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

Michael Andrews†

October 8, 2017

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 "losing finalist" locations that were strongly considered to become the site of a new college but were ultimately not chosen for plausibly exogenous reasons. The losing finalists are very similar to the winning college counties along observable dimensions. Using the losing finalists as counterfactuals, I find that the establishment of a new college caused 33% more patents per year in college counties relative to the losing finalists. To determine the channels by which colleges increase patenting, I use a novel dataset of college yearbooks and individual-level census data to learn who the additional patents in college counties come from. A college's alumni account for about 10% of the additional patents, while faculty account for less than 1%. Knowledge spillovers to individuals unaffiliated with the college or living in the college county prior to the establishment of the new college also account for less than 1% of the additional patents. Migration is the primary channel by which colleges affect local invention, as controlling for county population accounts for 40-65% of the increase in patenting in college counties relative to the losing finalists. In spite

*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, Joe Ferrie, Shawn Kantor, Joel Mokyr, Sarada, and Nicolas Ziebarth, as well as from seminar participants at the University of Iowa, Kauffman Foundation, Northwestern University, NBER DAE Summer Institute, the University of Wisconsin, Auburn University, Economic History Association Annual Meetings, NBER Productivity lunch, and Illinois Economic Association. All errors are my own. The most recent version of this paper is available at http://sites.google.com/site/michaeljeffreyandrews/research.

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

of this, the presence of geographic spillovers suggests that colleges cause an overall net increase in patenting, although I find no evidence that colleges are better at promoting invention than other policies that lead to similar levels of urbanization.

1 Introduction

How do colleges affect innovation? Answering this question is vital in light of recent debates about supporting higher education in the U.S. (for instance, Press (2013), Science News Staff (2017), New York Times Editorial Board (2017)). The literature has proposed four channels by which colleges can increase invention. First, colleges provide human capital to students; more educated alumni can go on to become inventors. Second, faculty research may have direct commercial applications. Third, knowledge may spill over from alumni or faculty to increase the inventiveness of nearby individuals whoa re not directly affiliated with a college. Fourth, inventive people may migrate to live close to colleges. While each of these channels is plausible, estimating the relative importance of each has proven difficult. In this paper, I use historical data on the establishment of U.S. colleges to understand the role of each channel in promoting local invention.

To understand the importance of each of these channels, it is first necessary to determine if colleges have any causal effect on invention, and if so how much. A large literature documents cross-county correlations between higher education and innovation, productivity, and growth (Barro (1991), Benhabib & Spiegel (1994), Krueger & Lindahl (2001), Furman et al. (2002), Cohen & Soto (2007), Barro & Lee (2013), Valero & Van Reenen (2016)). Another large body of work notes that patenting tends to be more intense in places that are geographically close to colleges and universities, and that this relationship is strongest near the most research-intensive institutions (the seminal paper is Jaffe (1989)). Determining the extent to which colleges cause these increases in patenting has proven difficult, although several recent studies attempt to estimate local causal effects (Aghion et al. (2009), Whalley & Hicks (2014), N. Hausman (2013)). The reason for this difficulty is clear: colleges are not located at random. As Gumprecht (2003) points out, college towns differ from other places both in ways that are easily detected in economic statistics and in ways that are more subtle, making comparisons difficult. As N. 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...Of course, in reality universities exist non-randomly from their locations, and areas with universities differ from those without" (N. Hausman, 2013, p. 10).

I approximate this ideal experiment using data on the establishment of new colleges in the U.S., spanning the years 1839-1954. By exploring the narrative historical record, I am able to identify "losing finalist" sites that were strongly considered to become the site of a new college, adopting a methodology pioneered by Greenstone et al. (2010) to study the site selection decisions of large manufacturing plants.[1] Using the losing finalists as counterfactuals for the winning college counties, I find that establishing a new college causes 33% more patents per year in the winning county relative to the losing finalists.

The key idea behind this approach is that, when selecting where to locate a major investment, be it a college or a manufacturing plant, thousands of possible candidates 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 "losing finalists" 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 and the finalists are 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. I further exclude any cases in which losing finalists are very different from winning counties along observable dimensions.

A concrete example of a college site selection experiment with as-good-as-random assignment of the college may be useful. Georgia School of Technology, better known today as Georgia Tech, was founded in Atlanta in 1886. At the time, prominent citizens in Georgia wanted a technical college, but there was no consensus about where to put it. A number of cities vied to receive the new school. Two of the main rival sites were Atlanta and Macon. Both were known primarily as railway depots located in the interior of the state; the two cities also looked similar along a number of observable dimensions. In October 1886, a site

[1]Kantor & Whalley (2016), Malmendier et al. (2016), and Helmers & Overman (2017) apply similar methodologies in other contexts.

selection committee assembled to vote on the location of the college. For the first 23 ballots, neither Atlanta nor Macon obtained the requisite majority of votes. 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. In an even more striking case, representatives from North Dakota towns drew lots to determine the location of the University of North Dakota and North Dakota State University (Geiger, 1958, p. 13-27). The cases of Georgia Tech and the North Dakota universities were not isolated incidents: while the decisions were occasionally less dramatic, these kinds of college site selection experiments occurred all across the United States, in large towns and small, during the second half of the nineteenth century and first half of the twentieth.

The augmented losing finalist methodology is a valuable tool for estimating the causal effect of establishing a new college. Most previous studies that specifically focus on the establishment of new colleges assumes that colleges are located at random (Currie & Moretti (2003), Moretti (2004), Andersson et al. (2004), Andersson et al. (2009), Frenette (2009), Cowan & Zinovyeva (2013), and Toivanen & Väänänen (2016)). I show that such an assumption greatly overstates the effect of colleges on local invention.

Having documented that colleges cause an increase in local patenting, I next turn to understanding the channels through which this occurs. Using rich individual-level data, I am able to calculate which share of the additional 33% of patents in college counties come from each of the channels mentioned above.

Perhaps the most obvious way in which colleges can increase local invention is by providing human capital to their students. Several recent papers show that college-educated individuals are more likely to invent (Jung & Ejermo (2014), Bell et al. (2016), Aghion et al. (2016), Akcigit et al. (2017)). While seemingly obvious, this fact is difficult to document without detailed inventor-level data which has been unavailable until recently. While these studies show that educated people are more likely to patent, they do not link college graduates to their alma maters, so it is impossible to know how particular colleges affect invention by their alumni. I overcome this problem by transcribing college yearbooks from 1880 to 1940 to create a novel dataset of the names of college graduates. I then match these names to the patent record. Surprisingly, college graduates account for only 7% of all patents and about 10% of the additional patents in the counties of their alma maters. To clarify, this finding does not suggest that college graduates failed to play a major role in driving American invention; rather, if college graduates did go on to patent, they did not do it in the counties from which they obtained their degrees.

College faculty may also account for a large share of the patents in college counties if their research has commercial applications. The college yearbooks typically contain information about the faculty at each college, so I can also match by names of faculty to the patent record as well. Faculty members account for less than 1% of the additional patents in college counties. This result is consistent with findings that patenting by college faculty was exceedingly rare until recent decades (Mowery & Sampat (2001), Sampat (2006), Berman (2008), Lee (2013)).

I next investigate whether new colleges produce large knowledge spillovers to other individuals living in a college town that did not attend college. Most prior research investigating knowledge spillovers considers any increase in patenting by firms located close to a college to be due to knowledge spillovers (for example, Jaffe (1989), Zucker, Darby, & Brewer (1998), Adams (2002), Furman & MacGarvie (2007), N. Hausman (2013)). One challenge, as Zucker, Darby, & Armstrong (1998) point out, is that nearby firms may still employ college alumni or consult with college faculty, and therefore increases in their patenting may not be true spillovers. Estimating knowledge spillovers thus entails the opposite challenge from identifying patenting by alumni or faculty: one must find individuals who are *not* affiliated with a particular college. To do this, I observe how patenting changes for individuals who were living in college counties prior to the establishment of the new college, but were too old to attend the new school. These individuals account for a very small share of the extra patents, although this result is very preliminary.I also consider patenting by blacks living close to segregated colleges and women living close to non-coeducational colleges. While it looks like these groups increase their patenting after the establishment of a college, they can account for only a very small share of the additional patents.

To further clarify the role of alumni human capital, faculty research, and knowledge spillovers, I investigate patenting by different college types. If these three channels were very important, then colleges that focus on more "practical" skills (for example, land grant colleges that had to emphasize agriculture and the mechanical arts, or engineering schools) should see larger increases in patenting than colleges that do not have such a focus. Instead, I find very small differences in the increase in patenting across different types of colleges. Moreover, patents in colleges counties do not become more concentrated in patent classes that correspond to the local curriculum (for instance, agricultural patents near land grant colleges), but instead patenting becomes more diverse in the college counties relative to the losing finalists.

The fourth and final channel that I examine is migration. One of the few papers to examine how higher education directly influences migration is Aghion et al. (2009), who show both theoretically and empirically using political shocks that highly educated people are likely to migrate to live close to other highly educated people. They suggest that highly educated people may even leave college counties if it is not too costly to move to an urban area with even more highly skilled individuals, echoing results in Bound et al. (2004). Several other papers study the international migration patterns of inventors and find that inventive people are highly mobile and respond to changes in local conditions (for example, Kerr & Lincoln (2010), Akcigit et al. (2016)). A separate strand of the literature attempts to measure local spillovers from academics 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 highly productive researcher helps to attract other productive researchers (Azoulay et al. (2010), Waldinger (2012), Borjas & Doran (2012), Dubois et al. (2014)).[2] I likewise find that establishing a new college leads to large increase in local population and urbanization. Controlling for population can explain the majority of the increase in patenting in college counties, on the order of 40-65%, depending on the specification used.

In light of these results, an important question is whether establishing a college causes a net increase in patenting, or simply moves inventive people from one location to another. While data limitations make estimating the full general equilibrium effect of establishing a college on invention difficult to measure, I search for other evidence of positive externalities from colleges. The evidence is mixed. On one hand, the results of human capital, faculty

[2] While Azoulay et al. (2010) find that the death of a superstar academic has strong negative spillovers on their network of collaborators, there is no measurable negative effect for other academics that are geographically co-located with the superstar.

research, and knowledge spillovers presented above suggest that colleges must produce some net increase in patenting, even if the size is modest. On the other hand, I find no evidence that agglomeration economies are stronger in college counties than in control counties with comparable population. I do find that, rather than simply attracting inventive individuals from nearby areas, colleges tend to increase patenting in nearby locations, providing further evidence of positive spillovers from colleges. I thus conclude that colleges do lead to a net increase in patenting, although the results on migration and agglomeration suggest that the net increase may be modest in size.

This paper is organized as follows. Section 2 describes the data, including an in depth explanation of the college site selection experiments as well as the patent and other data sources used, and presents the empirical specification. Section 3.1 presents the baseline results for the total effect of the establishment of a new college on patenting. Section 4 conducts additional analysis to understand the channels by which colleges increase invention. Section 4.1 documents patenting that results from the human capital of college alumni. Section 4.2 details the low aggregate number of patents produced by college faculty. Section 4.3 shows how many patents can be accounted for by knowledge spillovers to individuals not affiliated with the college. In Section 4.4, I present additional evidence that human capital, faculty research, and knowledge spillovers account for modest shares of patents in college counties. Section 4.5 shows that most of the increase in patenting comes from individuals who migrate to the college county. In Section 5, I discuss whether colleges lead to a net increase in patenting or simply reallocate inventors across space. Section 6 concludes.

2 Data and Empirical Model

2.1 The College Site Selection Experiments

The mid-19th to mid-20th centuries saw an explosion in the number of colleges and universities in the U.S. Goldin & Katz (2008) refer to the 20th century as the "human capital century" due to the large increase in schooling. In other work, they calculate that roughly 630 new colleges were opened from 1890 to 1940 (Goldin & Katz, 1999). The six decades prior to 1890 also saw the establishment of numerous new colleges and universities, most notably the establishment of public universities in most states, a process supplemented and accelerated by the passage of the Morrill Land Grant Colleges Act in 1862.

To study the local effects of creating a new college, it is important to identify counterfactual locations that did not receive a collage but are otherwise similar to college counties. Several authors claim that the locations of new colleges were essentially random. For instance, Moretti (2004, p. 190-191), focusing exclusively on land grant colleges, argues 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." While it may be difficult to determine why colleges are in their current locations by looking at modern data, a great deal of thought went into the site selection decision at the time they occurred. Horace Bushnell, a theologian who played a central role in locating both the University of California and the University of Illinois, summed up how thoughtfully the site selection decision was approached: "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)). Moreover, many localities wanted to secure a new college, and any prestige and economic benefits that went along with it, for themselves, ensuring that the site selection decision often became quite contentious. Further complicating the site selection decision is the fact that new colleges 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. There was also significant tension between locating new colleges in rural areas, where it was thought that the mostly uneducated agricultural families could most benefit from easy access to higher education, versus cities with ready access to urban amenities. As one author writes, many questioned "the naive American assumption that small towns in the forests were as suitable for the life of the mind as large cities." (P. Miller, 1961, p. xxiv).

To identify counterfactual sites, I find historical information, often in narrative form, regarding the college site selection process. From studying the site selection process, it is often possible to identify losing "finalist" counties that did not receive the college. Throughout, I refer to counties that received a new college as "treatment" counties and the losing finalists as "control" counties. Kantor & Whalley (2016) adopt a similar approach to identify finalist sites to land grant colleges in the Northeast and Midwest. 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. To mitigate this problem, I only include cases in which the site selection decision is plausibly exogenous; I refer to these as "high quality" college selection experiments.

The Georgia Tech and North Dakota examples described in the Introduction above illustrate two such high quality experiments. As another example of high quality site selection experiments, in several instances states solicited bids from localities, with the college going to the place with the highest bid. If two counties submit nearly identical bids, this is strong evidence that their citizens value a college roughly equally and they have similar capacity to support the school. In other cases, quirky historical events conspired to locate a college in one location rather than another. For example, Ezra Cornell and Andrew White would go on to establish Cornell University, New York's land grant college. They knew they wanted to establish the college in one of their home towns. Ezra Cornell was from Ithaca, while Andrew White was from Syracuse. However, Cornell had been robbed of his wages as a young man in Syracuse and he refused to put locate the college there. Consequently, Cornell University is located in Ithaca.[3]

Restricting the sample to only the high quality college site selection experiments excludes a large number of potential experiments. Often, losing finalists were not strong contenders to receive the college. For example, while several counties submitted bids to receive Michigan State University, the legislature had always intended to locate the college close to Lansing, both to be near the state capital and as a compromise among the numerous competing

[3]I have been unable to find any evidence that Syracuse either had higher crime rates than Ithaca or that it tended to have citizens of a lower moral characters; as far as I can tell, Syracuse and Ithaca were very similar along observable dimensions before the establishment of Cornell University. Syracuse would, of course, get its own university several years later.

interests. In this case, the losing bidders do not appear very similar to East Lansing and do not form valid counterfactuals. In the Historical Appendix, I describe each college site selection experiment in detail, including the rationale for classifying the experiment as high versus low quality.[4] I include results using both the high and low quality experiments as a robustness check in Table 3. Obviously, the decision of whether or not a particular experiment is high or low quality is somewhat subjective. Appendix B.4 discusses the general types of experiments that tend to occur in many of the high quality cases; I show that the results are not sensitive to looking at a particular kind of experiment.

The study most similar in spirit to the process of identifying high quality losing finalists used here is Greenstone et al. (2010), who identify the winning and losing counties in contests to receive new large manufacturing plants. The authors argue that, because manufacturing firms considered a very large number of potential sites, by the time they whittled their list down to a few finalists, these finalist locations are likely to be very similar. In the college site selection experiments, in contrast, it need not be the case that a very large number of initial sites were considered before a list of vetted finalists was compiled. This is one reason why it is important to restrict attention to the high quality experiments: it ensures that the winning and losing counties are similar along both observable and unobservable dimensions. I next discuss the other data sources used, before describing the winning and losing counterfactual counties in more detail.[5]

2.2 Patent and County Data

Patent data for the years 1836-2010 come from four sources. The Annual Reports of the Commissioner of Patents covers a plurality of the years, from 1870 to 1942. The annual reports provide information on every patent granted by the United State Patent and Trade-

[4]Both high and low quality experiments are described in the Historical Appendix. It is my hope that, although the low quality experiments may not be useful for identifying causal effects, compiling a broad collection of college site selection experiments may be of independent historical interest.

[5]Patrick (2016) raises several challenges to the identification strategy employed by Greenstone et al. (2010). I believe my study avoids these critiques and is in fact better suited for the use of the losing finalists strategy. 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 winning and losing finalists are similar in terms of observables; Figures 1 and 2 show that the winning and losing finalists evolved similarly as well.

mark Office (USPTO) in a given year. For each granted patent, the report includes the inventor's first and last name, town and state of residence, and the invention name, patent number, and issue date. The annual reports were digitized by Google, and while the transfer to digital format is overall very good, it is not perfect. Hence, not all patents listed in the physical copies of the annual reports make it into my dataset. However (Sarada et al., 2017) show that there are no systematic errors in the digital recreation of this data; any missing patents appear to be randomly distributed.

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.[6] The years 1942 to 1975 come from the HistPat dataset compiled by Petralia et al. (2016a).[7] This dataset, constructed from digitized Google Patents documents, contains the same information as the Annual Reports.

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, taken a commonly used abbreviation from the authors, I refer to as the InvPat dataset. Unlike other commonly used modern datasets such as the NBER Patent Dataset (Hall et al., 2001), which was built with an eye towards linking patents to the names of firms and so focuses on inventions that include assignees, the InvPat data focuses on ensuring the quality of the names of inventors for all patents from 1975-2010.

Together, these datasets provide information on the location and inventor names of every patent granted to a U.S. inventor from 1836 to 2010. The fact that different years use different patent datasets does not pose a problem for this analysis, as every regression specification below includes year effects, which control for any change in the propensity to patent driven by the source of patent data. The results would be biased if, for instance, one patent dataset systematically recorded more patents from college counties than control counties. While Andrews (2017) documents that the HistPat data undersamples inventors in rural areas relative to the Annual Reports data, the fact that the control counties appear very similar to control counties along observable dimensions, including population and urbanization,

[6]See J. Miller (2016a) and J. Miller (2016b) for more information on how this dataset is compiled.
[7]See Petralia et al. (2016b) for details on the construction of this dataset.

minimizes the risks that this is a problem. 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. While efforts have been made to rebuild a record of early patenting from other sources (these are enticingly known as the "X-patent" datasets; see Andrews (2017) for more information), it is unknown how complete these data are or whether they represent a random sample of all pre-1836 patents.

For each year, I sum up all patents granted in each town. Using U.S. decennial censuses, I associate each town and state with its county.[8] Summing the number of patents for each town within a county produces the county-level patent counts. In Appendix B.5, I show results using other sources of patent data and alternative techniques to match town names to counties; in all cases the results are qualitatively similar to the baseline results presented below.

County-level data comes from the National Historic Geographic Information System (NHGIS), which provides decennial census data aggregated at the county level. 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 and 1880 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 Hornbeck (2010) and Perlman (2015). Data on residential segregation at the county level is from Logan & Parman (2017).

[8] In several 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 below use the mean patent count, but results are nearly identical when using the upper and lower bounds.

2.3 The Winning and Losing Counties

In total, I examine histories of 219 colleges from all 50 U.S. states. Of these, in 136 cases I am able to identify losing finalist locations. I consider 73 of these to be high quality experiments. These high quality experiments form the baseline sample used in the analysis below. The data are summarized in Table 1, which lists every college experiment in the 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.[9] A majority of the college experiments involve land grant colleges. Three experiments involve technical colleges, three involve normal schools, two involve HBCUs, and three involve military academies. Six public colleges are classified as "other," while one private college is classified as such. There are on average slightly less than two control counties for each treatment county.

Table 2 compares the treatment and control counties and shows that the control counties are a better match for the treatment counties than are the non-experimental counties. Columns 1 and 2 display the mean and standard deviations of the treatment and control counties, respectively, in the last U.S. census year before the college was established.[10] Column 3 subtracts the mean of the control counties from the mean of the treatment counties, and shows the standard errors of the difference. For a battery of patenting, demographic, and economic variables, the means of the treatment and control counties are statistically indistinguishable and remarkably similar in magnitude.

[9]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. Note that all results are insensitive to reclassifying colleges.

[10]I use census years because most of the demographic and economic variables are collected with the decennial census. So if, for example, a college was established in 1874, the results in Table 2 reflect the state of counties in 1869-1870, when the census was collected.

Column 4 shows the mean and standard deviation for the non-experimental counties, which are the counties in each state that are not classified as either treatment or control counties. Column 5 shows the differences in means and corresponding standard error between the treatment and non-experimental counties. The treatment and non-experimental counties also tend to be similar along most dimensions, making Moretti's (2004) claim that colleges were located "close to random" understandable. But the treatment counties do have a statistically larger population, are more urbanized, have a larger share of interstate migrants, more manufacturing workers, greater manufacturing product and wages, and higher farm wages. While not statistically different from zero, the non-experimental counties do have a much larger number of patents but a smaller $log(Num.Patents + 1)$ than the treatment county. This reflects the fact that the non-experimental counties also contain a number of outlier counties, namely large cities that acted as early innovation hubs and especially skew results regarding patenting; getting rid of these outliers is a major benefit of the losing finalists methodology. Appendix B.1 provides additional balance checks and placebo tests to verify that the losing finalists are valid counterfactual counties.

In Appendix B.2, I provide further evidence that restricting attention to very similar control counties is important for the results. For several college site selection experiments, it is possible to create ordinal rankings of the finalists, either by using the valuations of each finalists' bid or by the number of votes each finalist receives. I show that including only the losing finalists that submit the second-highest bids or receive the second-most votes look the most similar to the winning counties along a number of dimensions. Furthermore, in regressions that include lower ranked finalists, estimated effects of establishing a new college are larger, suggesting that including lower quality losing finalists inflates the observed effect of establishing a new college.

