The Role of Universities in Local Invention: Evidence from the Establishment of U.S. Colleges*

Michael Andrews†

July 6, 2018

Abstract

I exploit historical natural experiments to study how establishing a new college affects local invention. Throughout the nineteenth to the mid-twentieth century, many new colleges were established in the U.S. I use data on the site selection decisions for a subset of these colleges to identify "runner-up" locations that were strongly considered to become the site of a new college but were ultimately not chosen for reasons that are as good as random assignment. The runner-up counties are similar to the winning college counties along observable dimensions. Using the runner-up counties as counterfactuals, I find that the establishment of a new college caused 40% more patents per year in college counties relative to the runners-up. However I also conclude that colleges are not a necessary condition to promote local invention: establishing other types of institutions leads to a similar increase in local patenting. Reconciling these results, the primary channel by which colleges cause an increase in local invention is through migration. I cannot reject the hypothesis that, after controlling for population, colleges have no independent effect on local patenting. Furthermore, when linking patent data to a novel dataset of historical college yearbooks and to individual-level census data, only a small share of patents come from alumni or faculty of sample colleges, while most patents are by migrants to the college county.

*I am very grateful to the Kauffman Foundation, NSF Doctoral Dissertation Grant #1661421, the Balzan Foundation, and the Northwestern Center for Economic History for financial support. I would like to thank Enrico Berkes, Chelsea Crain, Claudia Goldin, Shawn Kantor, John Parman, Elisabeth Perlman, Heyu Xiong, and Yiling Zhao for sharing data and/or code. I am also indebted to the librarians in the Interlibrary Loan Office at the University of Iowa and Northwestern University, who have responded to a seemingly endless series of requests for old books about even older universities. This paper benefited from conversations with Ufuk Akcigit, Wes Cohen, Joe Ferrie, Roberto Fontana, Walker Hanlon, Ben Jones, Shawn Kantor, Joel Mokyr, Sarada, and Nicolas Ziebarth, as well as from seminar and conference participants at the University of Iowa, Kauffman Foundation, Northwestern University, NBER DAE Summer Institute, University of Wisconsin, Searle Center Conference on Innovation Economics, Auburn University, Economic History Association Annual Meetings, NBER Productivity lunch, Illinois Economic Association, Chicago Fed, Duke University, Utah State University, Bureau of Economic Analysis, Cliometric Society Annual Conference, and LSE Innovation conference. All errors are my own. The most recent version of this paper is available at https://www.m-andrews.com/research.

†Department of Economics, Northwestern University. *Email*: michael.andrews1@northwestern.edu.

1 Introduction

Policymakers, researchers, and business executives are all obsessed with finding "the next Silicon Valley" (Forbes Technology Council, 2017), but identifying specific policies and institutions to promote an innovation hub is not easy. A consensus has emerged that a successful college or university is a necessary condition to promote local invention (Davis, 2016).[1] For example, Moretti (2012, p. 197) argues that "proximity to a research university is important...to form a sustainable cluster of innovative companies," although he is careful to note that a university by itself is not sufficient to create such a cluster. O'Mara (2005, p. 6) refers to colleges and universities as "the economic development engine" at the heart of innovative cities. And Florida (2002, p. 291-292) concludes that "the presence of a major research university is a basic infrastructure component" of innovation hubs, even more important than physical infrastructure like bridges and railroads. For their part, colleges themselves frequently tout (likely inflated) estimates of their importance for promoting local entrepreneurship, employment, and innovation (Siegfried, Sanderson, & McHenry, 2007). An even starker way to gauge how colleges view their role in their local economies is to note that almost 15% of U.S. university mottoes explicitly refer to improving the local economy.[2] A sizable academic literature, starting with Jaffe's (1989) seminal paper, gives some credence to these claims, documenting that innovative firms tend to co-locate with colleges and that invention by these firms is correlated with changes in college spending.

This paper makes three contributions. First, I show that, as claimed by the sources above, creating a new college does cause an increase in local invention, as measured by patents. I do this by exploiting a historical natural experiment that is cleaner than previous strategies that attempt to establish causality. Second, I investigate the claims that colleges are a necessary condition for promoting local invention. Using the same natural experiments, I find that the creation of other types of institutions lead to increases in patenting indistinguishable from that caused by colleges, suggesting that colleges are not necessary. Finally, I reconcile the first two results by showing the mechanism by which colleges increase local invention.

[1] Throughout this paper, I use the terms "college" and "university" interchangeably.

[2] Author's calculation based on university mottoes listed on `https://en.wikipedia.org/wiki/List_of_university_mottos#United_States`. As examples, North Dakota State University's motto is "For the land and its people," while the University of Missouri is dedicated to "The welfare of the people."

The primary channel through which establishing a new college increases local invention is through migration. In fact, after controlling for changes in county population, I cannot reject the null hypothesis that colleges have no effect on local patenting. I further link the patent data to a novel dataset of historical college yearbooks and to the 100% decennial census and document that only a small share of patents come from alumni and faculty of new schools; the majority of patents come from migrants to the college counties.

Several recent studies attempt to estimate the causal effect of colleges on the local economy (e.g., Furman and MacGarvie (2007), Aghion, Boustan, Hoxby, and Vandenbussche (2009), Kantor and Whalley (2014), Hausman (2017)), although finding clear and convincing identification has proven difficult. The reason for this difficulty is clear: colleges are not located at random. As Hausman puts it: "To understand local industry effects of universities, one would ideally like to randomly allocate universities to locations and measure related industry activity in those locations after the universities arrived relative to before." (Hausman, 2017, p. 11). To the frustration of researchers, such an experiment is infeasible to conduct today. In this paper, I approximate this ideal experiment using historical data on the establishment of new colleges in the U.S., spanning the years 1839-1954. By exploring the narrative record, I am able to identify "runner-up" sites that were strongly considered to become the site of a new college, adopting a methodology used in Greenstone, Hornbeck, and Moretti (2010) to study the site selection decisions of large manufacturing plants.[3] The key idea behind this "runner-up" methodology is that, when selecting where to locate a new college, dozens of possible candidate locations are considered and iteratively eliminated; by the time only a few finalists sites are left they are likely similar along both observable and unobservable dimensions. While this methodology works well in the context of Greenstone et al. (2010), the identifying assumption can fail if only a small number of locations were ever considered, especially if these finalist locations appear very different from one another. To account for this, I refine the methodology by restricting the sample to cases in which I can verify that the site selection decision approximates random assignment.

A concrete example of a case with as-good-as-random assignment of the college may

[3]Malmendier, Moretti, and Peters (2016) and Helmers and Overman (2017) apply similar methodologies in other contexts.

be useful. In 1882, the state of North Dakota drew lots to determine where to locate the University of North Dakota and North Dakota State University; locations were literally randomly assigned (Geiger, 1958, p. 13-27). Not surprisingly, literal random assignment is rare, but many instances were close. As another example, in 1886 the citizens of Georgia wanted a technical college, but there was no consensus about where to put it. The two main rival sites were Atlanta and Macon. Both were known primarily as railway depots located in the interior of the state and looked similar along a number of observable dimensions. A site selection committee assembled to vote on the location of the college. For the first 23 ballots, neither Atlanta nor Macon obtained the requisite majority. Finally, on the 24th ballot, Atlanta won over Macon by one vote (McMath Jr. et al., 1985, p. 24-32). It is thus easy to believe that Georgia Tech University could have been located in Macon instead of Atlanta. The cases of the North Dakota universities and Georgia Tech were not isolated incidents: while the decisions were occasionally less dramatic, these kinds of college site selection experiments occurred all across the United States during the second half of the nineteenth century and first half of the twentieth.

Using the augmented runners-up methodology to identify counterfactual sites for the college counties, I show that the winning and runner-up counties are similar along observable dimensions. In contrast, most previous studies that focus on the establishment of new colleges assume that colleges are located at random (Currie and Moretti (2003), Moretti (2004), Andersson, Quigley, and Wilhelmsson (2004), Andersson, Quigley, and Wilhelmsson (2009), Frenette (2009), Cowan and Zinovyeva (2013), and Toivanen and Väänänen (2016)).[4] I show that such an assumption results in comparing colleges to dissimilar sites and greatly overstates the effect of colleges on local invention.

Using the runners-up as counterfactuals for the winning college counties, I examine how colleges affect the production of local patents. I find that establishing a new college causes 40% more patents per year in the winning county relative to the runners-up. But are colleges necessary to accomplish this increase? In some college site selection experiments, runner-up

[4]Moretti (2004, p. 190-191), focusing exclusively on land grant colleges, writes that, "Land-grant colleges were often established in rural areas, and their location was not dependent on natural resources or other factors that could make an area wealthier. In fact, judged from today's point of view, the geographical location of land-grant colleges seems close to random."

counties received other types of institutions, such as prisons or insane asylums. I refer to these as "consolation prizes." There is no measurable difference in patenting between the counties with colleges and counties with consolation prizes after establishing the new colleges, falsifying the claim that colleges are a necessary condition for promoting local invention. I also show that having a college that focuses on technical skills is not a necessary condition for promoting local invention by comparing "practical" colleges with curricula that focus on things like agriculture, engineering, and mining, to "classical" colleges that focus on subjects such as divinity, the law, and the classics. I find no difference between these different types of colleges either.

The first two findings, that colleges appear to increase local patenting on average but are no better than other types of institutions, may appear contradictory at first but are compatible if the primary channel by which colleges increase local invention is through migration. Indeed, while consolation prize institutions like prisons and insane asylums do not produce human capital like colleges do, throughout much of U.S. history they served as anchor institutions and attracted population similarly to colleges. When controlling for population in the baseline results, the coefficient on establishing a new college is reduced by about two-thirds and is no longer statistically different from zero: I cannot rule out that colleges play *no* role in promoting local invention outside of their effect on population. Moreover, I find no evidence that agglomeration economies are stronger in college counties than in the runner-up counties.

One benefit of using patent data is that it is possible to see which individuals are creating new inventions.[5] I therefore match the names of patentees to a novel dataset of historical college yearbooks to determine which share of patents are created by individuals who are directly affiliated with a newly established college. Less than 20% of patents in a college county after the college is established come from either alumnus or faculty of that college. Almost 75% of patents come from individuals who migrate to the college county, reaffirming that the most important channel through which colleges affect patenting is through migration.

[5] Other benefits are that the patent data are available for the entire U.S. over a long time period and a wide range of technology classes, allowing me to estimate the long-run effects of establishing a new college. Of course, college data have some drawbacks as well: not all innovations are patented, and not all patents are for meaningful innovations, making patents a less than perfect proxy for innovation. See Griliches (1990) and Nagaoka, Motohashi, and Goto (2010), as well as the results below, for a discussion of these issues.

This paper is organized as follows. Section 2 describes the data, including an in depth explanation of the college site selection experiments. Section 3 presents the baseline results for the effect of the establishment of a new college on local patenting. In Section 4, I show that colleges are not a necessary condition for promoting local invention. Section 5 shows that the primary channel through which colleges increase local invention is through migration, and that furthermore most patents are invented by migrants to the college county. Section 6 puts the results into a broader context and concludes.

2 Data and Empirical Model

2.1 The College Site Selection Experiments

U.S. history provides an ideal laboratory to study the establishment of new colleges. As Goldin and Katz (1999) and Xiong and Zhao (2017) point out, the mid-19th to mid-20th centuries saw an explosion in the number of colleges and universities in the U.S., providing many potential college site selection experiments. In addition, because most colleges were built more than a century ago, it is possible to trace the long-term effects of colleges on invention. Moreover, many of these colleges were built at the beginning of what Goldin and Katz (2008) call the "human capital century" and Gordon (2016) refers to as "the golden age" of America's technological leadership.

To estimate the causal effect of establishing a college, the first step is to identify valid counterfactuals to college sites. A large program evaluation literature uses losing applicants as counterfactuals.[6] A potential drawback of this methodology is that rejected applicants may be very different from the winners. For example, in the context of political candidates, Borgschulte and Vogler (2016) show that examining candidates in close elections, when the decision was closer to "random," gives very different results than investigating all candidates. Greenstone et al. (2010), in a study estimating the agglomeration externalities

[6] See, for example, Aizer, Eli, Ferrie, and Lleras-Muney (2016), von Wachter, Song, and Manchester (2011), and Bound (1989) for rejected applicants to various social insurance programs; Dale and Krueger (2002) for rejected applicants to selective colleges; Olenski, Abola, and Jena (2015), Borgschulte (2014), and Olshansky (2011) for losing political candidates; and Busso, Gregory, and Kline (2013) for rejected sites for place-based policies.

from large manufacturing plants, overcome this challenge by using "runner-up" locations as counterfactuals for winning sites. This ensures that both the winning and runner-up sites considered themselves valid to receive treatment, and so did the site selection committee, which strongly considered both the winners and runners-up. I adopt a similar approach to identify counterfactuals to college sites.[7]

A great deal of thought went into each college site selection decision throughout U.S. history. Horace Bushnell, a theologian who played a central role in locating both the University of California and the University of Illinois, articulated the weight of these decisions: "The site of a university is to be chosen but once. Once planted, it can never be removed; and if any mistake is made, that mistake rests on the institution as a burden to the end of time" (quoted in Ferrier (1930, p. 162)). Many localities wanted to secure a new college, and any economic benefits that went along with it, for themselves. This ensured that site selection decisions often became quite contentious. Further complicating the site selection decision is the fact that new colleges often had particular infrastructure needs. In the case of land grant universities, for example, the Morrill Act of 1862 explicitly prohibited states from using their land grant fund to construct buildings. This forced states to locate land grant colleges in towns with unused buildings large enough for a college or in localities willing to raise the funds for construction.

To find runner-up sites, I consult institutional histories, what Washburn (1979) calls "the driest of dry forms of historiography", for information on the college site selection process. I consult histories for 432 colleges, including nearly every prominent U.S. college.[8] For 193 of

[7]Patrick (2016) raises several challenges to the identification strategy employed by Greenstone et al. (2010). I believe my study avoids these critiques. First, because I study the site selection decisions of colleges rather than for-profit businesses, there is little strategic reason for colleges to hide their list of finalist locations from competitors. Second, I provide a great deal of institutional detail that shows that the site selection decision was indeed close to random. Finally, in Section 2.3 I show that college and runner-up sites are similar in terms of observables; Figures 8 and 3 show that the colleges and runners-up evolved similarly as well.

[8]I thus examine the histories of almost 15% of the 3,039 degree-granting four-year post-secondary schools in the U.S. as of 2013 (`https://nces.ed.gov/fastfacts/display.asp?id=84`). I investigate every national university ranked by the 2018 U.S. News and World Report Best Colleges ranking (`https://www.usnews.com/best-colleges/rankings/national-universities`), as well as the 25 best liberal arts colleges in the corresponding ranking (`https://www.usnews.com/best-colleges/rankings/national-liberal-arts-colleges`); every land grant college; the first public university founded in each state; the flagship university of a each state's public university system if this is different from either the land grant or first public university; every state technical school and mining college; every federal military academy; and every university belonging to a Power Five athletic conference. When data was available, I

these colleges, I am able to find information on the candidate locations that were considered. One drawback to this approach is that it identifies all finalists, regardless of how similar the winning and losing sites are or how close the site selection process was to random assignment. For instance, three different counties submitted bids to the Ohio State legislature to receive the new Ohio State University, but there does not appear to be any serious discussion in the legislature: the college was always intended to be located at the state capital in Columbus. Moreover, Columbus is very different from these other localities along observable dimensions. To mitigate this problem, I further restrict the sample to only include cases in which the site selection decision is plausibly exogenous; I refer to these as "high quality" college selection experiments. I consider 76 of the college cases to be high quality experiments. Four of these high quality experiments take place prior to the start of the patent data in 1836, so I exclude them. The remaining 72 high quality college site selection experiments form my baseline sample.

The high quality site selection experiments can grouped in four ways. First, a vote among candidate locations may be exceptionally close; the case of Georgia Tech described in the Introduction is one example of this. Second, candidate locations frequently submitted bids to boards of trustees or state legislatures to receive a new college. When two bids are similar, this is evidence that the localities valued receiving the school roughly equally, and the decision makers were largely indifferent between the two sites. Third, in some instances a new college had specific infrastructure needs, such as existing vacant buildings of a suitable size; in a few of these cases, only two or three such sites within the state possessed the required infrastructure. Finally, some site selection experiments involve quirky random events that are difficult to otherwise classify. The random assignment of the University of North Dakota and North Dakota State University is an example of this. Cornell University provides another example. Ezra Cornell and Andrew White, the fathers of Cornell, wanted to establish the college in one of their home towns but could not decide on which. Ezra Cornell was from Ithaca, while Andrew White was from Syracuse. But Cornell had been

also investigated historically black colleges and universities (HBCUs) and private colleges, with a focus on the private colleges that have been historically noteworthy or are currently considered prestigious. For a handful of states, I also investigated each normal school established in that state. Over time, normal schools typically evolved to become "directional" state universities (for example, the Michigan State Normal College became Eastern Michigan University). Any further sample reductions are due exclusively to data availability.

cheated of his wages as a young man in Syracuse and refused to locate the college there. Consequently, Cornell University is located in Ithaca.[9] In Appendix C.1, I show that the results are not sensitive to discarding any of these groups of the high quality experiments. The Historical Appendix describes each college site selection experiment in detail, including the low quality experiments. The results are also not sensitive to reclassifying marginal cases as either high or low quality.

One potential concern with this identification strategy, as with any methodology in which runners-up are used as a control group, is that even if the final site selection decision is contentious, the winning and runner-up counties may look nothing like one another. In Section 2.3, I verify that the winners and runners-up are similar along observable dimensions. But there may still be unobservable dimensions along which they differ. For instance, consider a case in which a site selection committee weighs a number of characteristics that are unobservable to the econometrician. The winning location could be strong in one characteristic but weak in others, while the runners-up could be strong in other characteristics. If the characteristic that is prevalent in the winning county is also correlated with patenting, then this would bias the results. This is unlikely for two reasons, however. First, runner-up sites were typically those that met all the conditions the site selection committees viewed as essential to receive a college, such as access to transportation, access to potable water, or aesthetic beauty; any remaining differences are likely to be negligible. Second, as long as the final site selection decision truly is close to random, as I verify in the high quality experiments, then which characteristic is prevalent in the winning versus losing counties should also be randomly distributed across the college site selection experiments.

Augmenting the runner-up methodology by using narrative history to exclude cases in which selection is not as good as random assignment is, as far as I know, novel in the literature.[10] In a paper studying agricultural experiment stations, Kantor and Whalley

[9] I have been unable to find any evidence that Syracuse tended to have citizens of a lower moral character than did Ithaca. Syracuse and Ithaca were furthermore similar along observable dimensions before the establishment of Cornell University. Syracuse would, of course, get its own university several years later.

[10] Liu (2015), Bonander, Jakobsson, Podestà, and Svensson (2016), and J. Lee (2018) use synthetic control methodologies to study the economic impact of establishing or expanding colleges. A synthetic control methodology is less appropriate in a historical context because data for several desired predictors are not available for most locations in most pre-treatment years. For instance, J. Lee (2018) argues that real estate prices are an important predictor to understand the demand for land in winning and losing locations. In

(2018) compare a subset of land grant colleges to all contending locations as a robustness check. They do not, however, restrict attention to those cases in which the winning site is as good as randomly assigned; for instance, Ohio State forms one of the colleges in their sample. In the analysis below, I show that failing to exclude these low quality experiments overstates the effect of establishing a college.

2.2 Patent and County Data

The data on patents covers 1836-2010.[11] This patent data come from four sources, with different sources available for different years. For the years 1836-1870, I use patent data collected in the Subject-Matter Index of Patents for Inventions Issued by the United States Patent Office from 1790 to 1873 (Leggett, 1874), compiled by Dr. Jim Shaw of Hutchinson, KS.[12] I use the Annual Reports of the Commissioner of Patents for the years 1870 to 1942. See Sarada, Andrews, and Ziebarth (2017) for details on cleaning, parsing, and preparing this dataset. The years 1942 to 1975 come from the HistPat dataset compiled by Petralia, Balland, and Rigby (2016a); see Petralia, Balland, and Rigby (2016b) for details on the construction of this data. Finally, for the years 1975 to 2010, contemporary digitized patent data sources can be used. I utilize the data created for Li et al. (2014) which contains cleaned inventor names. Because all future analysis will include year effects, there is no concern with the fact that different years make use of different patent data sources. Each of these datasets contains, for every granted U.S. patent, the names and residence of all inventors.[13] For the

addition, in most cases unobservable factors, such as the enthusiasm of the local population for education or specific infrastructure, were crucial both in becoming a finalist site and in the later production of innovations. These factors are taken into account in the current methodology but are neglected in any methodology that matches on observables.

[11] Patent data from before 1836 is not useful for analysis, as 1836 marked a major change in the U.S. patent system, essentially changing from a registration system to an examination system. In addition, a major fire at the U.S. Patent Office in 1836 destroyed most of the patents from the early United States.

[12] See Miller (2016a) and Miller (2016b) for more information on how this dataset is compiled.

[13] The fact that each patent dataset used in this paper reports the names of individual inventors is important for matching patentees to other datasets, namely college yearbook data or the U.S. population censuses. Other commonly used patent datasets, such as the NBER patent data and its supplements (Hall, Jaffe, & Trajtenberg, 2001), only include patents that are assigned to firms or other institutional entities and do not include the names of inventors. Moreover, the name listed on each patent is legally requird to be the name of the "first and true inventor" of a particular invention rather than, for instance, the owner of the firm in which the inventor is employed. Failure to accurately list the inventors on a patent can result in loss of patent rights, providing confidence that recorded inventor names are accurate up to transcription and character recognition errors; see Khan (2005) for more details.

results in this paper, I aggregate all patents to the county level, matching towns to counties using the U.S. censuses as described in Appendix B. In Appendix C.1, I show results using other sources of patent data and various techniques to match town names to counties; in all cases the results are qualitatively similar to the baseline results presented below. See Andrews (2017) for a more in-depth discussion of the differences and similarities between the various patent datasets.

I merge by patent number and/or inventor name to other datasets that include additional patent information. The U.S. Patent and Trademark Office's Historical Patent Data Files (Marco, Carley, Jackson, & Myers, 2015) contain information on patent classes for historical patents. Enrico Berkes graciously provided data on patent citations and patent claims; see Berkes (2018) for details.

County-level data comes from the National Historic Geographic Information System (NHGIS) (Manson, Schroeder, Riper, & Ruggles, 2017). The NHGIS data allows me to compare counties along a number of useful dimensions including population; composition of the county population along racial, gender, immigration, and age dimensions; urbanization; and wages and production in both agricultural and manufacturing sectors. I also use data on the total number of accredited colleges at the college level. These are found in Reports of the Commissioner of Education, several years of which have been transcribed: 1870, 1875, 1880, 1885, 1890, 1895, 1900, 1905, 1910, and 1914 by Heyu Xiong and Yiling Zhao; and 1897, 1924, and 1934 by Claudia Goldin. Because county names and boundaries change over time, I aggregate counties to their largest historical boundaries, adopting a method similar to Atack, Jaremski, and Rousseau (2014).

2.3 The College and Runner-Up Counties

Table 1 lists each of the 72 high quality college site selection experiments in the final sample as well as the year in which the experiment took place and the college type. To give a sense of the type of colleges involved in the study, I classify colleges into one of seven mutually exclusive groups: land grant colleges, technical colleges, normal schools, historically black colleges and universities (HBCUs), military academies, other public colleges, and other private colleges.[14] A plurality of the college experiments involve land grant colleges. 10% of the experiments involve technical colleges, 17% involve normal schools, 4% involve HBCUs, and 4% involve military academies. 17% of the colleges are classified as "other" public colleges, while 7% are classified as "other" private colleges.

Table 2 further summarizes the college experiment data. In addition to presenting the share of patents belonging to each experiment type, I show that each college site had on average 2.83 runner-up sites. The runner-up sites are on average about 140 km ($\approx$87 miles) away from the college towns, with the median runner-up 93 km ($\approx$58 miles) away, using geodetic distances. This is far enough that the college and runner-up sites are typically in different labor markets, but close enough to be affected similarly by region-wide shocks. While colleges were established throughout the entire period from 1839 to 1954, the mean and median college is established in the mid-1880s, with the median college beginning to admit students four years after determining where the school is to be located. Desegregation and co-education, not surprisingly, tended to happen much later for most colleges in the sample, although these dates are not available for all colleges. Figure 1 is a map of the college and runner-up counties throughout the U.S., providing visual verification that the college and runner-up counties vary in their distance from one another. The map also shows that the entire continental U.S. is represented in the sample and that colleges were not simply built near existing major population centers.

