The Location of Agricultural Research and the Direction of Agricultural Innovation*

Michael Andrews†

May 13, 2019

Abstract

I analyze the importance of the local environment on the direction and subsequent diffusion of agricultural innovations. By comparing counties that are near and far from land grant colleges using a variety of distance measures, I show that proximity is more important for agricultural productivity and output than for other measures of innovation in other sectors. To shed light on how widely innovations from land grant colleges diffuse, I exploit data on the histories of new wheat varieties introduced in the U.S. before 1920 and find that only 10-17% of wheat acreage planted in varieties developed since the establishment of land grant colleges is planted in varieties developed at land grant colleges. To present direct evidence that the local environment affects the direction of innovation, I use data on publications by researchers affiliated with land grant colleges to show that, even more than a century after the land grant colleges were established, land grant research is biased towards crops that were initially most prevalent in land grant college counties, rather than those that were most prevalent in the rest of the state. Finally, I show that alumni of land grant colleges with agricultural degrees were more likely to live near their alma maters than were alumni with other majors, which I interpret as evidence that agricultural human capital is more location-specific than other forms of human capital.

*I am grateful to Lauren Cohen, Jeff Furman, Scott Stern, Nicolas Ziebarth, and seminar participants at the NBER Conference on the Economics of Research and Innovation in Agriculture for thoughtful comments. All mistakes are my own.

†National Bureau of Economic Research. 1050 Massachusetts Ave., Cambridge, MA 02138. *Email*: mandrews@nber.org.

I Introduction

When, where, and why is geography important for innovation? Two strands of literature have explored the importance of geographic proximity and innovation, largely in isolation: a first strand focuses on the importance of creative individuals co-locating for the creation of new ideas, while a second strand focuses on how local connections between individuals helps existing new ideas to spread. There is now little doubt that localized knowledge flows are important for both the creation (Jaffe, Trajtenberg, and Henderson (1993), Ganguli, Lin, and Reynolds (2019), Andrews (2019a)) and adoption of new ideas (Bandiera and Rasul (2006), Conley and Udry (2010), Beaman and Dillon (ming)).

Geography may be important for innovation for another reason as well: an individual's local environment may affect the types of innovations an individual pursues, or likewise the types of innovations an individual is willing or able to adopt. The way in which the local environment affects the direction of innovation is difficult to study. Indeed, most of the work on geography and innovation seeks to remove these common environmental effects from the analysis (Manski, 1993).[1]

It is not immediately obvious that we should expect the local environment to have a measurable effect on the direction of invention, beyond exposing potential inventors or adopters to ideas from their neighbors. Glaeser and Hausman (2019), for example, emphasize the fact that for many innovations, the location in which the innovation occurs may be quite different from the location where the innovation is actually adopted. They go on to argue that this can be extremely useful as a policy tool: innovation can take place where creative individuals cluster, taking advantage of knowledge spillovers and other advantages of density, and then be applied in less economically advantaged areas. The U.S. land grant system was designed, at least implicitly, with a similar idea. The presence of a college centralizes innovation by causing high human capital individuals to co-locate, while the extension service brings the innovations from a potentially distant college to the farmers who are the ultimate end users

[1] For one example, Foster and Rosenzweig (2010) write: "Does the correlation [of high-yield variety (HYV) adoption] with neighbors' prior adoption behavior reflect learning externalities, or is it simply a reflection of common unobservables that make HYV returns higher for the farmer and his neighbors?" They then go on to discuss numerous attempts to identify learning externalities, but spend no time discussing the importance of these common unobservables for the rate and direction of innovation.

of these innovations.

But agriculture as a sector poses a distinct challenge when trying to separate the location of creation and adoption of new ideas. This is because in agriculture the local environment is likely to have an outsized effect on what innovations can be adapted.[2] Techniques for growing a crop in a town in Iowa, for example, are likely to be useful in the next town over but are much less likely to be adopted in, say, California or Florida. Likewise, a technique useful in northern California may be completely irrelevant in southern California. Therefore, if the local environment influences the direction of innovation, then clustering innovators at a land grant college may not produce innovations desired by the broad but geographically distant end user farmers. The influence of the local environment on the direction of innovation may in fact be particularly pernicious in the case of land grant colleges and other public institutions. Private researchers may have an incentive to ignore the local environment and focus on innovations that will be adopted by the largest possible market, while these incentives are likely less pronounced in public research.