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

$$\begin{aligned}\log(NumPat_{ijt}+1) =& \delta_0 + \delta_1 College_{ij} * PostCollege_{jt} + \delta_2 PostCollege_{jt} \\ &+ \alpha_i + \gamma_t + \epsilon_{ijt}, \end{aligned} \quad (1)$$

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 if year t is after college j has been founded, α_i is a county fixed effect, γ_t is a year effect, and ϵ_{ict} is a county-college-year varying error term.[11] In this context, the treatment is receiving the new college. With only a single experiment, the term $PostCollege_{jt}$ would be redundant because the post-treatment 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-treatment 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 a state after establishing a college within its borders. In the graphs that follow, for each new college, I normalize the year in which the college is founded to year 0. I use logged patents as the independent variable to limit the impact of outlier counties that have a large number of patents and to ease in the interpretation of the coefficient. Because many counties have zero patents in a given year, I add one to the number of patents before taking the log.

Equation (1) provides an easy-to-interpret estimate of the mean difference in logged patents per year in treatment counties relative to control counties following the establishment of a new college. I also estimate several variations of Equation (1) to probe the robustness of the baseline results. When an alternative estimating equation is used, I describe it in the text below. In all cases, standard errors are clustered at the county level unless otherwise noted.

[11] It may also be desirable to include an experiment-specific fixed effect, λ_j. In practice, however, there are very few counties that appear in multiple experiments. The collinearity between the λ_j and α_i terms are thus very strong, so the λ_js are omitted from most regressions that follow.

3 The Effect of Establishing a New College on Patenting

In this section, I present the results of the college site selection experiments. I first estimate Equation (1) as well as a number of alternative specifications. I next show that the effect of colleges increases over time, but is not driven solely by recent trends. Finally, I present evidence for geographic spillovers between college counties and their neighboring counties. In Section 4, I probe these results even further and bring in additional data to understand the mechanisms by which colleges affect patenting.

3.1 Baseline Results

Figure 1 plots the raw patent data for treatment, control, and "non-experimental" counties separately. The non-experimental counties are counties in the same state as a new college but which are neither the college county nor one of the nearly-chosen finalists used as a control. The year in which a new college is established is normalized to be year 0 for all experiments. Figure 2 presents the same data after smoothing the data and controlling for time effects.[12] Both figures contain a balanced set of college experiments in the sense that counties are included only if they have at least 20 years of pre-treatment and 80 years of post-treatment data available. Graphs using all college experiments are nearly identical.

Three results are immediately clear. First, new colleges do not appear to be randomly located; there is a large difference between the treatment and control 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 treatment and control counties patented similarly in pre-treatment years, sug-

[12] Figure 2 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 (see Fan & Gijbels (1996)). 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.

gesting that the experimental design is valid. Third, after the establishment of a new college, the treatment and control counties diverged, with treatment counties patenting more. The difference between the treatment and control counties is especially pronounced after several decades; there is no evidence from these figures that patenting in treatment and control counties converge as more time passes since the establishment of a college.

Figures 1 and 2 also present some more intriguing patterns. First, while treatment and control counties are both increasing their patenting at approximately the same rate in the twenty or more years before the establishment of a new college, after the college is established the increase in patenting in the control counties levels off. The cause for this is unclear, but there are at least two possibilities. First, this can be taken as evidence that site selection committees were indeed targeting rapidly growing counties. A college allowed winning counties to continue their growth, while the losing finalists reverted to a mean patenting rate. Second, the pattern could be due to inventive individuals from previously fast growing control counties moving to the winning treatment counties, and hence that the establishment of a new college served primarily to reallocate inventors rather than produce new ones. Such a result is less plausible since there is no noticeable increase in the growth rate of patenting in the treatment counties. In Section 5.2 I present additional suggestive evidence that new colleges increased patenting rather than simply moving it around. A second intriguing pattern is the decline in patenting that occurs 30-60 years after the establishment of a new college. At present I am aware of no likely candidate explanations for such a pattern.

Table 3 formalizes the intuition in Figures 1 and 2. The columns show different regression specifications. For all tables in the paper, each estimated coefficient is presented in two ways: first by a percentage change in patenting (/100), and second by the change in the number of patents.[13] Standard errors for each are corrected using the delta method. To get the change in number of additional patents when the dependent variable is $\log(NumPatents_{ijt}+1)$, the percentage increase is multiplied by the average number of patents per county in the U.S. in the year 1870, the first census year after the passage of the 1862 Land Grant Colleges Act and

[13] More precisely, because the variables of interest are indicators that are either equal to zero or one, the estimated coefficient must be adjusted to give the percentage increase in patenting using the equation $\%ChangeinPatents = e^{\hat{\delta_1}-\frac{1}{2}Var(\hat{\delta})}-1$, where δ_1 is the coefficient of interest from Equation (1); see Halvorsen & Palmquist (1980) and Kennedy (1981). This adjusted coefficient is presented in the table.

during a wave of new college openings.[14] When the dependent variable is $NumPatents_{ijt}$, the change in the number of patents can be observed directly; the percentage change is simply calculated as $\frac{\Delta NumPatents}{NumPatents_{1870}}$, again using 1870 as a baseline year. For all columns, standard errors are clustered at the county level.[15] The coefficient of interest is displayed in Row 1 and shows the percentage and number of estimated additional patents generated in the treatment county relative to the control county resulting from the establishment of the new college.

Column 1 shows the results of estimating Equation (1). Treatment counties have about 33% more patents per year than control counties, or roughly 1.2 additional patents every year in 1870. This result is significant at the 1% level.

Column 2 shows results using an alternative calculation of logged patents as proposed by Blundell et al. (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 16% more patents per year, or about an extra half patent per year.

Column 3 uses the absolute number of patents as the dependent variable. 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 Columns 3 and 4. I find that establishing a new college leads to 2.25 additional patents per year in 1870, a 42% increase over an 1870 baseline. This is larger than the baseline estimate, although it is only statistically significant at the 10% level. Column 4 uses the fact that the number of patents takes on integer values and presents

[14] Using a later baseline year, when average patenting in the U.S. was higher, leads to a larger estimated change in the number of patents. Likewise, earlier baseline years results in smaller estimated changes in the number of patents. In all cases, the story is qualitatively the same. I choose 1870 because it represents the state of patenting in America just before the beginning of the golden age of U.S. higher education (Goldin & Katz, 2008).

[15] I also cluster at the state, experiment, and county×experiment levels. I additionally cluster at multiple levels as proposed in A. C. Cameron et al. (2011) using the estimator described in Correia (2016): 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 as is done in the tables produces the largest standard errors, but the standard errors are virtually identical at every level and none of the inferences change. Additionally, 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 more discussion on the best level at which to cluster standard errors, see C. A. Cameron & Miller (2015).

estimates of a negative binomial regression.[16] In this specification, establishing a new college leads to a 53% increase in patenting, larger than the other baseline estimates.

In Column 5, I include all of the patent data from the U.S. for each year. This includes the "low-quality" experiments which are not included in the baseline results as well as every county that is not included in an experiment; these other counties are included as "non-experimental" counties. Instead of estimating Equation (1), I now estimate a triple-difference equation of the form

$$\begin{aligned}\log(NumPat_{ijt}+1) = &\beta_0 + \beta_1 College_{ij} * HighQuality_{ij} * PostCollege_{jt} \\ &+ \beta_2 PostCollege_{jt} + \beta_3 HighQuality_{ij} * PostCollege_{jt} \\ &+ \beta_4 College_{ij} * PostCollege_{jt} \\ &+ \gamma_t + \alpha_i + \epsilon_{ict}. \end{aligned} \tag{2}$$

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 list of high quality experiment j counties used for the baseline results, and zero otherwise.

In this new equation, the coefficient of the triple-interaction term β_1 measures how much larger the difference-in-differences estimator between high quality treatment and control counties is compared to the difference-in-differences estimator between all other treatment and all other counties. This coefficient is negative, although not statistically significant, indicating that if anything, there is negative selection into the college experiments; that is, the difference between the treatment and control counties is smaller for high quality experiments than for counties not included in the baseline results. This result shows why restricting atten-

[16] I use negative binomial regression in this setting because the variance in patenting is much larger than the mean. The standard approach to fixed effects negative binomial regression is to estimate using conditional maximum likelihood as proposed by J. Hausman et al. (1984). In Stata, this is done with the xtnbreg command. However several authors (Allison & Waterman (2002), Greene (2005), Guimaraes (2008)) note that this procedure does not actually control for cross-sectional time-invariant effects; more precisely, the overdispersion term is demeaned, but the full conditional mean is not. To correct for this, I estimate a negative binomial model using Stata's nbreg command and include a dummy variable for each cross-sectional observation. Allison & Waterman (2002) show in simulation studies that the incidental parameters problem does not appear to be a problem in negative binomial regressions. Estimates using other count data models, including Poisson and zero-inflated Poisson give qualitatively similar results. The listed R^2 is the McFadden pseudo-R^2.

tion to high quality experiments is important: otherwise, the effect of colleges on patenting would be overstated.β_2 has the same interpretation as before and simply measures the increase in patenting after establishment of a new college. β_3 estimates the change in patenting in high quality treatment and control counties after the establishment of a college relative to low quality treatment and control counties. Finally, β_4 estimates the increase in patenting in *all* treatment counties relative to *all* control 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 β_4 is positive and significant, so the qualitative conclusions of the baseline specification in Column 1 would still be true even if the low quality experiments were included.

The increase in patenting in high quality treatment counties over high quality control counties after establishment of a new college (that is, the same quantity as estimated by δ_1 in Equation (1)) is given by $\beta_1 + \beta_4$.[17] Combining these coefficients reveals that high quality treatment counties increased patenting by 37% (standard error= .082), which is significant at the 1% level. This estimate is very similar to the baseline coefficients presented in Column 1.

Figure 3 estimates Equation (1) with a separate interaction term for each experiment, orders the coefficients from lowest to highest, and then plots the coefficients. Formally, I estimate

$$\begin{aligned} \log(NumPat_{ijt} + 1) = &\delta_0 + \sum_{j \in J} [\delta_{1j} College_{ij} * PostCollege_{jt} + \delta_{2j} PostCollege_{jt}] \\ &+ \alpha_i + \gamma_t + \epsilon_{ijt}, \end{aligned} \tag{3}$$

where J is the number of college site selection experiments. In about 60% of the experiments, the estimated coefficient is positive, and in more than 70% of these, the coefficient

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

$$\begin{aligned} (y_{Coll.,HighQual.,Post} - y_{Coll.,HighQual.,Pre}) - (y_{Cont.,HighQual.,Post} - y_{Cont.,HighQual.,Pre}) &= [\beta_0 + \beta_1 + \beta_2 + \beta_3 + \beta_4] - [\beta_0] - [\beta_0 + \beta_2 + \beta_4] + [\beta_0] \\ &= \beta_1 + \beta_4. \end{aligned}$$

is statistically different from zero at the 5% level. Even when the coefficient is negative, it tends to be close to zero in magnitude, and in 40% of these cases the estimated coefficient is not statistically different from zero. In a majority of the college site selection experiments, the estimated coefficients are in line with the coefficients estimated in Table 3.

Appendix B presents numerous other robustness checks, including re-estimating the baseline results with different patent datasets, using town- instead of county-level data, and examining the results separately by different types of college site selection experiments. Beyond these, I verify that the baseline results are insensitive to using alternative dates for the college experiments (for instance, using the date on which classes began rather than the date when the college site was selected), using various assumptions about the patent grant delay (the length of time between when a patent is filed and granted), using different regression specifications (for instance, looking at the extensive margin of patenting or using various other count data models), and using numerous subsets of the college experiments. All these results are available upon request.

3.2 Dynamics

As Figures 1 and 2 make clear, the effect of a college changed over time. In Figure 4 I interact the effect of a college by ranges of years. More precisely, I estimate

$$\begin{aligned}\log(NumPat_{ijt}+1) =& \delta_0 + \sum_{\tau \in T} [\delta_{1\tau} College_{ij} * TimeBin_{j\tau} + \delta_{2\tau} TimeBin_{j\tau}] \\ &+ \alpha_i + \gamma_t + \epsilon_{ijt},\end{aligned}$$

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. $Time_{j\tau}$ is an indicator variable that is equal to one if $t = \tau$ and 0 otherwise. Results are nearly identical using different groupings of years and beginning or ending year cutoffs. Each plotted coefficient represents the percentage increase in patenting in the treatment counties relative to the control counties in the respective bin of years since the college establishment.

Confirming the intuition shown in the raw data, there is no significant difference between the treatment and control 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 establishment of a new college, the coefficients begin increasing dramatically.

One concern, highlighted by the fact that the treatment counties tend to out-patent the control counties the most 70-80 years after the establishment of a new college, is that the baseline results are largely driven by changes in the patent law that differentially encourage patenting in college counties relative to the controls. Indeed, just such a change occurred in 1980 with the passage of the Bayh-Dole Act. Prior to the passage of the Bayh-Dole Act, any patent obtained while the inventors were funded by a federal grant or research contract had to be assigned to the U.S. federal government. The Act instead gave ownership of these patent rights to the inventors or their institutions, including universities. As Sampat (2006) and N. Hausman (2013) show, this led to a dramatic increase in patents assigned to universities. The results after dropping all patent data from 1980 or later are actually larger than the baseline results presented in Table 3, suggesting that if anything the difference between treatment and control counties was even larger in the years before Bayh-Dole.

Other authors have suggested different years in which the relationship between colleges and invention change. Bound & Turner (2002) and Goldin & Katz (1999) point to the end of World War II as a time when the demand for technical skills to be taught in colleges increased, in part driven by the 1944 G.I. Bill. Furman & MacGarvie (2007) claim that the onset of World War I necessitated new chemical and biological innovations for the military, leading to the creation of biochemistry and similar departments in many universities, and that a second wave of biochemical departments opened in universities in order to support the pharmaceutical industry beginning in 1927. I verify that discarding all data after 1944, 1927, and 1917 also does not materially alter the baseline findings.[18] These results are available upon request.

[18]When all post-1917 data is discarded, the sample is much smaller and therefore, not surprisingly, the result is no longer statistically significant. But the magnitudes of the estimated coefficients are in line with the baseline estimates.

4 How Do Colleges Affect Patenting?

Having determined that establishing a new college leads to a large and persistent increase in local patenting, I now seek to understand the channels by which colleges cause this increase. In Section 4.1, I calculate the share of patents attributable to college alumni and thus are plausibly due to college-generated human capital. Section 4.2 documents patenting that results from faculty research. Section 4.3 explores the role of knowledge spillovers to individuals living near colleges but who are not directly affected by them. Each of these channels appears to account for a limited share of the observed increase in patenting near colleges. In Section 4.4 I present further indirect evidence that the role of human capital, faculty research, and knowledge spillovers are modest. In Section 4.5 I show that, in contrast, migration can explain most of the observed increase in patenting.

4.1 Human Capital

The most obvious channel by which colleges can increase patenting is by providing students with an advanced education that enables those students to come up with new inventions. Unfortunately, the direct effect of obtaining a college education on invention is typically difficult to test. One reason for this is that there is no way to link an inventor to a college in the historical data. To overcome this problem, I transcribe college yearbook data for several of the colleges in the sample. The college yearbooks are available from `ancestry.com` and contain full scans of historical college yearbooks, which include full student names, hometowns, and nicknames, all of which can be used to aid in matching 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.

Appendix A.1 discusses the construction of the college yearbook dataset in more detail. Essentially, I match the alumni by name to the patent record to obtain an alumni patenting rate based on the years in which a yearbook is available. I then use the total count of alumni for these years to impute the total number of alumni for the years in which yearbooks are unavailable. I then obtain the number of patents expected to be obtained by alumni by multiplying the calculated patenting rate by the imputed number of members of each group. Finally, I calculate the share of total patenting in the college county that can be accounted for by alumni.

Results are presented in Table 4. I present results separately for the entire county population, for college undergraduate alumni, and for college graduate alumni. Column 1 lists the total population for each group in the counties for which yearbook data is available. Column 2 lists the share of the total county population belonging to each group. Columns 3, 4, and 5 list the number of patents, patenting rate, and share of patents, respectively, attributable to each group. The patenting rate for each group g is given by $Pat.Rate_g = \frac{Num.Patents_g * 1000}{CountyPop._g}$. The share of patents is simply calculated as

$$SharePatents_g = \sum_{j \in J} \frac{Num.Patents_{gj}}{Num.Patents_j},$$

where J is the set of all college counties for which yearbooks are available. This is calculated and averaged for each year after the establishment of a new college. Column 6 lists the increase in patenting caused by the establishment of the new college, as listed in Equation (1). Because this result is computed only for the counties in which yearbook data are available, the estimated increase is different from the baseline estimate of β_1 presented in Table 3. As there were no alumni or faculty patents prior to the establishment of the new college, these cells are empty except for Column 1 with results for the overall county. Column 7 lists the share of the "extra patents" caused by establishing the new college that are attributable to each group. This share is calculated by

$$ShareExtraPatents_g = \sum_{j \in J} \frac{Num.Patents_{gj}}{(1 - \frac{1}{1+\beta_{1j}})Num.Patents_g},$$

where β_{1j} is the estimated coefficient for the interaction term from Equation (3). Note that because this result is calculated by first calculating the additional patents in each county (given by β_{1j}) rather than simply using the aggregate estimate in Column 6, it is not equal to simply the result in Column 6 divided by the share in Column 5.

This analysis reveals several surprising findings. First and most notably, alumni account for a small share of the patents in college counties: about 7% of overall patents, or 10% of the additional patents in college counties. I am unable to find any patents belonging to graduate student alumni. This is surprising because graduate students are likely to be the most technically skilled alumni in their respective fields. There are relatively few graduate students in the yearbooks, however, and graduate students may be more mobile and thus more likely than undergraduate alumni to be living in a county different from the college. Undergraduates patent at a lower rate (0.361 patents per 10,000 alumni) than does the overall population of the college county (0.643 patents per 10,000 capita). The calculated patenting rate, while suggestive, is particularly sensitive to assumptions made about alumni migration after graduation and imputations of college attendance for the years without yearbooks. As such, they should be taken with caution.

These results likely overstate the increase in total patenting caused by patenting by alumni and faculty. As documented in more detail in Appendix A.1, the techniques used to interpolate and extrapolate the number of alumni and faculty for years in which the yearbooks are missing likely produces an upper bound on the number of these individuals. Additionally, I assume that any patent matching the name of an alumnus or faculty member is attributable to the member of those groups. To the extent that alumni or faculty have common names, this also overstates the number of patents belonging to these groups. Finally, these results are not causal: if particularly intelligent, created, or driven individuals are both more likely to attend college and more likely to invent independently of education, then the 10% of additional patents accounted for by alumni overstates the role of college-produced human capital in explaining local invention. In contrast, Bianchi & 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 an important topic for future work.

It is important to emphasize that these results only document the *local* patenting by college alumni and faculty. Thus, while I find that college alumni and faculty account for a small share of the patents in their college's county, this result is silent on the overall patenting behavior of these groups. It is possible that college educated individuals patent actively, but if this is true, these patents are granted after alumni move away from the counties of their alma maters. Indeed, Bound et al. (2004) and Sumell et al. (2008) show with contemporary data that alumni frequently leave the counties where they attended college, but the rate varies greatly with local conditions. Tracking the patenting behavior of college alumni as they move geographically is beyond the scope of the present paper but is also an interesting topic for future work.

Who are these alumni inventors? In addition to names, the college yearbooks typically provide the major of each college senior. I group each college major from the yearbooks into the 15 fields used by the American Community Survey.[19] In Panel A of Figure 5, I plot the number of patents coming from students with each major. The highest number of alumni patents by major field come from engineers. Many of the colleges for which I have yearbook data have especially large engineering programs, however. In Panel B, I plot the patenting rates by major field, $\frac{Num.Patents_m}{Num.KnownSeniors_m}$ for each major field m. Engineers are now only the fourth most frequent inventors. Science-related majors, which include a wide range of majors from architecture and nursing to math teachers and pharmacy, are the most frequent inventors, with about 3.5 patents per 10,000 science-related majors.[20] Surprisingly, business majors patent more frequently than engineers, and nearly as frequently as physical science majors. Even majors that have little contact with technology, namely liberal arts and history, patent at more than one third the rate of engineers. Clearly, the curriculum one is exposed to in college is not the only determinant of later-life invention.

[19]See `www.census.gov/prod/2012pubs/acs-18.pdf`. These major fields are: Education; Science- and engineering-related fields; Psychology; Literature and languages; Visual and performing arts; Communications; Multidisciplinary studies; Social sciences; Liberal arts and history; Biological, agricultural, and environmental sciences; Business; Computers, mathematics, and statistics; Physical and related sciences; Engineering. Some of these fields were virtually nonexistent during the period for which I have yearbook data. In particular, I find no patents from individuals with multidisciplinary studies majors, so I do not report them in the following figures.

[20]The full list of science related majors according to the ACS is: Architecture; Computer programming; Data processing; Computer teacher education; Mathematics teacher education; Science teacher education; Engineering technologies; Engineering and industrial management; Electrical engineering technology; Industrial production technologies; Mechanical engineering related technologies; Miscellaneous engineering technologies; Applied biotechnology; Nuclear and industrial radiology technologies; General medical and health services; Communication disorders sciences and services; Medical assisting services; Medical technologies technicians; Health and medical preparatory programs; Nursing; Pharmacy pharmaceutical sciences and administration; Treatment therapy professions; Community and public health; Energy and biologically based therapies; Miscellaneous health medical professions.

4.2 Faculty Research

If faculty research has direct commercial applications, then college faculty may likewise account for a non-trivial share of the patents created in college counties. Mowery & Sampat (2001) show that university patenting was very rare until the early 1970s, on the order of roughly 20 university patents per year for the entire U.S. from 1920 to 1940, and fewer than 125 per year until 1970. After that, patenting by universities took off, particularly following the passage of the Bayh-Dole Act in 1980. A sizable literature examines the effects of the Bayh-Dole Act and concludes that it caused universities to expand their patenting activity (Mowery et al. (2001), Mowery & Ziedonis (2002), Sampat (2006), Berman (2008), Lee (2013)).

Studies of university patenting attribute a patent to a university if the patent is legally assigned to the university. These statistics may be misleading if university faculty patent but do not assign their inventions to their university employers. To understand the role of faculty, I instead match by the names of faculty to the patent record. Faculty names are collected from the college yearbooks in the same way as the alumni; see Appendix A.1.2 for more details.