Figure 2 compares the college and runner-up counties and shows that the runners-up are a better match for the college counties than are the "non-experimental" counties, which are all other counties in a college's state that are not either college or runner-up counties. The black diamonds display the difference in the mean between the college and runner-

[14] Technical colleges include schools focused on engineering, mining, and industrial arts. Normal schools are colleges focused on teacher training; many of these have evolved to become directional state universities. Other public and private universities include all public and private, respectively, schools that do not fit into any of the other classifications. For instance, the University of Texas is classified as an "other public" college in the sample; Texas also has two other state-wide (that is, not "directional states" targeted to a particular region within Texas) public universities, a land grant college (Texas A&M) and a technical college (Texas Tech), both of which are also in my sample. In some cases, a college may fall into multiple categories. For example, many HBCUs are also state land grant colleges. For clarity, in Table 1, I place each college into its "best" category. All results are insensitive to reclassifying colleges.

up counties in the last U.S. census year before the college was established on a number of economic, demographic, and educational variables. The black lines show 95% confidence intervals of a simple t-test of the difference in means. I use census years because most of the demographic and economic variables are collected with the decennial census. The means of the college and runner-up counties are statistically indistinguishable and remarkably similar in magnitude. The green circles show the difference in the mean between the college and the non-experimental counties, which are the counties in each state that are not classified as either college or runner-up counties. The green lines show 95% confidence intervals for the t-test. The college and non-experimental counties also tend to be similar along some dimensions, making Moretti's (2004) claim that colleges were located "close to random" understandable. But relative to the non-experimental counties, the college counties do have a statistically larger population, are more urbanized, have a larger share of interstate migrants, greater manufacturing output, and are more likely to already have an existing college. Appendix A.1 displays mean values in addition to the difference in means for these and a number of additional variables to further verify that the runners-up appear similar to the college counties.

2.4 Empirical Model

I estimate a straightforward differences-in-differences equation with grouped observations. That is, in county i associated with college j at time t, the number of patents is given by

$$\begin{aligned} PatentMeasure_{ijt} = &\delta_1 College_{ij} * PostCollege_{jt} + \delta_2 PostCollege_{jt} \\ &+ \alpha_i + \lambda_j + \alpha_i * \lambda_j + \gamma_t + \epsilon_{ijt}, \end{aligned} \tag{1}$$

where $College_{ij}$ is an indicator variable equal to one if county i associated with college experiment j receives the college, $PostCollege_{jt}$ is an indicator variable equal to one in years t after college j has been established, α_i is a county fixed effect, λ_j is an experiment fixed effect, γ_t is a year effect, and ϵ_{ict} is a county-college-year varying error term.[15] With

[15]While I include fixed effects for experiments, λ_j, counties, α_i, and counties-by-experiment, $\alpha_i \times \lambda_j$, most counties appear in only one experiment. In these cases λ_j and $\alpha_i \times \lambda_j$ are redundant and are omitted.

only a single experiment, the term $PostCollege_{jt}$ would be redundant because the post-college dummy is perfectly co-linear with the year effects. There are multiple experiments in the dataset, however, with each college being established in different years, and so each group j will be in the post-college period in different years. The year effects therefore control for nationwide time-variant changes in patenting, while $PostCollege_{jt}$ controls for changes that occur within all experiment j counties after establishing college j.

3 The Effect of Establishing a College on Local Patenting

Figure 3 plots smoothed $log(NumPat_{ijt} + 1)$ for college, runner-up, and non-experimental counties separately.[16] The year in which a new college is established is normalized to be year 0 for all experiments. Two results are immediately clear. First, new colleges do not appear to be randomly located; there is a large difference between the college and runner-up counties on one hand and the non-experimental counties on the other, both in the level and growth rate of patenting. It appears that, in choosing potential sites for a new college, the desire to locate the college where new ideas grew rapidly outweighed any accessibility concerns that might lead a site selection committee to place the college in backwater areas without much invention. Second, the college and runner-up counties patented similarly in pre-college years, suggesting that the experimental design is valid. Third, after the establishment of a new college, the college and runner-up counties diverge, with college counties patenting more. This divergence is especially pronounced after several decades.

Table 3 formalizes the intuition in Figures 3. The columns show different regression specifications. For all tables in the paper, coefficients are presented as a proportional change

[16] Figure 3 is constructed by regressing $log(NumPat_{ijt}+1)$ on year effects γ_t and then plotting the residuals using local mean smoothing with an Epanechnikov kernel function. Removing time effects is useful because, as Griliches (1990) shows, there has been a secular increase in patenting overtime as well as country-wide cyclical fluctuations in patenting that coincide with business cycles and changes in the administration of the Patent Office; failure to control for these factors makes interpreting the graph more difficult. The figure contains a balanced set of college experiments, including only counties with at least 20 years of pre-college and 80 years of post-college data available. The graph is nearly identical when using an unbalanced panel instead.

in patenting.[17] The coefficient of interest is displayed in Row 1 and shows the percentage of estimated additional patents generated in the college county relative to the runner-up county after establishing the new college. For all columns, standard errors are clustered at the county level.[18]

Column 1 shows the results of estimating Equation (1) where the dependent variable is $\log(Num.Pat_{ijt}+1)$. College counties have about 40% more patents per year than runner-up counties. As the table shows, the average county had about 4.8 patents in 1880, around the time the median new college is established. This translates to roughly two additional patents in 1880. By 2010, the average county had about 37.5 patents per year, so the college causes just over 16 additional patents per year in 2010. This result is statistically significant at the 1% level.

In Column 2, I use the inverse hyperbolic sin of patenting, $\log(Num.Pat_{ijt}+(Num.Pat.^2+1)^0.5)$, as the dependent variable. The benefit of the inverse hyperbolic sin is that it can take on zero values, yet still has the same simple interpretation as in Column 1. Here, establishing a new college causes about 48% more patents per year in the college counties relative to the runners-up, similar in magnitude to the results in Column 1.

Column 3 shows results using an alternative calculation of logged patents as proposed by Blundell, Griffith, and Reenen (1995). Rather than adding a positive constant before taking the log of patents, this alternative method uses $log(patents)$ as the dependent variable. Whenever $patents = 0$, a dummy variable is set to one and $log(0)$ is replaced with 0. In this specification, establishing a new college leads to a roughly 26% more patents per year, or about an extra 1.25 patent per year.

[17]More precisely, because the variables of interest are indicators that are either equal to zero or one, when the dependent variable is $\log(Num.Patents + 1)$ the estimated coefficient must be adjusted to give the percentage increase in patenting using the equation $\%Change\ in\ Patents = e^{\hat{\delta_1}} - 1$, where δ_1 is the coefficient of interest from Equation (1). This adjusted coefficient is presented in the table. When other transformations of the coefficient are necessary, standard errors are corrected using the delta method.

[18]I also cluster at the state, experiment, and county×experiment levels. I additionally cluster at multiple levels as proposed in A. C. Cameron, Gelbach, and Miller (2011): I cluster at the county and year; state and year; experiment and year; and county, state, experiment, and year levels. Clustering at the county level produces the largest standard errors, but the standard errors are virtually identical at every level and none of the inferences change. Clustering at the county level is preferred because in a small number of cases, the same county may appear as a control for multiple experiments; clustering at the county rather than experiment or county×experiment level ensures that multiple cross sectional appearances of the same county are not treated as independent of one another. For a discussion of the most appropriate level at which to cluster standard errors, see C. A. Cameron and Miller (2015).

Column 4 uses the fact that the number of patents takes on integer values and presents estimates of a negative binomial regression. To convert this coefficient to a proportional change, I simply divide the estimated change in the level of patenting $\hat{\delta_1}$ by the number of patents in 1880; the years 1880 is chosen as the baseline year because it is close to the year in which the median new college is established as shown in Table 2. Using a later baseline year, when patenting is secularly higher, leads to a smaller estimate of the percentage change. Because of the strong influence of outliers when raw counts of patents are used, I Winsorize the top 5% of counties by yearly patenting. In this specification, establishing a new college leads to an 86% increase in patenting, about twice as large as the baseline estimate in Column 1.

In Figure 4 I interact the effect of a college by ranges of years. More precisely, I estimate

$$\begin{aligned}\log(NumPat_{ijt} + 1) = &\sum_{\tau \in T} [\delta_{1\tau} College_{ij} * TimeBin_{j\tau} + \delta_{2\tau} TimeBin_{j\tau}] \\ &+ \alpha_i + \lambda_j + \alpha_i * \lambda_j + \gamma_t + \epsilon_{ijt},\end{aligned}$$

where $\tau \in T$ represent "bins of years" (i.e., 0-10 years after the college is established, 10-20 years after the college is established, etc.) and $TimeBin_{j\tau}$ is an indicator variable that is equal to one if $t \in \tau$ and 0 otherwise. Each plotted coefficient represents $\delta_{1\tau}$ in the respective range of years τ since the college establishment. Results are nearly identical using different groupings of years. Confirming the intuition shown in the raw data, there is no significant difference between the college and runner-up counties in any of the years before the establishment of each college and the estimated coefficients are very close to zero in magnitude. After the first decade, the difference between the college and runner-up counties is statistically significant and stays roughly constant in magnitude for the next thirty years or so (although some decades are not individually statistically significant). This is the first, albeit highly suggestive, evidence that the results are not driven by the human capital created within the schools: most colleges begin with a very small student and faculty population, and moreover it takes many more years for a substantial number of students to matriculate and begin potentially inventive careers. But Figure 4 shows that college counties have a measurable increase in patenting relative to the runners-up relatively quickly, and that the magnitude does not increase as the number of graduates increases with time. This is consistent with a story in which the colleges have little direct effect on patenting, but colleges act as an anchor to attract population, a hypothesis I explore in more detail below. After 50 years, the difference between college and runner-up counties begins to increase. For the average college, built during the 1880s, this timing corresponds to the massive increase in federal funding of university research in the post-World War II era.[19]

In Table 4, I repeat the analysis in Columns 1-3 of Table 3 but include data from all colleges and counties for which runner-up sites can be identified. This includes the "low-quality" experiments as well as other runner-up counties in the high quality experiments that were nevertheless not as good as randomly assigned and so are excluded from the baseline sample. Instead of estimating Equation (1), I now estimate a triple-difference equation of the form

$$\begin{aligned} PatentMeasure_{ijt} =& \delta_1 College_{ij} * HighQuality_{ij} * PostCollege_{jt} \\ & + \delta_2 College_{ij} * PostCollege_{jt} + \delta_3 HighQuality_{ij} * PostCollege_{jt} \\ & + \delta_4 PostCollege_{jt} + \alpha_i + \lambda_j + \alpha_i * \lambda_j + \gamma_t + \epsilon_{ijt}. \end{aligned} \tag{2}$$

The indices mean the same as in the previous equations. Now $HighQuality_{ij}$ is equal to one if county i is included in the original baseline sample for high quality college experiment j, and zero otherwise. In Column 1, I estimate equation (2) using $\log(NumPat_{ijt} + 1)$ as the dependent variable. In Column 2, I use as the dependent variable the inverse hyperbolic sin of patenting, while in Column 3 the dependent variable is the alternative $\log(NumPat_{ijt})$ measure that includes a dummy equal to one if a county has zero patents in a particular year. In the new regression specifications, the coefficient of the triple-interaction term δ_1 measures how much larger the difference-in-differences estimator between high quality college and runner-up counties is compared to the difference-in-differences estimator between all college counties (high and low quality) and all runner-up counties (not just the high quality runners-up). This coefficient is negative and statistically significant, indicating that there is

[19]The baseline results are robust to excluding all post-1940 years, so the results are not driven by the post-war change in federal policy towards university research.

positive selection into the low quality college experiments; that is, the difference between the college and runner-up counties is smaller for high quality experiments than for counties not included in the baseline results. This result shows why restricting attention to high quality experiments is important: otherwise, the effect of colleges on patenting would be overstated. δ_2 estimates the increase in patenting in *all* college counties relative to *all* runner-up counties after establishing a new college; this is analogous to the interaction term in Equation (1) if the low quality experiments and non-experimental control counties were included in those regressions. The estimate of δ_2 is positive and significant, so the qualitative conclusions of the baseline specification in Table 3 are still be true even when the low quality experiments are included, although the coefficients are two to three times larger when attention is not restricted to the high quality experiments. The increase in patenting in high quality college counties over high quality runner-up counties after establishment of a new college (that is, the same quantity as estimated by δ_1 in Equation (1)) is given by $\delta_1 + \delta_2$.[20] Combining these coefficients reveals that high quality college counties increase patenting by amounts slightly larger than, but qualitatively similar to, the findings in Columns 1-3 of Table 3. All the combined coefficients are still statistically significant. δ_3 estimates the change in patenting in high quality college and runner-up counties after the establishment of a college relative to low quality college and runner-up counties. Finally, δ_4 has the same interpretation as before and simply measures the increase in patenting after establishment of a new college.

3.1 Robustness Checks

Appendix C.1 presents numerous robustness checks. In particular, I show that the baseline results are robust to a battery of additional specifications, including different combinations of fixed effects, alternative measures of the dependent variable, and additional count-data models. I further show that the results are robust to using different subsets of the college site

[20]Let $y = \log(NumPatents + 1)$. Then, abusing notation and ignoring the fixed effects and error terms, the coefficient of interest is

$$\begin{aligned}(y_{Coll.,HighQual.,Post} - y_{Coll.,HighQual.,Pre}) - (y_{RunUp,HighQual.,Post} - y_{RunUp,HighQual.,Pre}) &= [\delta_1 + \delta_2 + \delta_3 + \delta_4] - [0] - [\delta_3 + \delta_4] + [0] \\ &= \delta_1 + \delta_2.\end{aligned}$$

selection experiments, so the results are not driven by the specific choice of colleges used, and to using different historical patent datasets. Finally, I present results from placebo tests that use other years as the "experiment date" and find no effect on patenting, suggesting that the results are indeed driven by the creation of the new college.

3.2 Patent Classes

One concern is that the baseline results capture a shift in the types of innovations that occur in college counties towards those that can be patented, rather than an actual increase in innovation. In Appendix C.2, I check this in two ways: first by controlling for the distribution of patents in each class, and second by estimating the model at the patent class-by-county-by-year level and including class and class-by-year fixed effects. While the estimated difference-in-differences coefficient is slightly smaller in these two models, suggesting some substitution towards patentable technologies may be taking place, establishing a new college still causes a statistically significant increase in patenting even after controlling for patent class.

Moreover, instead of college counties simply specializing in one or two patentable technologies after the college is established, I find that the diversity of patent classes increases in the college counties relative to the runners-up after the establishment of a new college. This increase in diversity is true for all different types of colleges. For instance, the share of agricultural patents declines in colleges with land grant colleges relative to their runners-up, as does the share of mining patents in counties with mining colleges. This is further evidence that the observed effects are not driven by the direct effect of human capital taught in these colleges: if human capital were the primary factor, then colleges that specialize in particular fields should see a relative increase in those types of patents, whereas I find the opposite result.

3.3 Patent Quality

As Trajtenberg (1990) makes clear, using patent counts without correcting for patent quality can produce highly misleading results. Following Hall, Jaffe, and Trajtenberg (2005), the literature typically uses citation-weighting to measure patent quality. In Appendix C.3, I show that the change in lifetime citations per patent is statistically indistinguishable from zero in the college counties relative to the runners-up after establishing the new college; the coefficient is a precisely estimated zero after controlling for changes in patent classes.

Patent citations are a less-than-ideal measure of patent quality for historical patents because citations were only required after 1947. Thus, for patents granted during the 19th century, most patents have zero lifetime citations and there is relatively little variation across patents. To overcome this challenge, I also use the length of a patent's first claim as an alternative measure of patent quality, as proposed by Kuhn and Thompson (2017). Patents with longer first claims are narrower patents, while shorter claims represent broader patent breadth. Again, there is no difference in first claim length between the college and runner-up counties after the establishment of a new college. Thus, creating a new college causes more patents, but has no measurable effect on patent quality. For the remaining results I therefore use logged patent counts as the dependent variable unless otherwise noted.

3.4 Geographic Spillovers

The estimates in Table 3 will be biased upwards if patenting individuals from runner-up counties migrate to college counties following the establishment of a new college, violating the stable unit treatment value assumption. If such migration is occurring on a large scale, and if individuals are more likely to move to nearby areas, then the difference between college counties and nearby runners-up should be larger than the difference between college counties and far away runners-up. In Appendix C.4, I show that this is not the case. In fact, I find that the difference between college counties and nearby runner-up counties is virtually non-existent after establishing the college, whereas the difference between college counties and geographically distant runner-up counties is large. This suggests the presence of geographic spillovers from college counties to nearby areas, consistent with a large literature on the geography of knowledge flows (e.g., Jaffe, Trajtenberg, and Henderson (1993), Thompson (2006), Belenzon and Schankerman (2013), Crescenzi, Nathan, and Rodgríguez-Pose (2016)) and implying that, if anything, the baseline estimates are underestimates of the causal effect of colleges on local invention.

4 Are Colleges Necessary to Promote Local Invention?

In this section, I propose tests that can falsify the claim that colleges are a necessary condition for the promotion of local invention. Specifically, I show that the creation of non-college institutions can lead to increases in local patenting indistinguishable from the increases caused by colleges. I also provide evidence against the even stronger claim that technically-focused universities are a necessary condition by showing that the effect of establishing a college on patenting does not appear to depend on the type of school established.

4.1 "Consolation Prizes"

In some of the college site selection experiments in my sample, runner-up counties were not truly "losers": while they may not have obtained a college, they did obtain some other type of institution. I refer to these as "consolation prizes." Consolation prizes are especially common in western states that were largely unsettled and achieved statehood after the passage of the Morrill Act in 1862. In these states, typically several state institutions were allocated at the same time, including the state capital, the state prison, the state hospital, or the state insane asylum. While numerous localities may have been lobbying to get a state institution, which locality ended up with which institution was as good as random. In one famous example, the Tucson delegation set out for Prescott for the Arizona territorial legislature in 1885 intent on getting the state mental hospital. But flooding on the Salt River delayed the delegation. By the time they reached Prescott, the mental hospital had already been spoken for; Tucson was stuck with state university.[21]

Table 5 shows results that compare college counties to the consolation prize runners-up. The coefficient is a statistically insignificant 19%, less than half as large as the baseline estimate of 40%. In other words, college counties do not have many more patents per year than counties that received prisons, hospitals, or insane asylums. Panel (a) in Figure 5 presents these results graphically, analogously to Figure 3, and shows that the college and consolation prize counties evolve remarkably similarly over many decades. Panel (b) suggests why this

[21]For more details on the site selection decision of the University of Arizona, see Martin (1960, p. 21-25), Wagoner (1970, p. 194-222), and Cline (1983, p. 2-4).

may be the case: population in the consolation prize counties grows nearly identically to population in the college counties. In Column 2, I control for logged population and again compare the college and consolation prize counties. In this case, the difference between the college and consolation prize counties is still statistically insignificant and even closer to zero in magnitude. Consistent with these findings, Column 3 shows that when consolation prize counties are excluded from the sample, a new college increases patenting by about 48%, slightly larger than the 40% baseline estimate. None of these results are qualitatively changed by excluding any particular type of consolation prize county; the results are not driven simply by types of consolation prizes that might attract high human capital individuals such as state capitals or hospitals. To make sense of these results, it is important to remember the context in which these experiments took place. Instead of repelling highly mobile workers, as prisons or asylums might today, the consolation prizes gave small towns an identity and acted as anchors to attract more people to the area.

4.2 College Types

I next test whether colleges that focus on technical skills are necessary for promoting local invention. Since different types of colleges had very different curricula, if human capital or faculty research are the primary channels by which colleges promote invention, then some types of colleges should cause much larger increases in local invention than others. I classify colleges by type as described in Section 2.3. From these college types, I further classify each college as either a "practical" or a "classical" college. Practical colleges are land grant colleges or technical schools. Classical colleges are normal schools and other private and public colleges. Land grant colleges were required by law to provide instruction on "agricultural and mechanical arts", and technical colleges explicitly focused on skills such as engineering, mining, or industry. At the same time, normal schools trained public school teachers, and so typically devoted less, if any, attention to technical skills. Other public and private colleges tended to have a less practical focus, providing instruction in classes like the classics or Latin.[22] The results are presented in Column 4 of Table 5. It does appear

[22]For some types of colleges, there is much more ambiguity regarding whether or not the college should be classified as practical or classical. In Appendix D.1 I use alternative classifications of practical and classical

that practical colleges increase patenting by more than classical colleges, but the difference between the two coefficients is modest and not statistically different from zero. The practical colleges caused 39% more patents per year (statistically significant at the 5% level), while the classical colleges caused 35% more patents per year (statistically significant at the 10% level).

In Column 5, I exclude all years after 1940 from the data. This is because following the explosion in demand for higher education after World War II, the curricula across colleges largely began to converge; see Alon (2017) for details on this convergence. In the pre-1940 years, the difference in estimated coefficients between practical and classical colleges is again very small, although both coefficients are smaller than in Column 4: practical colleges saw about 15% more patents per year relative to their runner-up counties, while classical colleges saw 7% more; neither coefficient is individually statistically significant.

One possible reason why patenting appears larger in the practical colleges is because they may be larger schools than the average classical college, attracting more students, faculty, and general economic activity to the county. To check this, in Column 6 I re-estimate the effect of practical and classical colleges but include a control for the logged number of students, using data on student populations from the Commissioner of Education reports described in Section 2.2. After controlling for the number of students, the coefficients are no longer statistically significant but are not much changed in magnitude from Column 4, although the gap between the practical and classical colleges is smaller. In Column 7, I again exclude all post-1940 data while controlling for logged student population. Again neither coefficient is statistically significant, but in this case the classical colleges actually have a larger point estimate than the practical colleges.

In sum, in no cases is the increase in patenting in practical college counties substantially larger than the increase in patenting in classical colleges. The two coefficients are never statistically different from one another, and depending on the sample of years used and whether or not student population is controlled for, the coefficient for classical colleges may even be larger than that for practical colleges. In short, there is no evidence that the types of colleges that focus on technical skills produce a much larger increase in local invention.

colleges and show that the results are similar.

Appendix D.1 explores differences between types of colleges in more detail.

5 How Do Colleges Affect Patenting?

In this section, I reconcile the results in Sections 3 and 4 by exploring the channels by which colleges increase local patenting. Since all of the above results are consistent with colleges affecting local patenting through their effect on the local population, I begin by re-estimating the baseline regressions while including controls for county population. I next directly observe how many patents come from migrants and how many come from a college's alumni and faculty by matching the patent data to census and college yearbook data.

5.1 Controlling for County Population

A growing literature is recognizing the importance of migration in explaining changing geographic patterns of invention over time. Aghion et al. (2009) show both theoretically and empirically using political shocks that highly educated people are likely to migrate to live close to other highly educated people. Moretti and Wilson (2014) show that while state subsidies can be effective at increasing the number of local star scientists, the observed effects are largely driven by relocation. Several other papers find that inventors are highly mobile and respond to changes in local conditions (for example, Kerr and Lincoln (2010), Akcigit, Baslandze, and Stantcheva (2016), Akcigit, Grigsby, and Nicholas (2017b), Moretti and Wilson (2017)). A separate strand of the literature studies intra-academia knowledge production and concludes that, while it is difficult to detect productivity spillovers from one researcher to another located in the same department, the presence of a star scientist helps to attract other productive scientists (Azoulay, Graff-Zivin, and Wang (2010), Waldinger (2012), Borjas and Doran (2012), Dubois, Rochet, and Schlenker (2014)).

To examine whether or not migration is the primary driver of the observed results in this study, I re-estimate Equation (1) while controlling for a changing population. Because population variables are collected from the decennial U.S. population censuses, in this section I restrict the data to observations that occur only in the census years: 1840, 1850, 1860, etc. Thus the "time" variable no longer represents the number of years since a college site selection experiment, but rather the number of decades. In Column 1 of Table 6, I reproduce the baseline result on patenting using only patenting in census years. The estimated coefficient is almost identical to the baseline coefficient estimated in Column 1 of Table 3. Column 2 estimates the effect of a new college on logged county population. I find that college counties are about 38% larger than the runner-up counties after establishing a college; this effect is statistically significant at the 10% level. As the percentage increase in population from Column 2 is close to the percentage increase in patenting in Column 1, it is intuitively not surprising that population explains most of the observed change in patenting.

In Columns 3 and 4, I re-estimate Equation (1) but include controls for county population, using two alternative functional forms. In Column 3, I include $log(TotalPop)$ as a control. Not surprisingly, county population is highly predictive of county patenting (a one percent increase in population leads to a .4% increase in patenting). When including $log(TotalPop)$, the coefficient on the interaction term of interest is only 41% of the baseline estimate, decreasing from 40% more patents per year in the baseline to a statistically insignificant 18% more patents per year. In Column 4, I include $TotalPop$ and $(TotalPop)^2$ as controls instead of $log(TotalPop)$, scaling both variables by 10,000. An addition of 100,000 people increases patenting by about 1.5%; the quadratic term is negative but extremely small in magnitude. As in Column 3, after controlling for $TotalPop$ and $(TotalPop)^2$, the interaction term is smaller than in Column 1 and statistically insignificant: colleges have about 14% more patents per year than the runners-up, or about one-third of the original estimate. In sum, controlling for population explains 60% to two-thirds of the observed increase in patenting, and I cannot reject the null that colleges have no effect on patenting in college counties relative to the runners-up.

