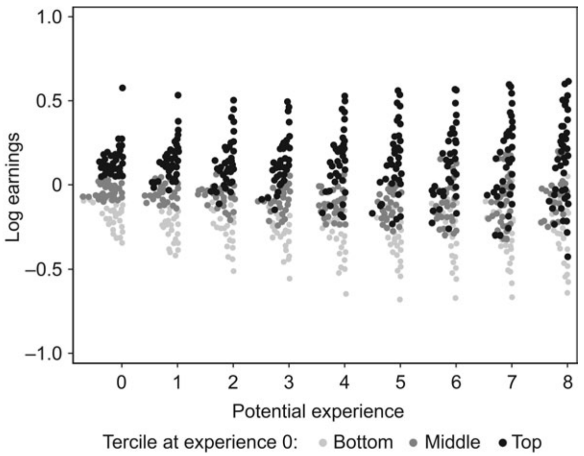

Fig. 4.5 Log earnings by potential experience

Notes: The sample includes 2003–12 graduates with earnings measured at 0–8 years of potential experience, defined as earnings year minus graduation year. Dots depict average log earnings at the 128 colleges in our sample with at least 10 earning observations for each experience level. Log earnings are demeaned by graduation year and experience. We group colleges into three terciles based on experience zero earnings and add horizontal spacing to improve visibility.

leges that we observe at all experience levels, where we demean earnings by graduation cohort and year. We group the 128 colleges into three terciles of different shadings based on their average earnings at experience zero and hold these terciles constant for all experience levels.

Figure 4.5 shows that the variance in average earnings across colleges increases with worker experience, a result first documented by MacLeod et al. (2015). At experience zero, nearly all colleges have average earnings within 30 percent of the mean, while many colleges lie outside this range after eight years. Further, there is substantial mixing of the terciles over time such that some colleges with low initial earnings ultimately have mean earnings above those of top tercile colleges. These two findings show that both the magnitude and the ordering of differences in earnings across colleges can change substantially depending on when one measures earnings.

Table 4.8 formalizes this point by showing how the correlation of earnings with initial measures of college productivity evolves with worker experience. For this table, we calculate earnings measures analogous to those above using the same students and colleges as in figure 4.5. Panel A displays the raw mean measures (from equation [1]), and panel B depicts residuals from a regression on college mean entrance exam scores (equation [4]).[25]

25. We do not present individual entrance exam residual measures in table 4.8 because we do not observe the full vector of individual exam scores for all 2003–12 graduates. For this reason, we also do not use a first-step regression to net out individual characteristics in calculating the college-level measures (see section 4.4.2.4).

Table 4.8 **Correlations by potential experience**

		Panel A: Raw means		*Panel B: College-level residuals*	
		Log earnings at exp. 0	Field exit score	Log earnings at exp. 0	Field exit score
Correlations	Log earnings at exp. 0	1.00***	0.44***	1.00***	0.04
	Log earnings at exp. 2	0.93***	0.63***	0.92***	0.16*
	Log earnings at exp. 4	0.88***	0.68***	0.85***	0.17*
	Log earnings at exp. 6	0.83***	0.70***	0.78***	0.15*
	Log earnings at exp. 8	0.76***	0.69***	0.67***	0.10
Differences from earnings at exp. 0	Log earnings at exp. 2	–0.07*	0.20	–0.08**	0.11
	Log earnings at exp. 4	–0.12***	0.25*	–0.15**	0.12
	Log earnings at exp. 6	–0.17***	0.26**	–0.22***	0.11
	Log earnings at exp. 8	–0.24***	0.26**	–0.33***	0.06

Notes: The top panel displays coefficients from linear regressions of earning measures at different experience levels on experience zero earning measures and the field-specific learning measures. The sample is the same as that for figure 4.5. All regressions have 128 observations with weights equal to each college's number of graduates. The bottom panel shows the difference between each of the experience 2–8 coefficients and the experience zero earnings coefficient. Asterisks indicate statistical significance with robust standard errors.