The second to last row of Table 4 presents results for patenting by faculty. The final row presents "residual" patenting that cannot be accounted for by either alumni nor faculty. The last row is computed using the results for the entire county in Row 1 and subtracting out results for alumni and faculty.

Broadly speaking, my results confirm the previous studies cited above: patenting by college faculty was rare through the first three quarters of the 20th century. Faculty account for about 0.5% of the total patents or roughly 1% of the additional patents in the college counties. While faculty patent at a slightly higher rate (0.677 patents per 10,000 faculty) than the overall county (0.643 patents per 10,000 capita), there are so few faculty members that this higher patenting rate does not contribute much to overall county patenting.

4.3 Knowledge Spillovers

The previous two sections show that alumni and faculty of a particular college do not account for a large share of that county's patents. Colleges may also increase patenting in less direct ways, by producing sizable knowledge spillovers. Many authors cite the fact that commercial patenting increases close to colleges as evidence of spillovers (for example, Jaffe (1989), Jaffe et al. (1993), N. Hausman (2013)). But these may not be true spillovers if commercial patents are produced by firms that employ college alumni or consult with faculty, as recognized in, for instance, Zucker, Darby, & Armstrong (1998) and Leten et al. (2014). Identifying knowledge spillovers entails the opposite challenge from that involved in identifying patents by alumni and faculty: instead of identifying individuals who attended or were employed by a particular college, one must identify individuals who were not directly affiliated with the college.

To isolate the effects of spillovers, I restrict attention to three groups for which it can be determined that they are not graduates or faculty of that county's college: people who lived in college counties prior to the opening of the college and were too old to subsequently attend; blacks living near segregated colleges; and women living near non-coeducational colleges. These results can be interpreted as within-county knowledge spillovers as they represent an increase in invention caused by a new college but not completed by individuals who have a direct affiliation with the college.

I first use the 100% decennial census data to identify a list of individuals who lived in the college or control counties and were more than 30 years old at the time the new college was established. These individuals were almost certainly too old to have attended the new college as undergraduates. Therefore, any increase in patenting for this group in the college county relative to the control is likely due to interactions with college-affiliated individuals rather than direct college knowledge. I refer to this group of individuals as the "townie" sample.

I calculate the results for townie patents only for those states for which college yearbook data is available as described in Section 4.1.[21] Results are presented in Column 1 of Table 5.

[21] I do this to speed up the computation time required to match patentees to the census and match individuals across censuses. Future versions of this paper will have townie results for all states. Since these results are calculated using a relatively small subset of the sample counties, they should be interpreted with caution.

While the coefficient is slightly positive, it is very close to zero in magnitude and not statistically significant, suggesting that the townies that were successfully matched to the census do not increase their patenting after establishing a new college. Even taking the mean value of the estimated coefficient at face value, knowledge spillovers to the townies can account for less than 1% of the additional patents in college counties.

As an additional test of within-county knowledge spillovers, I next exploit the fact that large classes of people were historically systematically prevented from attending public universities for many years. In particular, many colleges in my dataset were segregated when they first opened. This means that any increase in black patenting in these treatment counties relative to control counties must be due to knowledge spillovers; it cannot come from blacks attending college. Likewise, many colleges were initially open only to men, and so increases in female patenting are likely due to spillovers. When estimating changes in black patenting, I keep an experiment in my dataset from the year the college opens until the college desegregates. If a school is desegregated in its first year, it is excluded from the sample. Additionally, HBCUs are excluded from this sample, since blacks always attended these schools. For the results on female patenting, I keep an experiment in my dataset from the year the college was founded until the year it becomes coeducational. Goldin & Katz (2011) provide much more detail on the history of coeducation in American colleges.

To determine whether or not a patentee is white or black, I utilize first names from the 100% census data. For each state and each census, I calculate the probability of being black for each first name. Briefly, I impute the probability that each name belongs to a black or male individual, and then calculate the inferred race or gender of the patentee. This procedure is essentially a split-sample IV procedure (Angrist & Krueger, 1992). I describe this procedure in more detail in Appendix A.2. Because first name data for all individuals in the U.S. census is only available through 1940, I drop all more recent patent data from these regressions.

Column 2 of Table 5 shows the results for blacks. Opening a new segregated college led to about 7.5% more patents by blacks per year in treatment counties relative to control counties and is statistically significant at the 5% level. Thus, the increase in patents for blacks is roughly 20% as large as the increase in overall patenting estimated in Table 3. Column 3 shows results for women. The estimate for women is smaller and imprecisely estimated, but the magnitude suggests that opening a new male-only college led to about 3% more patents by females per year. In spite of these relatively large estimated percentage increases, blacks and women both patented at very low rates for most of the sample period.[22] Using the above method to count patents by blacks and females, in 1870 blacks were granted only 0.16 patents per year in the average county, while women were granted just over a third of a patent per year.[23] Thus, the establishment of a new college adds only about 0.01 additional patents by black inventors per year and 0.02 patents from female inventors per year. Recall from Table 3 that the establishment of a new college led to about 1.2 additional patents per year in treatment counties relative to control counties. These results suggest that blacks and women who were excluded from attending college together account for about 3% of the overall number of additional patents. Due to the difficulty of predicting race and gender by first name and the imprecision of the estimates, this result should be treated with caution, however.

4.4 College Types and Patent Classes

In this section, I further probe the magnitude and direction of human capital and local knowledge spillovers. Since different types of colleges had very different curricula, if human capital or knowledge spillovers are sizable, then the direction of patenting in college counties should change in different ways depending on the type of college. I classify colleges by type as described in Section 2.3. 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. If the technical skills taught at college spill over and affect local invention,

[22] See Sarada et al. (2017) for more detail on patenting by marginalized groups in the historical United States.

[23] Identifying the overall number of black patents by first name can be especially fraught with error, as African American inventors might attempt to "pass for white" to avoid discrimination against their inventions. See Cook (2014), Hobbs (2014), and Jaspin (2007).

then normal schools and other public and private colleges should produce less patenting than land grant and technical colleges.

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. For some types of colleges, there is much more ambiguity regarding whether or not the college is practical or classical. In Appendix B.7 I use alternative classifications of practical and classical colleges and show that the results are similar. The results are presented in Column 1 of Table 6. It does appear 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 had 28% more patents per year (significant at the 5% level), while the classical colleges had 24% more patents per year (statistically insignificant).

In Column 2, I exclude all years after 1940 from the data, because following the explosion in the demand for higher education after World War II, the curricula across colleges largely began to converge. In the pre-1940 years, the differences between practical and classical colleges is almost nonexistent: practical colleges saw about 23% more patents per year relative to their control counties, while classical colleges saw 22% more; neither coefficient is individually statistically significant. The fact that the standard errors are typically larger for the classical colleges could simply be indicative of the larger variance in terms of curricula that goes on across these types of colleges. In Appendix B.7, I present results for each college type separately. Because there are very few college experiments for HBCUs, military academies, normal schools, and other private colleges, however, there is insufficient power to draw conclusions from these effects. Even with a finer classification of colleges, the results do not conform to the simple intuition that establishing a more "technical" colleges, however measured, produces more patents. An additional concern is that certain types of colleges appear to have a larger effect on patenting simply because these colleges are typically larger, either in the sense of graduating more students or by attracting a larger number of migrants to the county; I discuss some of these effects in Section 4.5.

Even if the type of college does not have a large effect on the number of patents, it might be expected to alter the composition of patented technologies. To get a sense of patent technology type, I use the patent classes assigned to historical patents by Marco et al. (2015). I do not observe any significant differences in the fraction of patents assigned to NBER one-digit classes between treatment and control counties nor across different types of colleges.[24] Even when looking at more specific patent classes that one might expect to be particularly related to a given college type, I do not observe measurable effects in the predicted direction. Column 4 of Table 6 shows that the fraction of agricultural patents does not increase in land grant treatment counties relative to non-land grant treatment counties after establishing the college; in fact, it appears as if the fraction of agricultural patents plummets in counties that get land grant colleges relative to their controls, while there is almost no change in the fraction of agricultural patents in non-land grant treatment counties. 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.[25] In Appendix B.7, I use an alternative definition of agricultural patents and get nearly identical results. While a naive comparison of land grant and non-land grant colleges reveal that land grant college counties have more agricultural patents, these results show that this is not causal and that establishing a land grant college may actually cause a county to shift away from agricultural patenting. In column 6, I repeat this exercise but using mining patents and comparing technical schools to non-technical schools.[26] The establishment of a technical school does not lead to a higher fraction of mining patents than does the establishment of other types of colleges; in fact, the the estimated magnitude is negative and very close to zero for technical schools, while it is positive for non-technical schools.

[24]The NBER one-digit patent classes are: chemical, communications, medical, electric, mechanical, other, no class, and missing class.

[25]The one-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 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".

[26]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".

In Table 7, I examine how patenting across all classes changes. In short, there is no evidence that college counties become increasingly specialized following the establishment of colleges. In fact, instead of becoming more concentrated in particular fields, patenting becomes more diverse in the college counties after the establishment of a new college. 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:

$$Pat.Concent_{it} = \sum_{c \in C_{it}} \Big(\frac{Num.Pat_c}{\sum_{k \in C_{it}} Num.Pat_k} \Big)^2 \tag{4}$$

where C_{it} is the set of all patent classes in county i at time t. I construct this index using three different patent classifications: the one-digit NBER patent classes, the two-digit NBER patent classes, and the three-digit USPTO patent classes. Results are presented in Columns 1, 3, and 5 of Table 7. The estimated coefficient has the same sign and similar magnitudes across all three specifications: a new college causes concentration to fall by 0.10, which is a 40-80% decline in the 1870 baseline concentration measure depending on the patent classifications used. The estimated coefficients are all statistically significant at the 5 or 10% levels. In Columns 2, 4, and 6, 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 increased after the creation of a new college; the extra patents produced are not just in the same fields as previous patents in the treatment 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.

4.5 Migration

In this section, I document how much of the increase in patenting can be explained by individuals migrating to live in the college counties after the establishment of the new colleges. A great deal of research suggests that population size and density is crucial for the generation and dissemination of new ideas.[27]

Just as the 100% decennial census data can be used to determine which individuals lived in college and finalist counties prior to the establishment of a new college, they can conversely be used to determine which individuals migrated to the college or finalist counties after the establishment of a new college. Column 1 of Table 3 shows that college counties have about 16% more patents from these post-establishment inter-county migrants than do the losing finalist counties. Because it doesn't make sense to calculate a difference-in-differences estimator (all counties have 0 inter-county migrants in the years before a college is established by construction), I estimate a first difference between college and finalist counties with OLS, controlling for year effects and clustering standard errors by county. While not statistically different from zero at conventional levels, the difference is migrant patenting between the college and finalist counties is economically large and consistent with other possible channels being small in magnitude. The estimated difference amounts to slightly less than one additional patent in college counties coming from migrants each year. In this sample of the data used to compute these results, establishing a new college leads to 2.4 additional patents per year, so migrants account for about 40% of the observed increase in patenting.[28]

An alternative way to estimate the importance of migration to college counties is to simply control for county population in the baseline difference-in-differences estimates. Because population variables are collected from the decennial U.S. population censuses, 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 2 of Table 8, I reproduce the baseline result on patenting using only patenting in census years. The estimated coefficient is similar to the baseline coefficient estimated in Column 1 of Table 3, although slightly smaller, finding that establishing a new college leads to about 23% more patents per year. I then document that establishing new colleges increasing both the population and urbanization in

[27]For instance, see Duranton & Puga (2004) and Glaeser & Gottlieb (2009) for surveys of the theoretical literature on population and innovation.

[28]Recall from Section 4.3 that results using the 100% census data only use data from the states for which yearbook data is available.

college counties in Columns 3 and 4 of Table 8.

Column 3 estimates the effect of a new college on logged county population. If colleges succeeded in driving a large number of people into the treatment county, then any effects on patenting could simply be due to the college county having more people. I find that college counties are 43% larger than the control counties after establishing a college; this effect is statistically significant at the 10% level. In Column 4, I check whether a new college also increases urbanization in treatment counties relative to controls. Following the establishment of a new college, the fraction of a county living in urban areas increases by 8 percentage points in treated counties relative to controls, a large increase that is significant at the 5% level.

Examining the effect of colleges on population in more detail is telling. I estimate the effect of establishing a new college on county population, using the analog to Equation (3). Figure 6 plots the estimated coefficients for $\log(TotalPop.)$ against the estimated coefficients for $\log(Num.Patents + 1)$ for each college site selection experiment. The 45 degree line is plotted as the gray dashed line in the figure. The estimated slope coefficient is a statistically significant 0.3. Counties that grow more are thus more likely to be the same counties that see a larger increase in patenting. Notably, however, this estimated slope is less than one, so there is no evidence that patenting exhibits increasing returns in population.

Building on the results in Figure 6, I account for population's effect on patenting in various ways in Columns 5-7. In Column 5, I estimate the effect of establishing a new college on patenting per capita, which is calculated as the number of patents divided by county population for each county and each census year. Although the estimated coefficient is positive, it is not statistically different from zero (establishing a new college leads to about one additional patent per million people). In Column 6, I re-estimate the baseline regression with logged patents as the dependent variable, but 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 .25% increase in patenting). When including $log(TotalPop)$, the coefficient on the interaction term of interest is only 35% as large as in the baseline estimate, decreasing from 34% in the baseline estimate to 12%, and is no longer statistically significant. Thus changes in county population can plausibly explain roughly two-thirds of the observed increase in patenting in college counties relative to their controls. In Column 7, I include $TotalPop$ and $(TotalPop)^2$ as controls instead of $log(TotalPop)$, but the conclusions are qualitatively similar: controlling for population in this way explains roughly 45% of the observed increase in patenting in college counties. In sum, controlling for population in various ways still leaves between one third and 60% of the increase in patenting to be explained.

As further evidence that increased population and urbanization are major drivers of the observed increase in patenting in college counties, in Appendix B.9 I compare counties that receive new colleges to other counties that receive other state institutions at the same time. Such situations were especially common in western states, when the state institutions were formed after the passage of the 1862 Morrill Land Grant Act. These states would typically decide on the location of their land grant colleges, other public universities, state capitals, hospitals, prisons, or insane asylums at the same time. Towns that were unable to receive their desired institutions were often given another as a consolation prize. The places that receive new colleges are not statistically different from the locations that receive other types of institutions; while the estimated coefficient is imprecisely estimated, it is positive in magnitude but much smaller than the baseline estimate. In addition, population grew almost identically between the college counties and the counties that received other institutions. These results are robust to excluding types of institutions that might be expected to attract individuals with particularly high levels of human capital, such as the state capital or hospitals.

5 Do Colleges Cause a Net Increase in Patenting?

The fact that migration appears to explain such a large share of the additional patents in college counties raises an important question: do colleges simply shift inventive people around geographically, or do they actually lead to a net increase in patenting? As documented in Sections 4.1-4.3, there is evidence that human capital, faculty research, and knowledge spillovers all account for some of the increase in patenting observed in college counties, although the size of these effects are modest. In this section, I test for other evidence of spillovers generated by colleges. Evidence of such spillovers would further confirm that colleges lead to a net increase in invention rather than simply shuffling around the location where invention happens.

5.1 Agglomeration

Section 4.5 shows that controlling for population can explain the majority of the increase in patenting in college counties. But just because population explains most of the increase does not mean that colleges are not generating substantial spillovers: if agglomeration economies are present, then concentrating creative individuals into one place will produce more patents than if the same individuals lived apart.

As outlined in Kline & Moretti (2014a) and Kline & Moretti (2014b), estimating agglomeration economies requires estimating an elasticity of invention with respect to population both in the destination and source counties for each migrant. While such an analysis is possible using the 100% census data and tracking migrants across space and time, it is computationally difficult and therefore beyond the scope of the present paper; I will address this question in future work. While I cannot test whether or not establishing a college leads to an increase in the overall concentration of the population, and therefore an increase in agglomeration, I conduct a simpler test in the same spirit to determine whether the college itself causes larger agglomeration economies. I estimate whether the elasticity of invention with respect to patenting is greater in college counties than in the counterfactual counties. A positive estimated coefficient would indicate that the presence of the college makes it more valuable for people to congregate in one location relative to co-locating in a place without a college.

Results are presented in Table 9. Surprisingly, there is no evidence that agglomeration economies are larger in college counties. I interact $\log(TotalPop.)$ with the $PostCollege$ dummy and $CollegeCounty \times PostCollege$ interaction term. If anything, the elasticity of invention with respect to population appears to be smaller in college counties, as shown in Column 1. In Column 2, I interact $TotalPop.$ and $(TotalPop.)^2$ with the college dummy and interaction term. For readability, I scale the population terms by 10,000. The interaction term with the linear population term is negative, consistent with the elasticity estimates.

The quadratic term is positive but very small in magnitude, suggesting that there may be increasing returns but that they lead to more patents in college counties relative to controls only for the very largest counties.

5.2 Geographic Spillovers

Another potential source of spillovers by which colleges can increase net patenting is by increasing the inventiveness of neighboring geographic areas. In contrast, 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.

In Table 10, I present results by the distance from the treatment county to control counties. Column 1 compares treatment counties to control counties that are "far away" from the treatment county in the sense that they are in the same state but do not share a common border. The treatment counties increase their patenting by 41% relative to these far away controls, larger than the baseline estimate in Table 3. In Column 2, I compare treatment counties to counties that do share a common border, which I refer to as adjacent counties. In this column, the estimated increase in patenting from the establishment of a new college is only less than 40% of the magnitude of the estimated increase in column 1 and is not statistically different from zero. These results show that the treatment counties increase patenting much more than distant controls, 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 treatment counties to *all* non-treatment counties; I no longer restrict attention to the control counties. Column 3 compares the treatment counties to far away counties in the same state and finds that treatment counties increase patenting by 51%. Column 4 compares the treatment counties to all counties that share a border; in this case, the treatment counties only increase patenting by 39%. Thus, even when attention is not restricted to the counterfactual sites, which may not be randomly distributed across a state, it appears that the treatment counties grow somewhat similarly to their neighbors, but increase patenting by much more than far away locations.

6 Conclusion

In this paper, I exploit a natural experiment to identify valid counterfactuals to counties that received a new college in the historical United States. I find that establishing a new college caused roughly 33% more patents per year in college counties relative to control counties. The difference between treatment and control counties tended to get larger over time, and the benefits in treatment counties appear to spill over and affect neighboring counties as well.

I also isolate the channels by which colleges increase patenting. Alumni and faculty of a particular college can account for only about 10% of the increase in patenting in that county. Knowledge spillovers to groups that cannot attend college explain about another 1% of the increase in patenting. Depending on how population is controlled for, recruitment of inventive people to college counties can explain the majority of the increase in patenting, between 50 and 65%.

One important issue is whether establishing a new college leads to a net increase in patenting, or whether all of the increase is simply the result of inventive individuals relocating to be close to colleges. This question is especially important because recruitment appears to explain a large portion of the additional patenting near colleges. The fact that I find measurable knowledge spillovers suggests that there is at least a small positive net impact from colleges. Alumni and faculty also account for a non-negligible share of the observed increase in patenting, and while I do not estimate how many patents these alumni and faculty would have obtained in the absence of their college education, it is likely that colleges played a large role in promoting invention by these groups. Finally, I show that colleges increased the inventiveness of neighboring areas, suggesting that instead of attracting all the nearby inventors, college counties make nearby places more inventive. Future versions of this paper will look more explicitly for evidence of agglomeration economies occurring near colleges.

These findings raise a number of additional questions to be addressed in future work.

While the above analysis suggests that migrants account for the bulk of additional patents in college counties, it is less successful in showing what kinds of individuals these migrants are. The large net increase in population may mask the fact that the additional patents come from a small "knowledge elite" that is attracted by the new college. Mokyr (2002), Mokyr (2005), and Squicciarini & Voigtländer (2015) argue that agglomeration of particularly talented individuals was an important driver of innovation during the Industrial Revolution in England and France, even while promoting human capital more generally had little effect (Mitch, 1999).

Beyond these issues, the novel datasets I construct for this paper can be used to address several additional questions. I introduce two new datasets: a list of college site selection experiments, and detailed college yearbook data that includes information on students and faculty at various colleges. It is my hope that in the future these data will be used to explore the effects of colleges throughout U.S. history more generally, including their effect on labor markets, innovation, inequality, and economic growth.

References

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

Acharya, A., Blackwell, M., & Sen, M. (2016, August). Explaining causal findings without bias: detecting and assessing direct effects. *American Political Science Review*, *110*(3), 512-529.

Adams, J. D. (2002, July). Comparative localization of academic and industrial spillovers. *Journal of Economic Geography*, *2*(3), 253-278.

Aghion, P., Akcigit, U., Hyytinen, A., & Toivanen, O. (2016). *Living the American dream in Finland: the social mobility 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.

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. (2017). *The birth of American ingenuity: innovation and inventors of the golden age.* (Unpublished, University of Chicago)

Allison, P. D., & Waterman, R. P. (2002). Fixed-effects negative binomial regression models. *Sociological Methodology*, *32*.

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)

Angrist, J. D., & Krueger, A. B. (1992). The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples. *Journal of the American Statistical Association*.

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

Barro, R. J. (1991, May). Economic growth in a cross section of countries. *Quarterly Journal of Economics*, *106*(2), 407-443.

Barro, R. J., & Lee, J. W. (2013, September). A new data set of educational attainment in the world, 1950-2010. *Journal of Development Economics*, *104*(C), 184-198.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., & Van Reenen, J. (2016). *The lifecycle of inventors.* (Unpublished, Harvard University)

Bellemare, M. F. (2016). *'Metrics Monday: testing for mechanisms (and possibly ruling out all other mechanisms) (updated).* (Marc F. Bellemare, website, `http://marcfbellemare.com/wordpress/12165`, accessed May 22, 2016)

Benhabib, J., & Spiegel, M. M. (1994, October). The role of human capital in economic development: evidence from aggregate cross-country data. *Journal of Monetary Economics*, *34*(2), 143-173.

Berman, E. P. (2008, December). Why did universities start patenting? Institution-building and the road to the Bayh-Dole Act. *Social Studies of Science*, *38*(6), 835-871.