The goal of this paper is to begin to shed light on the importance of the local environment on the direction of innovation. I begin by presenting suggestive evidence that proximity to the source of an innovation is especially important in the agricultural sector. I present two pieces of evidence. First, I show that after the establishment of a new college, areas that are "close" to the county that received a college improve on a number of agricultural metrics relative to areas that are "far away," while metrics unrelated to agriculture, such as manufacturing or urbanization, show less of a decline with distance. I consider several measures of distance, including geodesic and latitudinal distance. Moreover this effect is only present after the establishment of land grant colleges, suggesting that the effect is related to agricultural research. Second, I use a historical USDA report that contains detailed data on the geographic origin of U.S. wheat varieties. In spite of the extensive research conducted at land grant colleges, a majority of the new wheat varieties used by farmers over this period of time were developed elsewhere, suggesting that land grant research has a limited reach.

[2]Others have pointed this out as well. For instance, Alston and Pardey (1996, p. 237) write: "This feature (site-specific R&D) is not so prevalent in the development and diffusion of nonbiological technologies; it tends to be peculiar to agriculture." But I am unaware of any quantitative work attempting to compare the importance of the local environment on the creation and diffusion of innovations across sectors.

It is difficult to draw many conclusions about the reasons why proximity may be so important in agriculture without having more detailed and disaggregated data about the types of innovations coming out of land grant colleges and the types of innovations demanded by farmers. The wheat varieties origin data is a first step towards this, but it is limited in many dimensions, which I discuss below. I next compile direct evidence that land grant colleges are producing innovations to cater to their local environment, rather than serving the needs of the much broader market of farmers in their state. To do this, I isolate the subset of land grant colleges that are located in counties that focus on different crops than the rest of the state. I show that, even more than a century after the establishment of the land grant college, researchers at these colleges are much more likely to publish papers about the local crop than they are the crop that is more important for the rest of the state. As a second piece of evidence, I compile data on the alumni of land grant colleges. I show that alumni with agricultural degrees are much more likely to be living geographically close to their alma maters than are alumni with other degrees, which I interpret as evidence that the human capital agricultural students acquire is more location-specific than the human capital in other fields.

These findings are valuable for at least two reasons. First, by using this data in a novel and geographically disaggregated way, I am able to examine the success or the failure of the U.S. land grant college system in the diffusion of agricultural innovation through a new lens. In describing the success of the "grand mission" of the land grant colleges, Wright (2012, p. 1725) concludes that "[i]n the United States, unusually wide geographic dispersion of farmers across diverse environments generated advantages in local research that would accrue in large part to local farmers, and federal subsidies encouraged state research that generated spillovers." While Wright concludes that the land grant mission has been a success, the evidence presented in this paper raises questions of whether the spillovers from land grant colleges were wide enough and if more could have been done to ensure they spilled over to an even larger area.

Second, a number of studies in the economics and innovation literatures specifically study the agricultural sector. As early as the 1950s, Zvi Griliches was studying hybrid corn to draw conclusions about the diffusion of new technologies (Griliches (1957), Myers and Rhode

(2019)) and the returns to public and private R&D (Griliches, 1958). The tradition of examining the agricultural sector to study technological diffusion continues, with Feder, Just, and Zilberman (1985) and Foster and Rosenzweig (2010) providing surveys. More recently, Kantor and Whalley (2018) document that the productivity benefits of land grant research are highly geographically concentrated, interpreting this as the result of knowledge spillovers. Moser and Rhode (2012) and Moscona (2019) examine agricultural innovation to draw conclusions about the effects of intellectual property rights. In light of all these studies, it is important to understand how "special" agriculture is relative to other sectors, and how the conclusions about knowledge spillovers and diffusion in agriculture do or do not apply in other contexts.

This paper also contributes to three other literatures. First, I build on the literature on the determinants of the direction of innovation. While previous researchers have shown that the legal regime (Moser, 2005), changes in factor or input prices (Hanlon (2015), Juhász (2018)), or major events like wars (Khan, 2015) can influence the types of innovations produced, I show that local environmental conditions can have a similar effect. Second, I contribute to the literature on the importance of user input on innovation (von Hippel (1988), Chatterji and Fabrizio (2012)) by showing how geographic distance between users and innovators can harm adoption when innovators are sensitive to the local environment. Finally, by highlighting how public researchers conducting applied research may be particularly influenced by local conditions, I add a small wrinkle to the large literature on the proper role of universities in basic versus applied research (Nelson (1959), Stokes (1997), Li, Azoulay, and Sampat (2017)).

This paper is organized as follows. Section II presents suggestive evidence that proximity is more important in agriculture than in other fields, first in a differences-in-differences framework studying the establishment of new colleges and second using data on the origins of wheat varieties. Section III uses data on publications from researchers affiliated with land grant colleges to show that land grant college research is focused on crops of local interest rather than crops of more general interest. Section IV uses data on alumni of land grant colleges to argue that agricultural human capital is more location-specific than other types of human capital. Section V is currently blank, but will eventually conclude the paper.

II How Important Is Proximity for the Diffusion of Agricultural Innovations?