If the presence of knowledge spillovers make agglomeration economies especially large near colleges, then a marginal increase in population should have a larger effect on patenting in college counties than in runner-up counties. Formally, I estimate

$$\begin{aligned}
\log(NumPat_{ijt}+1) = \sum_{k\in K} \Big[& \delta_{1k} College_{ij} * PostCollege_{jt} * Pop^k_{it} \\
& + \delta_{2k} PostCollege_{jt} * Pop^k_{it} + \delta_{3k} Pop^k it \Big] \\
& + \delta_3 College_{ij} * PostCollege_{jt} + \delta_4 PostCollege_{jt} \\
& + \alpha_i + \lambda_j + \alpha_i * \lambda_j + \gamma_t + \epsilon_{ijt}, \qquad (3)
\end{aligned}$$

where $K = \{\log(Pop.)\}$ or $\{Pop., Pop^2\}$.

Results are presented in Columns 5 and 6 of Table 6. For readability, I omit the δ_{2k} estimates. There is no evidence that agglomeration economies are larger in college counties. In Column 5, I use $\log(Pop.)$ as the measure of population. A 10% in population actually reduces the difference-in-differences estimate by a statistically insignificant 0.5%. In Column 6, I use $Pop.$ and Pop^2 as measures of population and obtain similar results.

5.2 Who Do the Patents Come From?

In this section, I use the names of patentees to directly check what share of patents in college counties are invented by individuals with a direct affiliation with the college, either as alumni or faculty, and which are migrants to the college county. Finding that a large share of patents are from migrants, while only a small share are from alumni or faculty, is additional evidence that the primary channel through which colleges affect innovation is by encouraging migration.[23] I next describe the datasets used to determine which patents come from which groups, and then describe how the datasets are matched to one another and the obtaining results.

[23] The opposite finding, that a large share of patents come from alumni and faculty of the focal college, while suggestive, would still not be strong evidence that human capital or academic research channels are economically meaningful; what is a needed is a counterfactual estimate of how many patents the alumni and faculty would obtain in the absence of the college, which is much more difficult to estimate. Nevertheless, the results in this section can falsify the claim that human capital and faculty research are major channels through which a college promotes local invention.

5.2.1 College Yearbooks

The prior literature has struggled to convincingly identify the channels through which colleges promote local patenting. Dating to Jaffe (1989), the literature typically observes how patenting by firms co-located with a university co-vary with some university activity, and if any change in firm patenting is observed, this is counted as a "knowledge spillover" from the university to these firms. The issue, as noted by Zucker, Darby, and Armstrong (1998) and Leten, Landoni, and Looy (2014), is that changes in nearby patenting may not be true spillovers if the individuals within the firm have a direct affiliation with the college, for instance as either alumni or as consulting faculty members. Data issues usually prevent researchers from discovering which individuals have a direct affiliation with a college or university, however.[24] To overcome this problem, I construct a novel dataset of historical college yearbooks, which contain names of students (who go on to become alumni) and faculty members, which can then be linked to the patent data.

The yearbook data are described in much more detail in Appendix E. The college yearbooks are available from `ancestry.com` and contain full scans of historical college yearbooks, which include full student names, which can be used to match students from yearbooks to other data sources such as the patent record or the US decennial censuses. The college yearbooks also contain a wealth of other interesting information, including students' majors, sports and clubs, and fraternities and sororities. The yearbooks usually include names of faculty members as well. I collect yearbooks from 20 different colleges, roughly 28% of the colleges in my sample, covering 249 yearbooks from 1879 to 1940 and including records for 70,106 undergraduate seniors and 28,743 faculty members.[25]

To compile a list of candidate alumni to match to the decennial census records, for each college and each year, I combine the names of all that college's seniors from the previous 60 years worth of yearbooks. Essentially, this assumes that alumni who graduate at age 20 may

[24]Several recent papers document that college-educated individuals are more likely to invent (Jung and Ejermo (2014), Bell, Chetty, Jaravel, Petkova, and Van Reenen (2017), Aghion, Akcigit, Hyytinen, and Toivanen (2017), Akcigit, Grigsby, and Nicholas (2017a)), but they cannot link a college graduate to a *particular* college. What is needed to estimate local spillovers from, say, the University of North Dakota, is a way to determine if an inventor in Grand Forks, ND is an alumnus of the University of North Dakota.

[25]Note that these are not *unique* faculty members; in most cases the same faculty member is listed in multiple yearbook years. In rare cases, yearbooks do not record names of college seniors but do list names of juniors. In these cases, I record the names of juniors instead.

obtain patents until they are 80 years old. Such an assumption appears innocuous, as studies conclude that very few inventors are older than 80.[26] To compile a list of candidate faculty members, I use the names of the faculty members listed in the current year's yearbook for each college.

5.2.2 Matching to the U.S. Census Data

To determine which patentees are alumni or faculty and which are migrants to a college county, I merge both the patent and yearbook data to the U.S. 100% decennial population census records, transcribed by `ancestry.com` and the Minnesota Population Center and hosted by the NBER. I proceed in five steps.

First, I prepare the census data for each census from 1850 to 1940. The 1890 census manuscripts were destroyed by fire, so I am left with nine censuses. I restrict attention to males.[27] For each county in census, I then fuzzy match by first and last name to the same county in the previous census, using a matching procedure similar to Ferrie (1996) but including common names.[28] Doing this for all censuses allows me to identify the earliest year in which a particular name appears in a particular county; individuals who first appear in a college or runner-up county after the college is established are potential migrants.

Second, I match by first name, last name, state, and county from the patent record to the census record. This creates a list of all patents in each county for which personal information about the patentees can be known. See Sarada et al. (2017) for more details on the patent-census matching procedure. To minimize concerns about individuals moving from their census-recorded locations, I only match in the full years before and after the census is enumerated, as well as the years for which the census is enumerated. Typically, censuses are enumerated in the year before, and occasionally during the year that, they are released. So, for example, for the 1900 census, I match patentees from 1898, 1899, 1900, and 1901.

[26]For examinations of the age distribution of U.S. inventors prior to 1940, see Sarada et al. (2017) and Akcigit et al. (2017a). Papers that document that ages of more recent inventors include Jones (2009), Jung and Ejermo (2014), and Acemoglu, Akcigit, and Celik (2014).

[27]I restrict attention to males for two reasons. First, women are likely to change their names between the time they show up in the yearbook data and when they patent later in life. Second, the majority of women were not a part of the labor force during the sample period, and so occupational scores are not informative for them.

[28]See Appendix E for more details on the matching procedure.

Third, I match the lists of potential alumni and current faculty to the census, again matching on first name, last name, county, and state. I again match yearbooks only from the full years before and after the census is enumerated and the years during which it is enumerated.

Fourth, I use the matched census-patent-yearbook data to determine which patentees are alumni, faculty, and migrants. A patentee is recorded as an alumnus if there is a positive match between the individual's name from the alumni list and a name in the same county in the census and that individual is also linked to a patent. An individual is recorded as a faculty member if the individual is not recorded as an alumnus and there is a positive match between his name from the faculty list and a patent-matched name in the same county in the census. An individual is recorded as a migrant if the individual is neither an alumnus nor a faculty member and he does not match to a name in the same county in a prior census before the college was established. All other patent-matched names are assumed to be non-migrants, that is, individuals who lived in the county prior to the establishment of the new college and who are neither alumni nor faculty.

Fifth and finally, an adjustment must be made because yearbook data are not available for all years. This means that the list of potential alumni is too small. To correct for this, I interpolate the number of students attending the college in the years in between collected yearbooks. I then increase the size of the potential alumni list by that number of students for each successive year. Using the matched alumni patentees, I calculate an alumnus patenting rate. I then multiply the size of the new potential alumni list by the calculated alumni patenting rate to get the corrected number of patents by alumni, decreasing patent counts by migrants by the corresponding increase.

This procedure therefore gives the share of patents in a college county coming from alumni, faculty, migrants, and non-migrants, after correcting for missing yearbook data. Note that these categories are constructed so as to be exhaustive and mutually exclusive. As Bailey, Cole, Henderson, and Massey (2018) show, such a simple fuzzy matching procedure can produce a large number of false positive matches. Thus, these results likely to overstate the share of patents belonging to alumni and faculty and, because an individual is recorded as a migrant only if his name fails to match to the census prior to the college establishment, to understate the share of migrants and overstate the share of non-migrants.

5.2.3 Results

After matching to the census, I find that 14.6% of patents in college county belong to alumni of that college. 4.9% of the patents belong to faculty members of the college. Non-migrants, that is, individuals who were living in the college county at the time the college was established and were not directly affiliated with the college, account for 7.3% of patents. Finally, migrants to the college county account for 73.2% of patents. Consistent with changes in population explaining the bulk of the increase in patenting, migrants are by far the largest group of patentees.

As mentioned above, these results likely overstate the increase in total patenting caused by alumni and faculty and understate patents by migrants. These results are additionally an overstatement of the direct effect of colleges if particularly intelligent, creative, or driven individuals are both more likely to attend college and more likely to invent independently of education, as seems plausible. In contrast, Bianchi and Giorcelli (2017) argue that attending college causes talented individuals to go into careers like public administration that patent at low rates. Understanding the causal effect of college attendance on patenting for individuals at different points in the skill distribution is therefore an important topic for future work.

6 Discussion and Conclusion

In this paper, I document that establishing a new college causes 40% more patents per year in counties that receive a new college. At the same time, colleges are not a necessary condition to promote local invention, as other institutions increase local patenting by similar amounts. It should not be surprising that it is difficult to determine which policies are necessary to create "the next Silicon Valley,", as the historiography of Silicon Valley itself does not reach a consensus on the role of Stanford University and the University of California in developing the region. While some authors argue that these universities, and Stanford in particular, played a crucial role (Lowen, 1997), others give the credit to spending by the federal government (Richards (1990), Lécuyer (2007)). Gordon Moore credits luck with the fact that a few key anchor firms located in the what would become Silicon Valley (Moore & Davis, 2001).[29] It is likely that each of these stories played an important role by helping to create anchor institutions: I show that a larger population can explain most of the observed increase in patenting, but which type of institution acts as an anchor to attract population does not make much of a difference.

Three important caveats to this study are important to keep in mind. First, while I find no evidence that colleges had any effect on local invention after controlling population, it is possible that the nature of invention has changed since the colleges in my sample were established. It may be more important for researchers today "to be physically close to frontier academic research in order to remain on the cutting edge" (Moretti, 2012, p. 182) than in decades past. Certainly the share of patentees with a college degree has increased over time, as has the share of patents assigned to universities (Mowery and Sampat (2001), Mowery and Ziedonis (2002), Sampat (2006), P. Lee (2013)). And the results in Figures 3 and 4 show that the difference between college and runner-up counties increased in recent decades, following massive increases in federal funding of higher education. But even today, college graduates are highly mobile (Bound, Groen, Gézdi, and Turner (2004), Sumell, Stephan, and Adams (2008), Zolas et al. (2015)) so colleges may still find it difficult to retain their alumni to become innovation hubs. Moreover, Figure 5 shows that there has been no increase in the difference between college and consolation prize counties over time, so other areas have also been growing in importance. Perhaps one change is that colleges have become increasingly aware of their roles as anchor institutions. For instance, O'Mara (2005, p. 4) argues that, during the Cold War era, university expansion "was actually a process of city building" and that university administrators embraced the role of "urban planner and political actor" (p. 2). Further research is needed to understand precisely how the relationship between higher education and invention has changed over time.

Second, while I find that colleges played no role in promoting local invention beyond their effects on population, this does not mean that colleges had no *global* effect on invention. Because alumni are so mobile, it is possible, even likely, that alumni left the counties of

[29] These descriptions are, of course, a simplification. Lowen (1997), Lécuyer (2007), and to some extent Moore and Davis (2001) all give nuanced accounts of various factors that helped to shape Silicon Valley as it exists today.

their alma maters to pioneer breakthrough innovations in other locations. Anecdotes of entrepreneurs leaving places like Urbana-Champaign, IL and Madison, WI to find success in Silicon Valley show that this occurs, although it is unclear whether it is more or less frequent now than in earlier decades. In future work, I plan to track college alumni across time and space using the decennial census data to determine where alumni move after they graduate and where, and if, they invent.

Finally, it is important to note that promoting innovation is clearly not the only, nor even perhaps the primary, purpose of colleges and universities. Nevertheless, to the extent that policymakers wish to create inventive hubs, the results in this paper suggest that attracting migrants, rather than building colleges per se, is the policy to pursue.

References

Acemoglu, D., Akcigit, U., & Celik, M. A. (2014). *Young, restless and creative: openness to disruption and creative innovation.* (NBER Working Paper)

Aghion, P., Akcigit, U., Hyytinen, A., & Toivanen, O. (2017). *The social origins of inventors.* (Unpublished, University of Chicago)

Aghion, P., Boustan, L., Hoxby, C., & Vandenbussche, J. (2009). The causal impact of education on economic growth: evidence from the United States. *Brookings Papers on Economic Activity*, 1-73.

Aizer, A., Eli, S., Ferrie, J., & Lleras-Muney, A. (2016, April). The long-run impact of cash transfers to poor families. *American Economic Review*, *106*(4), 935-971.

Akcigit, U., Baslandze, S., & Stantcheva, S. (2016, October). Taxation and the international mobility of inventors. *American Economic Review*, *106*(10), 2930-2981.

Akcigit, U., Grigsby, J., & Nicholas, T. (2017a). *The birth of American ingenuity: innovation and inventors of the golden age.* (Unpublished, University of Chicago)

Akcigit, U., Grigsby, J., & Nicholas, T. (2017b, May). Immigration and the rise of American ingenuity. *American Economic Review*, *107*(5), 327-331.

Alon, T. (2017). *Schools, skills, and curricula in the 20th century.* (Unpublished, Northwestern University)

Andersson, R., Quigley, J. M., & Wilhelmsson, M. (2004, August). University decentralization as regional policy: the Swedish experiment. *Journal of Economic Geography*, *4*(4), 371-388.

Andersson, R., Quigley, J. M., & Wilhelmsson, M. (2009, July). Urbanization, productivity, and innovation: evidence from investment in higher education. *Journal of Urban Economics*, *66*(1), 2-15.

Andrews, M. (2017). *Comparing historical patent datasets.* (Unpublished, University of Iowa)

Atack, J., Jaremski, M., & Rousseau, P. L. (2014, December). American banking and the transportation revolution before the Civil War. *Journal of Economic History*, *74*(4), 943-986.

Azoulay, P., Graff-Zivin, J., & Wang, J. (2010, May). Superstar extinction. *Quarterly Journal of Economics*, *125*(2), 549-589.

Bailey, M., Cole, C., Henderson, M., & Massey, C. (2018). *How well do automated linking methods perform in historical samples? Evidence from new ground truth.* (Unpublished, University of Michigan)

Belenzon, S., & Schankerman, M. (2013, July). Spreading the word: geography, policy, and knowledge spillovers. *Review of Economics and Statistics*, *95*(3), 884-903.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., & Van Reenen, J. (2017). *Who becomes an inventor in America? The importance of exposure to innovation.* (Unpublished, Harvard University)

Berkes, E. (2018). *Comprehensive universe of U.S. patents (CUSP): data and facts.* (Unpublished, Northwestern University)

Bianchi, N., & Giorcelli, M. (2017). *Scientific education and innovation: from technical diplomas to university STEM degrees.* (Unpublished, Northwestern University)

Blundell, R., Griffith, R., & Reenen, J. V. (1995, March). Dynamic count data models of technological innovation. *The Economic Journal*, *105*(429), 333-344.

Bonander, C., Jakobsson, N., Podestà, F., & Svensson, M. (2016, September). Universities as engines for regional growth? Using the synthetic control method to analyze the effects of research universities. *Regional Science and Urban Economics*, *60*, 198-207.

Borgschulte, M. (2014). *The effect of presidential service on life expectancy.* (Unpublished, University of California Berkeley)

Borgschulte, M., & Vogler, J. (2016). *The effect of executive service on life expectancy: evidence from US governors.* (Unpublished, University of Illinois)

Borjas, G. J., & Doran, K. B. (2012, August). The collapse of the Sovier Union and the productivity of American mathematicians. *Quarterly Journal of Economics*, *127*(3), 1143-1203.

Bound, J. (1989, June). The health and earnings of rejected disability insurance applicants. *American Economic Review*, *79*(3), 482-503.

Bound, J., Groen, J., Gézdi, G., & Turner, S. (2004, July-August). Trade in university training: cross-state variation in the production and stock of college-educated labor.

Journal of Econometrics, *121*(1-2), 143-173.

Busso, M., Gregory, J., & Kline, P. (2013, April). Assessing the incidence and efficiency of a prominent place based policy. *American Economic Review*, *103*(2), 897-947.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011, April). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, *29*(2), 238-249.

Cameron, C. A., & Miller, D. L. (2015, Spring). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, *50*(2), 317-372.

Cline, P. (1983). *Mountain campus: the story of Northern Arizona University*. Flagstaff, AZ: Northland Press.

Cowan, R., & Zinovyeva, N. (2013, April). University effects of regional innovation. *Research Policy*, *42*(3), 788-800.

Crescenzi, R., Nathan, M., & Rodgríguez-Pose, A. (2016, February). Do inventors talk to strangers? On proximity and collaborative knowledge creation. *Research Policy*, *45*(1), 177-194.

Currie, J., & Moretti, E. (2003, November). Mother's education and the intergenerational transmission of human capital: evidence from college openings. *Quarterly Journal of Economics*, *118*(4), 1495-1532.

Dale, S. B., & Krueger, A. B. (2002, November). Estimating the payoff to attending a more selective college: an application of selection on observables and unobservables. *Quarterly Journal of Economics*, *117*(4), 1491-1527.

Davis, B. (2016, December 12). There's an antidote to America's long economic malaise: college towns. *The Wall Street Journal*. (`https://www.wsj.com/articles/theres-an-antidote-to-americas-long-economic-malaise-college-towns-1481558522?ns=prod/accounts-wsj`, accessed July 5, 2018)

Dubois, P., Rochet, J.-C., & Schlenker, J.-M. (2014). Productivity and mobility in academic research: evidence from mathematicians. *Scientometrics*, *98*, 1669-1701.

Ferrie, J. P. (1996, Fall). A new sample of males linked from the public use microdata sample of the 1850 US federal census of population to the 1860 US federal census manuscript schedules. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, *29*(4), 141-156.

Ferrier, W. W. (1930). *Origin and development of the University of California.* Berkeley, CA: The Sather Gate Book Shop.

Florida, R. (2002). *The rise of the creative class: and how it's transforming work, leisure, community and everday life.* New York: Basic Books.

Forbes Technology Council. (2017, October 17). Where's the next Silicon Valley? *Forbes*. (`https://www.forbes.com/sites/forbestechcouncil/2017/10/17/wheres-the-next-silicon-valley/#31d5ddab37f0`, accessed July 4, 2018)

Frenette, M. (2009, June). Do universities benefit local youth? Evidence from the creation of new universities. *Economics of Education Review*, *28*, 318-328.

Furman, J. L., & MacGarvie, M. J. (2007, August). Academic science and the birth of industrial research laboratories in the U.S. pharmaceutical industry. *Journal of Economic Behavior & Organization*, *63*(4), 756-776.

Geiger, L. G. (1958). *University of the northern plains: a history of the University of North Dakota, 1883-1958.* Grand Forks, ND: The University of North Dakota Press.

Goldin, C., & Katz, L. F. (1999, Winter). The shaping of higher education: the formative years in the United States, 1890 to 1940. *Journal of Economic Perspectives*, *13*(1), 37-62.

Goldin, C., & Katz, L. F. (2008). *The race between education and technology.* Cambridge, MA: The Belknap Press of the Harvard University Press.

Gordon, R. J. (2016). *The rise and fall of American growth: the U.S. standard of living since the Civil War.* Princeton, NJ: Princeton University Press.

Greenstone, M., Hornbeck, R., & Moretti, E. (2010, June). Identifying agglomeration spillovers: evidence from winners and losers of large plant openings. *Journal of Political Economy*, *118*(3), 536-598.

Griliches, Z. (1990, December). Patent statistics as economic indicators: a survey. *Journal of Economic Literature*, *28*(4), 1661-1707.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005, Spring). Market value and patent citations. *RAND Journal of Economics*, *36*(1), 16-38.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). *The NBER patent citation data file: lessons, insights and methodological tools.* (NBER Working Paper)

Hausman, N. (2017). *University innovation and local economic growth.* (Unpublished, Hebrew University)

Helmers, C., & Overman, H. G. (2017). My precious! The location and diffusion of scientific research: evidence from the Synchrotron Diamond Light Source. *The Economic Journal*, *Forthcoming*.

Jaffe, A. B. (1989, Dec). Real effects of academic research. *American Economic Review*, *79*(5), 957-970.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993, August). Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics*, *108*(3), 577-598.

Jones, B. F. (2009, January). The burden of knowledge and the "death of the renaissance man": is innovation getting harder? *Review of Economic Studies*, *76*(1), 283-317.

Jones, B. F. (2010, February). Age and great invention. *Review of Economics and Statistics*, *92*(1), 1-14.

Jung, T., & Ejermo, O. (2014, July). Demographic patterns and trends in patenting: gender, age, and education of inventors. *Technological Forecasting and Social Change*, *86*, 110-124.

Kantor, S., & Whalley, A. (2014, March). Knowledge spillovers from research universities: evidence from endowment value shocks. *Review of Economics and Statistics*, *96*(1), 171-188.

Kantor, S., & Whalley, A. (2018). Research proximity and productivity: long-term evidence from agriculture. *Journal of Political Economy*, *Forthcoming*.

Kerr, W., & Lincoln, W. (2010, July). The supply side of innovation: H-1B visa reforms and US ethnic invention. *Journal of Labor Economics*, *28*(3), 473-508.

Khan, B. Z. (2005). *The democratization of invention: patents and copyrights in American economic development, 1790-1920.* New York: Cambridge University Press.

Kuhn, J. M., & Thompson, N. (2017). *The ways we've been measuring patent scope are wrong: how to measure and draw causal inferences with patent scope.* (Unpublished, MIT)

Kuhn, J. M., Younge, K. A., & Marco, A. C. (2017). *Patent citations reexamined: new data* *and methods.* (Unpublished, UNC)

Lécuyer, C. (2007). *Making Silicon Valley: innovation and the growth of high tech, 1930-1970.* Cambridge, MA: MIT Press.

Lee, J. (2018). *The local economic impact of a large research university: evidence from UC Merced.* (Unpublished, Korea Development Institute)

Lee, P. (2013, October). Patents and the university. *Duke Law Journal*, *63*(1), 1-87.

Leggett, M. D. (1874). *Subject-matter index of patents for invention issued by the United States Patent Office from 1790 to 1873, inclusive* (Vol. 1-3). Washington, DC: Government Printing Office.

Leten, B., Landoni, P., & Looy, B. V. (2014, October). Science or graduates: how do firms benefit from the proximity of universities? *Research Policy*, *43*(8), 1398-1412.

Li, G.-C., Lai, R., D'Amour, A., Doolin, D. M., Sun, Y., Torvik, V. I., ... Fleming, L. (2014, July). Disambiguation and co-authorship networks of the U.S. Patent Inventor Database (1975-2010). *Research Policy*, *43*(6), 941-955.

Liu, S. (2015, May). Spillovers from universities: evidence from the land-grant program. *Journal of Urban Economics*, *87*, 25-41.

Lowen, R. S. (1997). *Creating the Cold War university: the transformation of Stanford.* Berkeley, CA: University of California Press.

Malmendier, U., Moretti, E., & Peters, F. (2016). *Winning by losing: evidence on the long-run effects of mergers.* (Unpublished, University of California Berkeley)

Manson, S., Schroeder, J., Riper, D. V., & Ruggles, S. (2017). *IPUMS national historical geographic information system: version 12.0.* Minneapolis, MN: University of Minnesota. (`http://doi.org/10.18128/D050.V12.0`)

Marco, A. C., Carley, M., Jackson, S., & Myers, A. F. (2015, June). *The USPTO historical patent data files: two centuries of invention.* (Unpublished, USPTO Economic Working Papter No. 2015-1)

Martin, D. D. (1960). *The lamp in the desert: the story of the University of Arizona.* Tucson, AZ: The University of Arizona Press.

McMath Jr., R. C., Bayor, R. H., Brittain, J. E., Foster, L., Giebelhaus, A. W., & Reed, G. M. (1985). *Engineering the new south: Georgia Tech, 1885-1985.* Athens, GA: The

University of Georgia Press.

Miller, J. (2016a). *Jim Shaw historical patent data.* (electronic correspondence, October 24, 2016)

Miller, J. (2016b). *Jim Shaw historical patent data.* (electronic correspondence, November 1, 2016)

Moore, G., & Davis, K. (2001). *Learning the Silicon Valley way.* (Unpublished, Stanford Institute for Economic Policy Research)

Moretti, E. (2004, July-August). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*, *121*(1-2), 175-212.

Moretti, E. (2012). *The new geography of jobs.* New York: Houghton Mifflin Harcourt Publishing Company.

Moretti, E., & Wilson, D. J. (2014, January). State incentives for innovation, star scientists and jobs: evidence from biotech. *Journal of Urban Economics*, *79*, 20-38.