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.

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., 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.

Bound, J., & Turner, S. (2002, October). Going to war and going to college: did World War II and the G.I. Bill increase educational attainment for returning veterans? *Journal of Labor Economics*, *20*(4), 784-815.

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.

Celik, M. A. (2015). *Does the cream always rise to the top? The misallocation of talent in innovation.* (Unpublished, University of Pennsylvania)

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

Cohen, D., & Soto, M. (2007, March). Growth and human capital: good data, good results. *Journal of Economic Growth*, *12*(1), 51-76.

Cook, L. D. (2014, June). Violence and economic activity: evidence from African American patents, 1870-1940. *Journal of Economic Growth*, *19*(2), 221-257.

Cook, L. D., Logan, T. D., & Parman, J. M. (2014). Distinctively black names in the American past. *Explorations in Economic History*, *53*, 64-82.

Correia, S. (2016). *Linear models with high-dimensional fixed effects: an efficient and feasible estimator.* (Unpublished, Board of Governors of the Federal Reserve System)

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

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.

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

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In J. V. Henderson & J.-F. Thisse (Eds.), *Handbook of urban and regional economics* (Vol. 4). Amsterdam: Elsevier.

Fan, J., & Gijbels, I. (1996). *Local polynomial modelling and its applications*. New York: Chapman & Hall.

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

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.

Furman, J. L., Porter, M. E., & Stern, S. (2002, August). The determinants of national innovative capacity. *Research Policy*, *31*(6), 899-933.

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.

Glaeser, E. L., & Gottlieb, J. D. (2009, December). The wealth of cities: agglomeration economics and spatial equilibrium in the United States. *Journal of Economic Literature*, *47*(4), 983-1028.

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.

Goldin, C., & Katz, L. F. (2011, Winter). Putting the "co" in education: timing, reasons, and consequences of college coeducation from 1835 to the present. *Journal of Human Capital*, *5*(4), 377-417.

Greene, W. (2005). Functional form and heterogeneity in models for count data. *Foundations and Trends in Econometrics*, *1*(2), 113-218.

Greenstone, M., Hornbeck, R., & Moretti, E. (2010, June). Identifying agglomeration spillovers: evidence from winners and loser 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.

Guimaraes, P. (2008, February). The fixed effects negative binomial model revisited. *Economics Letters*, *99*(1), 63-66.

Gumprecht, B. (2003, January). The American college town. *Geographical Review*, *93*(1), 51-80.

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

Halvorsen, R., & Palmquist, R. (1980, June). The interpretation of dummy variables in semilogarithmic equations. *American Economic Review*, *70*(3), 474-475.

Hausman, J., Hall, B. H., & Griliches, Z. (1984, July). Econometric models for count data with an application to the patents-R&D relationship. *Econometrica*, *52*(4), 909-938.

Hausman, N. (2013). *University innovation, local economic growth, and entrepreneurship.* (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*.

Hobbs, A. (2014). *A chosen exile: a history of racial passing in American life*. Cambridge, MA: Harvard University Press.

Hornbeck, R. (2010, March). Barbed wire: property rights and agricultural development. *Quarterly Journal of Economics*, *125*(2), 767-810.

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.

Jaspin, E. (2007). *Buried in the bitter waters: the hidden history of racial cleansing in America*. New York: Basic Books.

Joffe, M. M., & Greene, T. (2009, June). Related causal frameworks for surrogate outcomes. *Biometrics*, *65*(2), 530-538.

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). Age and great invention. *Review of Economics and Statistics*.

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. (2016). *Research proximity and productivity: long-term evidence from agriculture.* (Unpublished, Florida State University)

Kennedy, P. E. (1981, September). Estimation with correctly interpreted dummy variables in semilogarithmic equations. *American Economic Review*, *71*(4), 801.

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.

Kline, P., & Moretti, E. (2014a, February). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority. *Quarterly Journal of Economics*, *129*(1), 275-331.

Kline, P., & Moretti, E. (2014b, August). People, places, and public policy: some simple welfare economics of local economic development programs. *Annual Review of Economics*, *6*, 629-662.

Krueger, A. B., & Lindahl, M. (2001, December). Education for growth: why and for whom? *Journal of Economic Literature*, *39*(4), 1101-1136.

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). Science or graduates: how do firms benefit from the proximity of universities? *Research Policy*.

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.

List, J. A., Sadoff, S., & Wagner, M. (2011, November). So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design. *Experimental Economics*, *14*(4), 439-457.

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

Logan, T. D., & Parman, J. M. (2017, March). The national rise in residential segregation. *Journal of Economic History*, *77*(1), 127-170.

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

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)

Miller, P. (1961). Introduction. In *America: a sketch of its political, social, and religious character.* Belknap/Harvard.

Mitch, D. (1999). The role of education and skill in the British Industrial Revolution. In J. Mokyr (Ed.), *The British Industrial Revolution: an economic perspective* (2nd ed., p. 241-279). Boulder, CO: Westview Press.

Mokyr, J. (2002). *The gifts of Athena: historical origins of the knowledge economy.* Princeton, NJ: Princeton University Press.

Mokyr, J. (2005, June). The intellectual origins of modern economic growth. *Journal of Economic History*, *65*(2), 285-351.

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.

Mowery, D. C., Nelson, R. R., Sampat, B. N., & Ziedonis, A. A. (2001, January). The growth of patenting and licensing by US universities: an assessment of the effects of the Bayh-Dole Act of 1980. *Research Policy*, *30*(1), 91-119.

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.

New York Times Editorial Board. (2017, March 27). The Trump administration's war on science. *New York Times*.

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

Perlman, E. R. (2015). *Dense enough to be brilliant: patents, urbanization, and transportation in nineteenth century America.* (Unpublished, Boston University)

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)

Press, W. H. (2013, November). What's so special about science (and how much should we spend on it?). *Science*, *342*(6160), 817-822.

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)

Science News Staff. (2017). A grim budget day for U.S. science: analysis and reaction to Trump's plan. *Science*.

Squicciarini, M. P., & Voigtländer, N. (2015, November). Human capital and industrialization: evidence from the Age of Enlightenment. *Quarterly Journal of Economics*, *130*(4), 1825-1883.

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.

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

Valero, A., & Van Reenen, J. (2016). *The economic impact of universities: evidence from across the globe.* (Unpublished, NBER Working Paper 22501)

Vansteelandt, S. (2009, November). Estimating direct effects in cohort and case-control studies. *Epidemiology*, *20*(6), 851-60.

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.

Whalley, A., & Hicks, J. (2014, January). Spending wisely? How resources affect knowledge production in universities. *Economic Inquiry*, *52*(1), 35-55.

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

Zucker, L. G., Darby, M. R., & Brewer, M. B. (1998, March). Intellectual human capital and the birth of U.S. biotechnology enterprises. *American Economic Review*, *88*(1), 290-306.

Graphs

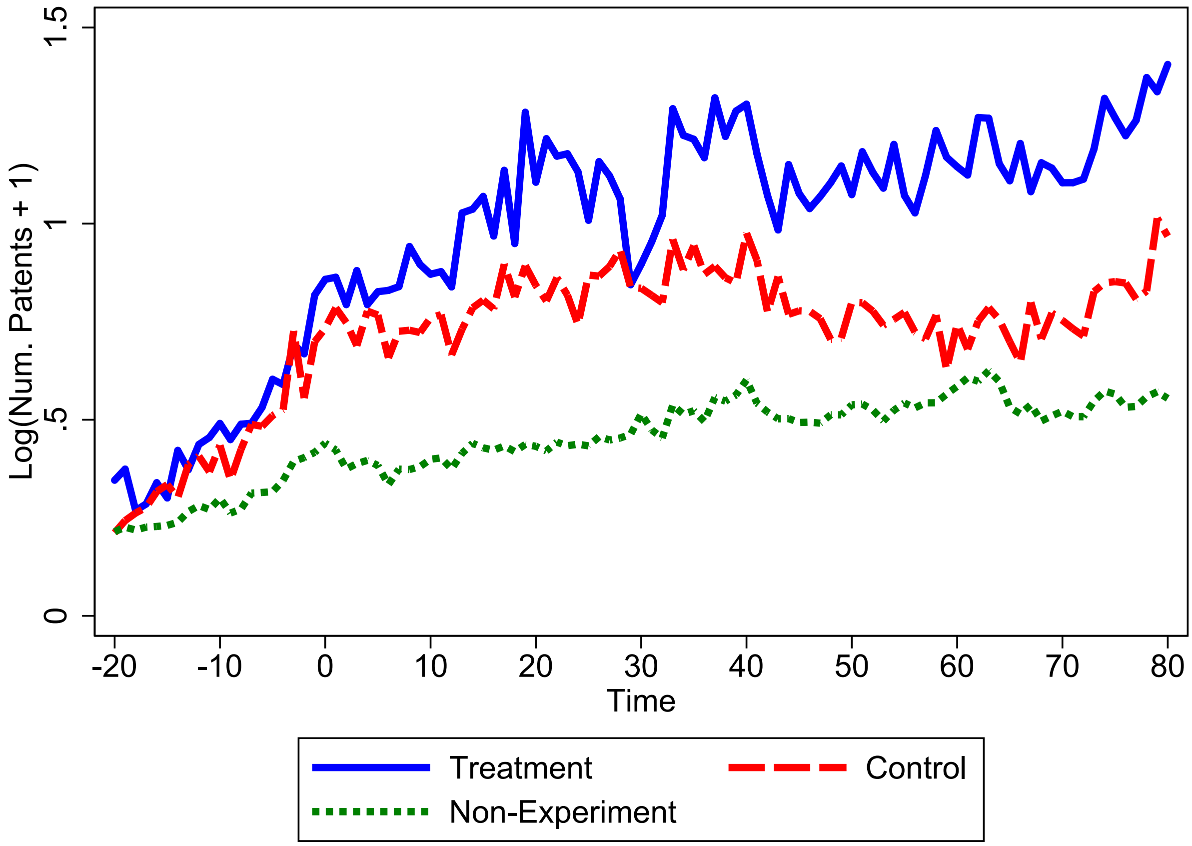


Figure 1: Unconditional mean patenting in treatment and control counties. The x-axis shows the number of years since the college experiment. The year of the college experiment is normalized to year 0. Everything left of year 0 shows pre-treatment means; everything to the right shows post-treatment means. The y-axis shows $log(Patents + 1)$. The treatment counties are represented by the blue solid line. The control 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.

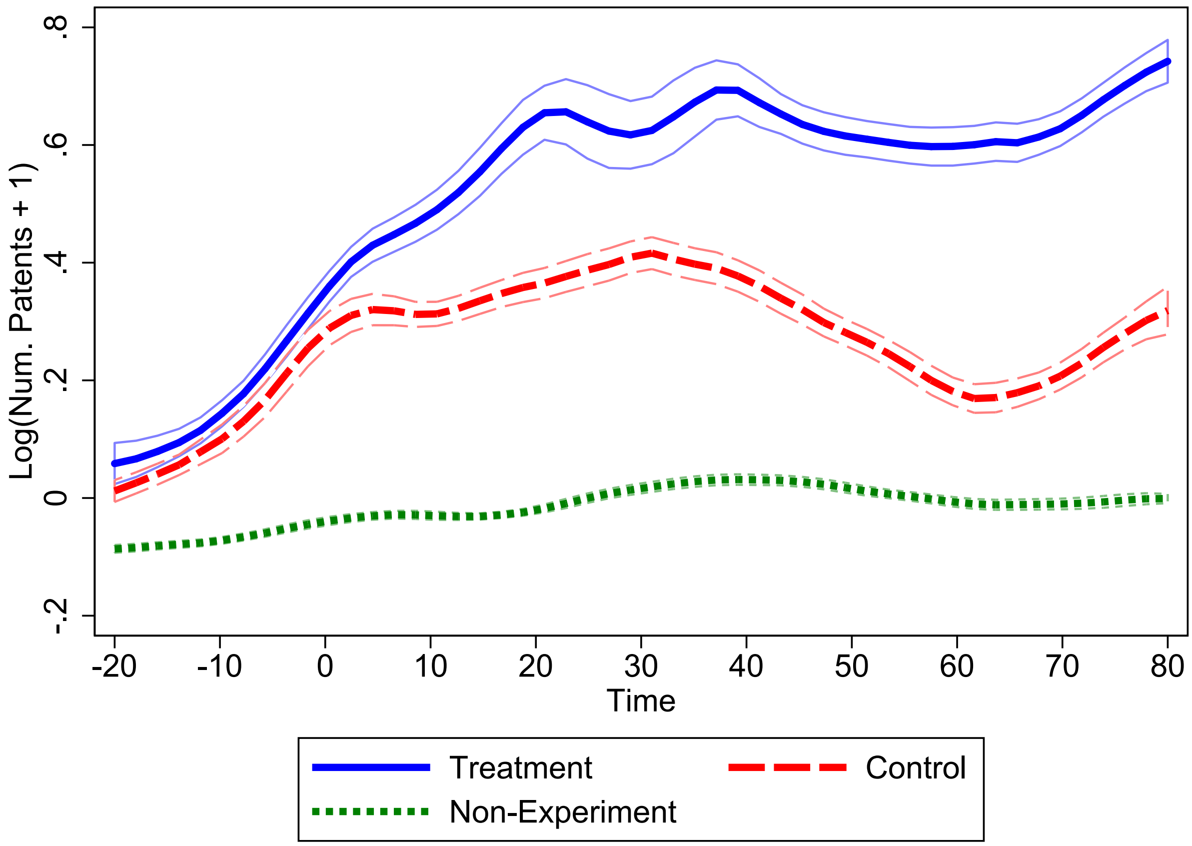


Figure 2: Mean patenting in treatment and control counties after controlling for year effects. Counties are excluded if less than five years of pre-treatment patent data is available. The x-axis shows the number of years since the college experiment. The year of the college experiment is normalized to year 0. Everything left of year 0 shows pre-treatment means; everything to the right shows post-treatment 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 treatment counties are represented by the blue solid line. The control 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.

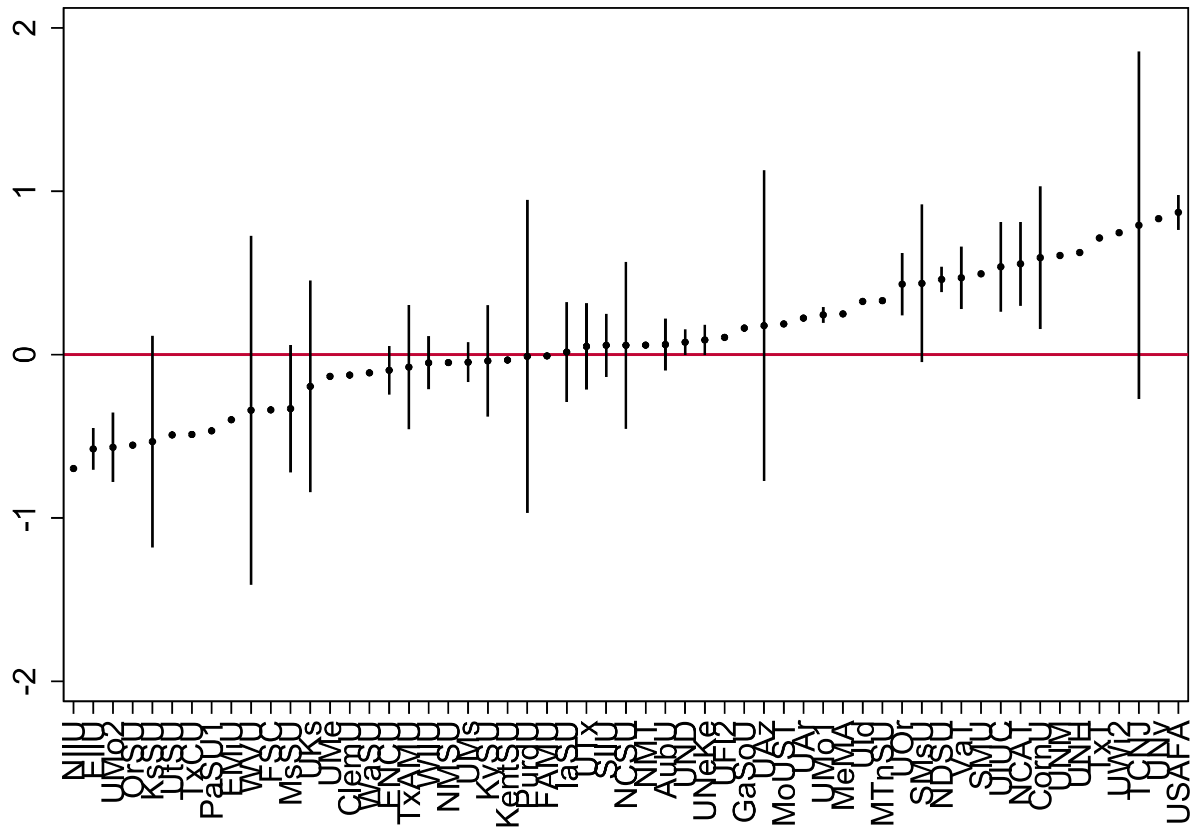


Figure 3: Estimated coefficient of the level shift in patenting in treatment counties relative to control counties after establishments of a new college with a separate interaction term estimated for each college experiment. Coefficients are ordered from smallest to largest and then plotted, along with 95% confidence bands. Results are for high quality experiments only.

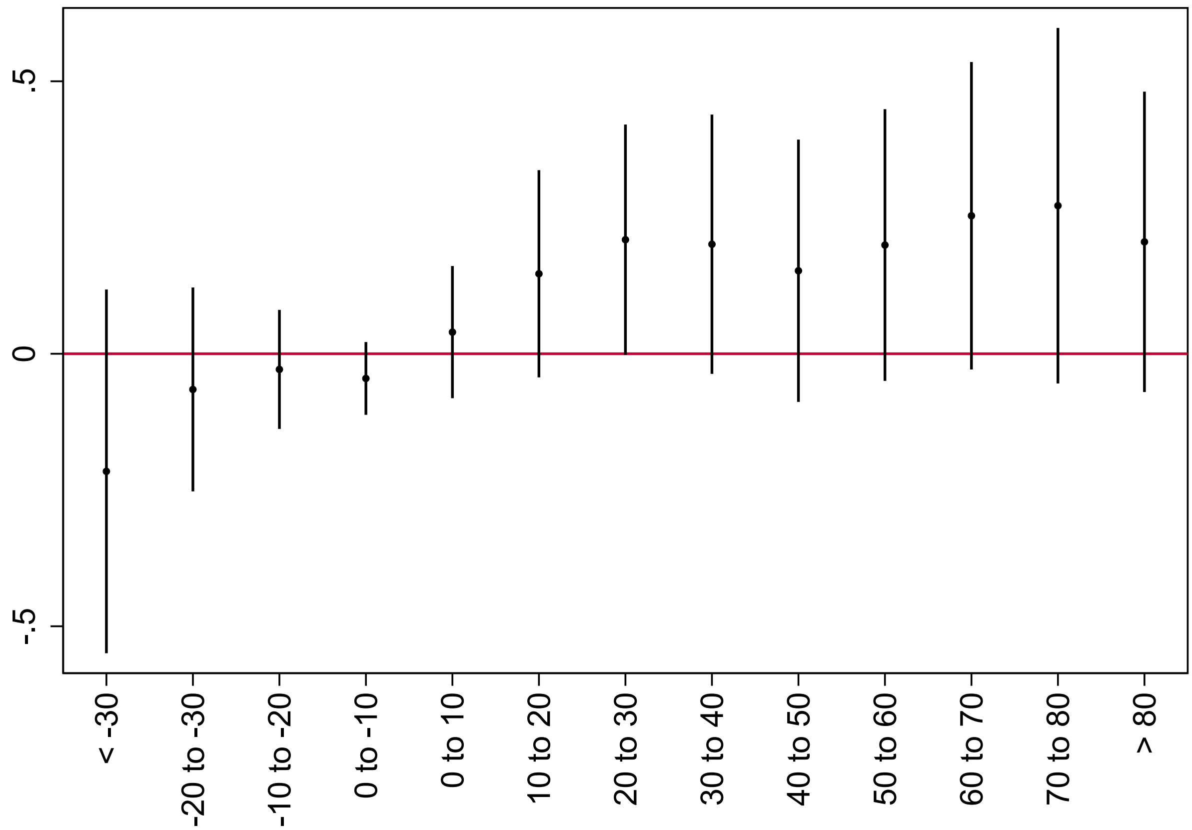
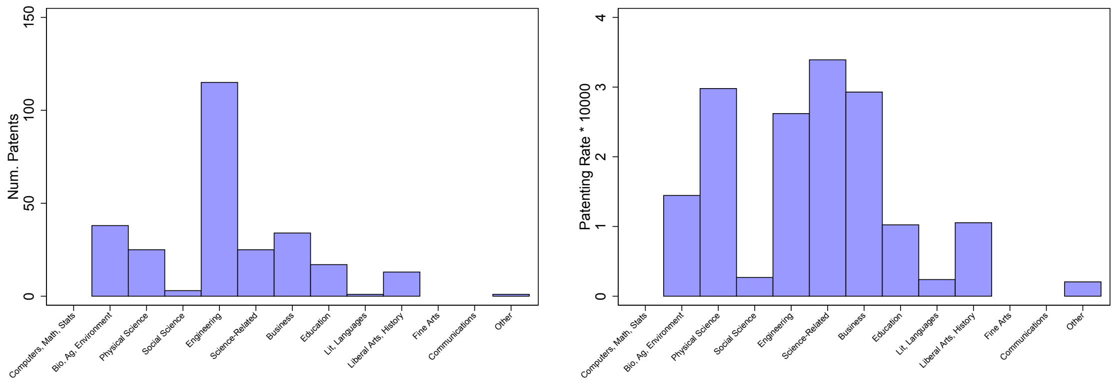


Figure 4: Estimated coefficient of the level shift in patenting in treatment counties relative to control counties after establishments of a new college with a separate interaction term estimated for each time bin, along with 95% confidence bands. Time bins are are dummy variables that are equal to one for treatment counties in every ten year period before and after the establishment of the new college. Results are for high quality experiments only.