I begin by documenting that areas proximate to the location of agricultural research benefit more than geographically distant areas in terms of agricultural output and productivity. Others have documented the importance of proximity to research for agricultural productivity before, including most recently Kantor and Whalley (2018). I contribute to this literature by showing that proximity is *especially* important in agriculture: when examining the effect in other sectors or for general innovation or population, proximity appears far less important. While these results show in aggregate statistics that proximity is valuable, in Section II.B I present evidence from disaggregated wheat varieties to show that adoption of specific land grant-originating innovations is limited.

II.A The Diffusion of Innovation Following the Establishment of New Colleges

II.A.1 Data

I use data on county-level agriculture from the U.S. agricultural censuses covering the years 1840-2012, compiled by Haines, Fishback, and Rhode (2018). These agricultural censuses were originally conducted on the census years, but more recently are conducted more frequently. The questions asked in the censuses vary substantially from year to year, making constructing long-run time series of some variable difficult. However, the agricultural censuses do provide fairly consistent measures of the number of farms, various measures of farm acreage, the market value of broad output categories such as crops and livestock/livestock products sold, and either output (typically in bushels) or input (in acreage planted) measures for several key cereal crops; in some years, in particular starting around 1930, these data become much more detailed and include data on a much wider variety of crops and other agricultural products as well as detailed measures of other inputs such as fertilizer, irrigation, and hired labor. I supplement the agricultural census data with data on manufacturing and county demographics from the National Historic Geographic Information System (Manson, Schroeder, Riper, and Ruggles, 2018). I use data on county-level patenting as in Andrews (2019c).

As the agricultural census data is at the county level, I calculate the distance from each county centroid to the land grant college located in the same state. I measure distance in two main ways. First, and most obviously, I calculate the geodesic distance between county centroids and the town of the land grant college.

Geodesic distance may not be the most meaningful proxy for how difficult it is for an agricultural innovation to diffuse, as nearby areas may have markedly different environments. Environments are particularly likely to vary with latitude. Indeed, Diamond (1997) argues that the ease of adapting crops and livestock along east-west axes relative to north-south axes played a crucial role in world history by shaping where agriculture and the domestication of animals was able to flourish. Steckel (1983) shows that "latitude-specific investments in seeds...induced migrants to move along east-west lines" in America during the period of westward expansion, citing in particular the photoperiodic sensitivity of corn varieties.[3] Steckel (1983, p. 20-24) uses data from the University of Illinois Agricultural Experiment Station at Champaign (University of Illinois, 1894) to show that yields of seeds from the experiment station declined rapidly when with distance north and south of Champaign, but yields varied little with east-west distance. Olmstead and Rhode (2008, p. 70) cite similar results from the Arkansas Agricultural Experiment Station (Newman, 1899). I therefore also calculate the latitudinal distance, calculated as the difference in latitude between each county centroid and the land grant college in that state. In Appendix A I present results using several alternative distance measures as well.

II.A.2 Baseline Results

I use this data to estimate whether increasing distance from the site of a land grant college decreases agricultural productivity or output more after a college is build; this is essentially

[3]See Thomas and Vince-Prue (1997) for the importance of photoperiodism. Doyle, Sung, and Amasino (2002) is a more recent review of the factors affecting the timing of plant flowering, mentioning photoperiodism as well as other environmental factors such as vernalization (acceleration of flowering resulting from prolonged exposure near-freezing temperatures).

a differences-in-differences estimator with a continuous treatment. More formally, I estimate

$$\begin{aligned} Dep.Var._{it} = &\beta_1 DistanceToCounty_i \times PostCollege_{it} + \beta_2 PostCollege_{it} + \\ &+ County_i + College_i + Year_t + \epsilon_{it}, \end{aligned} \tag{1}$$

where $DistanceToCounty_i$ is the distance between the county centroid for county i and the land grant college in the same state. I omit the college's county in all analysis. $PostCollege_{it}$ is a dummy variable equal to one in the years after the land grant college in i's state is established and is zero otherwise. $County_i$, $College_i$, and $Year_t$ are county, college, and year fixed effects, respectively.

I estimate this regression for a battery of dependent variables: county patents, the fraction of the county population living in urban areas, manufacturing productivity, agricultural productivity, crop productivity, agricultural output, the value of crops sold, the value of livestock and livestock products sold, and improved farm acreage.[4] If the local environment is particular influential in agriculture, then the estimated coefficient on $DistanceToCounty_i \times PostCollege_{it}$ should be negative and large in magnitude for agricultural dependent variables, and closer to zero in magnitude for dependent variables that are not related to agriculture.