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

The top panel of table 4.8 shows the correlation of earnings measured at different experience levels with earnings at experience zero and with our field-specific earnings measure from above. The bottom panel shows how the correlations at each experience level change relative to those at experience zero. The results show that the correlation of earning measures with initial earnings declines substantially over time and that this holds for both the raw and residual methods. By contrast, the earning measures become *more* correlated with the field-specific exit scores over time, though the differences are not significant for the residual measures.

The main takeaway from figure 4.5 and table 4.8 is that one can arrive at very different conclusions for a college's earning productivity depending on when one measures earnings. This highlights the fact that colleges do not have complete control over their graduates' earnings, which also depend on the postschooling actions of workers and employers. This leaves open the possibility that learning measures do a better job of isolating a college's contribution to students' human capital.

4.5.3 Learning and Earning across Majors

Our final set of results concern one way in which colleges might be able to influence these productivity measures: their choice of which majors to offer. To explore how our measures vary across majors, we repeat the four procedures described in section 4.4.2, but instead of calculating productiv-

Table 4.9 **Average institution/major rank by major area**

		Panel A: Entrance exam residuals		*Panel B: College/major-level residuals*	
Major area	Proportion of grads	Field exit score	Log earnings	Field exit score	Log earnings
Business/economics	0.35	0.50	0.53	0.53	0.59
Engineering	0.29	0.51	0.60	0.45	0.59
Law	0.14	0.48	0.81	0.43	0.75
Social sciences	0.14	0.55	0.41	0.51	0.33
Health	0.07	0.52	0.66	0.54	0.68
Education	0.06	0.55	0.27	0.57	0.36
Fine arts	0.05	0.50	0.46	0.41	0.27
Agronomy	0.02	0.52	0.35	0.47	0.37
Natural sciences	0.02	0.75	0.62	0.55	0.50

Notes: This table includes all college/major pairs with at least 20 graduates in our sample, where majors are defined by the program name at each college. The Ministry of Education records aggregate these majors into the nine listed "areas." The first column shows the proportion of graduates from each major area, and the remaining columns display percentile ranks of college/major pairs using the learning and earning measures in the column header. For these, we sort college/majors according to each measure and then calculate average ranks within the major areas. Averages are weighted by each college/major's number of graduates.

ity at the institution level, we do so at the institution/major level. In other words, we calculate separate learning and earning productivity measures for each major offered by each college.[26] We then sort the roughly 1,100 college/major pairs according to each measure and calculate each college/major's percentile rank. This is analogous to the procedure used to calculate institution ranks in table 4.5.

Table 4.9 summarizes the resulting ranks using nine broader major "areas" defined by the Ministry of Education.[27] The first column displays the proportion of all graduates in our sample in each major area. More than half of all graduates are in majors related to business and engineering, which are offered by almost all colleges in the country. Majors related to fine arts and natural sciences are less popular and are offered by only a small number of colleges.

The other columns in table 4.9 show the average ranks from the 1,100 college/major pairs using different learning and earning measures. Panel A presents ranks based on the entrance exam residuals method, and panel B displays ranks based on the college/major-level residual method. Using either method, the results show that some majors—such as those in engi-

26. We include only institution/major pairs that have at least 20 graduates in our sample.
27. The ministry's categorization actually combines social sciences and law, but we split these major groups because they have vastly different properties with respect to our productivity measures.