(a) Num. Patents **(b)** Patenting Rate

Figure 5: Patents by college major. Panel A plots the number of patents from alumni with each major field. Panel B plots the patenting rate of alumni in each major field, presented as the number of patents per 10,000 alumni with a particular major.

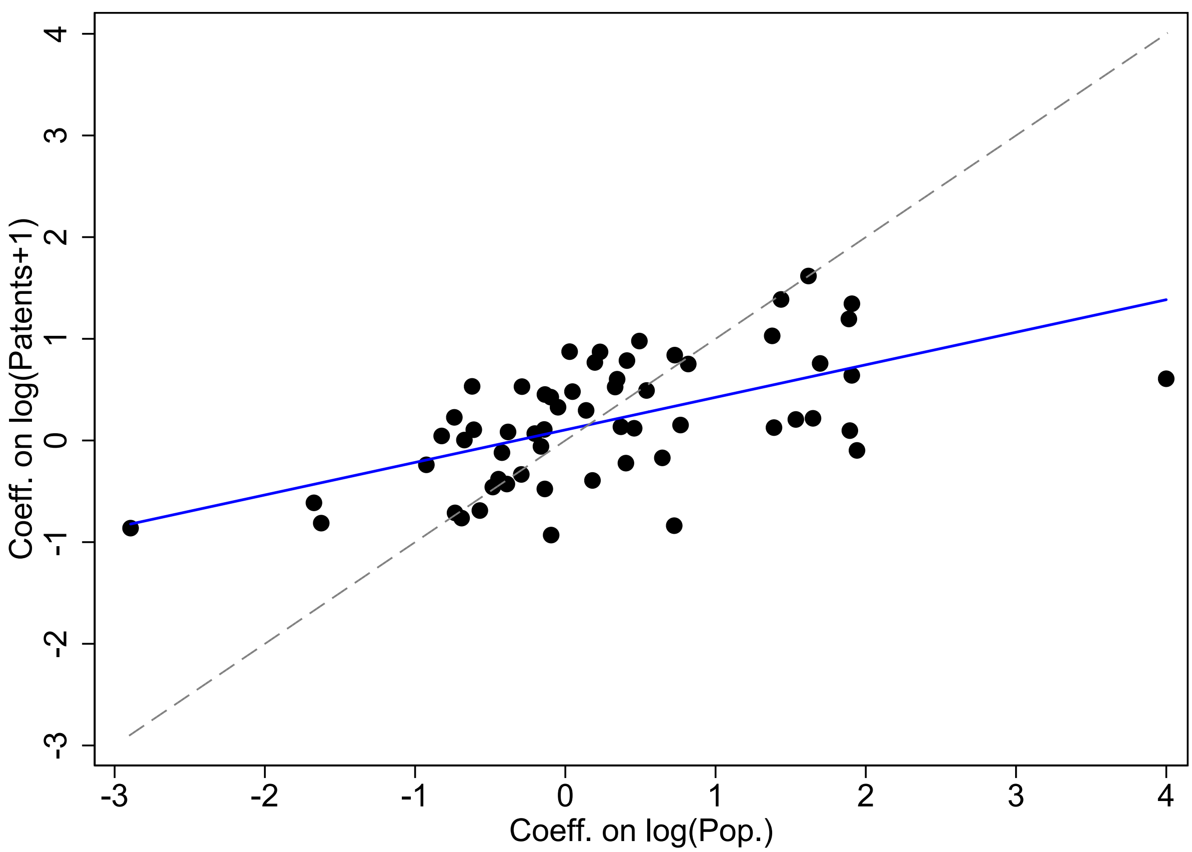


Figure 6: Scatter plot of the estimated coefficients for $log(NumPatents + 1)$ compared to estimated coefficients for $log(TotalPopulation)$ for each college site selection experiment. The blue solid line is the line of best fit. The gray dashed line is the 45 degree line.

Tables

	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	Michigan State Normal College	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	UCa	Alameda	California	1857	Land Grant
7	Iowa State University	IaSU	Story	Iowa	1859	Land Grant
8	Kansas State University	KsSU	Riley	Kansas	1863	Land Grant
9	University of Kansas	UKs	Douglas	Kansas	1863	Other Public
10	Cornell University	CornU	Tompkins	New York	1865	Land Grant
11	University of Wisconsin	UWi2	Dane	Wisconsin	1866	Land Grant
12	University of Maine	UMe	Penobscot	Maine	1866	Land Grant
13	University of Illinois	UIUC	Champaign	Illinois	1867	Land Grant
14	West Virginia University	WVU	Monongalia	West Virginia	1867	Land Grant
15	Oregon State University	OrSU	Benton	Oregon	1868	Land Grant
16	Purdue University	PurdU	Tippecanoe	Indiana	1869	Land Grant
17	University of Tennessee	UTn2	Knox	Tennessee	1869	Land Grant
18	Southern Illinois University	SIlU	Jackson	Illinois	1869	Normal School
19	Louisiana State University	LaSU	Eastbatonrouge	Louisiana	1870	Land Grant
20	University of Missouri	UMo2	Boone	Missouri	1870	Land Grant
21	Missouri University of Science and Technology	MoUST	Phelps	Missouri	1870	Land Grant
22	University of Arkansas	UAr	Washington	Arkansas	1871	Land Grant
23	Texas A and M University	TxAMU	Brazos	Texas	1871	Land Grant
24	Virginia Polytechnic Institute	VaT	Montgomery	Virginia	1872	Land Grant
25	Auburn University	AubU	Lee	Alabama	1872	Land Grant
26	University of Oregon	UOr	Lane	Oregon	1872	Other Public
27	University of Colorado	UCo	Boulder	Colorado	1874	Other Public
28	Mississippi State University	MsSU	Oktibbeha	Mississippi	1878	Land Grant
29	University of Texas Austin	UTx	Travis	Texas	1881	Other Public
30	University of North Dakota	UND	Grandforks	North Dakota	1883	Other Public
31	North Dakota State University	NDSU	Cass	North Dakota	1883	Land Grant
32	University of Arizona	UAz	Pima	Arizona	1885	Other Public
33	University of Nevada	UNv	Washoe	Nevada	1885	Land Grant
34	Arizona State University	AzSU	Maricopa	Arizona	1885	Land Grant
35	North Carolina State University	NCSU	Wake	North Carolina	1886	Land Grant
36	Georgia School of Technology	GaT	Fulton	Georgia	1886	Technical School
37	Kentucky State University	KySU	Franklin	Kentucky	1886	HBCU
38	Florida Agricultural and Mechanical University	FAMU	Leon	Florida	1887	HBCU
39	Utah State University	UtSU	Cache	Utah	1888	Land Grant
40	University of New Mexico	UNM	Bernalillo	New Mexico	1889	Other Public
41	University of Idaho	UId	Latah	Idaho	1889	Land Grant
42	New Mexico School of Mines	NMT	Socorro	New Mexico	1889	Technical School
43	New Mexico State University	NMSU	Donaana	New Mexico	1889	Land Grant
44	Clemson University	ClemU	Pickens	South Carolina	1889	Land Grant
45	Washington State University	WaSU	Whitman	Washington	1891	Land Grant
46	University of New Hampshire	UNH	Strafford	New Hampshire	1891	Land Grant
47	North Carolina A and T University	NCAT	Guilford	North Carolina	1892	HBCU
48	Eastern Illinois University	EIlU	Coles	Illinois	1895	Normal School
49	Northern Illinois University	NIlU	Dekalb	Illinois	1895	Normal School
50	Western Illinois University	WIlU	Mcdonough	Illinois	1899	Normal School
51	University of Nebraska at Kearney	UNeKe	Buffalo	Nebraska	1903	Normal School
52	University of Florida	UFl2	Alachua	Florida	1905	Land Grant
53	Georgia Southern College	GaSoU	Bulloch	Georgia	1906	Other Public
54	East Carolina University	ENCU	Pitt	North Carolina	1907	Technical School
55	Middle Tennessee State University	MTnSU	Rutherford	Tennessee	1909	Normal School
56	Kent State University	KentSU	Portage	Ohio	1910	Normal School
57	Southern Mississippi University	SMsU	Forrest	Mississippi	1910	Normal School
58	Texas Christian University	TxCU	Tarrant	Texas	1910	Other Private
59	Southern Methodist University	SMU	Dallas	Texas	1911	Other Private
60	Florida Southern College	FSC	Polk	Florida	1922	Other Private
61	Texas Tech	TxT	Lubbock	Texas	1923	Technical School
62	US Merchant Marine Academy	USMMA	Nassau	New York	1941	Military Academy
63	Maine Maritime Academy	MeMA	Hancock	Maine	1941	Military Academy
64	US Air Force Academy	USAFA	Elpaso	Colorado	1954	Military Academy

Table 1: 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, land grant status, or other uncertainty 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.

	Treatment	Controls	Treat. - Cont.	Non-Experiment	Treat. - Non-Exp.
log(Patents + 1)	0.68	0.61	0.0688	0.45	0.2306**
	(0.97)	(0.92)	(0.1442)	(0.85)	(0.1083)
Num. Patents	2.70	2.51	0.1868	4.39	-1.6870
	(5.92)	(6.52)	(0.9756)	(67.19)	(8.4667)
log(Total Pop.)	9.85	9.43	0.4238*	9.18	0.6689***
	(0.96)	(1.56)	(0.2193)	(1.37)	(0.1777)
Frac. Urban	0.15	0.14	0.0129	0.08	0.0767***
	(0.20)	(0.20)	(0.0324)	(0.17)	(0.0223)
Frac. Foreign Immigrant	0.09	0.11	-0.0154	0.13	-0.0341
	(0.09)	(0.10)	(0.0208)	(0.15)	(0.0259)
Frac. Male	0.52	0.51	0.0066	0.52	-0.0034
	(0.12)	(0.13)	(0.0207)	(0.11)	(0.0145)
Frac. White	0.78	0.77	0.0137	0.81	-0.0254
	(0.26)	(0.28)	(0.0474)	(0.24)	(0.0342)
Mean Age	29.77	29.04	0.7330	31.00	-1.2282
	(8.02)	(7.76)	(1.3722)	(6.16)	(0.8899)
Frac. Interstate Migrants	0.60	0.58	0.0163	0.50	0.0981**
	(0.35)	(0.35)	(0.0602)	(0.31)	(0.0446)
log(Manuf. Employment)	4.72	4.53	0.1908	3.84	0.8769**
	(2.35)	(2.45)	(0.4727)	(2.46)	(0.3917)
log(Value Manuf. Output)	12.55	11.98	0.5660	11.20	1.3455***
	(1.83)	(2.90)	(0.4395)	(3.51)	(0.4829)
log(Manuf. Wages)	9.10	8.83	0.2720	7.57	1.5302*
	(4.72)	(4.67)	(0.9181)	(4.98)	(0.7934)
log(Value Farm Product)	13.28	13.06	0.2238	13.02	0.2623
	(1.36)	(1.43)	(0.2448)	(2.29)	(0.3276)
log(Farm Wages)	11.50	11.15	0.3493	10.36	1.1338**
	(0.82)	(1.08)	(0.3219)	(1.79)	(0.4631)
log(Value Farms)	14.58	14.27	0.3018	14.18	0.3979
	(1.39)	(2.00)	(0.3063)	(2.26)	(0.3110)

Table 2: T-tests comparing the means of the treatment counties, control counties, and non-experimental counties. The first column lists the mean and standard deviation of treatment counties. The second column lists the mean and standard deviation of the control counties. The third column lists the difference in the mean between the treatment and control 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 treatment and the non-experimental counties, as well as the standard error of the difference. The treatment and control counties are from high quality experiments only. Stars in columns 3 and 5 indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

	log(Patents +1)	Alt. log(Patents)	Num. Patents	Neg. Binomial	log(Patents + 1)
Treat.County * HighQual. * PostTreatment: (% Change)	0.3274***	0.1601**	0.4197*	0.5343*	-0.1135
	(0.1231)	(0.0808)	(0.2188)	(0.2848)	(0.0836)
(# Change)	1.7505***	0.8560**	2.2443*	2.8570	-0.6068
	(0.6583)	(0.4321)	(1.1700)	(1.5229)	(0.4472)
PostTreatment: (% Change)	0.0277	-0.0779*	-0.1324	2.4805	-0.0057
	(0.0693)	(0.0409)	(0.1445)	(0.6459)	(0.0094)
(# Change)	0.1481	-0.4165*	-0.7080	13.2633	-0.0307
	(0.3703)	(0.2187)	(0.7726)	(3.4534)	(0.0505)
Zero Patents Dummy: (% Change)		-0.7403***			
		(0.0125)			
(# Change)		-3.9582***			
		(0.0669)			
HighQual. * PostTreatment: (% Change)					0.1407**
					(0.0575)
(# Change)					0.7525**
					(0.3075)
Treat.County * PostTreatment: (% Change)					0.4874***
					(0.0303)
(# Change)					2.6059***
					(0.1620)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes
Cnty-Year Obs.	34,194	34,194	34,194	34,194	763,106
# Counties	197	197	197	163	5,963
# Experiments	64	64	64	64	64
Adj. R-Sqr.	0.3428	0.6160	0.1773	0.2809	0.2148
Log-Likelihood	-37,578.0019	-28,390.8504	-133,060.8191	-44,025.3869	-676,385.9209

Table 3: Baseline regression results. Column 1 estimates the level shift in patenting in treatment counties relative to control 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. Column 5 includes all experiments and all non-experimental counties as additional controls and estimates a triple interaction equation. For each independent variable, the first row represents that percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

	Num.	Share of Pop.	Num. Patents	Patent Rate (x1000)	Share of County Patents	Estimated Perc. Increase	Share of Extra Patents
Entire County	96,960.1705	1.0000	5.5641	0.0643	1.0000	0.4202	1.0000
	(139,178.8184)	(1.0000)	(13.5182)	(0.0415)		(0.0468)	
Undergraduate Alumni	17,487.4548	0.2895	0.1653	0.0361	0.0703		0.1004
	(133,861.7162)	(0.2895)	(0.8210)	(0.1284)	(0.3669)		(0.4987)
Graduate Alumni	3,375.9311	0.0615	0.0000	0.0000	0.0000		0.0000
	(31,931.3216)	(0.0615)	(0.0000)	(0.0000)	(0.0000)		(0.0000)
Faculty	151.4930	0.0029	0.0147	0.0677	0.0051		0.0089
	(390.0103)	(0.0029)	(0.0668)	(0.2465)	(0.0288)		(0.0406)
County Less Alumni	69,880.6758	0.5983	4.4589	0.0481	0.9870		0.9813
	(227,818.6693)	(0.5983)	(13.1782)	(0.0348)	(0.0449)		(0.0572)

Table 4: 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 graduate student alumni. The fourth row lists statistics for college faculty. The fifth row lists statistics for the rest of the county after subtracting out aggregate numbers for alumni and faculty. The first column lists the average number of 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 estimated percentage increase in patenting after the establishment of the college; since there were no patents by alumni or faculty prior to the establishment of the college, this value is omitted for all rows except the first. The seventh column lists the share of the additional patents attributable to each group. Standard deviations are displayed in parentheses. Results are for college counties for which yearbook data is available.

	log(Townie Patents + 1)	log(Black Patents + 1)	log(Female Patents + 1)
Treat.County * HighQual. * PostTreatment: (% Change)	0.0053	0.0748**	0.0293
	(0.0032)	(0.0367)	(0.0599)
(# Change)	0.0002	0.0099**	0.0043
	(0.0001)	(0.0049)	(0.0089)
PostTreatment: (% Change)	0.0086**	0.0153	0.1003
	(0.0035)	(0.0277)	(0.0758)
(# Change)	0.0004**	0.0020	0.0149
	(0.0002)	(0.0037)	(0.0112)
County Fixed Effects	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes
Cnty-Year Obs.	35,967	18,705	7,003
# Counties	212	184	168
# Experiments	66	61	64
Adj. R-Sqr.	0.0073	0.1558	0.1854

Table 5: Regression results by individuals unaffiliated with colleges. The dependent variable in Column 1 is $log(TowniePatents + 1)$, where $Townies$ are individuals who lived in college or control counties and were at least 30 years old at the time the new college is established. The sample in Column 1 includes only those counties for which college yearbook data is also available. The dependent variable in Column 1 is $log(BlackPatents + 1)$. The dependent variable in Column 1 is $log(FemalePatents + 1)$. These results are for high quality experiments only. For each independent variable, the first row represents that percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

	Practical vs. Classical Colleges	Pre-1940 Practical vs. Classical Colleges	Frac. Ag. Patents	Frac. Mining Patents
Practical College Interaction: (% Change)	0.2799**	0.2270		
	(0.1305)	(0.1683)		
(# Change)	1.4966**	1.2137		
	(0.6977)	(0.8999)		
Classical College Interaction: (% Change)	0.2425	0.2150		
	(0.1883)	(0.1873)		
(# Change)	1.2965	1.1498		
	(1.0070)	(1.0015)		
Land Grant Interaction: (% Change)			-4.3555	
			(3.0384)	
(# Change)			-0.0296	
			(0.0206)	
Non-Land Grant Interaction: (% Change)			-0.1735	
			(1.2817)	
(# Change)			-0.0012	
			(0.0087)	
Technical School Interaction: (% Change)				-0.0163
				(0.2759)
(# Change)				-0.0003
				(0.0050)
Non-Technical School Interaction: (% Change)				0.1964
				(0.1679)
(# Change)				0.0036
				(0.0030)
County Fixed Effects	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Cnty-Year Obs.	34,194	20,544	10,345	10,345
# Counties	197	197	195	195
# Experiments	64	64	64	64
Adj. R-Sqr.	0.3407	0.3092	0.0190	0.0143

Table 6: Regression results by college type and patent class. The dependent variable in Columns 1 and 2 is $log(Num.Patents + 1)$. The dependent variable in Column 3 is the fraction of agricultural patents, $Num.AgriculturalPatents_{ijt}/Num.Patents_{ijt}$. The dependent variable in Column 4 is the fraction of mining patents, $Num.MiningPatents_{ijt}/Num.Patents_{ijt}$. These results use high quality experiments only. For each independent variable, the first row represents that percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

	NBER 1 Dig.	NBER 2 Dig.	USPTO 3 Dig.
Treat.County * HighQual. * PostTreatment: (% Change)	-0.7898*	-0.4162**	-0.4358**
	(0.4446)	(0.1988)	(0.2117)
(# Change)	-0.1126*	-0.1023**	-0.1058**
	(0.0634)	(0.0488)	(0.0514)
PostTreatment: (% Change)	0.1590	0.0928	0.1105
	(0.2218)	(0.1096)	(0.1156)
(# Change)	0.0227	0.0228	0.0268
	(0.0316)	(0.0269)	(0.0281)
County Fixed Effects	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes
Cnty-Year Obs.	7,884	8,173	8,173
# Counties	195	195	195
# Experiments	64	64	64
Adj. R-Sqr.	0.1495	0.1371	0.1238

Table 7: Regression results for changes in patent class diversity. The dependent variable is a measure of patent class concentration, similar to an HHI measure. More precisely, $Pat.Concent_{it} = \sum_{c \in C_{it}} \left(\frac{Num.Pat_c}{\sum_{k \in C_{it}} Num.Pat_k} \right)^2$ where C_{it} is the set of all patent classes in county i at time t. A higher number thus indicates that a county has more patents concentrated in a few classes. The dependent variable in Column 1 usees as the set of patent classes C_{it} the NBER one-digit classes. The dependent variable in Column 2 uses as the set of patent classes the NBER two-digit classes. The dependent variable in Column 3 uses as the set of patent classes the USPTO three-digit patent classes. These results are for high quality experiments only. For each independent variable, the first row represents the percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

	log(Migrant Patents + 1)	log(Patents + 1)	log(Total Pop.)	Frac. Urban	Patents per Capita	log(Patents + 1)	log(Patents + 1)
Treat.County: (% Change)	0.1708						
	(0.1199)						
(# Change)	0.9132						
	(0.6410)						
Treat.County * HighQual. * PostTreatment: (% Change)		0.2228*	0.4239*	0.9754**	0.0014	0.1199	0.1785*
		(0.1148)	(0.2317)	(0.3964)	(0.0020)	(0.0928)	(0.1060)
(# Change)		1.1914*	7,843.0510*	0.0758**	0.0077	0.6410	0.9544*
		(0.6137)	(4,286.2569)	(0.0308)	(0.0106)	(0.4961)	(0.5668)
PostTreatment: (% Change)		-0.0075	0.2970**	0.0687	0.0025	-0.0687	-0.0081
		(0.0737)	(0.1251)	(0.2547)	(0.0018)	(0.0624)	(0.0718)
(# Change)		-0.0400	5,494.8793**	0.0053	0.0132	-0.3674	-0.0431
		(0.3943)	(2,314.6039)	(0.0198)	(0.0094)	(0.3334)	(0.3841)
log(Total Pop.): (% Change)						0.2740***	
						(0.0340)	
(# Change)						1.4649***	
						(0.1816)	
Total Pop.: (% Change)							0.0112***
							(0.0042)
(# Change)							0.0597***
							(0.0223)
Total. Pop Squared: (% Change)							-0.0000***
							(0.0000)
(# Change)							-0.0000***
							(0.0000)
County Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cnty-Year Obs.	35,861	3,390	3,390	2,588	3,389	3,390	3,390
# Counties	212	204	204	204	204	204	204
# Experiments	66	66	66	66	66	66	66
Adj. R-Sqr.	0.2033	0.3213	0.5901	0.4908	0.1750	0.3489	0.3338