This is precisely what I find. I plot coefficients of $DistanceToCounty_i \times PostCollege_{it}$ for each outcome variable graphically in Panel (a) of Figure 1 when using geodesic distance and Panel (a) of Figure 2 when using latitudinal distance. With either distance measure, the agricultural output measures decline sharply with distance, although the large variance in agricultural output means that these declines are only sometimes statistically significant at the 5% level. The agricultural and crop productivity measures display smaller elasticities with respect to distance, but are statistically significant in the case of latitudinal distance. The improved farm acreage measures provides one hint why the magnitude of the agricul-

[4]For comparability and ease of interpretation, I log all variables. Manufacturing productivity is calculated as the value of manufacturing output divided by the number of agricultural workers. The agricultural output and crop output productivity measures are land productivity measures, calculated by dividing the value of output by the number of acres (either total acreage for the agricultural productivity measure or cropland acreage for the crop productivity measure). Unfortunately, data on agricultural labor is not available for most early censuses and, when it is available, is difficult to interpret; there is typically no count of workers from the farm family or proxy for hours worked, and it is unclear how to value slave labor or indentured servitude prior to the 1870 census. For these reasons, I am unable to construct time series of agricultural labor productivity that would be more comparable to the manufacturing productivity measure.

tural productivity estimates are small: counties close to the land grant colleges typically see increases in improved acreage, likely because agricultural innovations make marginal land profitable to farm which depresses average productivity but increases total output. In contrast to the agricultural variables, patenting, urbanization, and manufacturing productivity change very little with distance after establishing the land grant college.

As another basis for comparison, I repeat the analysis using the sample of non-land grant colleges identified by Andrews (2019c). The non-agricultural variables appear similar whether examining the land grant or the non-land grant colleges (although the larger number of non-land grant colleges means that these are estimated much more precisely). In the case of the agricultural variables, however, as distance to the non-land grant counties increase, the non-land grant counties experience very small changes which are typically positive. The positive sign could possibly be due to non-land grant colleges attracting resources in nearby counties away from agriculture, while counties near the non-land grant colleges benefit from positive spillovers.

II.A.3 Causal Effects of Colleges

One concern is that land grant colleges are not located at random. They may be located near the best agricultural land in their state, for instance. If the quality of agricultural land varies nearly continuously over distance, then this could explain the decrease in agricultural productivity and output as distance to the land grant college increases.

To address this concern, I restrict attention to the cases identified in Andrews (2019c) in which the final location for a new college was selected essentially at random from among a set of runner-up sites. Each of these finalist site locations was considered nearly equally suitable to become the site of the new college by state officials, boards of trustees, or site selection committees, and so were likely very similar in terms of agricultural quality. In one example, the location of North Dakota State University was literally selected at random when contesting towns could not agree among themselves. Andrews (2019c), and especially Andrews (2019b), provide much more detail on these college selection decisions, including providing evidence that the runner-up sites are similar to one another along observable dimensions.

Rather than compare *all* colleges in the land grant college's state, I now only consider the runner-up counties. The idea here is that the set of runner-up counties are similar in terms of agricultural suitability, growth potential, and other unobservable factors. But once the college site is selected, some runner-up counties will be geographically closer to the final college site for exogenous reasons.

I repeat the exercise on the battery of non-agricultural and agricultural controls using this new, restricted sample of counties. Results are plotted in Panel (b) of Figure 1 and Panel (b) of Figure 2 for geodesic and latitudinal distance, respectively. All results are remarkably similar to those using the full sample.

From this I conclude that proximity to an agricultural college causes an increase in agricultural productivity and output. This is consistent with recent findings by Kantor and Whalley (2018). I expand on this result by showing that there is no such distance gradient for non-agricultural outcomes, suggesting that care must be taken in extrapolating these conclusions to all sectors. Additionally, the importance of proximity is restricted to the institutions that actually conduct agricultural research and bestow agricultural human capital.

II.B Land Grant Colleges and the Adoption of New Wheat Varieties

The findings in the previous section suggest that the benefits of land grant colleges were quite geographically limited, and raise the question of whether the entire country could plausibly have benefited from their activities. This is impossible to answer definitively without direct data on how widely the innovations form land grant colleges were used. In this section, I focus on one specific type of innovation for which detailed data is available: new wheat varieties. Wright (2012, p. 1719) states that by 1920, "more than three-quarters of wheat acreage in the United States was planted with varieties that had been unavailable when the Morrill Act was passed." Such a statistic, however, does not show that land grant colleges themselves directly contributed to the creation of these new widely-adopted wheat varieties. I explore this issue here.

II.B.1 Data

To obtain information on the location in which various wheat varieties were first developed, I consult a USDA technical report (Clark, Martin, and Ball, 1922) that attempts to classify every variety of wheat grown in the United States as of 1922. Crucially, and exceedingly rare among agricultural studies, the authors also provide histories of each wheat variety, including how, when, and where each variety was developed and/or introduced to the United State. This allows me to investigate the extent to which land grant colleges directly contributed to innovation in the wheat sector. Clark et al. (1922) also conduct an extensive series of surveys from 1917-1919 to determine the geographic distribution of the cultivation of these wheat varieties. I describe the construction of the varietal histories and cultivation surveys in the Appendix B.