Moretti, E., & Wilson, D. J. (2017, July). The effect of state taxes on the geographical location of top earners: evidence from star scientists. *American Economic Review*, *107*(7), 1858-1903.

Mowery, D. C., & Sampat, B. N. (2001, August). University patents and patent policy debates in the USA, 1925-1980. *Industrial and Corporate Change*, *10*(3), 781-814.

Mowery, D. C., & Ziedonis, A. A. (2002, March). Academic patent quality and quantity before and after the Bayh-Dole Act in the United States. *Research Policy*, *31*(3), 399-418.

Nagaoka, S., Motohashi, K., & Goto, A. (2010). Patent statistics as an innovation indicator. In B. H. Hall & N. Rosenberg (Eds.), *Handbook of the economics of innovation* (Vol. 2, p. 1083-1127). Amsterdam: Elsevier.

Olenski, A. R., Abola, M. V., & Jena, A. B. (2015, December). Do heads of government age more quickly? Observational study comparing mortality between elected leaders and runners-up in national elections of 17 countries. *British Medical Journal*, *351*, 1-7.

Olshansky, S. J. (2011). Aging of US presidents. *The Journal of the American Medical Association*, *306*(21), 2328-2329.

O'Mara, M. P. (2005). *Cities of knowledge: Cold War science and the search for the next Silicon Valley.* Princeton, NJ: Princeton University Press.

Patrick, C. (2016, October). Identifying the local economic development effects of million dollar facilities. *Economic Inquiry*, *54*(4), 1737-1762.

Petralia, S., Balland, P.-A., & Rigby, D. L. (2016a, August). Data descriptor: unveiling the geography of historical patents in the united states from 1836-1975. *Nature: Scientific Data*, *3*, 1-14. (Article number: 160074)

Petralia, S., Balland, P.-A., & Rigby, D. L. (2016b). *Historical patent dataset.* Harvard Dataverse. (`http://dx.doi.org/10.7910/DVN/BPC15W`, accessed September 15, 2016)

Richards, E. (1990, April 29). Uncle Sam as venture capitalist. *The Washington Post*. (`https://www.washingtonpost.com/archive/business/1990/04/29/uncle-sam-as-venture-capitalist/e0848ef5-128c-4cea-b157-d2b18b01a1e8/?noredirect=on&utm_term=.535d0aa926cf`, accessed May 19, 2018)

Sampat, B. N. (2006, July). Patenting and US academic research in the 20th century: the world before and after Bayh-Dole. *Research Policy*, *35*(6), 772-789.

Sarada, Andrews, M., & Ziebarth, N. L. (2017). *The demographics of inventors in the historical United States.* (Unpublished, University of Iowa)

Siegfried, J. J., Sanderson, A. R., & McHenry, P. (2007, October). The economic impact of colleges and universities. *Economics of Education Review*, *26*(5), 546-558.

Sumell, A., Stephan, P., & Adams, J. (2008). Capturing knowledge: the location decision of new Ph.Ds working in industry. In *Science and engineering careers in the United States: an analysis of markets and employment* (p. 257-287). Chicago: University of Chicago Press.

Thompson, P. (2006). Patent citations and the geography of knowledge spillovers: evidence from inventor- and examiner-added citations. *Review of Economics and Statistics*.

Toivanen, O., & Väänänen, L. (2016, May). Education and invention. *Review of Economics and Statistics*, *98*(2), 382-396.

Trajtenberg, M. (1990, Spring). A penny for your quotes: patent citations and the value of innovations. *RAND Journal of Economics*, *21*(1), 172-187.

von Wachter, T., Song, J., & Manchester, J. (2011, December). Trends in employment and earnings of allowed and rejected applicants to the Social Security Disability Insurance Program. *American Economic Review*, *101*(7), 3308-3329.

Wagoner, J. J. (1970). *Arizona Territory, 1863-1912: a political history.* University of Arizona Press.

Waldinger, F. (2012, April). Peer effects in science - evidence from the dismissal of scientists in Nazi Germany. *Review of Economic Studies*, *79*(2), 838-861.

Washburn, W. E. (1979, July). Samuel Eliot Morison, historian. *The William and Mary Quarterly*, *36*(3), 325-352.

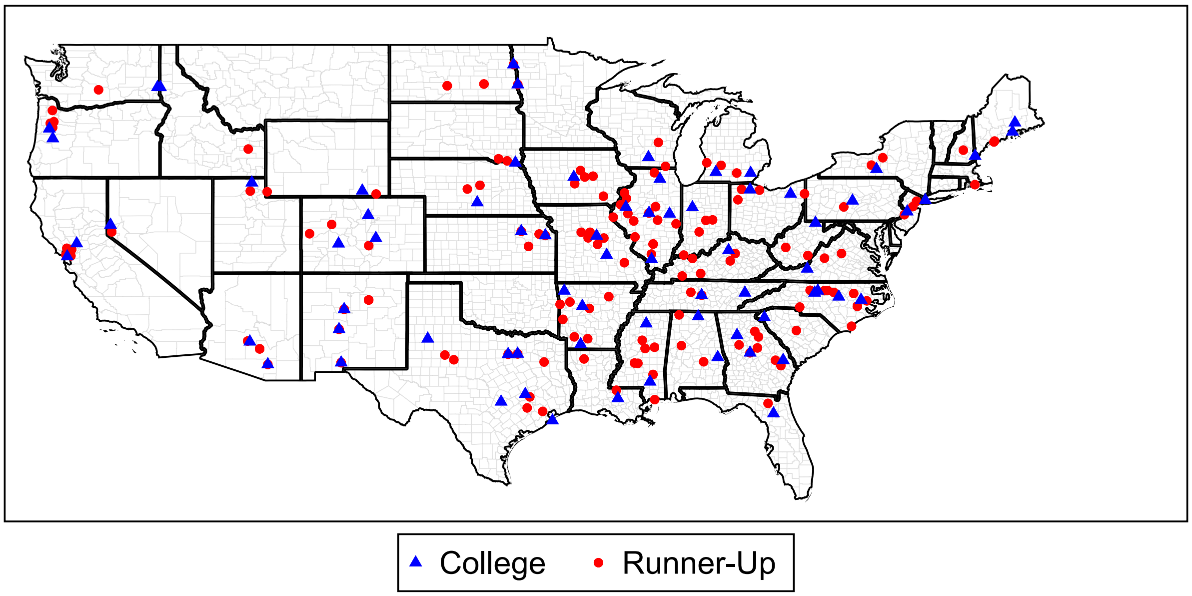
Xiong, H., & Zhao, Y. (2017). *Denominational competition and the market provision of human capital.* (Unpublished, Northwestern University)

Zolas, N., Goldschlag, N., Jarmin, R., Stephan, P., Owen-Smith, J., Rosen, R. F., ... Lane, J. I. (2015, December). Wrapping it up in a person: examining employment and earnings outcomes for Ph.D. recipients. *Science*, *350*(6266), 1367-1371.

Zucker, L. G., Darby, M. R., & Armstrong, J. (1998, January). Geographically localized knowledge: spillovers or markets? *Economic Inquiry*, *36*(1), 65-86.

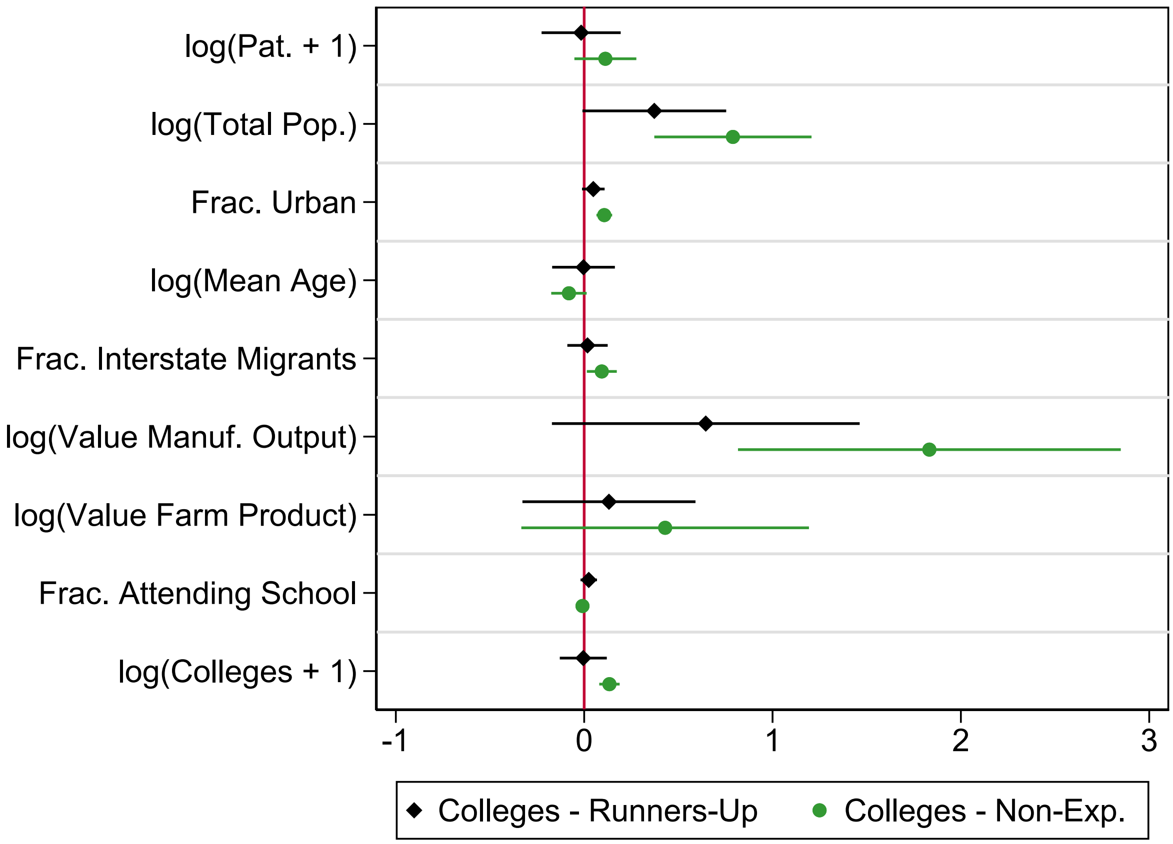
Graphs

Figure 1: Map of College and Runner-Up Sites



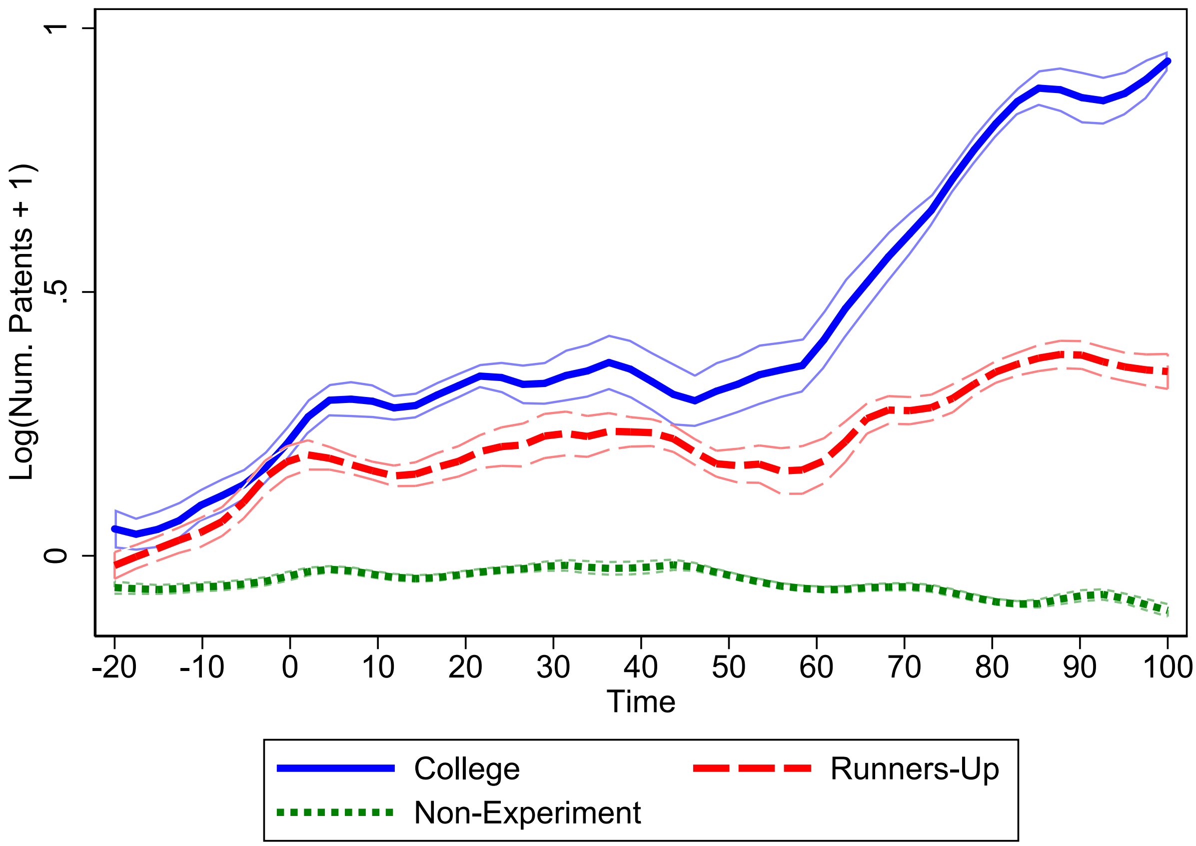
Notes: Map of the location of high quality college and runner-up sites in the sample. Colleges are represented by blue diamonds. The runner-up sites are represented by red circles.

Figure 2: Balance Checks



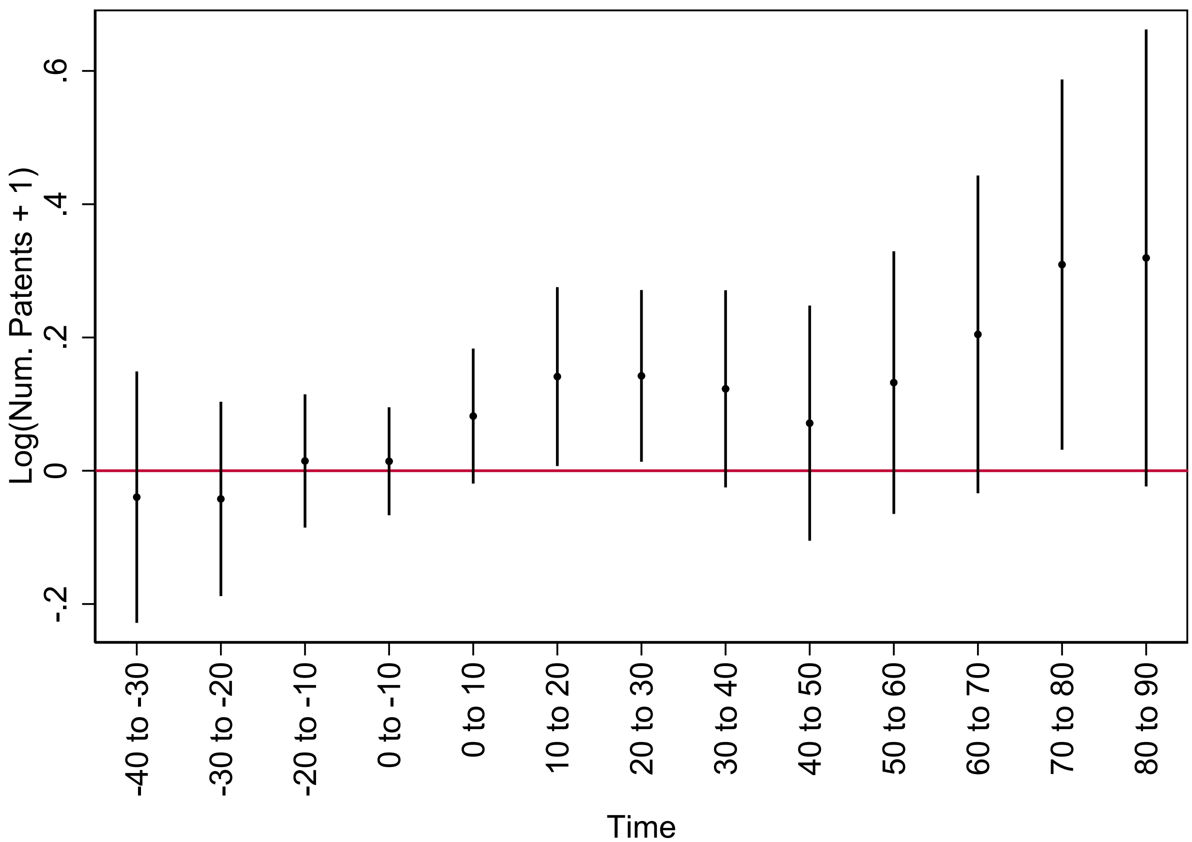
Notes: Results of t-tests comparing the means of the college counties, runner-up counties, and non-experimental counties. The black diamonds display the difference in the mean between the college and runner-up counties in the last U.S. census year before the college was established on a number of economic, demographic, and educational variables. The black lines show 95% confidence intervals of a simple t-test of the difference in means. The green circles display the difference in the mean between the college and non-experiment counties in the last U.S. census year before the college was established. The green lines show 95% confidence intervals.

Figure 3: Patenting in College, Runner-Up, and Non-Experimental Counties



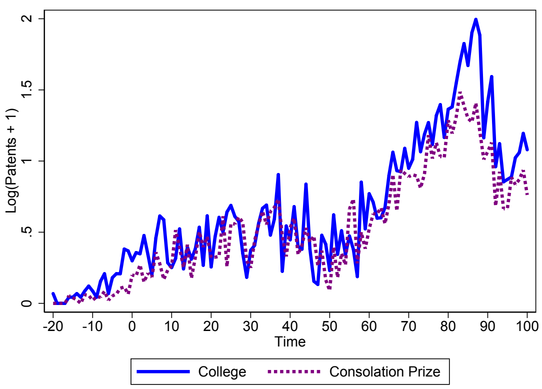
Notes: Mean patenting in college and runner-up counties after controlling for year effects. The x-axis shows the number of years since the college experiment. The year of the college experiment is normalized to year 0. Everything left of year 0 shows pre-college means; everything to the right shows post-college means. The y-axis shows smoothed $log(Patents+1)$. The smoothed patenting is constructed by regressing $log(Patents+1)$ on year effects and then plotting the residuals using local mean smoothing with an Epanechnikov kernel function. The college counties are represented by the blue solid line. The runner-up counties are represented by the red long-dashed line. The non-experimental counties are represented by the green short-dashed line. Data are for high quality experiments only.

Figure 4: Dynamics of Treatment Effect

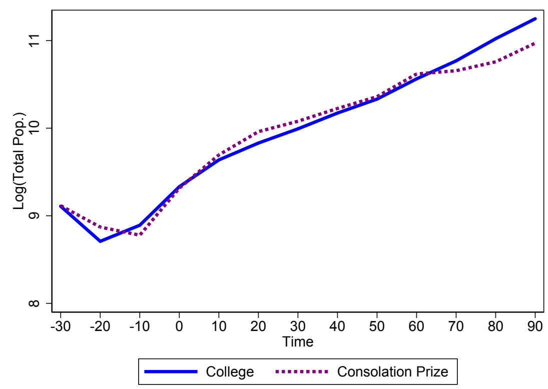


Notes: Estimated coefficient of the level shift in patenting in college counties relative to runner-up counties after establishment of a new college with a separate interaction term estimated for each time bin, along with 95% confidence bands. Time bins are are dummy variables that are equal to one for college counties in every ten year period before and after the establishment of the new college.

Figure 5: Patenting and Population in College and Consolation Prize Counties



(a) log(Num. Patents + 1)



(b) log(Total Pop.)

Notes: Logged patents and total population in the college and consolation prize counties. The x-axis shows the number of years since the college experiment. The year of the establishment of the new college is normalized to 0. Everything left of 0 shows pre-college results; everything to the right shows post-college results. In Panel (a), the y-axis shows log(Num. Patents + 1). In Panel (b), the y-axis shows $log(TotalPopulation)$. The college counties are represented by the blue solid line. The consolation prize counties are represented by the purple dashed line. Data are for high quality experiments only.

Tables

Table 1: List of College Site Selection Experiments

	College	Exp. Abbrev.	County	State	Experiment Year	College Type
1	University of Missouri	UMo1	Boone	Missouri	1839	Other Public
2	University of Mississippi	UMs	Lafayette	Mississippi	1841	Other Public
3	Eastern Michigan University	EMiU	Washtenaw	Michigan	1849	Normal School
4	Pennsylvania State University	PaSU1	Centre	Pennsylvania	1855	Land Grant
5	The College of New Jersey	TCNJ	Mercer	New Jersey	1855	Normal School
6	University of California Berkeley	UCaB	Alameda	California	1857	Land Grant
7	Iowa State University	IaSU	Story	Iowa	1859	Land Grant
8	University of South Dakota	USD	Clay	South Dakota	1862	Other Public
9	University of Kansas	UKs	Douglas	Kansas	1863	Other Public
10	Kansas State University	KsSU	Riley	Kansas	1863	Land Grant
11	Lincoln College (IL)	LincC	Logan	Illinois	1864	Other Private
12	Cornell University	CornU	Tompkins	New York	1865	Land Grant
13	University of Maine	UMe	Penobscot	Maine	1866	Land Grant
14	University of Wisconsin	UWi2	Dane	Wisconsin	1866	Land Grant
15	West Virginia University	WVU	Monongalia	West Virginia	1867	Land Grant
16	University of Illinois	UIUC	Champaign	Illinois	1867	Land Grant
17	Oregon State University	OrSU	Benton	Oregon	1868	Land Grant
18	Purdue University	PurdU	Tippecanoe	Indiana	1869	Land Grant
19	University of Tennessee	UTn2	Knox	Tennessee	1869	Land Grant
20	Southern Illinois University	SIlU	Jackson	Illinois	1869	Normal School
21	Louisiana State University	LaSU	East Baton Rouge	Louisiana	1870	Land Grant
22	Mercer University	MercU	Bibb	Georgia	1870	Other Private
23	Missouri University of Science and Technology	MoUST	Phelps	Missouri	1870	Technical School
24	Texas A and M University	TxAMU	Brazos	Texas	1871	Land Grant
25	University of Arkansas	UAr	Washington	Arkansas	1871	Land Grant
26	University of Oregon	UOr	Lane	Oregon	1872	Other Public
27	Auburn University	AubU	Lee	Alabama	1872	Land Grant
28	Virginia Polytechnic Institute	VaT	Montgomery	Virginia	1872	Land Grant
29	University of Colorado	UCo	Boulder	Colorado	1874	Other Public
30	University of Texas Medical Branch	UTxMB	Galveston	Texas	1881	Technical School
31	University of Texas Austin	UTxA	Travis	Texas	1881	Other Public
32	North Dakota State University	NDSU	Cass	North Dakota	1883	Land Grant
33	University of North Dakota	UND	Grand Forks	North Dakota	1883	Other Public
34	University of Nevada	UNv	Washoe	Nevada	1885	Land Grant
35	University of Arizona	UAz	Pima	Arizona	1885	Other Public
36	Arizona State University	AzSU	Maricopa	Arizona	1885	Land Grant
37	University of Wyoming	UWy	Albany	Wyoming	1886	Land Grant
38	North Carolina State University	NCSU	Wake	North Carolina	1886	Land Grant
39	Georgia Institute of Technology	GaT	Fulton	Georgia	1886	Technical School
40	Kentucky State University	KySU	Franklin	Kentucky	1886	HBCU
41	Utah State University	UtSU	Cache	Utah	1888	Land Grant
42	Clemson University	ClemU	Pickens	South Carolina	1889	Land Grant
43	University of Idaho	UId	Latah	Idaho	1889	Land Grant
44	New Mexico State University	NMSU	Dona Ana	New Mexico	1889	Land Grant
45	New Mexico Tech	NMT	Socorro	New Mexico	1889	Technical School
46	University of New Mexico	UNM	Bernalillo	New Mexico	1889	Other Public
47	University of New Hampshire	UNH	Strafford	New Hampshire	1891	Land Grant
48	Washington State University	WaSU	Whitman	Washington	1891	Land Grant
49	Alabama Agricultural and Mechanical University	AlAMU	Madison	Alabama	1891	HBCU
50	North Carolina A and T University	NCAT	Guilford	North Carolina	1892	HBCU
51	Northern Illinois University	NIlU	DeKalb	Illinois	1895	Normal School
52	Western Illinois University	WIlU	McDonough	Illinois	1899	Normal School
53	Western Michigan University	WMiU	Kalamazoo	Michigan	1903	Normal School
54	University of Nebraska at Kearney	UNeKe	Buffalo	Nebraska	1903	Normal School
55	University of Florida	UFl2	Alachua	Florida	1905	Land Grant
56	Georgia Southern College	GaSoU	Bulloch	Georgia	1906	Other Public
57	University of California Davis	UCaDav	Yolo	California	1906	Land Grant
58	East Carolina University	ENCU	Pitt	North Carolina	1907	Technical School
59	Western State Colorado University	WSCoU	Gunnison	Colorado	1909	Normal School
60	Middle Tennessee State University	MTnSU	Rutherford	Tennessee	1909	Normal School
61	Texas Christian University	TxCU	Tarrant	Texas	1910	Other Private
62	Bowling Green State University	BGSU	Wood	Ohio	1910	Normal School
63	Arkansas Tech University	ArTU	Pope	Arkansas	1910	Technical School
64	Kent State University	KentSU	Portage	Ohio	1910	Normal School
65	Southern Mississippi University	SMsU	Forrest	Mississippi	1910	Normal School
66	Southern Arkansas University	SArU	Columbia	Arkansas	1910	Other Public
67	Southern Methodist University	SMU	Dallas	Texas	1911	Other Private
68	High Point University	HPU	Guilford	North Carolina	1921	Other Private
69	Texas Tech	TxT	Lubbock	Texas	1923	Technical School
70	Maine Maritime Academy	MeMA	Hancock	Maine	1941	Military Academy
71	US Merchant Marine Academy	USMMA	Nassau	New York	1941	Military Academy
72	US Air Force Academy	USAFA	El Paso	Colorado	1954	Military Academy

Notes: List of all high quality college site selection experiment in the dataset in chronological order of the experiment date. Also included is the abbreviation of each experiment used in following results, the county and state of each college, the experiment year, and the college type of each experiment. The dates listed on this table are the date at which uncertainty over the college site location was resolved; these need not coincide with the official date of establishment for each college. In some cases, colleges have changed location, so the county listed need not be the current location or original location of the college. For colleges that changed location or were under consideration to change location, multiple experiments may be listed for the same college. For details on each site selection experiment, see the Historical Appendix.