Table 8: Regression results after controlling for changes in county population in various ways. The dependent variable in Column 1 is $log(Num.\ Patents\ by\ Inter-countyMigrants\ +1)$, where $Inter-county\ Migrants$ are individuals who were not present in the county at the time the college is established. Column 1 is estimated by OLS. Data for Columns 2-7 are from census years only. The dependent variable in Column 2 is $log(Patents+1)$. The dependent variable in Column 3 is $log(Population)$. The dependent variable in Column 4 is $Population\ Living\ in\ Cities/Total\ Pop$. The dependent variable in Column 5 is patents per capita, or $Num.Patents/TotalPop$. Column 6 re-restimates the baseline regression of Column 2 but includes an additional control for $log(TotalPop)$. Columns 7 re-restimates the baseline regression of Column 2 but includes additional controls for $TotalPop$ and $(TotalPop)^2$. For each independent variable, the first row represents that percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

	log(Patents + 1)	log(Patents + 1)
Treat.County * PostTreatment: (% Change)	0.0879	0.3674*
	(0.9640)	(0.1964)
(# Change)	0.4700	1.9647*
	(5.1543)	(1.0503)
PostTreatment: (% Change)	-0.8199***	0.0717
	(0.1056)	(0.0913)
(# Change)	-4.3840***	0.3836
	(0.5647)	(0.4882)
Treat.County * PostTreatment * log(Total Pop.): (% Change)	-0.0449	
	(0.0841)	
(# Change)	-0.2401	
	(0.4495)	
PostTreatment * log(Total Pop.): (% Change)	0.1617**	
	(0.0695)	
(# Change)	0.8648**	
	(0.3717)	
Treat.County * log(Total Pop.): (% Change)	0.0383	
	(0.0533)	
(# Change)	0.2047	
	(0.2848)	
log(Total Pop.): (% Change)	0.1630***	
	(0.0411)	
(# Change)	0.8713***	
	(0.2198)	
Total Pop.: (% Change)		-0.0000
		(0.0000)
(# Change)		-0.0000
		(0.0000)
Treat.County * PostTreatment * Total Pop.: (% Change)		-0.0000
		(0.0000)
(# Change)		-0.0000
		(0.0000)
PostTreatment * Total Pop.: (% Change)		0.0000
		(0.0000)
(# Change)		0.0000
		(0.0000)
Treat.County * HighQual. * Total Pop.: (% Change)		0.0000
		(0.0000)
(# Change)		0.0000
		(0.0000)
Treat.County * PostTreatment * Total Pop. Squared: (% Change)		0.0000***
		(0.0000)
(# Change)		0.0000***
		(0.0000)
PostTreatment * Total Pop. Squared: (% Change)		-0.0000
		(0.0000)
(# Change)		-0.0000
		(0.0000)
Treat.County * Total Pop. Squared: (% Change)		0.0421**
		(0.0196)
(# Change)		0.2250**
		(0.1050)
Total Pop. Squared: (% Change)		-0.0000***
		(0.0000)
(# Change)		-0.0000***
		(0.0000)
County Fixed Effects	Yes	Yes
Year Effects	Yes	Yes
Cnty-Year Obs.	3,390	3,390
# Counties	204	204
# Experiments	66	66
Adj. R-Sqr.	0.3522	0.3361

Table 9: Results for the effect of population on patenting interacted with the treatment status. The dependent variable for both columns is $log(Patents + 1)$. Column 1 estimates the effect of the level shift in patenting in treatment counties relative to control 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 2 estimates the effect of the level shift in patenting in treatment counties relative to control 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 The number of additional patents is calculated by multiplying the estimated percentage change in patenting by the average number of patents per town in the year 1900. The number in curly brackes beneath the number of additional patents is the estimated percentage change in patenting 100. These results are for high quality experiments only. For each independent variable, the first row represents that percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

	No Adjacent Controls	Adjacent Controls	No Adjacent Counties	Adjacent Counties
Treat.County * HighQual. * PostTreatment: (% Change)	0.4104***	0.1618	0.5108***	0.3898***
	(0.1450)	(0.1302)	(0.1276)	(0.1231)
(# Change)	2.1943***	0.8649	2.7311***	2.0843***
	(0.7755)	(0.6962)	(0.6824)	(0.6580)
PostTreatment: (% Change)	-0.0359	0.0979	-0.0017	-0.0363
	(0.0820)	(0.1038)	(0.0099)	(0.0459)
(# Change)	-0.1921	0.5235	-0.0093	-0.1944
	(0.4385)	(0.5551)	(0.0527)	(0.2455)
County Fixed Effects	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Cnty-Year Obs.	26,669	18,725	710,375	64,106
# Counties	154	107	5,589	439
# Experiments	64	64	64	64
Adj. R-Sqr.	0.3424	0.3751	0.2139	0.2611

Table 10: Regression results by distance to control counties. Column 1 compares treatment counties to control counties that do not share a border. Column 2 compares treatment counties to control counties that do share a common border. The dependent variable in both columns is $log(Patents + 1)$. The coefficient is the percentage increase in patenting caused by the treatment. These results are for high quality experiments only. For each independent variable, the first row represents that percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

A Dataset Construction

A.1 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. Currently, yearbook data has been transcribed for Auburn University, the University of Arizona, the University of Colorado, Georgia Tech University, Louisiana State University, Missouri University of Science and Technology, the University of Missouri, the University of Nevada, the University of New Hampshire, North Carolina Agricultural and Technology University, Clemson University, Utah State University, Virginia Tech University, and Washington State University.[29] 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.

[29] For Auburn University, yearbooks are available from 1916, 1932-1935, 1938-1940. For the University of Arizona, yearbooks are available for 1913, 1920, 1922, 1927, 1933, 1936-1938, 1940. For the University of Colorado, yearbooks are available from 1893, 1903, 1908-1912, 1917, 1919-1922, 1924-1929, 1931-1939. For Georgia Tech, yearbooks are available from the years 1917-191, 1922-1924, 1926-1933, 1936, 1938-1940. For Louisiana State University, yearbooks are available from 1927, 1933, 1936-1940. For the Missouri University of Science and Technology, yearbooks are available for 1911, 1913, 1916, 1920-1922, 1925-1926, 1928-1929, 1939-1940. For the University of Missouri, yearbooks are available for 1898, 1903, 1905-1906, 1908-1910, 1912-1914, 1916-1918, 1920-1937, 1939-1940. For the University of Nevada, yearbooks are available for 1901, 1904, 1908, 1913, 1931, 1934, 1940. For the University of New Hampshire, yearbooks are available from 1909-1910, 1914, 1917, 1921, 1925-1927, 1930, 1933, 1937, 1939-1940. For North Carolina Agricultural and Technology University, yearbooks are available from 1939. For Clemson University, yearbooks are available from 1915, 1934, 1936, 1939-1940. For Utah State University, yearbooks are available from 1911, 1928, 1930, 1932, 1939. For Virginia Tech University, yearbooks are available for 1898, 1901, 1907, 1909, 1914-1915, 1917, 1919-1920, 1922, 1925-1927, 1930, 1933-1935, 1939. For Washington State University, yearbooks are available for 1903, 1908, 1912, 1926, 1931, 1933-1936, 1938-1940. Due to budget limitations, and due to the fact that future work will link students 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.

A.1.1 Alumni

Figure 7 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. (2017) for recent exceptions), modern data shows that very few inventors are that old (Jung & 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.[30]

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.[31] For years before the first college yearbook 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

[30] 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.

[31] 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.

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,

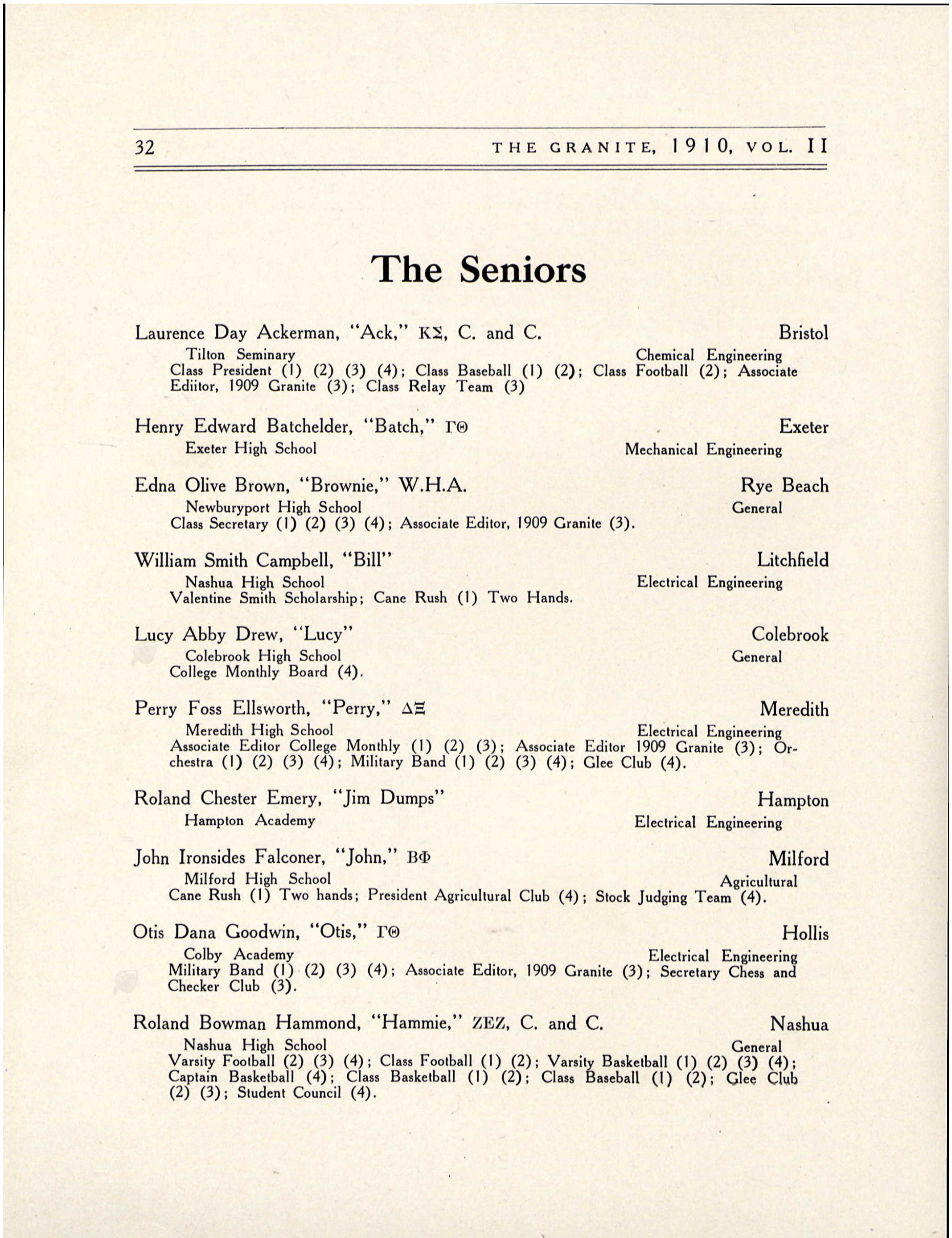
$$AlumniPatentRate_j = \frac{1}{T - t_{0j}} \sum_{t_{0j}}^{T} \frac{Num.Matched_{jt}}{Num.Alumni_{jt}}, \quad (5)$$

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, $Num.AlumniPatents_j = \sum_{t_{0j}}^{T} AlumniPatentRate_j * Num.\bar{A}lumni_{jt}$.

A.1.2 Faculty

Figure 8 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


32 THE GRANITE, 1910, VOL. II

The Seniors

Laurence Day Ackerman, "Ack," ΚΣ, 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," ΒΦ — Milford
Milford High School — Agricultural
Cane Rush (1) Two hands; President Agricultural Club (4); Stock Judging Team (4).

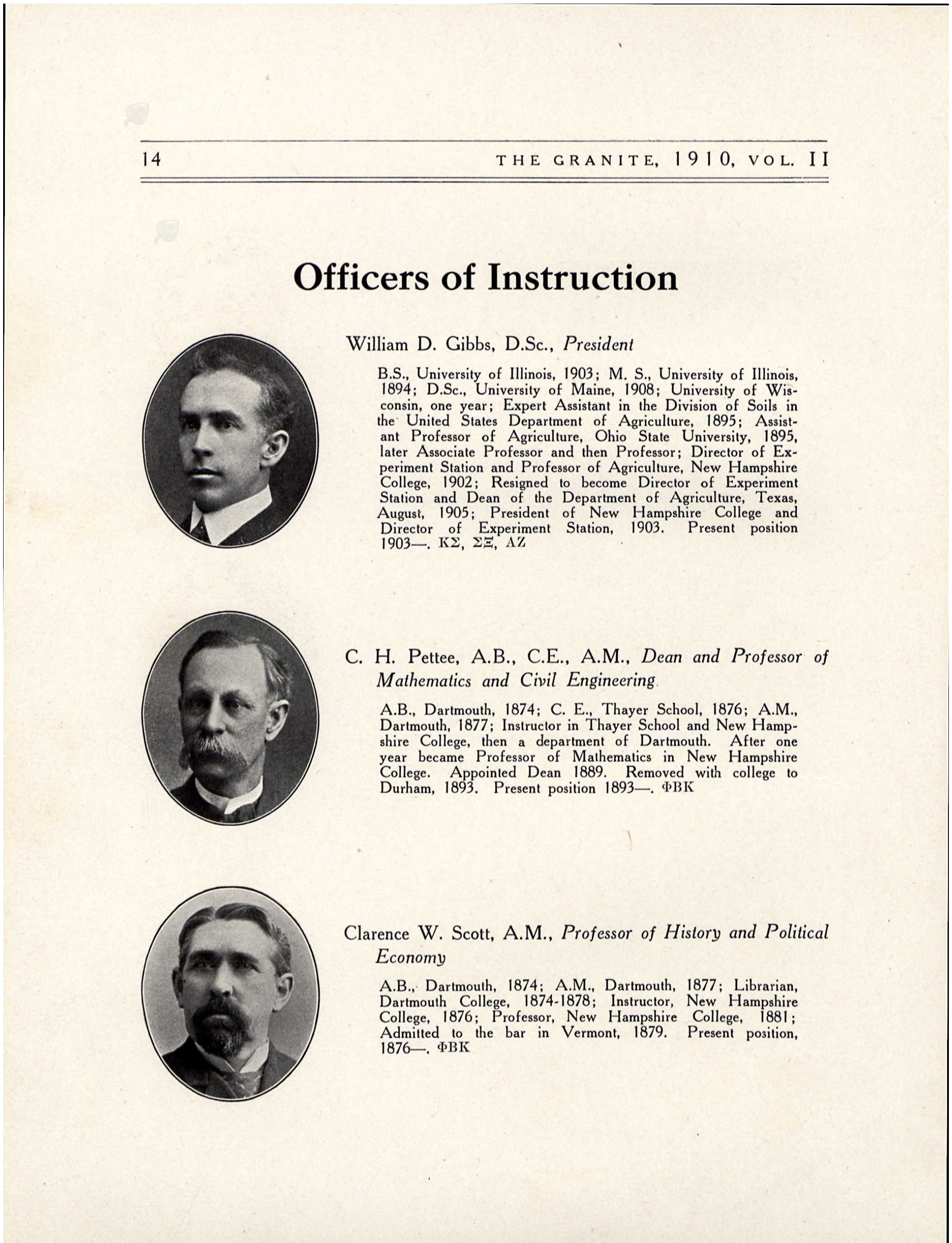
Otis Dana Goodwin, "Otis," ΓΘ — Hollis
Colby Academy — Electrical Engineering
Military Band (1) (2) (3) (4); Associate Editor, 1909 Granite (3); Secretary Chess and Checker Club (3).

Roland Bowman Hammond, "Hammie," ΖΕΖ, 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).

Figure 7: 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.

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 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.



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

Figure 8: 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.

A.2 Inferring Inventor Race and Gender

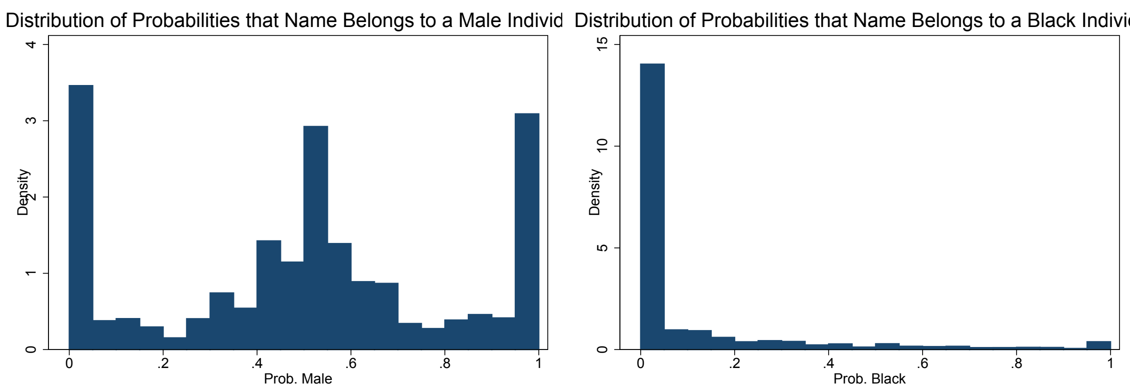
I use first names in the US decennial censuses to infer the race and gender of each patentee. The same technique is used in Sarada et al. (2017) to validate results on patenting by race and gender using direct patent-to-census name matching. Similar techniques have also been used in Cook et al. (2014) to identify patentee race, Jung & Ejermo (2014) to infer gender, Jones (2009) to infer age, and Celik (2015) to infer income.

To infer race and gender, for each first name in each county and year in the census, η_{it}, I calculate $p_{x,\eta_{it}} = Pr(x|\eta_{it})$ for characteristic $x = \{Black, Male\}$ in county i and year t. Then, for patentee j with name η_{jit}, I impute the probability of characteristic x using $p_{x,\eta_{jit}}$. At the county level, the inferred fraction of inventors with characteristic x is given by

$$\frac{1}{N}\sum_{\eta} N_\eta p_\eta$$

where N_η is the number of patentees with name η. This procedure is essentially a split-sample instrumental variables procedure as detailed in (Angrist & Krueger, 1992). Note that first name data for the 100% decennial censuses are only available through 1940.

To verify that this procedure is informative, Figure 9 plots the distribution of probabilities that names belong to either males or females. Panel (a) shows the distribution of probabilities that names belong to males, while panel (b) shows the distribution of probabilities that names belong to blacks. As can be seen, the distribution for each characteristic has multiple peaks: most names are either highly predictive of belonging to either males (blacks) or whites (females).



(a) Prob. Male

(b) Prob. Black

Figure 9: Distributions of probabilities that a patentee has a particular characteristic. Panel (a) plots the probabilities that a patentee is male. Panel (b) plots the probabilities that a patentee is black.

B Robustness Checks, Alternative Specifications, and Additional Results

B.1 Additional Balance Checks and Placebo Tests

Table 11 conducts t-tests for several other categories along which treatment and control counties can be measured. As in Table B.1, I show that the treatment and control counties are much more similar than the treatment and non-experimental counties in the same state.

I next conduct a placebo test to determine whether patenting changes differentially in treatment and control 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 12. If the treatment counties are up-and-coming places, then they should be growing faster then the control 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-treatment years as the placebo "treatment" year.

Finally, one may be concerned that the subjective nature by which an experiment is classified as high quality or low quality may lead to "cherry-picking" of experiments to achieve desired estimate. This is unlikely for two reasons. First, including all of the low quality experiments actually increases the estimated magnitude. Second, the relatively large number of college site selection experiments and the distribution of estimated experiment coefficients presented in Figure 3 make it unlikely that reclassifying a small number of experiments as either high or low quality will materially affect the results. Indeed, I verify this by 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.