II.B.2 Results

First, I examine a simpler question: did land grant colleges cause an increase in the level of new wheat varieties developed in college counties relative to the runner-up locations that did not receive the college. This analysis is independently interesting, since most research studying the effects of colleges on innovation uses other measures, such as patents or publications, to measure innovation. I am unaware of any previous studies using counts of agricultural innovations introduced. To conduct this analysis, I estimate a simple differences-in-differences regression in which the "control group" in this regression is the set of runner-up counties described in Section II.A.3:

$$\begin{aligned} NewWheatVarieties_{ict} = & \beta_1 LandGrantCounty_{ic} \times PostLandGrant_{ct} \\ & + \beta_2 PostLandGrant_{ct} + County_i + LandGrant_c \\ & + Year_t + LandGrant_c \times Year_t + \epsilon_{ict}, \end{aligned} \tag{2}$$

where i indexes the county, c indexes the land grant college, t indexes the year, $NewWheatVarieties_{ict}$ is a measure of new wheat varieties developed in county i associated with land grant college c (as either the college or the runner-up county) in year t, $LandGrantCounty_{ic}$ is an indicator equal to one if county i receives land grant college c from among the set of finalist counties, and $PostLandGrant_{ct}$ is an indicator equal to one if year t is after the year in which land grant college c is established. $County_i$, $LandGrant_c$, and $Year_t$ are county, college, and year fixed effects, respectively.

Results are presented in Table 1. In Column 1, the dependent variable is simply the number of additional wheat varieties. I find that establishing a land grant college causes about 0.008 new wheat varieties each year, or about one new variety every 125 years. In Column 2, I used the logged number of wheat varieties as the dependent variable, while in Column 3 I use the inverse hyperbolic sin transformation. Both produce similar results: establishing a new college causes an increase of about one-half of one percent more wheat varieties each year. While each of these results is statistically significant at the 10% level, they are very small in magnitude. This reflects the fact that counties that received land grant colleges were developing new wheat varieties before the Morrill Act, that runner-up counties continued to develop their own varieties even after land grant colleges were established in their states, and that development of a new wheat variety is relatively rare in the data in both the college and runner-up counties. In Column 4, I weight each variety by the share of national acreage planted in each variety in the Clark et al. (1922) survey, so that the dependent variable is $\sum_{n=1}^{N} ShareAcreage_{nit}$ where N is the number of new varieties introduced in county i and year t. If land grant colleges were developing varieties that were on average more valuable than those developed in other counties, then this measure should be positive and significant; instead, it is close to zero in magnitude and statistically insignificant. Column 5 investigates the average quality of each variety introduced, with the dependent variable $\frac{1}{N}\sum_{n=1}^{N} ShareAcreage_{nit}$. The results are similar, but if anything are even smaller in magnitude.

The last two columns suggest that counties with land grant colleges did not begin developing varieties that would eventually become more widely adopted after the college was established. To explore more fully the extent to which the entire country was able to benefit from the new wheat varieties developed at land grant colleges, I use the Clark et al. (1922) data to show what share of new varieties, across all varieties in the country (not just the college and runner-up counties), came from the land grant colleges. Table 2 presents results.

I restrict attention only to wheat varieties introduced since the passage of the Morrill Land Grant Act in 1862 to avoid counting varieties from before the land grant system could have had any effect. In Column 1, I count all varieties that Clark et al. (1922) indicate were introduced as a result of research at land grant colleges or state agricultural experiment stations.[5] About 30% of all new varieties introduced between 1862 and 1922 came from land grant research. In Column 2, in addition to the varieties attributed to land grant research in Column 1, I also include any varieties introduced in a county that had a land grant college. This thus captures any varieties for which Clark et al. (1922) erroneously fail to mention the role of land grant research, as well as some local knowledge spillovers from land grant colleges to local farmers. Including these additional varieties increases the share of varieties from land grant college counties to about 36% of all new varieties. In Columns 3 and 4, I weight each variety by acres planted.[6] Varieties introduced as a result of land grant research account for only 10% of acreage planted. All varieties from land grant counties account for 13% of acreage planted. Comparing the number of varieties introduced to the acreage results suggests that land grant research produced varieties that were, on average, less useful for American farmers. Unfortunately, the Clark et al. (1922) data does not contain county-level information on wheat acreage, so it is impossible to tell which counties are growing varieties developed by land grant colleges.