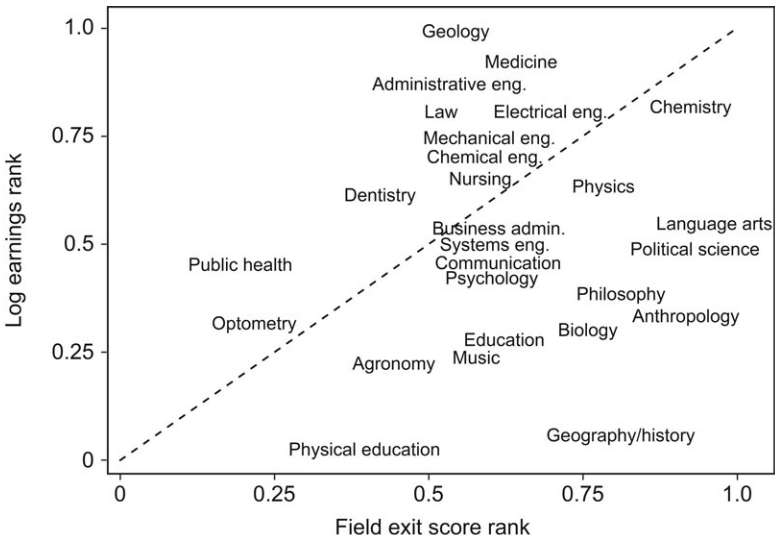

Fig. 4.6 Earning vs. field-specific learning ranks by major group

Notes: This figure plots percentile ranks for college/major pairs using the entrance exam, residual earning, and field-specific learning measures. We calculate these ranks as in panel A of table 4.9, but we display average ranks within a more granular categorization of majors into 51 groups defined by the Ministry of Education. Averages are weighted by each college's/major's number of graduates.

neering, business, and law—receive much higher ranks under the earning measures than under the learning measures. Conversely, majors related to education, fine arts, and social or natural sciences are much lower ranked using the earning measures.

Figure 4.6 elaborates on this result using a slightly more granular grouping of majors. The horizontal axis displays the average rank in each major group using the field-specific learning measure from panel A of table 4.9. The vertical axis depicts the average rank using the earning measure from the same procedure. Major groups that lie below the 45-degree line are ranked more highly on learning than on earning; these include many majors in social and natural sciences. Major groups above the 45-degree line, including many related to engineering and health, appear more favorable when rankings are based on earnings.

The results in table 4.9 and figure 4.6 suggest that the use of different productivity measures may create incentives for colleges to favor some majors over others. In particular, if policy makers primarily use earnings to measure performance, this could encourage college administrators to shift resources away from more specialized majors. Furthermore, in a separate analysis, we find that—holding fixed the measure of productivity—a college's ranking in one major is only moderately correlated with its ranking in another major.[28] Thus there is substantial scope for colleges to respond to accountability schemes by favoring some majors over others.

4.6 Conclusion

Increasingly, policy makers are looking to provide information on the outcomes that different colleges produce for their graduates. In many ways, this reflects a desire to extend school accountability to higher education. Casual observation suggests this desire is particularly prevalent in countries that have seen some combination of significant growth in access to college, growth of a substantial (and often relatively unregulated) private sector, and increasing amounts of student debt.[29] As with school accountability in K–12 education—despite its much longer history—questions remain as to the informational content and the ultimate effects of initiatives in this area.

Our goal here has been to contribute by calculating, for the country of Colombia, system-wide measures of college productivity in terms of earning and learning. While we do not claim that our measures isolate causal college value added, they allow for analyses beyond those that have been previously feasible. Our findings suggest that measures of college productivity on earning and learning are far from perfectly correlated.

A key implication of this is that the design of accountability systems will affect how these portray different types of colleges and potentially also how these colleges respond. For instance, we find that in the case of Colombia, top private colleges generally perform better under our earning measure, while selective public colleges appear more favorably under our learning measure.

In addition, in the earnings dimension, one can arrive at starkly different conclusions regarding colleges' relative productivity depending on when one measures earnings. This is problematic because the more chronologically removed the observation is from graduation, the more that factors extraneous to colleges—such as postschooling human capital investment decisions made by employers and employees—will have a chance to affect wages. This leaves open the possibility that learning measures do a better job of isolating a college's contribution to students' human capital. Of course, trade-offs abound, as shifting weight toward learning measures may induce gaming similar to that which has been observed around "No Child Left Behind" and analogous K–12 accountability initiatives.