	Treatment	Controls	Treat. - Cont.	Non-Experiment	Treat. - Non-Exp.
Total Pop.	31,849.28	24,773.03	7,076.2497	22,435.87	9,413.4085
	(54,724.02)	(40,259.65)	(7,219.1101)	(109,806.06)	(14,200.8323)
Frac. Rural	0.84	0.84	-0.0078	0.93	-0.0910***
	(0.21)	(0.21)	(0.0361)	(0.17)	(0.0230)
Segregation	0.34	0.34	0.0056	0.36	-0.0167
	(0.23)	(0.24)	(0.0583)	(0.21)	(0.0429)
Pop. per Sq. Mile	80.26	47.10	33.1579	49.45	30.8019
	(248.73)	(92.78)	(38.1282)	(634.17)	(117.8889)
Frac. Attending School	0.14	0.12	0.0183	0.15	-0.0153
	(0.08)	(0.08)	(0.0234)	(0.08)	(0.0186)
Manuf. Establishments	128.00	118.52	9.4839	114.16	13.8434
	(138.01)	(170.22)	(48.2134)	(584.29)	(141.8459)
Manuf. Employment	823.33	1,418.92	-595.5950	1,174.31	-350.9879
	(1,935.58)	(7,423.64)	(1,197.0870)	(11,152.15)	(1,764.0384)
Value Manuf. Output	1,432,135.91	3,262,587.35	-1,830,451.4479	4,391,171.67	-2,959,035.7603
	(3,772,832.04)	(22,766,052.45)	(3,154,735.5862)	(57,226,282.76)	(7,862,090.4114)
Manuf. Wages	416,109.72	992,501.92	-576,392.1950	993,785.23	-577,675.5067
	(1,094,275.31)	(6,529,775.53)	(1,042,206.1112)	(11,356,701.19)	(1,796,113.3422)
Value Farm Product	1,095,524.47	1,154,241.75	-58,717.2757	1,678,618.65	-583,094.1767
	(1,351,876.43)	(2,479,984.59)	(379,132.9680)	(3,978,590.63)	(568,990.1071)
Farm Wages	132,687.60	108,372.69	24,314.9077	76,953.60	55,734.0035*
	(105,706.01)	(100,856.57)	(33,274.1962)	(116,367.11)	(30,208.6285)
Value Farms	5,094,621.34	4,069,069.71	1,025,551.6293	4,883,984.25	210,637.0932
	(8,284,190.81)	(5,100,590.43)	(1,062,864.8802)	(8,333,384.34)	(1,153,409.5965)

Table 11: T-tests comparing the means of the treatment counties, control counties, and non-experimental counties. The first column lists the mean and standard deviation of treatment counties. The second column lists the mean and standard deviation of the control counties. The third column lists the difference in the mean between the treatment and control 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 treatment and the non-experimental counties, as well as the standard error of the difference. The treatment and control counties are from high quality experiments only. Stars in columns 3 and 5 indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

	log(Patents + 1)	log(Patents + 1)
Treat.County * HighQual. * PostTreatment: (% Change)	0.0637	-0.0611
	(0.0955)	(0.0895)
(# Change)	0.3409	-0.3268
	(0.5105)	(0.4785)
PostTreatment: (% Change)	-0.0685	-0.0396
	(0.0475)	(0.0441)
(# Change)	-0.3665	-0.2118
	(0.2538)	(0.2360)
Treat.County * Trend * PostTreatment: (% Change)		0.0091
		(0.0090)
(# Change)		0.0489
		(0.0483)
Trend * PostTreatment: (% Change)		0.0003
		(0.0055)
(# Change)		0.0014
		(0.0295)
Trend: (% Change)		0.0150***
		(0.0048)
(# Change)		0.0802***
		(0.0258)
County Fixed Effects	Yes	Yes
Year Effects	Yes	Yes
Cnty-Year Obs.	8,887	8,887
# Counties	197	197
# Experiments	64	64
Adj. R-Sqr.	0.2848	0.2883

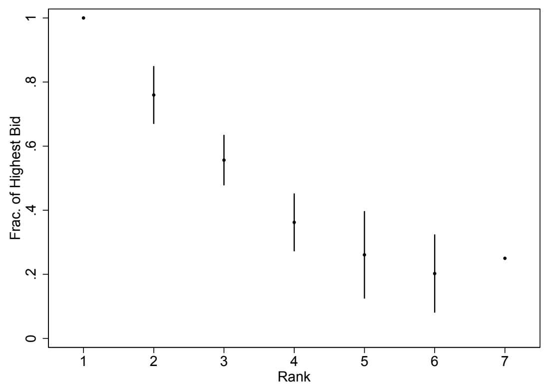
Table 12: Placebo tests. The baseline regression results are reproduced with all post-experiment data dropped. The experiment year is set to halfway between the initial year of patent data and the year prior to the original college site selection experiment. Column 1 estimates the level shift in patenting in treatment counties relative to control counties after establishment of a new college. Column 2 includes a linear trend term as well as coefficients estimating if the trend changes after the establishment of a new college and if the trend changes differently in treatment and control counties after establishment of the new college. The dependent variable for both columns is $log(Num_Patents + 1)$. The coefficient is the percentage increase in patenting caused by the treatment. Results are for high quality experiments only. For each independent variable, the first row represents that percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

B.2 College Rankings

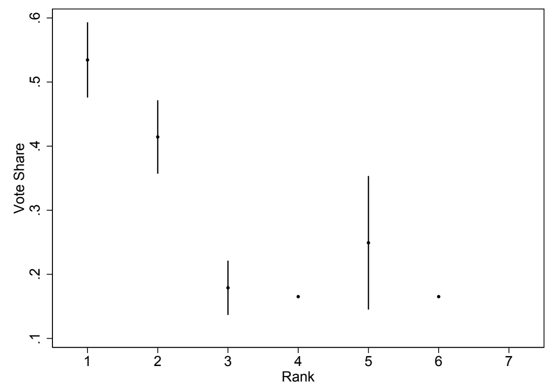
In certain instances, it is possible to rank the quality of each finalist county. In particular, I can rank the finalists when I either observe the bids that different counties offered or the number of votes that each received from site selection committees or state legislators. The finalist site ranked number one offers the highest bid; the lowest ranked county offers the lowest bid. For this exercise, I use data on all 136 college site selection experiments for which I am able to identify finalist locations; I do not restrict attention to only the high quality experiments. This allows me to verify that the "quality" of the counterfactual, in the sense of how close two given sites are in the rankings, is important when comparing winning to losing counties. There are 29 experiments for which I have data on either bids offered or votes received.

Figure 10 plots differences in a number of outcome variables across finalists of different ranks. In particular, I examine the probably that a county with a given rank wins the college. The highest ranked county has the highest probability of winning the college, but the probability is far from one, suggesting that in many cases other factors such as politics also play an important role in determining which county ultimately receives a college. To give a sense of how close the competitions were, I also compare the fraction of the highest bid by colleges of each rank. Finally, I compare colleges by the amount of patenting or county population in the last census year before the college site selection experiment. The first and second ranked counties appear very similar, with each additional ranking appearing less similar to the others.

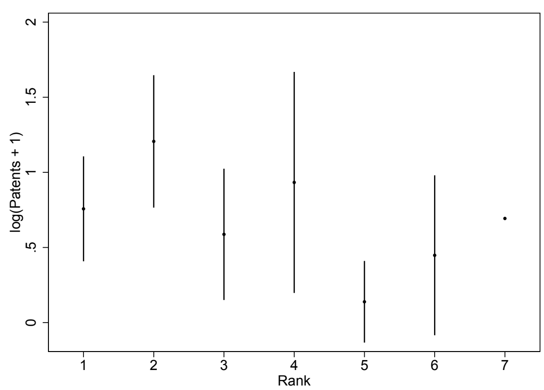
In Table 13 re-estimates the baseline specification but including different numbers of finalist sites. In Column 1, I include all 29 experiments for which I can rank the finalists. Due to the relatively small number of such experiments, these results are much noisier than the estimates using the full sample of colleges. In Column 2, I include only the first, second, third, and fourth ranked finalists. In Column 3, I include the first, second, and third ranked finalists. Finally, column 4 includes only the first and second ranked finalists. As one reads the table from left to right, the coefficients shrink in magnitude, until Column 4; Column 3 and 4 are almost identical in magnitude, suggesting that the three top finalists are good counterfactuals for one another. These results confirm the intuition discussed in Section 3.1 that including lower quality controls inflates the estimate of the effect of establishing a new college.



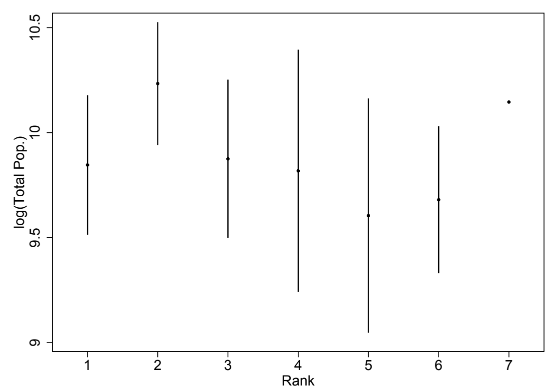
(a) Frac. of Highest Bid



(b) Vote Share



(c) $\log(Num.Patents + 1)$



(d) $\log(TotalPop.)$

Figure 10: Outcomes for finalists of various ranks. Colleges are ranked based on the size of bids submitted or the number of votes received. Panel (a) compares finalists by the probability of "winning" and receiving the college. Panel (b) compares finalists by the fraction of the highest bid submitted by each finalist, panel (c) the vote share, panel (d) logged patenting, and panel (e) logged total population. For each rank, the mean and 95% confidence intervals are plotted.

	All Ranked Counties	Top 4 Counties	Top 3 Counties	Top 2 Counties
Treat.County * Exp.County * PostTreatment: (% Change)	0.1826	0.1287	0.1012	0.1074
	(0.1733)	(0.1719)	(0.1770)	(0.1876)
(# Change)	0.9765	0.6882	0.5409	0.5742
	(0.9266)	(0.9189)	(0.9465)	(1.0033)
PostTreatment: (% Change)	0.0127	-0.0142	0.0115	0.1219
	(0.1170)	(0.1199)	(0.1328)	(0.1517)
(# Change)	0.0680	-0.0759	0.0616	0.6517
	(0.6258)	(0.6409)	(0.7100)	(0.8110)
County Fixed Effects	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Cnty-Year Obs.	15,575	13,825	12,600	10,500
# Counties	89	79	72	60
# Experiments	30	30	30	30
Adj. R-Sqr.	0.3623	0.3650	0.3581	0.3825

Table 13: Results using finalist counties that can be ordinally ranked. Counties can be ranked if the value of a bid or the number of votes received are recorded. Column 1 compares treatment counties to control counties when all ranked counties are included. Column 2 includes only the four top ranked counties. Column 3 includes only the three top ranked counties. Column 4 includes only the two top ranked counties. These results use both high and low quality experiments, as long as the counties can be ranked. For each independent variable, the first row represents that percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

B.3 Additional Regression Specifications

In this section, I estimate several additional regression specifications to demonstrate that the baseline results described in Section 3.1 are robust. Results are presented in Table 14.

In Column 1, I include a linear time trend in the regression and allow the trend to change after the establishment of a new college, allowing me to determine if the new college changed the rate of increase in patenting. More formally, this is estimated by

$$\begin{aligned}\log(NumPat_{ijt}+1) =& \beta_0 + \beta_1 College_{ij} * PostCollege_{jt} + \beta_2 PostCollege_{jt} \\ &+ \beta_3 time_j + \beta_4 time_j * PostCollege_{jt} \\ &+ \beta_5 time_j * College_{ij} * PostCollege_{jt} + \pi_i + \kappa_t + \xi_{ijt},\end{aligned} \tag{6}$$

where $time_j$ is a linear time trend, π_i are county fixed effects, and κ_t are year effects. So β_4 captures how the trend in patenting changed after a college was established. β_5 captures the change in the slope of patenting in treatment counties relative to control counties after a new college is established, while β_1 captures the change in intercept. Not surprisingly given Figure 2, the trend between the treatment and control counties is extremely close to zero. The estimated level change (β_1) is very similar to the baseline estimate in Table 3.

Column 2 generalizes the results in Column 1 by allowing the linear time trend to vary by county. More specifically, I estimate

$$\begin{aligned}\log(NumPat_{ijt}+1) =& \beta_0 + \beta_1 College_{ij} * PostCollege_{jt} + \beta_2 PostCollege_{jt} \\ &+ \sum_{i \in I} \beta_{3i} time_{ji} + \beta_{4i} time_{ji} * PostCollege_{jt} \\ &+ \beta_5 time_{ji} * College_{ij} * PostCollege_{jt} + \pi_i + \kappa_t + \xi_{ijt},\end{aligned} \tag{7}$$

where I is the set of all counties. The county-specific trends and interaction terms are omitted for readability. Even in this more demanding specification, establishing a new college has an effect that is nearly as large as in the baseline estimate (31% more patents per year) and is statistically significant at the 10% level.

In Column 3, I estimate 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 about ten percentage points more likely to have at least one patent in a given year. In 1870, each county had about a 40% probability of having at least one patent, so receiving a new college increased this probability by about 25%.

Column 4 shows how counties change their location in the ranking of counties by patents per year following the establishment of a new college. Establishing a new college raises the college county by about 3.5 percentiles relative to the controls.

Column 5 uses the square root of patents as the dependent variable. The square root of patents has many of the same benefits as does the modified *log* specification presented in Column 2 of Table 3; namely, square roots diminish the effect of outliers and can also handle zero values. The results from the square root regression are much smaller in magnitude than the baseline results and are no longer statistically significant, but still point in the same direction.

Column 6 displays the results using the number of patents as the dependent variable in a simple linear specification, identical to Column 3 in Table 3 with the exception that the number of patents has not been Winsorized. While the estimated percentage increase is huge (81% more patents per year in the college counties relative to the controls), it is not statistically significant, suggesting that the outlier counties add substantial noise to the estimation, and so it is more appropriate to use the Winsorized data.

Column 7 presents results from a simple fixed effects Poisson regression with non-Winsorized data. Column 8 presents results from a fixed effects negative binomial regression using identical techniques to those described for Column 4 of Table 3, with the only difference being that non-Winsorized data is used here. An α-likelihood ratio test shows that the patent count data is overdispersed and so a negative binomial specification is preferred to a Poisson regression. The preferred model is not sensitive to the use of Winsorized data. All models have qualitatively similar results that are in line with the baseline estimates, although the coefficients are typically not statistically different from zero unless the data is Winsorized.

	log(Patents + 1)	log(Patents + 1)	Any Patents	Percentile of Patents	Sqrt. Patents	Num. Patents	Poisson	Neg. Binomial	Zero-Inflated Neg. Binomial
Treat.County * HighQual. * PostTreatment: (% Change)	0.3125*	0.3093*	0.2532***	0.0539***	0.0463	0.8088	0.3548	0.3548	0.0768
	(0.1700)	(0.1691)	(0.0824)	(0.0153)	(0.0402)	(0.6331)	(0.5809)	(0.5809)	(0.1027)
(# Change)	1.6707*	1.6541*	0.1014***	0.0361***	0.2476	4.3247	1.8973	1.8973	0.4106
	(0.9093)	(0.9041)	(0.0330)	(0.0102)	(0.2148)	(3.3854)	(3.1061)	(3.1061)	(0.5491)
PostTreatment: (% Change)	0.0220	0.0538	0.2034**	0.0158	0.0003	-0.4442	4.0440	4.0440	-0.0051
	(0.0788)	(0.0822)	(0.0802)	(0.0150)	(0.0023)	(0.4473)	(2.0715)	(2.0715)	(0.0414)
(# Change)	0.1177	0.2877	0.0814**	0.0106	0.0017	-2.3749	21.6235	21.6235	-0.0274
	(0.4211)	(0.4398)	(0.0321)	(0.0100)	(0.0122)	(2.3919)	(11.0763)	(11.0763)	(0.2213)
Trend * Treat.County * PostTreatment: (% Change)	0.0001	0.0024							
	(0.0015)	(0.0039)							
(# Change)	0.0005	0.0131							
	(0.0081)	(0.0210)							
Trend * PostTreatment: (% Change)	-0.0006	0.0010							
	(0.0035)	(0.0051)							
(# Change)	-0.0035	0.0056							
	(0.0187)	(0.0273)							
Trend: (% Change)	0.0012	0.0449***							
	(0.0028)	(0.0047)							
(# Change)	0.0065	0.2400***							
	(0.0148)	(0.0249)							
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cnty-Year Obs.	34,194	34,194	34,194	34,194	34,194	34,194	34,194	34,194	34,194
# Counties	197	197	197	197	197	197	163	163	163
# Experiments	64	64	64	64	64	64	64	64	64
Adj. R-Sqr.	0.3428	0.4145	0.3760	0.5285	0.2467	0.0757		0.2488	
Log-Likelihood	-37,577.0635	-35,603.2930	-12,385.9727	30,277.0691	-62,445.6833	-156,960.2640	-119,432.9943	-49,340.8182	-36,326.3418

Table 14: Regression results using alternative specifications. Column 1 estimates the level shift in patenting in treatment counties relative to control counties after establishment of a new college when the dependent variable is $log(Num_Patents + 1)$ and a linear trend intercept and interaction term is included. Column 2 repeats the specification in Column 1 but allows for county-specific linear time trends. Column 3 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 4 is a county's percentile in rankings by patenting in a given year. The dependent variable in column 5 is $\sqrt{Num.Patents}$. The dependent variable in column 6 is the number of patents. Column 7 estimates a Poisson regression. Column 8 estimates a negative binomial regression. Column 9 estimates a zero-inflated negative binomial regression. Columns 6-9 use non-Windsorized counts of patenting. These results use high quality experiments only. For each independent variable, the first row represents that percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

B.4 Results by Experiment Type

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 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.[32] "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 control 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).

[32] 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.

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 15 shows the results. The coefficients for the interaction term are qualitatively the same over all experiment types. The coefficients for auctions, politics, and other are in line with the baseline regression results. The coefficients are much larger for cases in which states selected sites on the basis of existing infrastructure, although admittedly few of these cases occur in the data; removing these cases and re-estimating the baseline equation does not change in the results in any meaningful way, nor does estimating the results without any of the other experiment types, as Columns 2-5 show.

	log(Patents + 1)	No Auctions	No Politics	No Infrastructure	No Other
Auctions: (% Change)	0.2821				
	(0.2070)				
(# Change)	1.5082				
	(1.1066)				
Politics: (% Change)	0.3003**				
	(0.1478)				
(# Change)	1.6059**				
	(0.7904)				
Infrastructure: (% Change)	0.9136***				
	(0.3403)				
(# Change)	4.8851***				
	(1.8194)				
Other: (% Change)	0.3443				
	(0.4100)				
(# Change)	1.8410				
	(2.1920)				
Treat.County * HighQual. * PostTreatment: (% Change)		0.3665**	0.3563*	0.3022**	0.3158**
		(0.1429)	(0.1797)	(0.1260)	(0.1277)
(# Change)		1.9595**	1.9051*	1.6156**	1.6886**
		(0.7643)	(0.9611)	(0.6737)	(0.6831)
PostTreatment: (% Change)		-0.0531	0.0674	0.0594	0.0211
		(0.0890)	(0.0965)	(0.0726)	(0.0670)
(# Change)		-0.2840	0.3601	0.3174	0.1130
		(0.4759)	(0.5161)	(0.3882)	(0.3580)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes
Cnty-Year Obs.	34,194	18,235	20,369	32,619	31,359
# Counties	197	105	118	188	180
# Experiments	64	64	56	64	64
Adj. R-Sqr.	0.3466	0.3431	0.3577	0.3392	0.3403

Table 15: Regression results by experiment type. The dependent variable is $log(Patents+1)$. In column 1, the coefficient is the increase in patenting caused by the treatment interacted by experiment type. All other coefficients in the regression are suppressed for readability. Row 1 presents results for experiments decided by auction, row 2 for experiments decided by politics, row 3 for experiments decided by the presence of existing infrastructure, and row 4 for other site selection experiments. Columns 2-5 re-estimate the baseline results but excluding each experiment type in turn. These results are for high quality experiments only. For each independent variable, the first row represents that percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

B.5 Results Using Other Patent Data

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. Column 1 of Table 16 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.[33] These results are presented in Column 2. The coefficients are very similar to, and in fact slightly larger than, those presented in Table 3.

The same analysis could also be performed using alternative patent data altogether. While the data used above draws on annual reports compiled by the U.S. Patent Office, others have collected data on each patent individually. I also repeat the baseline estimates using HistPat data (Petralia et al., 2016b).[34] The HistPat data is collected from Google Patents, which were digitized in the 1980s and thus tend to be lower quality images. See Andrews (2017) for an in-depth discussion of the strengths and weaknesses of different historical patent datasets. The results using the HistPat data, presented in Column 3, are very 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.

[33] More precisely, Stata's reclink command is used, which performs a bigram string comparator that returns 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.

[34] Petralia et al. (2016a) describe the construction of the HistPat dataset in detail.

	Exact-Matched	Fuzzy-Matched	HistPat
Treat.County * HighQual. * PostTreatment: (% Change)	0.3274***	0.3848**	0.2789**
	(0.1231)	(0.1670)	(0.1173)
(# Change)	1.7505***	2.0573**	1.4911**
	(0.6583)	(0.8928)	(0.6270)
PostTreatment: (% Change)	0.0277	-0.0171	0.0860
	(0.0693)	(0.0778)	(0.0803)
(# Change)	0.1481	-0.0916	0.4600
	(0.3703)	(0.4161)	(0.4296)
County Fixed Effects	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes
Cnty-Year Obs.	34,194	34,194	34,160
# Counties	197	197	196
# Experiments	64	64	64
Adj. R-Sqr.	0.3428	0.2805	0.4174

Table 16: Regression results using different experiment dates. 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 $log(Patents + 1)$. The number of additional patents is calculated by multiplying the estimated percentage change in patenting by the average number of patents per county in the year 1900. The number in curly brackes beneath the number of additional patents is the estimated percentage change in patenting 100. These results are for high quality experiments only. For each independent variable, the first row represents that percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

B.6 Patenting at Other Geographic Levels

It may be the case that the county is the wrong geographical unit at which to measure the effects of a new college. Section 5.2 shows that the effects of colleges may spill over into neighboring counties. In Table 17, I examine the effect at an even smaller geographical unit: the town. When filing a patent application, inventors list their town of residence. There are numerous reasons to be skeptical of town-level patenting data. First, inventors may list neighborhoods instead of the larger city (for instance, listing the town of invention as Manhattan, or even the Upper East Side, rather than New York City; this can be checked for in larger cities where neighborhood names are well known, but this is infeasible for every city in the country). Second, there is greater ambiguity about town borders than county borders. Third, town borders or names are much more likely to change over time. Fifth, many areas are not within any incorporated town boundary; it is unclear how inventors in these areas record their town of residence. Sixth, many states have multiple towns of the same name; Section 2.2 discusses how this is handled at the county level, but at the town level this leads to an over-count in the number of patents. Seventh, for some experiments, it is possible to identify counterfactual counties but not towns; the Historical Appendix lists the treatment and control towns and counties when they are known.

Despite all these objections, the results in Table 17 are qualitatively similar to the county-level results presented in column 1 of Table 3, although smaller in magnitude. When using patents matched to towns, I find that establishing a new college increasing patenting by about 13% in the treatment towns relative to the control towns.

	log(Patents + 1)	log(Patents + 1)
Treat.Town * PostTreatment: (% Change)	0.1218**	0.1129
	(0.0547)	(0.1198)
(# Change)	0.6512**	0.6039
	(0.2925)	(0.6404)
PostTreatment Dummy: (% Change)	0.0638	0.0714
	(0.0604)	(0.0718)
(# Change)	0.3409	0.3817
	(0.3232)	(0.3841)
Treat.Town * Trend * PostTreatment: (% Change)		0.0001
		(0.0011)
(# Change)		0.0004
		(0.0061)
Trend * PostTreatment: (% Change)		0.0009
		(0.0016)
(# Change)		0.0046
		(0.0084)
Trend: (% Change)		-0.0013
		(0.0013)
(# Change)		-0.0069
		(0.0068)
County Fixed Effects	Yes	Yes
Year Effects	Yes	Yes
Cnty-Year Obs.	27,171	27,171
# Counties	192	192
# Experiments	64	64
Adj. R-Sqr.	0.2488	0.2490

Table 17: Regression results at the town level. Column 1 estimates the level shift in patenting in treatment town relative to control towns after establishment of a new college. Column 2 includes a linear trend term as well as coefficients estimating if the trend changes after the establishment of a new college and if the trend changes differently in treatment and control towns after establishment of the new college. The dependent variable for both columns is $log(Patents+1)$. The number of additional patents is calculated by multiplying the estimated percentage change in patenting by the average number of patents per town in the year 1900. The number in curly brackes beneath the number of additional patents is the estimated percentage change in patenting 100. These results are for high quality experiments only. For each independent variable, the first row represents that percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

B.7 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 18, I show how patenting differs between practical and classical colleges, using an alternative classification of practical and classical than described in Section 4.4. 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 larger when the alternative definitions are used than in the baseline results presented in Table 18, with practical colleges having 39% more patents per year compared to 24% for classical colleges. It is unclear why the inclusion of military academies changes the estimated effect of practical colleges by so much, especially since the graduates of those schools are commissioned and sent away from the college county, and so are unlikely to remain in the same county and patent.