In row 2, I repeat the exercise but keep only varieties introduced since the passage of the Hatch Act in 1887, which established and provided federal funding for agricultural experiment stations. When restricting attention to this period in which land grant research was on an even firmer financial footing and was conducted in a larger number of geographic locations, land grant colleges account for a slightly larger share of both varieties and acreage (35% and 11%, respectively). This is also true when including all varieties introduced in land grant counties (39% of varieties and 17% of acreage).

[5]In the calculations, I include wheat varieties developed outside the U.S. as long as Clark et al. (1922) can identify the location within the U.S. at which the variety is first introduced. Many (although not all) of the varieties attributed to land grant research were initially developed outside the U.S., lending support to the claims in Alston (2002) and Maredia, Ward, and Byerlee (1996) that federal support of agricultural innovation generated sizable international spillovers.

[6]See Appendix B for details on how acreage planted in each variety was determined.

II.B.3 Discussion

From these results, it is clear that the land grant college system played a non-trivial role in the introduction of new varieties, but they were far from the primary driver of agricultural innovation in the late 19th and early 20th centuries. Other locations continued to develop their own varieties even once the land grant system was in place, likely because the varieties developed at land grant colleges were not suitable for use everywhere. These findings support the conclusions from the previous section that proximity is more important for agriculture than for other sectors. A major caveat is that these data are only for the introduction of wheat varieties and only extend up to 1920; I have been unable to find comparably detailed data for other crops or for wheat for more recent years. It is possible land grant colleges played a larger role in the development of different species of crops or in the development of farm machinery, or that their role qualitatively changed in recent decades. It is also likely that land grant colleges played a substantial role in promoting the diffusion of wheat varieties developed elsewhere. Indeed, several of the descriptions of varieties indicate that agricultural experiment stations researchers scoured the country to discover varieties developed by obscure farmers.[7] But this hardly sounds like the situation Wright (2012) describes in which wheat farmers are afraid to develop new varieties for fear of expropriation: "Unlike hybrid corn, wheat was a self-pollinated plant that the farmer could replant for several years and sell extra seed to others. Given this competitive threat from potential customers, wheat breeding was privately unprofitable, and thus necessarily located mainly in the public sector" (p. 1719).

Wright's comparison to hybrid corn is, however, instructive. The development of hybrid corn is typically hailed as a triumph of the American land grant system. Indeed, histories of the development of hybrid corn illustrate the importance of knowledge spillovers among academic scientists for the development of innovative concepts (Crabb (1947), Nelson (1993)). Early research on hybrid corn breeding at agricultural experiment stations depended not on the applicability of a particular cross for American farmers but, not surprisingly, on the

[7] As one example, the Wyandotte variety was discovered by researchers from the Ohio Agricultural Experiment Station at Columbus being grown on a farm in Nevada, OH, although the variety's exact origins remain a mystery. The Indiana agricultural experiment station in Bloomington frequented Everlitt's O.K. Seed Store in Indianapolis to learn about new varieties from across the country.

manipulability of alleles or the observability of phenotypic traits. This work eventually led to the development of the concept of hybridization in a form that was easy enough for American farmers–the vast majority of whom did not have a background in maize genetics–to adopt. But when it came time to develop the particular varieties used throughout the Midwest, that role was primarily filled by private seedsmen such as those at Funk Farms and Pioneer Hi-Bred, who would canvass and interact with farmers from a wide geographic range in order to develop seeds they could sell in these large but distant markets.[8] This is not, of course, to say that the distinction between private and land grant research was clear or that credit for innovations may be easily divided among them; Funk Farms even became a USDA branch experiment station for a time, and Henry A. Wallace founded Pioneer Hi-Bred just two years after his father stepped down as the U.S. Secretary of Agriculture (Henry A. Wallace would himself hold the same position only seven years later). Unfortunately, I am aware of no history of the hybrid corn varieties in use that is of comparable quality or completeness to the wheat variety histories compiled by Clark et al. (1922). Even extremely detailed work on hybrid corn adoption, such as Griliches (1957), does not record the specific varieties grown in particular counties. It is thus impossible to know with certainty the location of origin of most hybrid corn grown in the U.S.

The story of hybrid corn, together with the realities of wheat variety adoption, suggest that public research may be most broadly beneficial when it is directed towards conceptual advancements such as the "idea" of hybrid corn, while the development of specific varieties is left to local farmers or to those with a greater incentive to sell to the largest market. By focusing on more basic science, the direction of public research is therefore less likely to be influenced by the local environment, while the applied science that may be more sensitive to local conditions is placed in the hands of researchers who have an incentive to ignore local conditions to sell to the largest customer base, which may be geographically distant.