Table 2: Summary Statistics of College Site Selection Experiments

	N	Mean	S.D.	Min	Median	Max
# Finalist Counties	72	2.83	1.45	1.00	3.00	7.00
Distance to Finalists	154	138.06	168.66	11.92	92.85	1,443.16
Experiment Year	72	1885.74	22.95	1839.00	1885.50	1954.00
Year of First Class	66	1887.53	26.15	1795.00	1889.50	1955.00
Year Desegregated	35	1941.80	28.97	1871.00	1953.00	1965.00
Year Co-Ed	45	1896.40	35.33	1804.00	1889.00	1976.00
Land Grant Colleges	72	0.42	0.50	0.00	0.00	1.00
Technical Schools	72	0.10	0.30	0.00	0.00	1.00
Normal Schools	72	0.17	0.38	0.00	0.00	1.00
HBCUs	72	0.04	0.20	0.00	0.00	1.00
Military Academies	72	0.04	0.20	0.00	0.00	1.00
Other Public Colleges	72	0.17	0.38	0.00	0.00	1.00
Other Private Colleges	72	0.07	0.26	0.00	0.00	1.00

Notes: Summary statistics for the high quality college site selection experiments. Column 1 lists the count of experiments or counties. Column 2 lists mean values, Column 3 the standard deviation, Column 4 the minimum value, Column 5 the median value, and Column 6 the maximum value. Row 1 lists the number of runner-up counties for each experiment. Row 2 lists the distance between college and runner-up sites. Row 3 lists the experiment year. Row 4 lists the year in which students began attending the college. Row 5 lists the year when the college became racially desegregated. Row 6 lists the year the college became coeducational. Rows 7-13 list the fraction of colleges that are of each college type.

Table 3: Baseline Regression Results

	log(Pat. +1)	log(Pat. + (Pat.^{2}+1)^0.5)	Alt. log(Pat.)	Neg. Binomial
Coll.County * PostColl.	0.401***	0.484***	0.264***	0.857**
	(0.143)	(0.177)	(0.100)	(0.382)
PostColl.	-0.032	-0.026	-0.084*	7.247***
	(0.062)	(0.074)	(0.047)	(1.596)
Zero Pat. Dummy			-0.748***	
			(0.013)	
County Fixed Effects	Yes	Yes	Yes	Yes
Exp. Fixed Effects	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Cnty-Year Obs.	39,580	39,580	39,580	39,580
# Counties	179	179	179	179
# Experiments	72	72	72	72
Mean in 1880	4.845	4.845	4.845	4.845
Adj. R-Sqr.	0.508	0.508	0.699	0.208
Log-Likelihood	-50,003.763	-56,897.077	-40,269.923	-63,080.296

Notes: Baseline regression results. Column 1 estimates the level shift in patenting in college counties relative to runner-up counties after establishment of a new college when the dependent variable is $log(Num.Patents+1)$. The dependent variable in column 2 is $log(Num.Patents)$, with values replaced with 0 if $Num.Patents = 0$ and a dummy variable for zero patents included. The dependent variable in column 3 is the number of patents. Column 4 presents results for a negative binomial regression. These results for columns 1-4 use high quality experiments only. Each coefficient is transformed into a percentage change in the dependent variable /100. When the model estimates changes in levels, the percentage change is calculated based on the baseline value of the independent variable in 1880. Results are for high quality experiments only. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table 4: Results with High and Low Quality College Site Selection Experiments

	log(Pat. +1)	log(Pat. + (Pat.^{2}+1)$^{0.5}$)	Alt. log(Pat.)
Coll.County * HighQual. * PostColl.	-0.326***	-0.341**	-0.329***
	(0.124)	(0.136)	(0.108)
Coll.County * PostColl.	1.086***	1.263***	0.931***
	(0.313)	(0.373)	(0.264)
HighQual * PostColl.	0.170***	0.215***	0.095**
	(0.064)	(0.079)	(0.046)
PostColl.	0.021***	0.027***	-0.019***
	(0.007)	(0.008)	(0.005)
Zero Pat. Dummy			-0.663***
			(0.005)
County Fixed Effects	Yes	Yes	Yes
Exp. Fixed Effects	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes
Cnty-Year Obs.	2,359,471	2,359,471	2,359,471
# Counties	1,958	1,958	1,958
# Experiments	179	179	179
Mean in 1880	4.845	4.845	4.845
Adj. R-Sqr.	0.570	0.568	0.718

Notes: Regression results using all of the college site selection experiments, including the low quality experiments and runner-up counties. Column 1 estimates the level shift in patenting in college counties relative to all runner-up counties after establishment of a new college when the dependent variable is $log(Num.Patents + 1)$. The dependent variable in column 2 is $log(Num.Patents)$, with values replaced with 0 if $Num.Patents = 0$ and a dummy variable for zero patents included. The dependent variable in column 3 is the number of patents. Each coefficient is transformed into a percentage change in the dependent variable /100. When the model estimates changes in levels, the percentage change is calculated based on the baseline value of the independent variable in 1880.Standard errors are clustered by county and shown in parentheses.Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table 5: Consolation Prize and College Type Results

	Cons. Prize	Cons. Prize Control for Pop.	No Cons. Prize	Practical vs. Classical Colleges	Pre-1940 Practical vs. Classical Colleges	Practical vs. Classical Colleges	Pre-1940 Practical vs. Classical Colleges
Coll.County * PostColl.	0.188	0.088	0.480***				
	(0.214)	(0.163)	(0.176)				
PostColl.	0.013	-0.150	-0.030				
	(0.135)	(0.109)	(0.072)				
log(Total Pop.)		0.253**					
		(0.109)					
Practical College Interaction				0.388**	0.147	0.369	0.348
				(0.187)	(0.091)	(0.720)	(0.744)
Classical College Interaction				0.348*	0.066	0.349	0.524
				(0.193)	(0.087)	(0.555)	(0.849)
log(Num. Students + 1)						0.009	0.005
						(0.083)	(0.083)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exp. Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cnty-Year Obs.	8,955	7,755	31,500	39,580	23,722	3,152	1,892
# Counties	34	34	150	179	179	19	19
# Experiments	16	16	60	72	72	19	19
Mean in 1880	4.845	4.845	4.845	4.845	4.845	4.845	4.845
Adj. R-Sqr.	0.490	0.500	0.517	0.508	0.450	0.574	0.401

Notes: Regression results under various assumptions about consolation prize counties and different types of colleges. Column 1 compares college counties to only runner-up counties that receive a consolation prize. Column 2 compares college counties to only runner-up counties that receive a consolation prize while controlling for $log(Population)$. Column 3 excludes all counties that receives a consolation prize and compares college counties to runner-up counties that do not receive a consolation prize. Column 4 includes all counties but shows results for practical colleges and classical colleges. Column 5 is identical to Column 4 but excludes all years after 1940. Column 6 is identical to Column 4 but includes a control for $log(Students+1)$ in each county. Column 7 is identical to Column 5 but includes a control for $log(Students+1)$ in each county. The dependent variable in all columns is $log(Patents+1)$. Each coefficient is transformed into a percentage change in the dependent variable /100. When the model estimates changes in levels, the percentage change is calculated based on the baseline value of the independent variable in 1880. Results are for high quality experiments only. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table 6: The Effect of Population on Patenting

	log(Pat. + 1)	log(Total Pop.)	log(Pat. + 1)	log(Pat. + 1)	log(Pat. + 1)	log(Pat. + 1)
Coll.County*PostColl.	0.430***	0.379*	0.178	0.141	0.216	0.269
	(0.161)	(0.211)	(0.160)	(0.148)	(2.893)	(0.348)
PostColl.	-0.029	0.265**	-0.005	0.071	-0.979***	-0.068
	(0.078)	(0.115)	(0.102)	(0.104)	(0.027)	(0.124)
log(Total Pop.)			0.396***		0.364***	
			(0.104)		(0.101)	
Total Pop.				0.152***		0.101**
				(0.044)		(0.051)
(Total Pop.)^2				-0.000*		-0.000*
				(0.000)		(0.000)
Coll.County * PostColl. * log(Total Pop.)					-0.005	
					(0.240)	
Coll.County * PostColl. * Total Pop.						-0.150
						(0.140)
Coll.County * PostColl. * Total Pop. Squared						0.000**
						(0.000)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Exp. Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
Cnty-Year Obs.	4,078	3,809	1,092	1,092	1,092	1,092
# Counties	179	179	176	176	176	176
# Experiments	72	72	72	72	72	72
Mean in 1880	4.845	20,154.983	4.845	4.845	4.845	4.845
Adj. R-Sqr.	0.453	0.715	0.199	0.211	0.216	0.218

Notes: Results for the effect of population on patenting. Column 1 estimates the level shift in patenting in college counties relative to runner-up counties after establishment of a new college when the dependent variable is $log(Num.Patents{+}1)$. Column 2 estimates the level shift in population in college counties relative to the runner-up counties after establishment of a new college when the dependent variable is $log(TotalPop.)$. The dependent variable for both Columns 3-6 is $log(Patents + 1)$. Column 3 re-estimates Column 1 but includes a control for $log(TotalPop.)$. Column 4 re-estimates Column 1 but includes controls for $TotalPop.$ and $(TotalPop.)^2$. Column 5 estimates the effect of the level shift in patenting in college counties relative to runner-up counties after establishment of a new college when controlling for $log(TotalPop.)$ and interacting $log(TotalPop.)$ with a dummy for college counties, a dummy for post-college years, and the interaction term. Column 6 estimates the effect of the level shift in patenting in college counties relative to runner-up counties after establishment of a new college when controlling for $TotalPop.$ and $(TotalPop.)^2$ and interacting both controls with a dummy for college counties, a dummy for post-college years, and the interaction term. Each coefficient is transformed into a percentage change in the dependent variable /100. When the model estimates changes in levels, the percentage change is calculated based on the baseline value of the independent variable in 1880. Results are for high quality experiments only. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table 7: Patents by Alumni and Faculty

	Num. People	Share of Pop.	Num. Patents	Patents per 10,000 Cap.	Share of County Patents	Share of Extra Patents
Entire County	40,319.943	1.000	2.465	0.052	1.000	1.000
	(45,735.840)	(0.000)	(6.740)	(0.042)		
Undergraduate Alumni	3,882.335	0.118	0.000	0.000	0.000	0.000
	(5,465.699)	(0.182)	(0.000)	(0.000)	(0.000)	(0.000)
Faculty	65.982	0.002	0.000	0.000	0.000	0.000
	(96.292)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Population and patenting results for college alumni and faculty. The first row lists statistics for the entire county. The second row lists statistics for college undergraduate alumni. The third row lists statistics for college faculty. The first column lists the average number of people in each group per county. The second column lists the share of the county's total population belonging to each group. The third column lists the number of patents attributable to each group. The fourth column lists the patenting rate for individuals in each group ($Num.Patents_j * 1000/Num.Members_j$ for members of group j). The fifth column lists the share of the county's total patents attributable to each group. The sixth column lists the share of the additional patents caused by the creation of the new college that can be attributable to each group. Standard deviations are displayed in parentheses. Results are for college counties for which yearbook data is available.

A More Information on the College Site Selection Experiments

Table 8 lists the runner-up counties for each college county in the high quality college sample. Additional information about the college and runner-up counties is provided in the subsections below.

A.1 Additional Balance Checks

In Table A.1 in Section 2.3, I compare college counties to runner-up counties along a number of dimensions by conducting a t-test for each dimension and find that no individual dimension predicts treatment status. Here, I verify that these dimensions do not jointly predict treatment status either. Unfortunately, for several of the dimensions considered, missing data is a major concern. This is because the data come from different censuses and particular data were not necessarily collected every decade; moreover, even when data is available for a particular census, it is occasionally missing for particular counties. Comparing only experiments in which data for all dimensions are available for all college and runner-up counties results in an extremely small sample size. I instead present results of joint tests with data that are available for most counties in the census year prior to the establishment of the new college.

Results of the joint tests are presented in Table 9. Column 1 estimates a linear probability model in which the dependent variable is a dummy variable taking the value of 1 when the county obtains the college and 0 otherwise. The F-test for the joint significance of all included regressors is 0.473, which is insignificant at conventional levels. Column 2 estimates a logit model with the same regressors. A likelihood ratio χ^2-test also concludes that the regressors do not jointly predict treatment status. Column 3 and 4 repeat Columns 1 and 2 but include additional regressors and hence have much smaller sample sizes; the coefficients are again not jointly significant. Results are similar with other combinations of regressors, although the sample size falls even further.

Table 10 conducts t-tests for several other categories along which college and runner-up

Table 8: Runner-Up Counties

	Experiment	County	State	Runner-Up Counties
1	AlAMU	Madison	Alabama	Montgomery
2	ArTU	Pope	Arkansas	Sebastian; Conway; Franklin
3	AubU	Lee	Alabama	Tuscaloosa; Lauderdale
4	AzSU	Maricopa	Arizona	Pinal; Pima
5	BGSU	Wood	Ohio	Sandusky; Van Wert; Henry
6	ClemU	Pickens	South Carolina	Richland
7	CornU	Tompkins	New York	Seneca; Schuyler; Onondaga
8	EMiU	Washtenaw	Michigan	Jackson
9	ENCU	Pitt	North Carolina	Beaufort; Lenoir; Edgecombe
10	GaSoU	Bulloch	Georgia	Tattnall; Emanuel
11	GaT	Fulton	Georgia	Bibb; Baldwin; Clarke; Greene
12	HPU	Guilford	North Carolina	Alamance
13	IaSU	Story	Iowa	Polk; Marshall; Hardin; Jefferson; Tama
14	KentSU	Portage	Ohio	Medina; Trumbull
15	KsSU	Riley	Kansas	Shawnee; Douglas; Lyon
16	KySU	Franklin	Kentucky	Fayette; Boyle; Warren; Christian; Daviess
17	LaSU	East Baton Rouge	Louisiana	Bienville; East Feliciana
18	LincC	Logan	Illinois	Macon; Edgar; Warrick
19	MTnSU	Rutherford	Tennessee	Montgomery
20	MeMA	Hancock	Maine	Sagadahoc
21	MercU	Bibb	Georgia	Spalding
22	MoUST	Phelps	Missouri	Iron
23	NCAT	Guilford	North Carolina	Alamance; New Hanover; Durham; Forsyth
24	NCSU	Wake	North Carolina	Lenoir; Mecklenburg
25	NDSU	Cass	North Dakota	Grand Forks; Burleigh; Stutsman
26	NIlU	DeKalb	Illinois	Winnebago
27	NMSU	Dona Ana	New Mexico	Socorro; San Miguel; Bernalillo
28	NMT	Socorro	New Mexico	San Miguel; Bernalillo; Dona Ana
29	OrSU	Benton	Oregon	Marion
30	PaSU1	Centre	Pennsylvania	Blair
31	PurdU	Tippecanoe	Indiana	Marion; Monroe; Hancock
32	SArU	Columbia	Arkansas	Ouachita; Polk; Hempstead
33	SIlU	Jackson	Illinois	Jefferson; Perry; Washington; Marion; Clinton
34	SMU	Dallas	Texas	Tarrant
35	SMsU	Forrest	Mississippi	Hinds; Jones
36	TCNJ	Mercer	New Jersey	Middlesex; Burlington; Essex
37	TxAMU	Brazos	Texas	Austin; Grimes
38	TxCU	Tarrant	Texas	Dallas
39	TxT	Lubbock	Texas	Scurry; Nolan
40	UAr	Washington	Arkansas	Independence
41	UAz	Pima	Arizona	Maricopa; Pinal
42	UCaB	Alameda	California	Contra Costa; Napa
43	UCaDav	Yolo	California	Solano; Contra Costa
44	UCo	Boulder	Colorado	Fremont
45	UFl2	Alachua	Florida	Columbia
46	UIUC	Champaign	Illinois	McLean; Logan; Morgan
47	UId	Latah	Idaho	Bonneville
48	UKs	Douglas	Kansas	Shawnee; Lyon; Riley
49	UMe	Penobscot	Maine	Sagadahoc
50	UMo1	Boone	Missouri	Saline; Howard; Cole; Callaway; Cooper
51	UMs	Lafayette	Mississippi	Winston; Monroe; Harrison; Attala; Rankin; Montgomery
52	UND	Grand Forks	North Dakota	Cass; Stutsman; Burleigh
53	UNH	Strafford	New Hampshire	Belknap
54	UNM	Bernalillo	New Mexico	Dona Ana; San Miguel; Socorro
55	UNeKe	Buffalo	Nebraska	Custer; Valley
56	UNv	Washoe	Nevada	Carson City
57	UOr	Lane	Oregon	Linn; Washington; Polk
58	USAFA	El Paso	Colorado	Walworth; Madison
59	USD	Clay	South Dakota	Bon Homme; Yankton
60	USMMA	Nassau	New York	Bristol
61	UTn2	Knox	Tennessee	Rutherford
62	UTxA	Travis	Texas	Smith
63	UTxMB	Galveston	Texas	Harris
64	UWi2	Dane	Wisconsin	Fond du Lac
65	UWy	Albany	Wyoming	Uinta; Laramie
66	UtSU	Cache	Utah	Weber
67	VaT	Montgomery	Virginia	Albemarle; Rockbridge
68	WIlU	McDonough	Illinois	Schuyler; Hancock; Mercer; Warren; Adams
69	WMiU	Kalamazoo	Michigan	Allegan; Barry
70	WSCoU	Gunnison	Colorado	Garfield; Mesa
71	WVU	Monongalia	West Virginia	Greenbrier; Kanawha
72	WaSU	Whitman	Washington	Yakima

Notes: List of runner-up counties for each high quality college county. Column 1 lists the college experiment. Columns 2 and 3 list the county and state, respectively, of the college county. Column 4 lists each runner-up county name.

counties can be measured. I show that the college and runner-up counties are much more similar than the college and non-experimental counties in the same state. The first column lists the mean and standard deviation of college counties. The second column lists the mean and standard deviation of the runner-up counties. The third column lists the difference in the mean between the college and runner-up counties, as well as the standard error of the difference. The fourth column lists the mean and standard deviation of the non-experimental counties. The fifth column list the difference in the mean between the college and the non-experimental counties, as well as the standard error of the difference.

Figure 6 shows that not only are the levels of a number of economic and demographic variables similar in college and runner-up counties prior to establishing a new college, but the evolve similarly as well. In Panel (a), I plot residual logged county population after controlling for year effects for several decades both before and the establishment of a the new college in the college, runner-up, and non-experimental counties. Panel (b) plots the residual fraction of the county population that lives in an urban area. Panel (c) plots the residual fraction of the county population that is black. Finally, Panel (d) plots residual logged manufacturing output. Plots for the other variables presented in Table 10 are similar. Confidence intervals are omitted in the figure for readability.

A.2 Low Quality Site Selection Experiments

While not used in the baseline results, data is collected on a number of “low quality” college site selection experiments as well. These are cases in which finalist sites can be identified, but the site selection decision does not approximate random assignment. Figure 7 shows the location of all college and runner-up counties, including both the high and low quality experiments. Table 11 is analogous to Table 2 but presents results only for the low quality experiments.

B Constructing Patent Data

For the baseline results, I match patents to their county. This is non-trivial because each patent lists the town and state of each inventor, but not the county. To match towns to

Table 9: Tests for Joint Significance of Covariates Predicting Whether a County Receives a College

	Linear Probability	Logit	Linear Probability	Logit
log(Pat. + 1)	-0.095*	-0.531	-0.030	-0.019
	(0.055)	(0.281)	(0.357)	(1.365)
log(Total Pop.)	0.105***	0.625	0.331	4.903
	(0.040)	(0.240)	(0.653)	(4.818)
log(Mean Age)	-0.099	-0.507	2.641	13.652
	(0.083)	(0.385)	(7.824)	(34.531)
Frac. Interstate Migrants	0.033	0.260	0.396	0.651
	(0.119)	(0.556)	(1.101)	(4.674)
Frac. Male	0.477	2.453	2.491	43.731
	(0.648)	(2.982)	(7.514)	(50.810)
Frac. White	0.103	0.513	0.161	2.504
	(0.189)	(0.886)	(1.230)	(5.871)
Frac. Urban			1.236	10.368
			(1.584)	(9.066)
log(Value Manuf. Output)			0.101	0.215
			(0.272)	(1.254)
log(Value Farm Product)			-0.054	-1.484
			(0.389)	(2.102)
Frac. Attending School			-1.004	-9.736
			(3.164)	(13.567)
log(Colleges + 1)			-0.524	-5.753
			(0.446)	(4.665)
# Counties	184	184	24	24
# Experiments	62	62	59	59
Adj. R-Sqr.	0.007		-0.473	
F-Stat	1.225		0.328	
F-Test p-Value	0.296		0.963	
LR Chi-Sqr. Stat		8.614		7.357
LR-Test p-Value		0.196		0.769

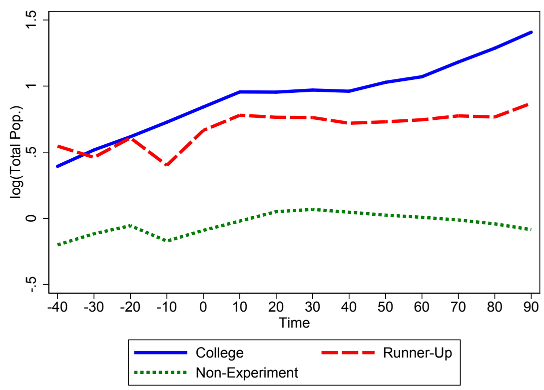
Notes: Joint tests for the significance of several covariates in predicting whether a county is a college county or a runner-up. Data are from the last census year before each college site selection experiment. The included covariates are those that are available for most counties in nearly every census. Column 1 presents results from a linear probability model. Column 2 presents results from a logit model. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table 10: Additional Balance Checks