In Column 2, I verify that counties with a land grant college do not increase their share of agricultural patents relative to control counties even when an alternative definition of agricultural patent is used. Here I use a slightly broader definition of agricultural patent: in addition to the patent classes listed in Section 4.4, I consider a patent to be agricultural if it belongs to one of the following 3-digit USPTO patent classes: 59 "Chain, staple, and horseshoe making"; 111 "Planting"; 131 "Tobacco"; 171 "Unearthing plants or buried objects"; 185 "Motors: spring, weight, or animal powered"; 231 "Whips and whip apparatus"; 256 "Fences"; 260 "Chemistry of carbon compounds"; 417 "Pumps"; 425 "Plastic article or earthenware shaping or treating: apparatus"; 426 "Food or edible material: processes, compositions, and products"; 435 "Chemistry: molecular biology and microbiology"; 452 "Butchering"; 800 "Multicellular living organisms and unmodified parts thereof and related processes"; and PLT "Plants". As in Table 18, establishing a land grant college leads to a decrease in the share of agricultural patents.

Column 3 shows the increase in patenting in treatment relative to control counties after college establishment with separate interactions for each type of college. Row 1 presents estimates land grant colleges, which make up most of the colleges in the sample. Row 2 presents results 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.

These results must be interpreted with extreme caution. As there are often only a few colleges of a particular type (for instance, HBCUs and other private colleges have an especially small number of colleges the sample), there is insufficient power to draw strong conclusions. Nevertheless, the results are suggestive, so I discuss them below.

The results for land grant colleges, technical colleges, and normal schools seem to conform to the intuition that college type influences the amount of patenting. Normal schools increase patenting very little, only 2% more than the control counties. Technical colleges, on the other hand, have a remarkable 97% more patents per year than their control counties. Land grant colleges are in the middle, with 22% more patents per year than their control counties. Surprisingly, however, other public universities had 33% more patents per year than their control counties, a larger estimated coefficient than land grant colleges which purportedly had a more technical focus.

HBCUs, shown in row 4, are an interesting case. The curriculum at HBCUs varies greatly across colleges, with some resembling normal schools and others providing a practical education in agriculture and machinery. The comparison to other college types is even more difficult because HBCUs tended to be established in areas with large African American populations, which tended to be poor and have a very low level of baseline patenting. HBCU counties have about 13% more patents per year than their control counties.

Military academies, shown in row 5, also form an interesting comparison group. Military academies focus on technical skills such as engineering and seamanship, but graduates from the academies are commissioned and dispersed to other locations; they cannot remain in the college county, eliminating one channel by which colleges can affect local invention. Moreover, because these colleges are focused on producing soldiers, sailors, and airmen, the technical skills taught may be less directed towards commercial applications and therefore less likely to obtain a patent. Nevertheless, counties with military academies produce far more additional patents per year than do land grant colleges. Military academy counties have 125% more patents per year than their control counties.

The one large outlier in these different college types are other private colleges, which remarkably saw 11% fewer patents per year than their control counties. As Table 1 shows, however, there are very few "Other Private" college experiment in the sample, so this result may reflect an idiosyncrasy associated with that particular experiment.

While there are certainly differences between the coefficients between different types of colleges, the limited number of each type of college makes drawing inferences from these estimates difficult. In particular, the coefficient for the effect of normal schools on patenting is statistically indistinguishable from the coefficient on land grant colleges, technical schools, and other public colleges. Even military academies, with clearly the largest coefficient, are statistically indistinguishable from technical schools and other public colleges. The only college type that is typically statistically different from all the rest is the private colleges (although even this is statistically identical to normal schools at conventional levels).

In all, these results make it difficult to paint a general picture about the effects of college type on patenting. Technical schools, which are expected to teach practical skills that easily translate into more invention, do indeed show a dramatic increase in patenting relative to control counties. But public universities that have a more classical focus ended up leading to more invention than did technically-focused land grant colleges. As expected, normal schools and private colleges saw the smallest increase in patents relative to their control counties. But while the results for normal schools and private colleges are consistent with the idea that the skills developed in college matter for patenting, it is unclear whether this is due to the fact that these schools do not teach skills conducive to patenting or whether these schools are simply very small.[35] I explore the effects of new colleges in driving county population directly in Section 4.5.

[35] A different way to think about this issue is to draw an analogy to the experimental literature. Because large state universities, land grant colleges, and state-sponsored technical schools tend to be larger colleges, the variance between the "treatment" administered to the college counties and the "lack of treatment" given to the controls is larger than when the treatment involves a smaller college. See, for example, List et al. (2011).

	Alt. Definitions Practical vs. Classical Colleges	Frac. Alt. Definition Ag. Patents	High Quality Experiments	All Quality Experiments
Practical College Interaction: (% Change)	0.3941**			
	(0.1788)			
(# Change)	2.1073**			
	(0.9563)			
Classical College Interaction: (% Change)	0.2354			
	(0.1694)			
(# Change)	1.2588			
	(0.9057)			
Land Grant Interaction: (% Change)		-2.0362	0.2026	0.1955
		(1.7582)	(0.1326)	(0.1360)
(# Change)		-0.0369	1.0835	1.0456
		(0.0318)	(0.7091)	(0.7270)
Non-Land Grant Interaction: (% Change)		-0.1617		
		(0.7601)		
(# Change)		-0.0029		
		(0.0138)		
Technical School Interaction: (% Change)			0.5875	0.4028
			(0.4346)	(0.3101)
(# Change)			3.1415	2.1536
			(2.3238)	(1.6578)
Normal School Interaction: (% Change)			0.0384	0.0404
			(0.2580)	(0.2812)
(# Change)			0.2056	0.2160
			(1.3793)	(1.5038)
HBCU Interaction: (% Change)			0.1272	-0.0164
			(0.3265)	(0.1804)
(# Change)			0.6801	-0.0875
			(1.7457)	(0.9648)
Military Academy Interaction: (% Change)			1.3104	-0.1691
			(1.8718)	(0.5050)
(# Change)			7.0068	-0.9041
			(10.0086)	(2.7003)
Public Other Interaction: (% Change)			0.3367*	0.2896
			(0.1898)	(0.4057)
(# Change)			1.8002*	1.5485
			(1.0150)	(2.1692)
Private Other Interaction: (% Change)			-0.1092	1.4456**
			(0.0766)	(0.7182)
(# Change)			-0.5841	7.7295**
			(0.4098)	(3.8402)
County Fixed Effects	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Cnty-Year Obs.	34,194	10,345	34,194	91,502
# Counties	197	195	197	531
# Experiments	64	64	64	126
Adj. R-Sqr.	0.3429	0.0363	0.3552	0.2962

Table 18: 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 and 4, the coefficient is the percentage increase in patenting caused by the treatment 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. The results in Columns 1-3 are for high quality experiments only. Column 4 uses both the high and low quality experiments. For each independent variable, the first row represents that percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

B.8 Population and Patenting, Controlled Direct Effects

The estimated effect of establishing a college on patenting after controlling for population, documented in Table 8 in Section 4.5, is likely to be biased. Any post-treatment confounding variable that affects both population and patenting will bias the estimated treatment effect, which Acharya et al. (2016) call "intermediate variable bias." Such bias is almost certain to occur here. Obviously, colleges vary greatly in terms of the number of faculty, number of students, and resources spent on research. Colleges with more faculty and students or more to spend on research likely attract more people to a college county and also produces more patents. Ignoring this will bias downward the estimated treatment effect.[36]

I correct for this potential source of bias in Table 19. In Column 1 I repeat the estimates from Column 6 of Table 8 in which I simply control for $log(TotalPop)$ in the baseline regression. I then obtain an estimate of the controlled direct effects of colleges on patenting, re-estimating the baseline regression while controlling for population using the sequential g-estimation proposed by Acharya et al. (2016), which in turn builds on techniques from the biology literature (Joffe & Greene (2009), Vansteelandt (2009)). The sequential g-estimator is a two-stage estimator. The first stage estimates the same model as in Column 1 and saves the estimated coefficient for $log(TotalPop)$ conditional on the treatment and fixed effects. Then, a "demediated" outcome variable $\log(Num.\tilde{P}atents + 1)$ is constructed by subtracting the fitted values for $log(TotalPop)$ from $\log(Num.Patents + 1)$. In the second stage, the demediated outcome is regressed on the treatment and fixed effects. The second stage estimate of the effect of establishing a new college is consistent because it does not condition on $log(TotalPop)$ or any intermediate variables.

The results are presented in Column 2.[37] The first stage results are already calculated in Column 1, so I do not reproduce them for Column 2. As expected, the coefficient on the treatment effect is larger than in Column 1, although the differences are extremely small. The bootstrapped standard errors are slightly smaller than in Column 1 as well, although still not quite statistically significant at conventional levels.

[36]Even if complete data on, for instance, research spending could be obtained over this time period, simply controlling for research spending still leads to biased results because it removes the causal channel through which colleges affect research spending which in turn affects patenting independently of population.

[37]Bellemare (2016) points out a small typo in Acharya et al.'s (2016) Stata code.

I repeat this exercise in Columns 3 and 4, controlling for total population and squared total population instead of $log(TotalPop)$ as in Column 7 of Table 8. Column 4 presents standard estimates simply controlling for these variables in the regression. The coefficient on the squared term is negative but extremely close to zero; as in Figure 6 I find no evidence that patenting exhibits increasing returns in population. Column 5 repeats the g-estimation procedure as in Column 8. In both columns, establishing a new college leads to roughly 19% more patents per year even after controlling for population and squared population, and the effect is statistically significant (at the 10% level in Column 4 and the 5% level in Column 5). Thus in this specification, controlling for population explains roughly 44% of the increase in patenting.

	OLS	Sequential-g	OLS	Sequential-g
Treat.County * HighQual. * PostTreatment: (% Change)	0.1199	0.1200	0.1785*	0.1788**
	(0.0928)	(0.0801)	(0.1060)	(0.0861)
(# Change)	0.6410	0.6416	0.9544*	0.9563**
	(0.4961)	(0.4284)	(0.5668)	(0.4606)
PostTreatment: (% Change)	-0.0687	-0.0687	-0.0081	-0.0081
	(0.0624)	(0.0537)	(0.0718)	(0.0601)
(# Change)	-0.3674	-0.3728	-0.0431	-0.0491
	(0.3334)	(0.2862)	(0.3841)	(0.3201)
log(Total Pop.): (% Change)	0.2740***			
	(0.0340)			
(# Change)	1.4649***			
	(0.1816)			
Total Pop.: (% Change)			0.0112***	
			(0.0042)	
(# Change)			0.0597***	
			(0.0223)	
Total. Pop Squared: (% Change)			-0.0000***	
			(0.0000)	
(# Change)			-0.0000***	
			(0.0000)	
County Fixed Effects	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Cnty-Year Obs.	3,390		3,390	
# Counties	204	204	204	204
# Experiments	66	66	66	66
Adj. R-Sqr.	0.3489	0.3527	0.3338	0.3527

Table 19: Results of population on patenting with different estimation methods. Data are from census years only. The dependent variable in all columns is $log(Num.Patents + 1)$. Column 1 uses OLS and includes an additional control for $log(TotalPop)$. Column 3 uses OLS and includes additional controls for $TotalPop$ and $(TotalPop)^2$. Columns 2 and 4 use the results from Columns 1 and 3, respectively, as a first stage and presents estimates of the controlled direct effect using sequential-g estimation as described in the text. For each independent variable, the first row represents that percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

B.9 Colleges versus Consolation Prizes

In this section, I turn my attention to a subset of the college site selection experiments in which the losing counties, while they did not receive a college, did experience many of the demographic changes described in Section 4.5. In some cases, losing finalist counties were not truly "losers": while they may not have obtained a state university, they did obtain some other state 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 many state institutions were allocated at the same time, including the state college, the state prison, the state hospital, and 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. For 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.[38]

Table 20 shows results that explicitly consider the consolation prize counties. In column 1, I compare patenting in the college counties to consolation prize counties. The coefficient is statistically insignificant 20%, smaller than the baseline estimate of 33%. This suggests that college counties do not increase their patenting much faster than counties that received prisons, hospitals, or insane asylums. Consistent with this, column 2 shows that when consolation prize counties are excluded from the sample, a new college increases patenting by about 35%, slightly larger than the 33% baseline estimate. Likewise, in column 3 the consolation prize counties are classified as treatment counties, and the estimated coefficient is positive and statistically significant: in this case, the "treatment" counties patent 30% more than control counties. This column can be thought of as presenting results from an experiment in which the treatment is receiving any state institution.

To make sense of these results, it is important to remember the context in which these experiments took place. Localities actively lobbied for prisons and insane asylums. Instead

[38]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).

of repelling highly mobile skilled workers, as these institutions might today, the consolation prizes gave small towns an identity and attracted more people to the area. Strikingly, the estimated coefficient in Column 1 is very similar to the coefficients in Columns 5 and 6 of Table 8, suggesting that the consolation prize counties appear so similar to the treatment counties because they have similar populations even after the college is established. Figure 11 plots the logged county population by decade before and after the establishment of a new college for both the college counties and the consolation prize counties. Both grew nearly identically both before and after the establishment of a new college. While it is possible, although not especially likely, that places like prisons and insane asylums increased invention directly through their activities, this figure is suggestive that the increase is largely due to a broader pattern of urbanization at work in these places. The fact that the pattern is so similar between colleges and the other consolation prize sites presents further suggestive evidence that broad urbanization is a major driver of patenting in the college counties.[39]

B.10 Demographic and Economic Effects of Colleges

Colleges may plausibly effect the counties that receive them in many ways beyond simply changing patenting and population. In Table 21, I investigate the effects of establishing a new college on other demographic and economic outcome variables. Some of these have been identified by other authors as outcomes that colleges are likely to affect (for instance, Liu (2015)), and all could serve as channels through which colleges affect invention. As each of the outcome variables discussed below is collected either directly from the US censuses or from the NHGIS, many variables of which are also compiled from the population census and nearly all of which are collected at decadal frequencies, I again present results using decadal rather than yearly data.

[39]The treatment and consolation prize counties continue to look very similar even after controlling for log county population. In these regressions, the coefficient on population is very large in magnitude and reduces the coefficient on $Treatment_{ij} * PostTreatment_{jt}$ to almost zero, qualitatively similar to the results controlling for population in Table 21. In the case of consolation prize experiments, the estimated interaction return reduces from a 21% increase to only a 6% increase after controlling for population. This is suggestive of the fact that the experiments in which several institutions were allocated at the same time occurred in particularly undeveloped and unpopulated areas, and that subsequent population growth explains most of the observed differences in these cases. Understanding how the effect of establishing a new college varies with the initial level of development of an area is an important area for future research.

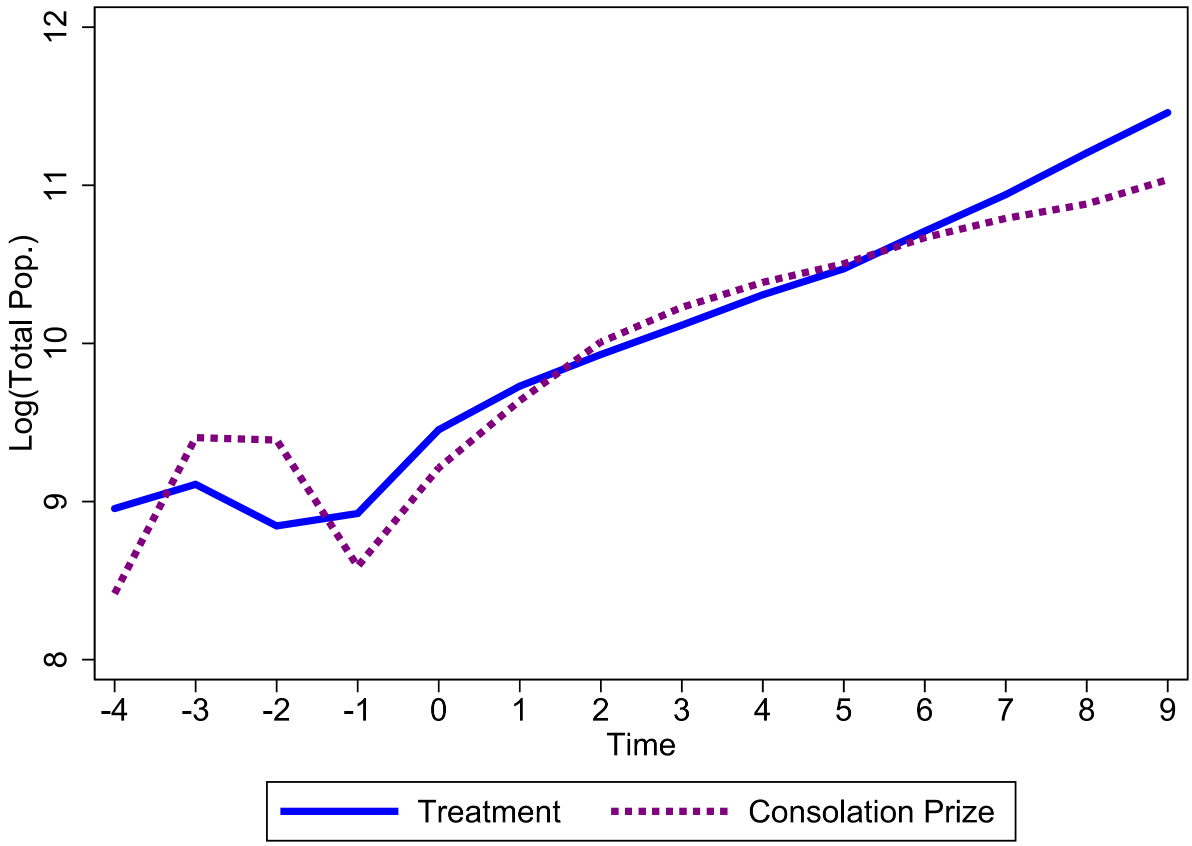


Figure 11: Logged total population by decade in the treatment (college) and consolation prize counties. The x-axis shows the number of decades since the college experiment. The ten year period immediately following the establishment of a new college is normalized to decade 0. Everything left of decade 0 shows pre-treatment population; everything to the right shows post-treatment population. The y-axis shows $log(TotalPopulation)$. The treatment counties are represented by the blue solid line. The consolation prize control counties are represented by the purple dashed line. Data are for high quality experiments only.

	Consolation Prize	No Consolation Prize	Cons. Prize as Treated
Treat.County * HighQual. * PostTreatment: (% Change)	0.2034	0.3488***	0.2913**
	(0.2005)	(0.1278)	(0.1151)
(# Change)	1.0877	1.8652***	1.5577**
	(1.0720)	(0.6834)	(0.6157)
PostTreatment: (% Change)	0.0274	0.0056	0.0159
	(0.1777)	(0.0709)	(0.0693)
(# Change)	0.1465	0.0302	0.0851
	(0.9502)	(0.3791)	(0.3705)
County Fixed Effects	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes
Cnty-Year Obs.	14,525	30,869	34,194
# Counties	83	178	197
# Experiments	64	64	64
Adj. R-Sqr.	0.3887	0.3393	0.3423

Table 20: Regression results under various assumptions about consolation prize counties. Column 1 compares treatment counties to only control counties that receive a consolation prize. Column 2 excludes all counties that receives a consolation prize and compares treatment counties to control counties that do not receive a consolation prize. Column 3 classifies control counties that receive a consolation prize as treatment counties and re-estimates the baseline regression. The dependent variable in all columns is $log(Patents+1)$. The coefficient is the percentage increase in patenting caused by the treatment. These results are for high quality experiments only. For each independent variable, the first row represents that percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Column 1 checks whether establishing a new college affects the average age in the treatment county by using mean age as the dependent variable in Equation (1); the result is insignificant and small in magnitude. Column 2 finds that treatment counties have about 1% more interstate migrant residents than the control counties after establishing the college, although this result is also not statistically significant. The dependent variable for Column 3 is the measure of residential segregation constructed by Logan & Parman (2017). This measure captures uses census records of the race of each household as well as its neighbors to calculate how segregated a county is relative to a random distribution of households ($Segregation = 0$) and complete segregation ($Segregation = 1$), and so the estimated coefficient can be interpreted as a percentage. While segregation appears to decrease after the establishment of a new college in college towns by 17%, the effect is not statistically different from zero. Column 4 uses manufacturing productivity as the dependent variable and finds an insignificant positive result.

	Mean Age	Frac. Interstate Migrants	Segregation	Manuf. Productivity
Treat.County * HighQual. * PostTreatment: (% Change)	-0.0006	0.0105	-0.1706	0.0248
	(0.0085)	(0.0552)	(0.1374)	(0.1145)
(# Change)	-0.0208	0.0067	-0.0382	1,618.8205
	(0.2842)	(0.0353)	(0.0307)	(7,463.4651)
PostTreatment: (% Change)	-0.0068	0.0708*	0.2547**	0.0069
	(0.0053)	(0.0372)	(0.1090)	(0.0714)
(# Change)	-0.2264	0.0452*	0.0570**	447.9622
	(0.1774)	(0.0237)	(0.0244)	(4,650.2522)
County Fixed Effects	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Cnty-Year Obs.	1,465	1,477	1,077	1,272
# Counties	197	197	194	196
# Experiments	64	64	64	64
Adj. R-Sqr.	0.9850	0.4635	0.3832	0.8659

Table 21: Baseline regression results using various dependent variables. Data are from census years only. The dependent variable in Column 1 is *Mean Age*. The dependent variable in Column 2 is *Frac. Interstate Migrant*. The dependent variable in Column 3 is a measure of residential segregation. The dependent variable in Column 4 *Manufacturing Productivity*. These results use high quality experiments only. For each independent variable, the first row represents that percentage change in patents 100 and the second row represents the change in the number of patents. Standard errors are clustered by county and shown in parentheses. Stars indicate statistical significance: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$