[8] Crabb is explicit on this point when describing the difficulties of early hybrid corn pioneers working at agricultural research stations, many of whom were located on the East Coast, in getting their breakthroughs adopted in the American corn belt: ""Had Donald Jones developed a new kind of automobile or a razor blade in New Haven, Connecticut, he could have taken his wares to Iowa or California and demonstrated their good qualities effectively. Unfortunately, he did not have that chance with his new double-cross hybrid, since the Burr-Leaming hybrid was adapted to the soil and gentle climate of southern New England, and it failed to make an impressive demonstration at any of the numerous places it was grown in the corn belt." (Crabb, 1947, p. 95)"

I emphasize that these conclusions are, at best, highly tentative. I provide more direct evidence for the influence of the local environment on the direction of public science in the following sections.

III The Local Environment and the Direction of Innovation

The above results are all consistent with a story in which nearby areas benefit much more from land grant research than do areas that are farther away, but there is still no direct evidence that land grant research is biased in particular directions. I present such evidence in this section. The idea is to see if researchers at land grant colleges are more likely to write papers related to local crops than to crops grown in the rest of the state.

III.A Data

I begin by calculating the "lead" cereal crop in each college county in the last agricultural census prior to the establishment of a land grant college, defined as the crop with the greatest output in bushels and again using the data from Haines et al. (2018). I construct the same measure for the "rest of the state," defined as the crop with the greatest combined output over all counties in the state (leaving out the college county). I only use cereal crops for this analysis because other important agricultural output, such as fruits or livestock, are not recorded consistently across censuses. Unfortunately, the agricultural census data are quite coarse, so I can only observe inputs and outputs by species of crop (e.g., corn vs. wheat) rather than variety.[9] Thus, most college counties and the rest of the state predominantly grow the same species of crop. But in a handful of cases, the college county and rest of the state specialized in different species. Based on the historical evidence compiled, I have found no evidence that the specific type of crops grown by local farmers played a meaningful role in determining the location of land grant colleges; to further provide evidence that this is the case, I restrict this analysis to the sample of colleges for which the location was assigned

[9]Some years report sub-species of crops. For instance, the 1870 agricultural census reports output separately for spring and winter wheat. But such a breakdown is not available for most years.

essentially at random, as in Section II.A.3. Appendix C provides details on the cases in which the college county and the rest of the state grew different lead crops.

Data on contemporary agricultural publications come from the Scopus Database (see Scopus (2017) or Burnham (2006) for broad overviews of Scopus). Scopus includes the title and author affiliation, among other fields, of 2,741 journals in the field of "Agriculture and Biological Sciences" from 1960 to the present. From this data, I compile a list of all papers published in an agricultural journal from 1960 to 2010 in which at least one of the authors lists a sample land grant college as part of their affiliation. I then parse the title of each paper to see whether it contains the name of the land grant college county's lead crop from the last year prior to the establishment of the college and/or the runner-up county's lead crop.

III.B Results

I first estimate a simple logit equation of the form

$$\ln\left(\frac{Pr(Paper_i = LandGrantCrop)}{Pr(Paper_i = RunnerUpCrop)}\right) = \beta_0 + \beta_1 EarlyPeriod_i, \quad (3)$$

where the left hand side is the log odds ratio of paper i mentioning the land grant county lead crop to paper i mentioning the runner-up county crop. e^{β_0} represents the baseline odds, and $Time_i$ is a time-of-publication fixed effect. I begin by examining only the baseline odds (that is, setting $\beta_1 = 0$). Results are presented in Table 3, Columns 1 and 3. Publications are 62% more likely to be about land grant college crops when parsing only on the crop's common English name, and 75% more likely when parsing on either English or Latin names for crops. In other words, there are about 1.6 to 1.75 papers mentioning the land grant college lead crop for every paper mentioning the lead crop of the rest of the state. In Columns 2 and 4, I estimate β_1 and set $EarlyPeriod_i$ as an indicator equal to one if a paper is published in the first half of the sample period (1960-1985).[10] While the magnitude of β_1 is large, suggesting

[10] The relatively rare instances of publication titles mentioning the lead crops in the early decades introduces substantial multicollinearity when using, for instance, time or decade fixed effects and effects are dropped in typical statistical software. Because I am interested in interpreting the time effect, I opt instead to use this much coarser measure.

that paper titles were much more likely to mention land grant lead crops, the estimate is not significant. The baseline odds remain large and significant and are of comparable magnitude to those in Columns 1 and 3.