Finally, our results illustrate that the use of different productivity measures may create incentives for colleges to favor some majors over others.

28. For example, the correlation between a college's business/economics ranking and its engineering ranking is 0.53 using field exit score residuals from equation (2) and 0.54 using log earnings residuals.

29. For instance, the United States, Chile, and Colombia fit some of these criteria.

For example, our findings suggest that they might encourage institutions to shift resources away from more specialized majors and toward areas such as business and engineering.

References

Altonji, J. G., and R. K. Mansfield. 2014. "Group-Average Observables as Controls for Sorting on Unobservables When Estimating Group Treatment Effects: The Case of School and Neighborhood Effects." NBER Working Paper no. 20781, Cambridge, MA.

Angrist, J., E. Bettinger, and M. Kremer. 2006. "Long-Term Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia." *American Economic Review* 96 (3): 847–62.

Arum, R., and J. Roksa. 2011. *Academically Adrift: Limited Learning on College Campuses*. Chicago: University of Chicago Press.

Chetty, R., J. N. Friedman, and J. Rockoff. 2014. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review* 104 (9): 2593–2632.

Cooke, M. 1910. *Academic and Industrial Efficiency*. New York: Carnegie Foundation for the Advancement of Teaching.

Council for Aid to Education. 2013. *Does College Matter? Measuring Critical-Thinking Outcomes Using the CLA*. Technical report. Accessed August 2016. http://cae.org/images/uploads/pdf/Does_College_Matter.pdf.

Dale, S. B., and A. B. Krueger. 2014. "Estimating the Effects of College Characteristics over the Career Using Administrative Earnings Data." *Journal of Human Resources* 49 (2): 323–58.

Hagedorn, L. S., E. T. Pascarella, M. Edison, J. M. Braxton, A. Nora, and P. T. Terenzini. 1999. "Institutional Context and the Development of Critical Thinking: A Research Note." *Review of Higher Education* 22 (3): 265–85.

Hastings, J., C. Neilson, and S. Zimmerman. 2013. "Are Some Degrees Worth More Than Others? Evidence from College Admissions Cutoffs in Chile." NBER Working Paper no. 19241, Cambridge, MA.

Hoekstra, M. 2009. "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach." *Review of Economics and Statistics* 91 (4): 717–24.

Hoxby, C., and G. Bulman. 2015. "Computing the Value Added of American Postsecondary Institutions." Mimeo, Stanford University.

Kirkeboen, L., E. Leuven, and M. Mogstad. 2016. "Field of Study, Earnings, and Self-Selection." *Quarterly Journal of Economics* 131 (3): 1057–1111.

Klein, S., J. Steedle, and H. Kugelmass. 2010. "The Lumina Longitudinal Study: Summary of Procedures and Findings Comparing the Longitudinal and Cross-sectional Models." Unpublished manuscript.

MacLeod, W. B., E. Riehl, J. E. Saavedra, and M. Urquiola. 2015. "The Big Sort: College Reputation and Labor Market Outcomes." NBER Working Paper no. 21230, Cambridge, MA.

Pascarella, E. T., and P. Terenzini. 2005. *How College Affects Students: A Third Decade of Research*. San Francisco: Jossey-Bass.

Saavedra, A., and J. E. Saavedra. 2011. "Do Colleges Cultivate Critical Thinking,

Problem Solving, Writing and Interpersonal Skills?” *Economics of Education Review* 30 (6): 1516–26.
Saavedra, J. 2009. “The Learning and Early Labor Market Effects of College Quality: A Regression Discontinuity Analysis.” Mimeo, Harvard University.
Stange, K. 2012. “Ability Sorting and the Importance of College Quality to Student Achievement: Evidence from Community Colleges.” *Education Finance and Policy* 7 (1): 74–105.
Sullivan, T., C. Mackie, W. F. Massy, and E. Sinha. 2012. *Improving Measurement of Productivity in Higher Education*. Online book. Washington, DC: National Academies Press.