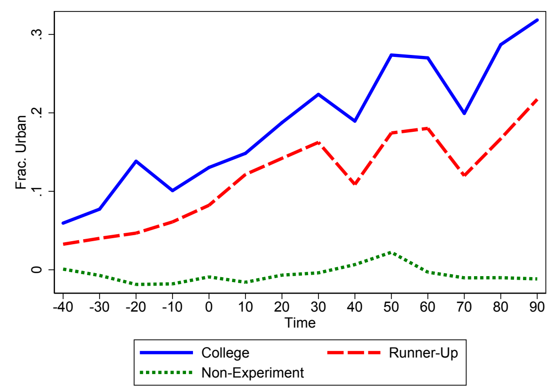
	Treatment	Controls	Treat. - Cont.	Non-Experiment	Treat. - Non-Exp.
log(Pat. + 1)	0.415	0.431	-0.016	0.303	0.112
	(0.739)	(0.747)	(0.107)	(0.702)	(0.084)
log(Total Pop.)	9.853	9.480	0.372*	9.063	0.790***
	(0.962)	(1.449)	(0.195)	(1.740)	(0.213)
Frac. Urban	0.185	0.137	0.048	0.078	0.107***
	(0.232)	(0.189)	(0.031)	(0.169)	(0.021)
log(Mean Age)	3.345	3.349	-0.004	3.426	-0.081*
	(0.605)	(0.506)	(0.085)	(0.366)	(0.048)
Frac. Interstate Migrants	0.613	0.595	0.018	0.519	0.093**
	(0.346)	(0.350)	(0.055)	(0.312)	(0.041)
log(Value Manuf. Output)	12.725	12.079	0.646	10.892	1.833***
	(1.953)	(2.823)	(0.417)	(3.907)	(0.518)
log(Value Farm Product)	13.172	13.040	0.132	12.742	0.430
	(1.395)	(1.365)	(0.234)	(2.749)	(0.390)
log(Colleges + 1)	0.179	0.184	-0.005	0.045	0.134***
	(0.337)	(0.366)	(0.064)	(0.186)	(0.028)
Num. Pat.	1.255	1.360	-0.105	2.093	-0.838
	(3.073)	(3.657)	(0.501)	(35.961)	(4.268)
Total Pop.	31,492.075	24,016.261	7,475.814	22,224.850	9,267.225
	(52,265.424)	(37,152.274)	(6,307.697)	(107,091.361)	(13,103.956)
Frac. Rural	0.800	0.846	-0.045	0.922	-0.121***
	(0.242)	(0.199)	(0.034)	(0.170)	(0.022)
Mean Age	30.047	29.547	0.501	30.971	-0.923
	(7.793)	(7.576)	(1.203)	(6.285)	(0.819)
Frac. Foreign Immigrant	0.093	0.104	-0.011	0.122	-0.029
	(0.085)	(0.095)	(0.019)	(0.147)	(0.026)
Frac. Male	0.524	0.519	0.005	0.526	-0.002
	(0.115)	(0.121)	(0.018)	(0.105)	(0.013)
Frac. White	0.819	0.805	0.014	0.824	-0.005
	(0.239)	(0.264)	(0.040)	(0.235)	(0.031)
Segregation	0.320	0.290	0.030	0.348	-0.028
	(0.244)	(0.249)	(0.053)	(0.221)	(0.039)
Pop. per Sq. Mile	77.646	42.033	35.613	61.159	16.487
	(225.232)	(81.185)	(29.696)	(888.896)	(148.225)
Pop. Attending School	8,546.739	5,368.396	3,178.343	5,298.191	3,248.548
	(18,494.064)	(11,427.919)	(3,568.528)	(28,552.059)	(5,967.935)
Frac. Attending School	0.141	0.117	0.024	0.150	-0.009
	(0.082)	(0.090)	(0.022)	(0.081)	(0.017)
Frac. Illiterate	0.140	0.125	0.015	0.152	-0.012
	(0.141)	(0.136)	(0.039)	(0.158)	(0.037)
Manuf. Establishments	121.944	107.784	14.161	92.351	29.593
	(135.001)	(158.352)	(43.466)	(456.389)	(107.690)
log(Manuf. Employment)	4.962	4.738	0.224	3.767	1.194***
	(2.379)	(2.425)	(0.452)	(2.548)	(0.391)
Manuf. Employment	1,108.279	1,395.107	-286.828	1,091.956	16.323
	(2,487.831)	(7,022.769)	(1,106.598)	(10,045.132)	(1,532.778)
Value Manuf. Output	2,577,168.298	3,379,563.752	-802,395.454	4,165,527.136	-1,588,358.838
	(8,913,310.353)	(21,603,565.227)	(2,985,143.029)	(53,796,135.208)	(7,127,825.416)
log(Manuf. Wages)	9.721	9.369	0.352	7.665	2.056***
	(4.421)	(4.493)	(0.838)	(4.949)	(0.759)
Manuf. Wages	628,071.628	960,679.464	-332,607.836	923,019.620	-294,947.992
	(1,751,901.201)	(6,175,432.565)	(962,603.198)	(10,737,956.268)	(1,638,096.216)
Value Farm Product	1,037,502.740	1,102,507.482	-65,004.742	1,526,697.672	-489,194.932
	(1,352,746.052)	(2,400,586.689)	(363,543.022)	(3,796,003.824)	(537,419.860)
log(Farm Wages)	11.641	11.351	0.290	10.547	1.094**
	(0.949)	(1.187)	(0.336)	(1.777)	(0.432)
Farm Wages	173,194.412	155,687.400	17,507.012	99,315.375	73,879.037*
	(173,285.347)	(194,111.125)	(56,759.849)	(156,920.089)	(38,354.234)
log(Value Farms)	14.522	14.284	0.238	13.968	0.555
	(1.426)	(1.969)	(0.291)	(2.691)	(0.357)
Value Farms	4,860,152.474	4,209,547.909	650,604.565	4,666,175.502	193,976.972
	(7,192,430.022)	(5,349,566.125)	(963,523.948)	(8,191,354.422)	(1,090,975.043)
Num. Faculty	3.915	3.758	0.157	1.813	2.102
	(12.693)	(11.722)	(2.149)	(14.321)	(2.098)
Num. Students	51.064	55.316	-4.252	23.067	27.997
	(159.714)	(173.282)	(30.128)	(161.350)	(23.663)
Frac. of Chemical Patents	0.002	0.007	-0.006	0.014	-0.013
	(0.010)	(0.034)	(0.005)	(0.096)	(0.015)
Frac. of Comm Patents	0.000	0.000	-0.000	0.002	-0.002
	(0.000)	(0.004)	(0.001)	(0.037)	(0.006)
Frac. of Medical Patents	0.000	0.013	-0.013	0.003	-0.003
	(0.000)	(0.111)	(0.017)	(0.045)	(0.007)
Frac. of Electric Patents	0.009	0.000	0.008*	0.004	0.005
	(0.042)	(0.001)	(0.005)	(0.037)	(0.006)
Frac. of Mechanical Patents	0.033	0.049	-0.016	0.038	-0.005
	(0.156)	(0.167)	(0.030)	(0.158)	(0.024)
Frac. of Other Patents	0.038	0.059	-0.021	0.061	-0.024
	(0.139)	(0.271)	(0.044)	(0.215)	(0.033)
Frac. of Tech Patents	0.043	0.056	-0.013	0.056	-0.013
	(0.164)	(0.180)	(0.033)	(0.199)	(0.030)
Frac. of Ag Patents	0.015	0.001	0.014	0.008	0.007
	(0.079)	(0.006)	(0.009)	(0.077)	(0.012)

Notes: T-tests comparing the means of the college counties, runner-up counties, and non-experimental counties. Data are from the last census year before each college site selection experiment. The first column lists the mean and standard deviation of college counties. The second column lists the mean and standard deviation of the runner-up counties. The third column lists the difference in the mean between the college and runner-up counties, as well as the standard error of the difference. The fourth column lists the mean and standard deviation of the non-experimental counties. The fifth column list the difference in the mean between the college and the non-experimental counties, as well as the standard error of the difference. The college and runner-up counties are from high quality experiments only. Stars in columns 3 and 5 indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

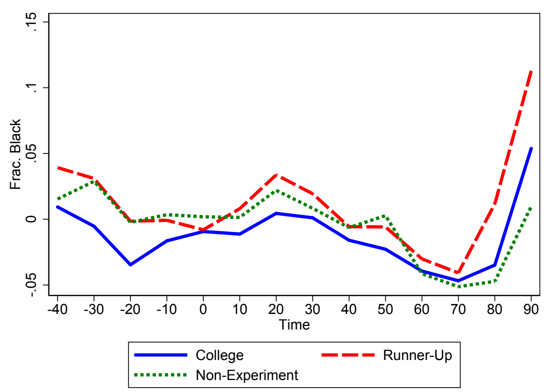
Figure 6: Time Series for Demographic and Economic Variables



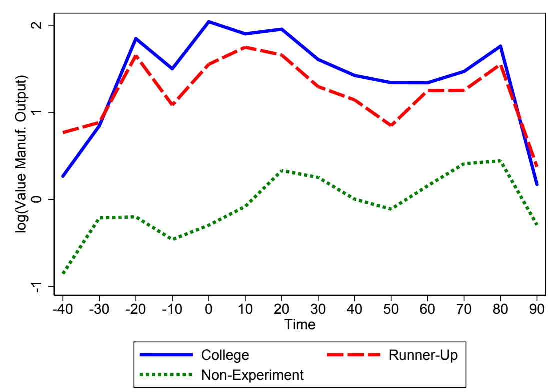
(a) log(Total Pop)



(b) Frac. Urban



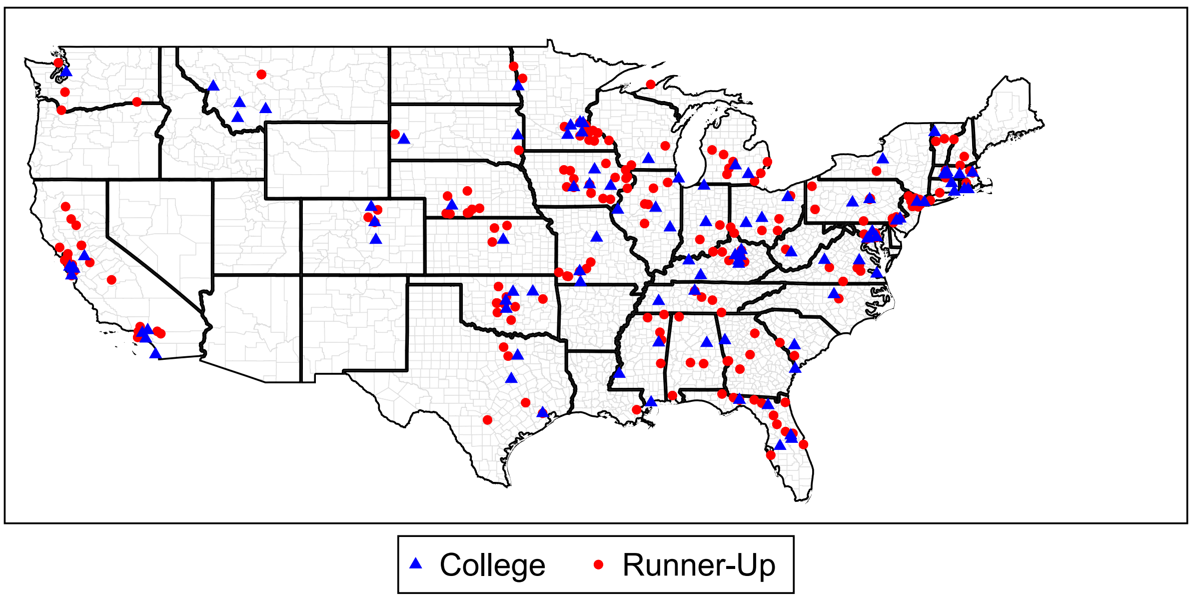
(c) Frac. Black



(d) log(Manuf. Output)

Notes: Time series for various demographic and economic variables in each census year. The year of the college experiment is normalized to year 0. Everything left of year 0 shows pre-college means; everything to the right shows post-college means. The college counties are represented by the blue solid line. The runner-up counties are represented by the red dashed line. The non-experimental counties are represented by the green short-dashed line. In each panel, the y-axis is a residual demographic or economic variable after controlling for year effects. Data are for high quality experiments only.

Figure 7: Map of Both High and Low Quality College and Runner-Up Sites



Notes: Map of the location of the low quality college and runner-up sites. Colleges are represented by blue diamonds. The runner-up sites are represented by red circles.

Table 11: Summary Statistics of Low Quality College Site Selection Experiments

	N	Mean	S.D.	Min	Median	Max
# Finalist Counties	98	2.95	1.90	1.00	2.00	10.00
Distance to Finalists	219	155.28	210.66	4.24	99.96	2,243.45
Experiment Year	108	1884.85	29.81	1836.00	1882.00	1963.00
Year of First Class	95	1887.66	32.87	1830.00	1886.00	1967.00
Year Desegregated	49	1921.90	40.71	1853.00	1945.00	1968.00
Year Co-Ed	57	1893.74	39.05	1846.00	1885.00	1997.00
Land Grant Colleges	108	0.22	0.42	0.00	0.00	1.00
Technical Schools	108	0.06	0.23	0.00	0.00	1.00
Normal Schools	108	0.09	0.29	0.00	0.00	1.00
HBCUs	108	0.06	0.23	0.00	0.00	1.00
Military Academies	108	0.07	0.26	0.00	0.00	1.00
Other Public Colleges	108	0.12	0.33	0.00	0.00	1.00
Other Private Colleges	108	0.36	0.48	0.00	0.00	1.00

Notes: Summary statistics for the low quality college site selection experiments. Column 1 lists the count of experiments or counties. Column 2 lists mean values, Column 3 the standard deviation, Column 4 the minimum value, Column 5 the median value, and Column 6 the maximum value. Row 1 lists the number of runner-up counties for each experiment. Row 2 lists the distance between college and runner-up sites. Row 3 lists the experiment year. Row 4 lists the year in which students began attending the college. Row 5 lists the year when the college became racially desegregated. Row 6 lists the year the college became coeducational. Rows 7-13 list the fraction of colleges that are of each college type.

their counties, I first standardize all town and county names by converting all characters to have consistent capitalization, removing all punctuation and non-alphabetic characters, and harmonizing common abbreviations, for instance changing "SAINT" to "ST" and "FORT" to "FT". I then obtain a list of all towns in each U.S. county in each decennial census year, compiled from the 100% censuses. I look for exact matches between town names in the patents and town names in the preceding decennial census. This means that, for instance, town names in 1883 patents are matched to town names in the 1880 decennial census. For 1890, the 100% decennial census was destroyed by fire, so I match town names to the 1990 census. The results are insensitive to matching to the closest census rather than the previous census. For all patents granted in 1950 or later, there is no declassified 100% decennial census from the previous decade to match to. In these cases, I first attempt to match to town names in the 1940 decennial census. For the remaining towns that are unmatched, I use zip code data from `https://www.unitedstateszipcodes.org/zip-code-database/` to match to any town name that is affiliated with a current U.S. zip code; the zip code database also contains the counties in which each town resides.

In some cases, a town's boundaries lie in several counties. Alternatively, there may be states with multiple towns of the same name. In these cases when a town is associated with multiple counties, I assume each patent has an equal probability of belonging to each county and divide the number of patents by the number of towns to find a mean number of patents. I also construct an upper bound, assuming that every patent belongs to a particular county, and a lower bound that assumes that no patents belong to a particular county. All results throughout the paper use the mean patent count, but results are nearly identical when using the upper and lower bounds.

Errors may also occur if spelling, transcription, or OCR errors occur in town names or if the patent data use slight variations of actual town names; there is no formal process to make town names uniform across patents. In the baseline results presented throughout the paper, I require standardized town names in the patent data to exactly match standardized town names in the town-county correspondence. I also match towns to counties using "fuzzy" matching techniques. These are bi-gram string comparators that return a "distance" between the town-state strings in each dataset. Using various different weights for the town and state strings in the distance function returned qualitatively similar results. See Andrews (2017) for more information on the differences between the exact and fuzzy matching between towns and counties. I present baseline results using the fuzzy matching in Appendix C.1; results are similar, reflecting the fact that standardizing town and county names eliminates most differences.

C Additional Baseline Results

Figure 8 plots logged patenting in the college, runner-up, and non-experimental counties using only the raw data; that is, I plot patenting without controlling for year effects or smoothing the data.

To show the heterogeneity of the estimated treatment effect across college site selection experiments, I estimate the baseline regression in Equation (1) with a separate interaction term for each experiment. Formally, I estimate

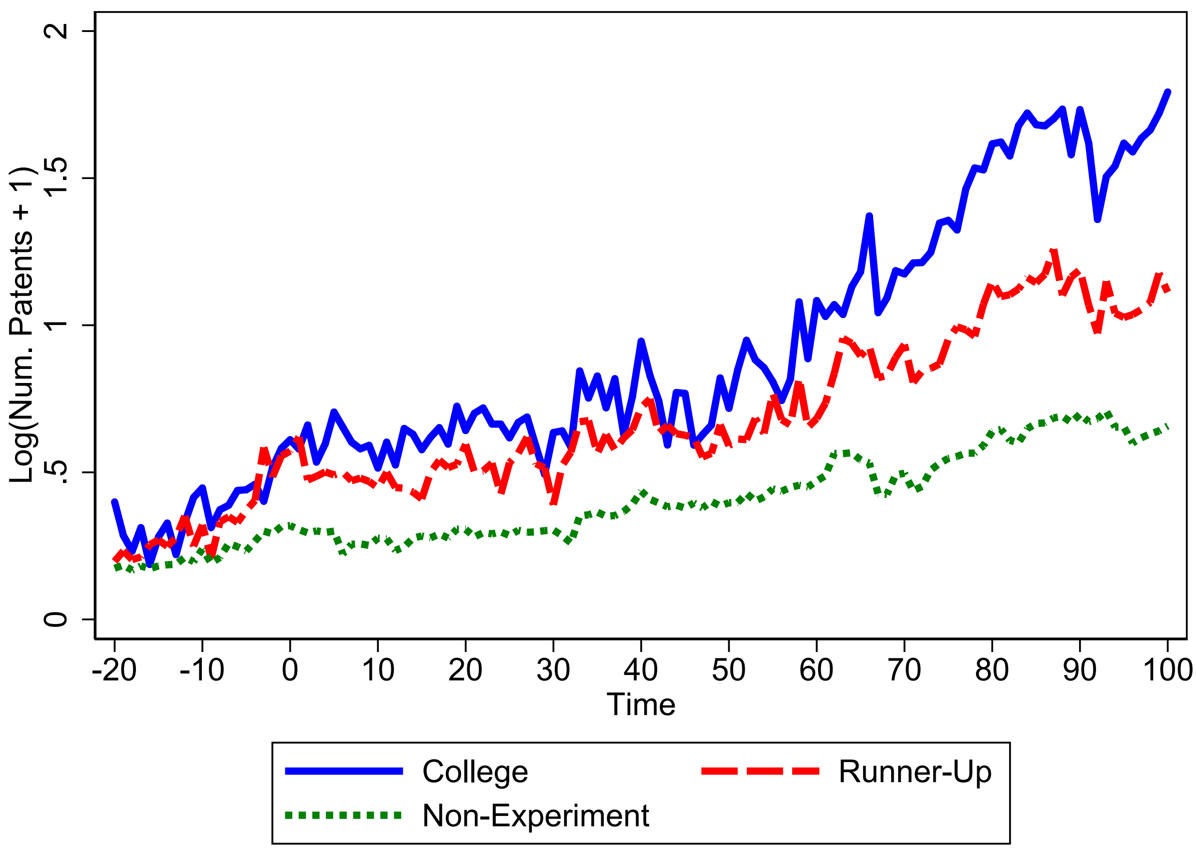
$$\begin{aligned}\log(NumPat_{ijt} + 1) = &\sum_{j \in J} [\delta_{1j} College_{ij} * PostCollege_{jt} + \delta_{2j} PostCollege_{jt}] \\ &+ \alpha_i + \lambda_j + \gamma_t + \epsilon_{ijt}, \end{aligned} \tag{4}$$

where J is the number of college site selection experiments. Each coefficient is plotted in a histogram in Figure 9. Estimated interaction terms for each individual experiment are available upon request. In about 60% of the experiments, the estimated coefficient is positive. Even when the coefficient is negative, it tends to be close to zero in magnitude. In a majority of the college site selection experiments, the estimated coefficients are in line with the coefficients estimated in Table 3.

C.1 Robustness Checks

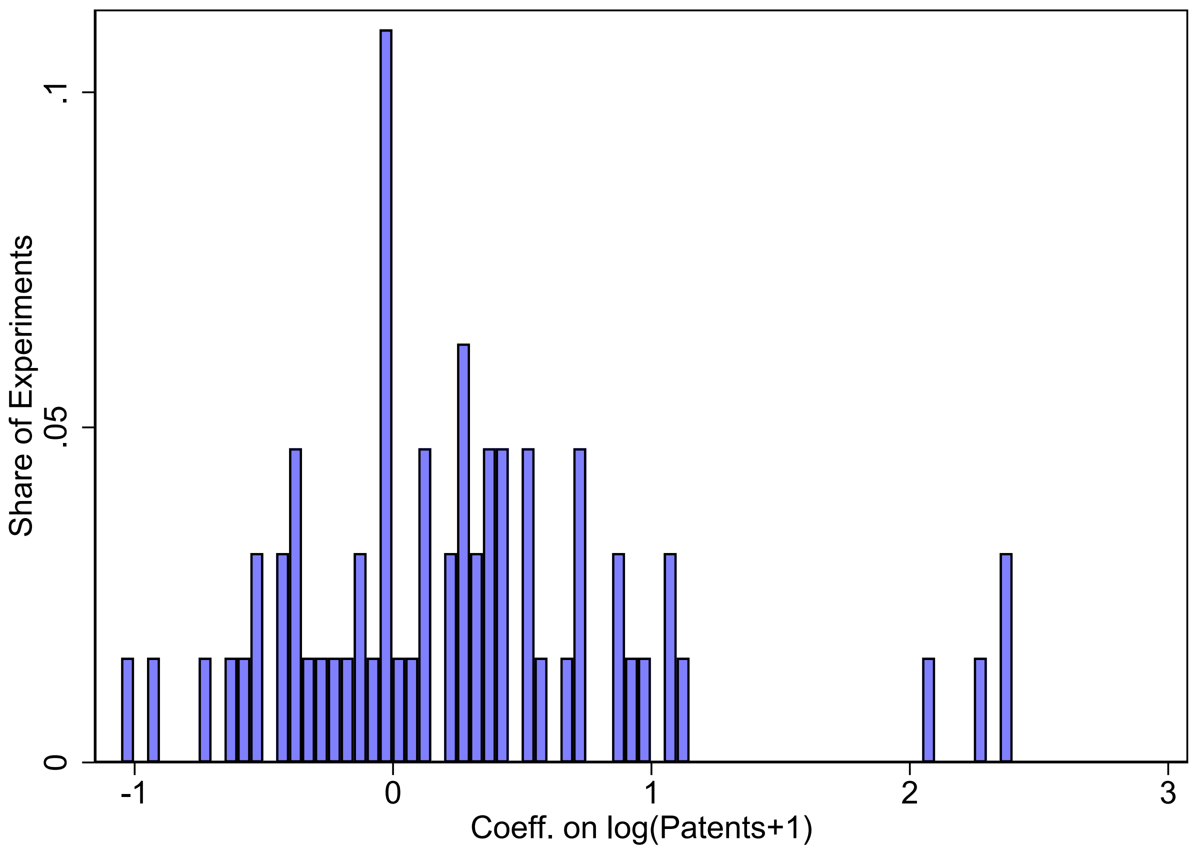
In this section, I estimate several additional regression specifications to demonstrate that the baseline results described in Section 3 are robust. Results are presented in Table 12.

Figure 8: Patenting in College, Runner-Up, and Non-Experimental Counties, Raw Data



Notes: Unconditional mean patenting in college and runner-up counties. The x-axis shows the number of years since the college experiment. The year of the college experiment is normalized to year 0. Everything left of year 0 shows pre-college means; everything to the right shows post-college means. The y-axis shows $log(Patents + 1)$. The college counties are represented by the blue solid line. The runner-up counties are represented by the red dashed line. The non-experimental counties are represented by the green short-dashed line. Data are for high quality experiments only.

Figure 9: Distribution of Treatment Effects Across College Experiments



Notes: Distribution of estimated treatment effects across each college site selection experiment. The coefficient of the level shift in patenting in college counties relative to runner-up counties after establishments of a new college is estimated with a separate interaction term for each college experiment. The histogram plots the share of estimated interactions falling into each bin of width .05.

In Column 1, I estimate the following modification of the baseline regression:

$$\begin{aligned} PatentMeasure_{ijt} =& \delta_1 College_{ij} * PostCollege_{jt} + \\ & + \alpha_i + \lambda_j + \alpha_i * \lambda_j + \gamma_t + \lambda_j * \gamma_t + \epsilon_{ijt}. \end{aligned} \tag{5}$$

Here, the $PostCollege_{jt}$ is not needed as it is colinear with the experiment-by-year fixed effect, $\lambda_j * \gamma_t$. δ_1 is similar to the baseline estimate in Column 1 of Table 3. Column 2 examines the extensive margin: do counties have a higher probability of obtaining at least one patent per year after receiving a new college. In this linear probability model, I find that establishing a new college makes a county 22% more likely to have at least one patent in a given year. Column 3 uses an alternative construction of logged patenting, $log(Num.Patents + 0.0001)$. These results are much larger than the baseline estimate. This is not surprising in light of the results in Column 2, since this specification penalizes having zero patents more heavily than the baseline specification that uses $log(Num.Patents + 1)$ as the dependent variable. Column 4 displays the results using the number of patents as the dependent variable in a simple linear specification. The estimated percentage increase is large (93% more patents per year in the college counties relative to the runners-up, using 1880 as the baseline year) but in line with the results using a negative binomial model in Table 3. Column 5 presents results from a simple fixed effects Poisson regression, again using Winsorized data. These results are similar in magnitude to Column 4 and to the negative binomial results presented in Table 3.

To further show that the results are not driven by the subjective classification of some experiments as either high or low quality, I also re-estimate the baseline regression excluding each high quality experiment, one at a time, and re-estimating the baseline regression. I also reclassify each low quality experiment as high quality, one at a time, and re-estimate the baseline regression. In all cases, the estimated coefficient is very similar to the baseline result and statistical significance is unchanged. These results are available upon request.

A related concern is that different types of college experiments may be systematically different from one another. I argue in Section 2.1 that the college site selection experiments are as good as random assignment. While each experiment is unique, they tend to fall into

Table 12: Additional Regression Specifications

	log(Patents + 1)	Any Patents	log(Pat. + 0.0001)	Num. Patents	Poisson
Coll.County * PostColl.	0.314***	0.216**	1.828*	0.927***	0.857**
	(0.103)	(0.086)	(1.063)	(0.351)	(0.382)
PostColl.		0.097	0.355	-0.411	7.247***
		(0.066)	(0.361)	(0.265)	(1.596)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes
Exp. Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes
Cnty-Year Obs.	39,580	39,620	39,580	39,580	39,580
# Counties	179	179	179	179	179
# Experiments	72	72	72	72	72
Mean in 1880	4.845	0.382	-5.519	4.845	4.845
Adj. R-Sqr.	0.649	0.377	0.439	0.311	
Log-Likelihood	-35,727.662	-18,781.531	-110,961.886	-166,492.392	-123,755.032

Notes: Regression results using alternative specifications. Column 1 includes experiment-by-year fixed effects whent he dependent variable is $log(Num.Patents+1)$. Column 2 estimates a linear probability model where the dependent variable is an indicator equal to one if a county has at least one patent in a given year and zero otherwise. The dependent variable in Column 3 is $log(Num.Patents+0.0001)$. The dependent variable in column 4 is the number of patents. Column 6 estimates a Poisson regression. Columns 5 and 6 use non-Windsorized counts of patenting. Each coefficient is transformed into a percentage change in the dependent variable /100. When the model estimates changes in levels, the percentage change is calculated based on the baseline value of the independent variable in 1880. Results are for high quality experiments only. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

groups in which the colleges were assigned with different general methods. It would be suspicious if one method of "random" assignment gave systematically different results from other such methods. In this subsection, I test this concern by grouping experiments by the method in which the college was assigned and then checking that the estimated coefficients are similar across different groups.

I use four broad groups: auctions, politics, infrastructure, and other. "Auctions" refer to all cases in which a board of trustees, state legislature, or other site selection body solicited bids from localities; identification comes from comparing very similar bids across different locations. "Politics" refers to cases where political maneuvering, involving things like quid pro quos, strategic timing of votes, or even outright bribery, secured the college for one location over another; identification rests on the assumption that these political schemes are uncorrelated with any other local factors that would affect the college location decisions.[30] "Infrastructure" refers to cases in which the college had specific infrastructure needs that could only be satisfied by a limited number of candidate locations. As an example, the Morrill Land Grant Colleges Act forbade the use of land grant funds to construct buildings, so many land grant colleges had to be located where there was an existing and available building large enough to be used for a college. In other cases, colleges had to be located near the center of a state, near viable drinking water or on navigable waterways, or close to railway lines. All of the runner-up counties in the were deemed to meet these infrastructure requirements by the site selection committee. Finally, "other" refers to all experiments that do not fit into one of the above descriptions. This can include pure random assignment (as in the case of the University of North Dakota), cases where weather played a pivotal role (as in the University of Arizona), or other bizarre circumstances (such as Cornell University). In several cases, an experiment could plausibly fit into several groups. For instance, in many cases bids were solicited only from localities that met certain infrastructure needs. I attempt to put each experiment into the most appropriate group; the results are not sensitive to reclassifying marginal experiments.

Table 13 shows the results. In Columns 1-4, I successively remove each of the above

[30]For this reason, I do not consider an experiment to be of high quality if the work of a governor or legislative leader was instrumental in deciding where to locate the college and represented the winning county as this may reflect longstanding political influence rather than a quasi-random event.

experiment types and show that the results are still qualitatively the same. In Column 5, I interact each experiment type with the dummy for college counties and the post-college dummy. All coefficients are qualitatively similar except for the interaction term for infrastructure, which is negative but close to zero and not statistically significant.

Table 13: Results by Type of College Site Selection Experiment

	No Auctions	No Politics	No Infrastructure	No Other	log(Pat. + 1)
Coll.County * PostColl.	0.449**	0.351*	0.423***	0.393***	
	(0.194)	(0.183)	(0.152)	(0.148)	
PostColl.	-0.089	-0.001	-0.042	-0.008	
	(0.091)	(0.085)	(0.064)	(0.062)	
Auctions					0.359*
					(0.210)
Politics					0.494**
					(0.233)
Infrastructure					-0.020
					(0.280)
Other					0.555
					(0.573)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes
Exp. Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes
Cnty-Year Obs.	19,833	24,227	38,005	36,675	39,580
# Counties	94	116	171	168	179
# Experiments	38	42	69	67	72
Mean in 1880	4.845	4.845	4.845	4.845	4.845
Adj. R-Sqr.	0.497	0.530	0.508	0.496	0.509

Notes: Regression results by experiment type. The dependent variable is $log(Patents + 1)$. Columns 1-4 re-estimate the baseline results but excluding each experiment type in turn. In column 5, the coefficient is the increase in patenting caused by the college interacted by experiment type. Row 3 presents results for experiments decided by auction, row 4 for experiments decided by politics, row 5 for experiments decided by the presence of existing infrastructure, and row 6 for other site selection experiments. Each coefficient is transformed into a percentage change in the dependent variable /100. When the model estimates changes in levels, the percentage change is calculated based on the baseline value of the independent variable in 1880. Results are for high quality experiments only. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

As noted in Section 2.2, the Annual Report and Jim Shaw patent data list each inventor's town and state of residence. The analysis above is conducted at the county level, so it is necessary to assign each patent to a county. In the analysis above, a town-state pair is placed into a county when the exact town-state pair is found in the U.S. census, which lists both the towns and counties of all residents. There are alternative ways to match town-state pairs to counties, however, as described in Appendix B. Column 1 of Table 14 recreates these results. I also experiment with the baseline estimate when a fuzzy matching algorithm is used to match town-state pairs in the patent data to town-state pairs in the census data. These results are presented in Column 2. The coefficients are similar to those in Column 1.

The same analysis could also be performed using alternative patent data altogether. The baseline results in the paper use the Annual Reports compiled by the U.S. Patent Office. In Column 3, I also repeat the baseline estimates using HistPat data (Petralia et al., 2016b) instead of the Annual Reports or Jim Shaw data for any years in which they overlap. Results are similar to those using the Annual Report and Jim Shaw data. These results provide confidence that the results presented above are not an artifact of the particular patent dataset used or the choices made to geo-locate patents.

Table 14: Results with Alternative Patent Data

	Exact-Matched	Fuzzy-Matched	HistPat
Coll.County * PostColl.	0.401***	0.351**	0.302**
	(0.143)	(0.151)	(0.141)
PostColl.	-0.032	0.002	0.058
	(0.062)	(0.068)	(0.076)
County Fixed Effects	Yes	Yes	Yes
Exp. Fixed Effects	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes
Cnty-Year Obs.	39,580	39,327	39,271
# Counties	179	179	179
# Experiments	72	72	72
Mean in 1880	4.845	4.845	4.845
Adj. R-Sqr.	0.508	0.577	0.606

Notes: Regression results using different patent data. Columns 1 uses town-state pairs from patents that are exactly matched to town-state pairs in the U.S. Census to obtain a patent's county. Columns 1 uses town-state pairs from patents that are fuzzily matched to town-state pairs in the U.S. Census to obtain a patent's county. Column 3 uses the HistPat data. The dependent variable for all columns is $log(Patents+1)$. Each coefficient is transformed into a percentage change in the dependent variable /100. When the model estimates changes in levels, the percentage change is calculated based on the baseline value of the independent variable in 1880. Results are for high quality experiments only. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

I next conduct a placebo test to determine whether patenting changes differentially in college and runner-up counties in the years leading up to the college site selection experiment. I drop all data for the years after and including the year in which the college was established; all the remaining data is for the pre-trend. I then artificially designate the halfway point between the first year of observations and the last pre-experiment year as the "experiment year" and re-run the baseline regressions. Results are presented in Table 15. If the college counties are up-and-coming places, then they should be growing faster then the runner-up counties in the years before the original college site selection experiment and the estimated coefficient ($College \times$ "$PostCollege$") should be significantly positive. Instead, none of the coefficients are statistically different from zero and, while slightly positive, the coefficients of interest are much smaller in magnitude than their counterparts in Table 3. I take this as further evidence that the college site selection experiment is valid. Results are very similar if I instead designate random pre-college years as the placebo "treatment" year.

C.2 Patent Classes

Establishing a new college may alter the composition of patented technologies in addition to changing the total number of patents. To get a sense of patent technology type, in Table 16 I use the patent classes assigned to historical patents by Marco et al. (2015) to examine how patenting across all classes changes after establishing a new college.[31] In Column 1, I include controls for the share of patents in each county that belong to each of the NBER patent classes (patents with missing classes is the omitted category). The difference-in-differences estimate is similar in magnitude to the baseline estimate and is significant at the 10% level.

In Column 2, I repeat the baseline estimate at the patent class-by-county-by-year level. That is, I estimate:

$$\begin{aligned} PatentMeasure_{ijct} = &\delta_1 College_{ij} * PostCollege_{jt} + \delta_2 PostCollege_{jt} \\ &+ \alpha_i + \lambda_j + \alpha_i * \lambda_j + \psi_c + \gamma_t + \psi_c * \gamma_t + \epsilon_{ijct}, \end{aligned} \tag{6}$$

[31] The NBER one-digit patent classes are: chemical, communications, medical, electric, mechanical, other, no class, and missing class. All results in this section are similar when using two-digit NBER patent classes, USPTO patent classes, or IPC classifications.

Table 15: Placebo Test

	log(Pat. + 1)
Coll.County * PostColl.	0.062
	(0.060)
PostColl.	-0.041
	(0.037)
County Fixed Effects	Yes
Exp. Fixed Effects	Yes
Year Effects	Yes
Cnty-Year Obs.	10,570
# Counties	228
# Experiments	72
Mean in 1880	4.845
Adj. R-Sqr.	0.186

Notes: Placebo tests. The baseline regression results are reproduced with all post-experiment data dropped. The experiment year is set to halfway between the initial year of patent data and the year prior to the original college site selection experiment. Each coefficient is transformed into a percentage change in the dependent variable /100. When the model estimates changes in levels, the percentage change is calculated based on the baseline value of the independent variable in 1880. Results are for high quality experiments only. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

for patent classes c. This specification thus includes patent class and patent class-by-year fixed effects, flexibly picking up the fact that certain types of technology may be more or less prevalent at different points in time. While the coefficient of δ_1 is a bit smaller than the baseline estimate, it is qualitatively similar and significant at the 10% level. Thus, changes in the composition of patents cannot explain all of the baseline results.

Nevertheless, the results in Columns 1 and 2 suggest that there is some shifting in the types of inventions patented in college counties after a new college is established. Are college counties becoming increasingly specialized in a few narrow technology areas that happen to be especially patent-prone? This does not appear to be the case. Instead, patenting becomes more diverse in the college counties after the establishment of a new college. To see this, I construct an index of patent concentration, essentially a Herfindahl-Hirschman Index that sums over each patent class the squares of the fraction of a county's patents belonging to that class:

$$Pat.Concent_{it} = \sum_{c \in C_{it}} \Big(\frac{Num.Pat_c}{\sum_{k \in C_{it}} Num.Pat_k} \Big)^2 \tag{7}$$

where C_{it} is the set of all patent classes in county i at time t. I construct this index using two-digit NBER patent classes, although results are similar with other patent class measures. Results are presented in Column 3 of Table 16. A new college causes concentration to falls by about 13% relative to the 1880 baseline concentration measure, statistically significant at the 5% level. In Column 4, I control for the overall number of patents granted in each county; counties with small numbers of patents will mechanically have higher concentrations. The results are quantitatively similar even when adding this additional control. These results suggest that the diversity of ideas patented increases after the creation of a new college; the extra patents produced are not just in the same fields as previous patents in the college counties. Such a result is inconsistent with most of the increase in patenting in college counties being driven by the skills taught or research conducted in colleges.

While overall patent diversity increases, it may still be the case that some types of colleges specialize in the types of inventions for which they also provide specialized human capital. For instance, land grant colleges might expect to see an increasing share of agricultural patents, or technical colleges might produce a growing fraction of patents in fields such as mining. I verify that the opposite is in fact true. Column 5 of Table 16 shows that the fraction of agricultural patents actually falls in land grant college counties relative to non-land grant college counties after establishing the college. I define a patent to be an agricultural patent if it belongs to a three-digit USPTO patent class that is likely affiliated with agriculture.[32] In column 6, I repeat this exercise but using mining patents and comparing technical schools to non-technical schools.[33] The fraction of mining patents falls, albeit not statistically significantly, in both technical and non-technical college counties after establishing a new college, although the drop is slightly smaller in magnitude in the technical college counties.

C.3 Patent Quality

As Trajtenberg (1990) makes clear, looking at raw patent counts without correcting for patent quality can produce misleading results. Ex ante, it is not clear whether patents in college counties should be expected to increase or decrease in average quality after establishing the college. On one hand, patents coming from more educated inventors might be expected to be of higher quality. On the other hand, more educated individuals, especially those trained in subjects like law, may have better access to the legal system and therefore patent more marginal inventions, leading to lower average quality. A third possibility is that the change in patenting is driven by shifts in the size of the population but not in the distribution of inventive abilities, in which case the distribution of patent qualities may

[32] The one- and two-digit NBER patent classes are much coarser than the USPTO patent classes, so excluding patents related to a specific industry like agriculture are difficult using NBER classes. The USPTO classes also have their issues, namely they are often criticized for being too narrow, not easily mapped to particular industries, and nonsensically organized (Hall et al., 2001). I consider a patent to be an agricultural patent if it belongs to the following USPTO classes: 47 "Plant husbandry"; 54 "Harness for working animal"; 56 "Harvesters"; 71 "Chemistry: fertilizers"; 119 "Animal husbandry"; 278 "Land vehicles: animal draft appliances"; 449 "Bee culture"; 460 "Crop threshing or separating"; or 504 "Plant protecting and regulating compositions". I also experiment with alternative definitions of agricultural patents and get nearly identical results.

[33] Most of the technical schools in the sample that were founded west of the Mississippi were explicitly mining colleges. This is not the case with technical schools in the east, such as Georgia Tech. In addition to mining, these colleges also taught subjects such as engineering. The curricula in eastern technical schools are still more likely to teach subjects similar to mining than are other colleges, however. I consider a patent to be a mining patent if it belongs to the following classes: 175 "Boring or penetrating the earth"; 299 "Mining or in situ disintegration of hard material"; 405 "Hydraulic and earth engineering"; or 507 "Earth boring, well treating, and oil field chemistry". I again explore different definitions of mining patents and get nearly identical results.

Table 16: Patent Classes

	Control for Patent Classes	By Patent Classes	Class Concentration	Class Concentration	Frac. Ag. Pat.	Frac. Mining Pat.
Coll.County * PostColl.	0.356*	0.258*	-0.129**	-0.124**		
	(0.185)	(0.138)	(0.057)	(0.056)		
PostColl.	-0.130	-0.103	0.016	0.014		
	(0.087)	(0.086)	(0.042)	(0.042)		
Num. Pat.				-0.000**		
				(0.000)		
Land Grant Interaction					-0.534	
					(0.599)	
Non-Land Grant Interaction					0.316	
					(0.528)	
Technical School Interaction						-0.065
						(0.354)
Non-Technical School Interaction						-0.228
						(0.202)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Exp. Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
Cnty-Year Obs.	17,449	109,125	20,166	20,166	21,879	21,879
# Counties	178	178	178	178	178	178
# Experiments	72	72	72	72	72	72
Mean in 1880	4.845	0.513	0.554	0.554	0.009	0.013
Adj. R-Sqr.	0.630	0.554	0.423	0.423	0.024	0.013

Notes: Regression results for patent scope. Column 1 estimates the change in the average number of words in a patent's first claim in the college counties relative to the runner-up counties after the establishment of a college. Column 2 estimates the change in the average logged number of words in a patent's first claim. Column 3 estimates the change in the fraction of a county's patents that are at or below the 10th percentile of patents with respect to the length of first claim in each year. Column 4 estimates the change in the fraction of a county's patents that are at or above the 90th percentile of patents with respect to the length of first claim in each year. Each coefficient is transformed into a percentage change in the dependent variable /100. When the model estimates changes in levels, the percentage change is calculated based on the baseline value of the independent variable in 1880. Results are for high quality experiments only. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

not change at all. Following Hall et al. (2001) and Hall et al. (2005), I check whether the number of patent citations and citations per patent change in college counties relative to the runners-up after the establishment of a new college. I thank Enrico Berkes for providing lifetime citation counts for the universe of patents (see Berkes (2018)).

In Column 1 of Table 17, I show that the absolute number of patent citations in college counties increases by 52% relative to the runner-up counties after establishing a new college. This is close in magnitude to the percentage change in the total number of patents granted in college counties. The next three columns confirm that, indeed, there is no measurable change in citations per patent after establishing a new college. Column 2 shows that citations per patent ($\frac{Citations_{it}}{Patents_{it}}$, where $Citations_{it}$ measures lifetime citations for all patents granted in county i in year t) declines in college counties relative to the runners-up after establishing a new college; while the decline is large in percentage terms relative to the 1880 average, it is not statistically significant, and many patents in this time had a very small number of citations so a small decline in citations per patent is a large percentage decline. In Column 2, I omit any counties with zero patents for which the number of patents in the denomination of $\frac{Citations_{it}}{Patents_{it}}$ is zero; in Column 3 I include these counties and code citations per patent to be zero in these cases. The coefficient is again negative but much closer to zero in magnitude and not statistically significant. In Column 3, I also control for the distribution of patent classes in each county as in Column 1 of Table 16. The coefficient is now very close to zero in magnitude and not statistically significant. Summing up, there is no measurable change in citations per patent.

Unfortunately, patent citations are only consistently available beginning in 1947, making them a less-than-ideal measure when using historical patent data. I therefore use an alternative measure to gauge patent quality. As suggested in Kuhn and Thompson (2017), the length of a patent's first claim is a remarkably informative measure of a patent's scope, and hence its quality. A patent's claims formally define the legal scope of an invention. The first listed claim is the most broad. A very short first claim therefore indicates a patent that is very broad in scope, while a long claim indicates a patent that is narrow in scope. Kuhn and Thompson (2017) and Kuhn, Younge, and Marco (2017) argue that patent claim length is in fact more informative of patent quality than citation-based measures. Additionally, unlike patent citations, claims are recorded in the body of a patent for all patents granted in the U.S. from 1836 onward. I use the patent body text and claim counts from Enrico Berkes. I again thank Enrico Berkes for graciously providing this data.

Results are presented in . In Column 5 of Table 17, I re-estimate the baseline regression specification using the average number of words in the first claim for all patents granted within each county in each year as the dependent variable. Column 2 uses the logged number of words in the first claim as the dependent variable. While the estimated coefficients is positive in one case and negative in the other, neither is statistically significant and both are small in magnitude; in both cases the average length of a patent's first claim changes by less than a word in the college counties relative to the runners-up.

One possible reason for the lack of a large average effect, as noted above, is that more high quality patents in college counties could also be offset by more marginal patents. To simply check for this, I estimate whether the share of patents falling in the tails of the distribution of first claim lengths changes in the college counties relative to the runners-up following the establishment of a new college. Column 3 estimates the change in the share of patents at or below the tenth percentile of the first claim length distribution, representing the very broadest patents granted in a particular year. Column 4 estimates the change in the share of patents at or above the 90th percentile, the narrowest patents. Again, neither coefficient is statistically significant or large in magnitude. These results suggest that, while counties that receive a college gain more patents overall, there is no measurable change in patent quality.

C.4 Geographic Spillovers

As mentioned in Section 3.4, if colleges only increase patenting by enticing inventive people to migrate, and if people are more likely to migrate from nearby areas, then the difference in patenting between college counties and nearby areas will increase after the establishment of a new college relative to the difference between college counties are far away areas. I test for this directly in this section.

In Table 18, I present results by the distance from the college county to runner-up counties. Column 1 compares college counties to runner-up counties that are "adjacent" from

Table 17: Patent Quality

	log(Citations + 1)	Citations per Patent	Citations per Patent	Citations per Patent	Length of 1st Claim	log(Length of 1st Claim)	Frac. 1st Claim <10th Pcntl	Frac. 1st Claim >90th Pcntl
Coll.County * PostColl.	0.523**	-0.601	-0.171	-0.015	-0.059	-0.029	-0.023	0.032
	(0.249)	(0.540)	(0.142)	(0.053)	(0.059)	(0.025)	(0.014)	(0.137)
PostColl.	-0.208**	0.275	-0.226	0.037	0.075*	0.026	0.008	0.121
	(0.101)	(0.544)	(0.178)	(0.074)	(0.042)	(0.022)	(0.011)	(0.102)
Num. Patents			-0.873***	0.319***				
			(0.132)	(0.104)				
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exp. Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cnty-Year Obs.	39,580	17,449	39,580	39,580	14,845	14,845	14,845	14,845
# Counties	179	178	179	179	178	178	178	178
# Experiments	72	72	72	72	72	72	72	72
Mean in 1880	77.172	9.153	9.153	9.153	44.411	44.411	0.729	0.089
Adj. R-Sqr.	0.536	0.173	0.106	0.540	0.381	0.445	-0.005	-0.004

Notes: Regression results for patent quality. Column 1 estimates the change in the average number of words in a patent's first claim in the college counties relative to the runner-up counties after the establishment of a college. Column 2 estimates the change in the average logged number of words in a patent's first claim. Column 3 estimates the change in the fraction of a county's patents that are at or below the 10th percentile of patents with respect to the length of first claim in each year. Column 4 estimates the change in the fraction of a county's patents that are at or above the 90th percentile of patents with respect to the length of first claim in each year. Each coefficient is transformed into a percentage change in the dependent variable /100. When the model estimates changes in levels, the percentage change is calculated based on the baseline value of the independent variable in 1880. Results are for high quality experiments only. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

the college county in the sense that they share a common border. The college counties have 23% more patents per year than these adjacent runner-up counties, which is smaller than the baseline estimate in Table 3 and is not statistically different from zero. In Column 2, I compare college counties to counties that are in the same state but do not share a common border. In this column, the estimated increase in patenting from the establishment of a new college is more than twice the magnitude of the estimated increase in Column 1 and is statistically significant, with college counties having 54% more patents per year than the far away runner-up counties. These results show that the college counties increase patenting much more than distant runners-up, but are statistically indistinguishable from their closer neighbors. This suggests that, instead of a new college having negative spillovers on neighboring areas by pulling all of the local talent away, colleges have positive geographic spillovers, benefiting neighboring areas as well as the county that actually receives the college.

In Columns 3 and 4, I extend this result to compare the college counties to *all* non-college counties; I no longer restrict attention to the runner-up counties. Column 3 compares the college counties to all adjacent counties and finds that college counties have 51% more patents per year after the establishment of the college. Column 4 compares the college counties to all non-adjacent counties that do not share a border; in this case, the college counties have 70% more patents per year. Thus, even when attention is not restricted to the counterfactual sites, which may not be randomly distributed across a state, it appears that the college counties grow somewhat similarly to their neighbors, but increase patenting by much more than far away locations.

Additionally, the college and runner-up sites vary in their location within counties: some are located very close to county borders, for instance, while some are close to the geometric midpoint of their respective counties. To address these concerns, I also use information on the precise latitude and longitude of college towns and runner-up towns and then compare cases in which the college and runner-up towns are closer and further from specific distance thresholds. Results using the geodesic distance between college and runner-up counties and various cutoffs between "nearby" and "far-away" runners-up produce results nearly identical to those using adjacent counties. These results are available upon request.

Table 18: Patenting by Geographic Distance

	Adjacent Runners-Up	No Adjacent Runners-Up	Adjacent Counties	No Adjacent Counties
Coll.County * PostColl.	0.240	0.506***	0.529***	0.725***
	(0.152)	(0.166)	(0.151)	(0.161)
PostColl.	-0.065	-0.103	-0.040	0.051***
	(0.102)	(0.073)	(0.052)	(0.008)
County Fixed Effects	Yes	Yes	Yes	Yes
Exp. Fixed Effects	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Cnty-Year Obs.	22,376	29,804	79,847	962,386
# Counties	117	138	370	1,722
# Experiments	72	72	72	72
Mean in 1880	4.845	4.845	4.845	4.845
Adj. R-Sqr.	0.490	0.513	0.518	0.540

Notes: Regression results for patenting rate and results by distance to runner-up counties. Column 1 compares college counties to runner-up counties that share a border. Column 2 compares college counties to runner-up counties that do not share a common border. Column 3 compares college counties to all counties that share a border. Column 4 compares college counties to all counties that do not share a common border. The dependent variable for all columns is $log(Patents+1)$. Each coefficient is transformed into a percentage change in the dependent variable /100. When the model estimates changes in levels, the percentage change is calculated based on the baseline value of the independent variable in 1880. Results are for high quality experiments only. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

D Additional Results on the Necessity of Colleges

D.1 Additional Results by Type of College

In this section, I further break down the results of different types of colleges on patenting. In Column 1 of Table 19, I show how patenting differs between practical and classical colleges, using an alternative classification of practical and classical than described in Section ??. Here, a college is considered a practical college if it is a land grant college, technical school, or military academy. Classical colleges are normal schools, other private and public colleges, and HBCUs. The difference between practical and classical colleges is qualitatively the same when the alternative definitions are used as in the baseline results presented in Table 19, with practical colleges having 41% more patents per year compared to 38% for classical colleges.

I also compare differences between each of the seven types of colleges: land grants, technical schools, military academies, normal schools, HBCUs, other public colleges, and other private colleges. Unfortunately, as Table 2 shows, there is only a small number of several types of experiments and so insufficient power to identify differences. Even simply comparing coefficients, however, paints a picture that does not conform to the naive intuition that colleges that focus on more practical skills should cause larger increases in patenting. For example, normal schools and land grant colleges produce nearly the same increase in patenting, while the former is focused on training primary and secondary school teachers and the latter has an explicit focus on very practical fields such as agriculture and machinery. These results are available upon request.

I next compare all public schools to all private schools. This involves reclassifying colleges, as some of the types described above may include both public and private colleges. For instance, the HBCUs may be either public or private. Cornell University, while officially New York's land grant university, is actually a private institution. I interact dummy variables for public or private status with the estimated college effect and display the results in Column 2. I find that public colleges have a large positive effect on patenting, while the effect for private colleges is less than half the magnitude and not statistically different from zero.

In Column 3, I check how the estimated treatment effect varies by college quality. Unfortunately, reliable data on college quality does not exist for most of each college's history. Instead, I proxy lifetime college quality with the 2018 national universities rankings in the U.S. New and World Reports (`https://www.usnews.com/best-colleges/rankings/national-universities`). This is problematic because current college rankings may be due in part to college's past patenting performance, but the measure may still be informative if rankings are highly persistent over time. I split colleges into four groups: those ranked 1-75, those ranked 76-150, those ranked 151-225, and those that do not have a 2018 U.S. News ranking. The estimated coefficient declines monotonically moving from the highest to lowest quality schools in the U.S. News rankings. These results suggest that higher quality colleges lead to more local patenting. It may be the case that better colleges are larger, and it is the size of the institution that drives patenting rather than measures of quality. To try and account for this, in Column 4 I control for $\log(Students + 1)$ using data from the Commissioner of Education reports. While the schools ranked 1-75 again have the largest increase in patenting, it is not statistically significant. Moreover, the schools ranked 75-150 see a sizable drop in patenting that is significant at the 10% level. It is thus difficult to conclude that higher quality schools monotonically lead to more patents. In Column 5, I control for $\log(TotalCountyPop.)$ instead. The coefficient again declines monotonically moving from highest to lowest quality schools, but the differences between the coefficients are much smaller in magnitude and no group of colleges produces an increase in patenting that is statistically different from zero.

E Yearbook Data

To determine whether a particular patentee is an alumni or faculty member of a particular college, I digitize historical college yearbooks to obtain names of individuals affiliated with each college. Scanned images of a large number of college yearbooks are available on `www.ancestry.com`. After obtaining the yearbook images, I transcribe them to obtain relevant information. Table 20 lists the colleges for which yearbook data has been transcribed, including the number of yearbooks available for each college, the first and last transcribed year, and the number of transcribed records for undergraduate alumni, graduate alumni, and faculty. Due to budget limitations, and due to the fact that future work will link students

Table 19: Additional Results by College Type

	Alt. Practical vs. Classical	Public vs. Private	College Rank	College Rank	College Rank
Practical College Interaction	0.411**				
	(0.205)				
Classical College Interaction	0.375*				
	(0.211)				
All Public Colleges		0.409***			
		(0.150)			
All Private Colleges		0.164			
		(0.315)			
Rank 1-75			0.866*	0.454	0.386
			(0.500)	(0.442)	(0.236)
Rank 76-150			0.474*	-0.215*	0.299
			(0.254)	(0.124)	(0.283)
Rank 151-225			0.354**	0.042	0.135
			(0.178)	(0.130)	(0.197)
Unranked			0.188	-0.012	0.095
			(0.262)	(0.130)	(0.190)
log(Students + 1)				0.041***	
				(0.012)	
log(Total Pop.)					0.450***
					(0.074)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes
Exp. Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes
Cnty-Year Obs.	39,580	39,580	39,580	2,964	36,378
# Counties	179	179	179	179	179
# Experiments	72	72	72	72	72
Mean in 1880	4.845	4.845	4.845	4.845	4.845
Adj. R-Sqr.	0.508	0.509	0.512	0.405	0.540

Notes: Regression results by college type. The dependent variable is $log(Patents + 1)$. In Column 1, the effect of establishing a new college is estimated separately for practical and classical colleges, using the alternate definition described in the text. The dependent variable in Column 2 is the fraction of agricultural patents, using the alternate definition of agricultural patents described in the text, $Alt.Num.AgriculturalPatents_{ijt}/Num.Patents_{ijt}$. In Column 3, the coefficient is the percentage increase in patenting caused by the college interacted by college type. All other coefficients in the regression are suppressed for readability. Row 1 presents results for the land grant college experiments, row 2 for technical colleges, row 3 for normal schools, row 4 for HBCUs, row 5 for military academies, row 6 for other public colleges, and row 7 for other private colleges. In Column 4, the coefficient is the percentage increase in patenting caused by the college interacted with whether a college is public or private. In Column 5, the coefficient is the percentage increase in patenting caused by the college interacted with each college's rank according to the 2018 U.S. News and World Report rankings. Each coefficient is transformed into a percentage change in the dependent variable /100. When the model estimates changes in levels, the percentage change is calculated based on the baseline value of the independent variable in 1880. Results are for high quality experiments only. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

in the yearbooks to individuals in the U.S. census and the 1940 census is the most recent that is available, no yearbooks have been transcribed for years more recent than 1940. Other colleges will be added to the data as the data is transcribed.

Because I handle the data for college alumni and faculty slightly differently, I describe them separately below.

Table 20: Yearbook Data Summary Statistics

	College	Num. Yearbooks	First Yearbook	Last Yearbook	Num. Undergrads	Num. Grad. Studs.	Num. Faculty
1	Auburn University	8	1916	1940	2573	7	202
2	Clemson University	5	1915	1940	1187	0	83
3	Cornell University	40	1879	1935	23339	2977	15473
4	Georgia Institute of Technology	14	1917	1940	3361	0	968
5	Iowa State University	3	1936	1939	1668	0	59
6	Louisiana State University	7	1927	1940	3528	713	83
7	Missouri University of Science and Technology	10	1911	1940	577	0	505
8	North Carolina A and T University	1	1939	1939	97	0	42
9	North Dakota State University	16	1908	1940	2956	0	310
10	Texas Tech	2	1937	1940	710	8	133
11	University of Arizona	9	1913	1940	1583	36	438
12	University of Colorado	19	1893	1939	3775	0	1562
13	University of Maine	28	1906	1940	4718	866	4343
14	University of Missouri	33	1898	1940	9792	574	1547
15	University of Nevada	6	1901	1934	408	0	165
16	University of New Hampshire	10	1909	1940	2164	0	1630
17	University of North Dakota	5	1906	1940	920	0	68
18	Utah State University	4	1911	1932	489	0	27
19	Virginia Polytechnic Institute	17	1898	1939	2125	50	788
20	Washington State University	12	1903	1940	4136	0	317

Notes: List of colleges for which yearbooks are transcribed. For each college, also listed is the total number of yearbooks transcribed, the earliest and the most recent transcribed yearbook, and the total number of transcribed records for undergraduate students, graduate students, and faculty.

E.0.1 Alumni

Figure 10 shows an example of a college yearbook page. This particular image shows college seniors from the University of New Hampshire's yearbook for 1910, but it is representative of the type of information available in a typical college yearbook. Unfortunately, the type of information available and formatting of each yearbook vary enormously from college to college or even by year within the same college. This makes analysis using particular types of information difficult, as it may not be available for most years. But almost all yearbooks include the names of college seniors along with their majors. Many also include seniors' hometowns, sports teams or clubs, fraternities or sororities, or professional organizations, and often this information is available for juniors or underclassmen as well.

Because I am interested in constructing a list of alumni from a particular college, I keep information only for college seniors. The assumption is that the vast majority of these individuals go on to become alumni in the following year; juniors will become seniors in the following yearbook, so ignoring them during their pre-graduation years saves on time and expense during the transcription process and prevents accidentally inflating the number of graduates from a particular year. I also record the number of graduate students, if available, for each year. Unfortunately, many yearbooks do not list their graduate students, so the data is somewhat limited. It is also typically impossible to know what year graduate students are expected to graduate; for instance an individual just beginning their PhD might remain a graduate student for another five years before becoming an alumni. Most graduate students belong to shorter programs, however, so I include all listed graduate students in each year when available.

I next compile a list of all past seniors and graduate students. I assume that patentees must be less than 80 years old. While information on the age distribution of historical inventors is sparse (see Sarada et al. (2017) and Akcigit et al. (2017a) for recent exceptions), modern data shows that very few inventors are that old (Jung and Ejermo (2014), Acemoglu et al. (2014)) and, if anything, the age of invention has been increasing in recent decades (Jones (2009), Jones (2010)). I further assume, for simplicity, that each college graduate is no less than 20 years old at time of graduation. For each year, I compile a list of alumni names in year t by combining each the seniors and graduate students from each yearbook from $t-60$ to t (since the assumptions mean that each alumni can patent for up to 60 years after graduating). Such a list consists of all alumni for whom a name is known for each college and each year. I drop all duplicate names from this list, which further alleviates problems from accidentally recording a student in a year before he or she graduates.[34]

I also construct a time series of the number of expected alumni in each year. To do this, I interpolate the number of seniors and graduate students in each year for which a yearbook is not available, using a cubic-spline interpolation.[35] For years before the first college yearbook

[34]As the alumni will be matched to the patent record by name, discarding duplicate names does not affect the number of patents attributed to alumni.

[35]I interpolate the number of graduate students for each year after the first year in which graduate students are observed. The assumption is that colleges very rarely discontinue their entire graduate programs; instead, no observed graduate students is likely simply due to a yearbook not recording the graduate students in a particular year.

or after the last college yearbook, I extrapolate the number of seniors and graduate students linearly; non-linear extrapolations lead to nonsensical predictions for several of the colleges. I set the number of seniors and graduate students to zero for all years before the establishment of the college. For years in which the extrapolation or interpolation predicts fewer seniors and graduate students than the smallest number observed in a yearbook, I replace that value with the smallest observed number. The time series of expected seniors and graduate students thus likely overstates the number of alumni in each year. In a few cases, the interpolation or extrapolation leads to implausibly large numbers of seniors or graduate students in a given year (namely, larger than the corresponding county population); to fix this, I drop the top 1% of observations by expected number of seniors and graduate students. I then sum up all the expected alumni for each of the previous 60 years as described above. This gives a list of the expected number of total alumni for every year, denoted by $Num.\bar{A}lumni_{jt}$ for college j in year t.

To determine whether a particular patent belongs to an alumnus, I match each individual in the alumni list by first, middle, and last name to the patent data in the college's county for the corresponding year. To clarify, this matching does not find *all* patents belonging to a particular alumnus, but rather only patents by the alumnus that occur in the county from which he or she obtained her degree. To search for name matches between the yearbooks and the census, I use a fuzzy matching algorithm as in Sarada et al. (2017). More specifically, I use Stata's reclink command, which is a modified bigram string comparator that returns a "distance" (match score) between two strings. Only matches with a sufficiently high match score are retained. Because at this point I am interested in the "most lenient" match of graduates to patents, I keep all plausible matches, regardless of the possibility that graduates living in a college county may share a name with a non-graduate living in the same county. Moreover, this procedure will attribute a patent to a college graduate if the graduate moves to another county but a different individual with the same name obtains a patent in the college county.

To calculate the alumni patenting rate, for each year I divide the number of patents matched to alumni to the total number of alumni with identifiable names,

$$AlumniPatentRate_j = \frac{1}{T - t_{0j}} \sum_{t_{0j}}^{T} \frac{Num.Matched_{jt}}{Num.Alumni_{jt}}, \tag{8}$$

for college j and years t_{0j} is the first year for which a yearbook for college j is available and $T = 2000$ (the last yearbook year, 1940, plus 60 years). I compute this patenting rate separately for alumni seniors and graduate students. Finally, I calculate the expected number of patents coming from alumni by multiplying the computed patenting rate by the expected number of alumni each year, $Num.AlumniPatents_j = \sum_{t_{0j}}^{T} AlumniPatentRate_j * Num.\bar{A}lumni_{jt}$.

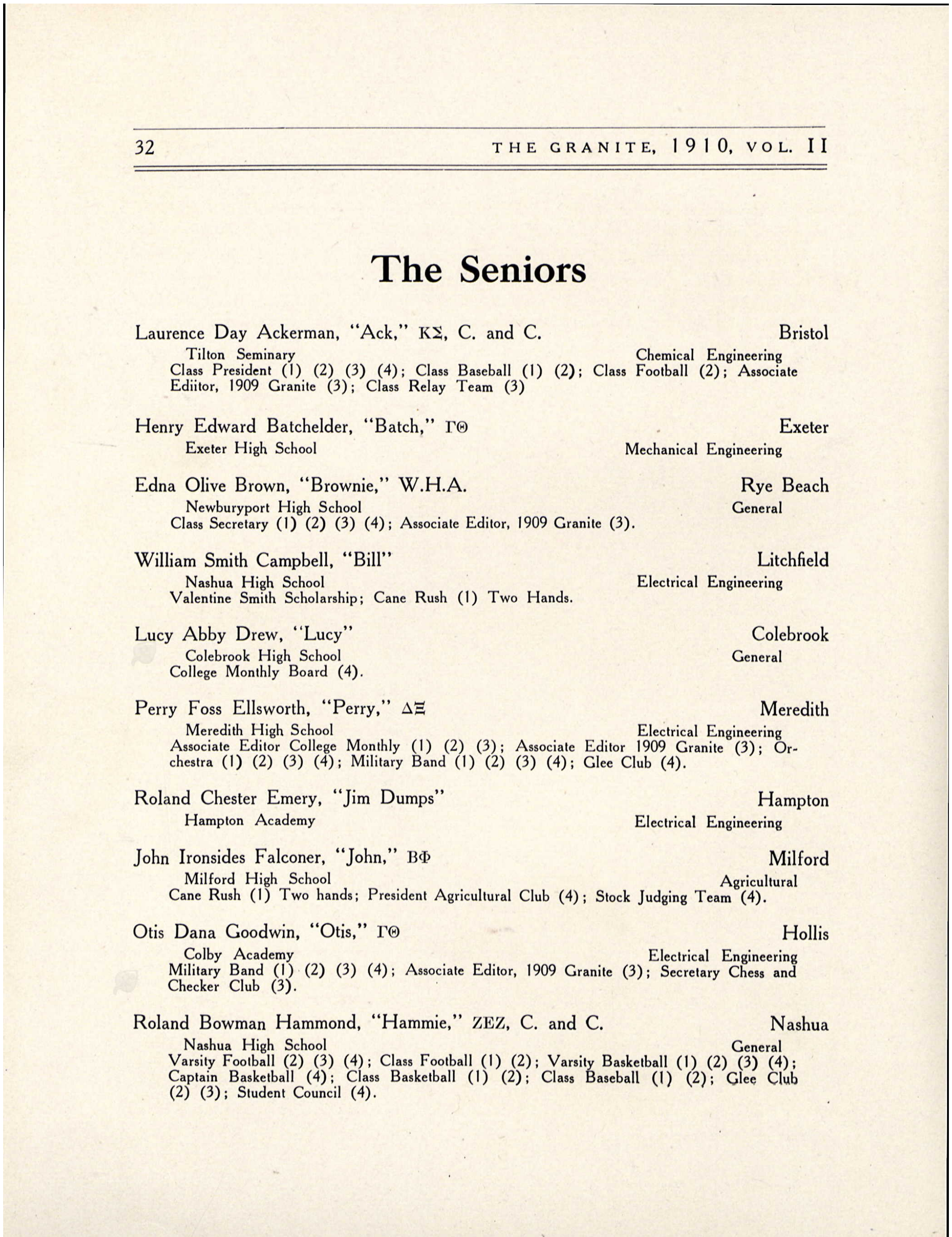
E.0.2 Faculty

Figure 11 shows an example of a college yearbook page with faculty information, also from the University of New Hampshire's 1910 yearbook. In this particular yearbook, each faculty member's name is listed along with his highest degree obtained, position and title at the university, and a biography that describes each member's academic subject and any previous academic positions held.

Unfortunately, the number of faculty members included and the information provided on each varies much more than does the alumni information. While every yearbook has a page dedicated to the university president, nearly all list administrative officers such as the registrar, and a majority of the yearbooks list the deans of the different schools within the college, many do not include a full list of the faculty.

I begin by transcribing all faculty information provided for each college and each year, just as in the case of the alumni. I discard any years with five fewer faculty members listed, as these are likely cases when only the university president or a handful of administrative officers are included. Since it is unlikely for faculty to cease serving at their college except through death or transfer to another college, I match faculty names to the patent record only for the year in which the faculty name appears in the yearbook. Matching uses the same information as in the alumni case above. Even in cases in which faculty last names are listed, frequently only the first and middle initials are included, making matching very

Figure 10: Example College Yearbook Page, Student Data



32 THE GRANITE, 1910, VOL. II

The Seniors

Laurence Day Ackerman, "Ack," KΣ, C. and C. — Bristol
Tilton Seminary — Chemical Engineering
Class President (1) (2) (3) (4); Class Baseball (1) (2); Class Football (2); Associate Ediitor, 1909 Granite (3); Class Relay Team (3)

Henry Edward Batchelder, "Batch," ΓΘ — Exeter
Exeter High School — Mechanical Engineering

Edna Olive Brown, "Brownie," W.H.A. — Rye Beach
Newburyport High School — General
Class Secretary (1) (2) (3) (4); Associate Editor, 1909 Granite (3).

William Smith Campbell, "Bill" — Litchfield
Nashua High School — Electrical Engineering
Valentine Smith Scholarship; Cane Rush (1) Two Hands.

Lucy Abby Drew, "Lucy" — Colebrook
Colebrook High School — General
College Monthly Board (4).

Perry Foss Ellsworth, "Perry," ΔΞ — Meredith
Meredith High School — Electrical Engineering
Associate Editor College Monthly (1) (2) (3); Associate Editor 1909 Granite (3); Orchestra (1) (2) (3) (4); Military Band (1) (2) (3) (4); Glee Club (4).

Roland Chester Emery, "Jim Dumps" — Hampton
Hampton Academy — Electrical Engineering

John Ironsides Falconer, "John," BΦ — Milford
Milford High School — Agricultural
Cane Rush (1) Two hands; President Agricultural Club (4); Stock Judging Team (4).

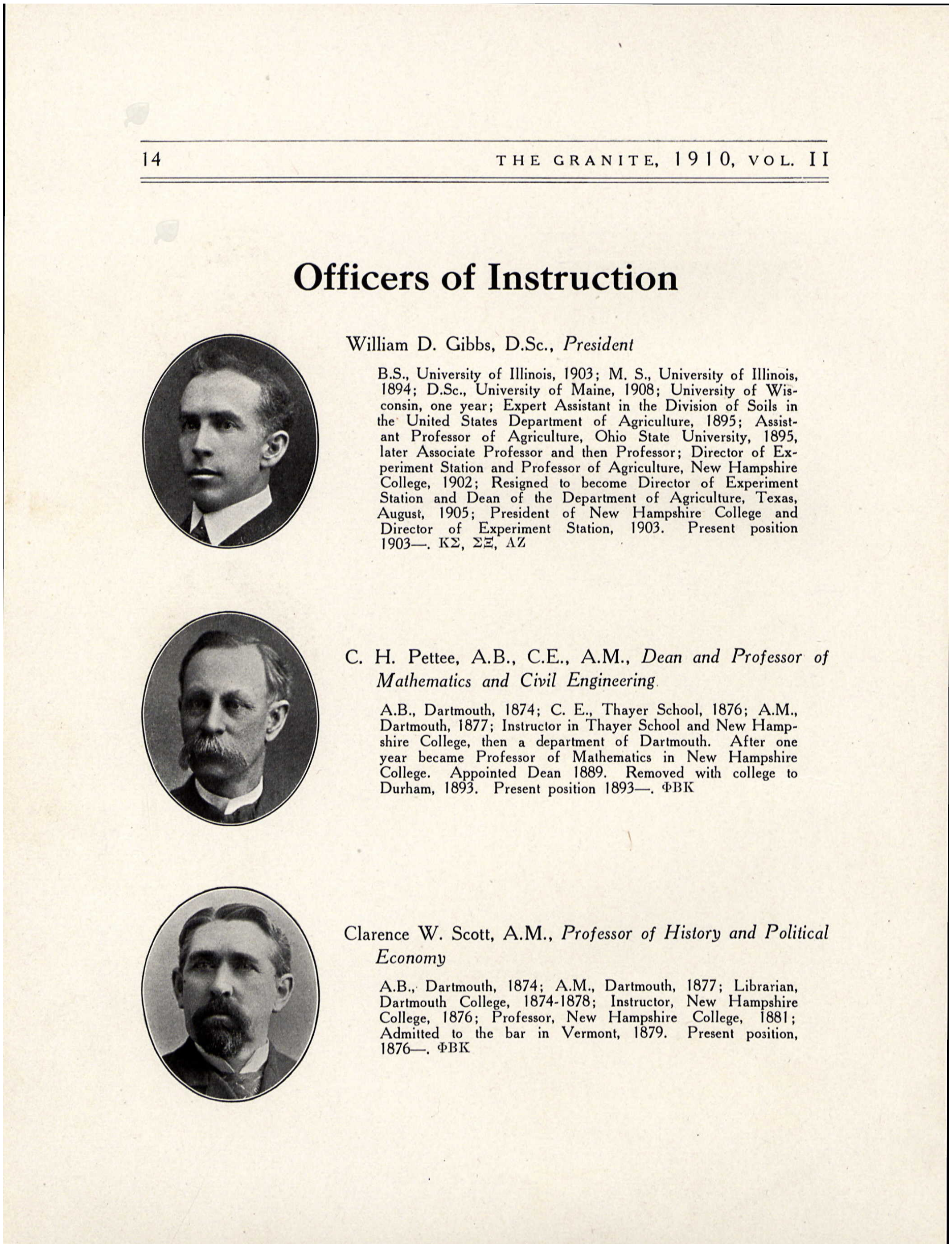
Otis Dana Goodwin, "Otis," ΓΘ — Hollis
Colby Academy — Electrical Engineering
Military Band (1) (2) (3) (4); Associate Editor, 1909 Granite (3); Secretary Chess and Checker Club (3).

Roland Bowman Hammond, "Hammie," ZEZ, C. and C. — Nashua
Nashua High School — General
Varsity Football (2) (3) (4); Class Football (1) (2); Varsity Basketball (1) (2) (3) (4); Captain Basketball (4); Class Basketball (1) (2); Class Baseball (1) (2); Glee Club (2) (3); Student Council (4).

Notes: An example page from one of the transcribed college yearbooks showing information on college seniors. This image is from the 1910 University of New Hampshire yearbook.

difficult. I discard these individuals, as including them would artificially lower the faculty patenting rate. Once the faculty are matched to the patent record, I calculate the patenting rate $FacultyPatentRate_j$ and $Num.FacultyPatents_j$ for each college j in the same way as I calculate those values for the alumni, described above.

Figure 11: Example College Yearbook Page, Faculty Data



14 THE GRANITE, 1910, VOL. II

Officers of Instruction

William D. Gibbs, D.Sc., *President*

B.S., University of Illinois, 1903; M. S., University of Illinois, 1894; D.Sc., University of Maine, 1908; University of Wisconsin, one year; Expert Assistant in the Division of Soils in the United States Department of Agriculture, 1895; Assistant Professor of Agriculture, Ohio State University, 1895, later Associate Professor and then Professor; Director of Experiment Station and Professor of Agriculture, New Hampshire College, 1902; Resigned to become Director of Experiment Station and Dean of the Department of Agriculture, Texas, August, 1905; President of New Hampshire College and Director of Experiment Station, 1903. Present position 1903—. ΚΣ, ΣΞ, ΑΖ

C. H. Pettee, A.B., C.E., A.M., *Dean and Professor of Mathematics and Civil Engineering*

A.B., Dartmouth, 1874; C. E., Thayer School, 1876; A.M., Dartmouth, 1877; Instructor in Thayer School and New Hampshire College, then a department of Dartmouth. After one year became Professor of Mathematics in New Hampshire College. Appointed Dean 1889. Removed with college to Durham, 1893. Present position 1893—. ΦΒΚ

Clarence W. Scott, A.M., *Professor of History and Political Economy*

A.B., Dartmouth, 1874; A.M., Dartmouth, 1877; Librarian, Dartmouth College, 1874-1878; Instructor, New Hampshire College, 1876; Professor, New Hampshire College, 1881; Admitted to the bar in Vermont, 1879. Present position, 1876—. ΦΒΚ

Notes: An example page from one of the transcribed college yearbooks showing information on college faculty. This image is from the 1910 University of New Hampshire yearbook.