I next repeat the same analysis in a multinomial logit framework, where I also consider the possibility that a paper title mentions neither land grant nor runner-up county lead crops. As papers mentioning either lead crop make up a small share of all paper published in agricultural journals by authors affiliated with land grant colleges, including the "neither" group greatly increases the sample size. More precisely, I estimate

$$\ln\left(\frac{Pr(Paper_i = LandGrantCrop)}{Pr(Paper_i = NeitherCrop)}\right) = \beta_0 + \beta_1 EarlyPeriod_i$$
$$\ln\left(\frac{Pr(Paper_i = RunnerUpCrop)}{Pr(Paper_i = NeitherCrop)}\right) = \beta_0 + \beta_1 EarlyPeriod_i, \quad (4)$$

where $Pr(Paper_i = NeitherCrop)$ is the probability that paper i's title mentions neither the land grant nor runner-up lead crop. As these results show, paper titles that mention either crop are quite rare. There are about 1.6 papers that mention the land grant county crop for every 100 agricultural papers that mention neither crop, and one paper that mentions the state lead crop for every 100 papers that mentions neither. The baseline probabilities are very similar when including an indicator for early periods. In the multinomial logit specification, land grant college crops are about 53% ($-0.52 = 0.48 - 1$) additionally less likely than papers that mention neither crop in the early period, although the indicator for the early period is only statistically significant at the 10% level, while the state lead crop is an additional 84% less likely ($-0.85 = 0.15 - 1$). These results are suggestive that both land grant and state crops become relatively more commonly mentioned in later years, with the state crops increasing more quickly but never catching up to become as common as the land grant county crops. Columns 3 and 4 use mentions of both the English and Latin names in paper titles and find qualitatively similar results. In the rows titled $College - StatewideConst. = 0$, I present χ-square test statistics and p-values for whether or not the estimates of the each multinomial logit model are statistically different from one another. In all columns, paper titles are statistically significantly more likely to mention land grant crops than state lead crops.

IV Location-Specific Agricultural Human Capital

As a second approach to show that the local environment has a larger impact of knowledge creation, and subsequently on the ability of knowledge to diffuse, in agriculture than in other sectors, I consider the mobility patterns of alumni from land grant colleges. The basic idea is that, if the human capital of agriculture students is more location-specific than the human capital acquired by students in other majors, then agriculture alumni will remain closer to their alma maters after graduating than will alumni with degrees in other majors.

IV.A Data

Data on the names of college seniors at land grant colleges, who go on to become alumni, come from college yearbooks available on `ancestry.com`. See Andrews (2019c) for more details on the college yearbook data. I have collected yearbooks from the following land grant colleges: Auburn University, University of Colorado, Iowa State University, Louisiana State University, University of Maine, University of Missouri, University of Nevada Reno, University of New Hampshire, Cornell University, North Dakota State University, Clemson University, Utah State University, Virginia Tech, and Washington State University. In addition to the names of college seniors, the yearbooks typically include seniors' major and hometown.

I keep records for every college senior for whom at least one major is recorded. I classify a student as an agriculture major if at least one of the student's majors is in agriculture or a closely related field such as agricultural engineering, agronomy, animal husbandry, crop science, dairy husbandry, farm husbandry or management, horticulture, or poultry husbandry. I also create an alternative classification that is more strict, excluding related fields such as botany, veterinary medicine, soils science, or zoology. Results are very similar regardless of which classification I use.

To determine the mobility of alumni, I "fuzzy" match by first and last name of these alumni to the 1940 population census data for the college's county. The 100% U.S. federal decennial population census data for 1940 is transcribed by `ancestry.com` and IPUMS and hosted by the NBER. If an alumnus record successfully matches to a census record in the college's county, I record that alumnus as still living in the college of his alma mater;

otherwise, the alumnus is recorded as being a cross-county migrant.

IV.B Results

Figure 3 plots the Lowess-smoothed share of agricultural and non-agricultural alumni living in the college county in the 1940 census as a function of their yearbook year (since these individuals are seniors in the yearbooks, the yearbook year is typically either the year they graduated or the year before they graduated). A greater share of agricultural alumni are living in the same county as their college, particularly among those who graduated two-to-four decades before the 1940 census. Not surprisingly, the shares of both alumni groups in the same county as the college converge as the yearbook year approaches 1940, as alumni have less time to make their migration decisions regardless of major.[11]

Table 5 formalizes the intuition in Figure 3. Column 1 shows the results of a simple linear probability model:

$$LiveInSameCountyAsCollege_i = \beta_0 + \beta_1 Ag.Alumni_i + \epsilon_i. \tag{5}$$

Agriculture alumni are about 1% more likely to remain in the same county as their alma mater in 1940, although this coefficient is not statistically significant. Because some colleges may be more desirable to live near than others, and because individuals are more likely to migrate as more time passes since they graduate, Column 2 includes fixed effects for the college and the yearbook year. In this specification, agriculture alumni are 1.6% more likely to remain in the same county, and the coefficient is statistically significant at the 5% level. Migration decisions may also be driven by a desire for alumni to return to their homes. In Column 3 I therefore also include a home state fixed effect and find that agriculture alumni are 1.8% more likely to remain in the same county, an effect that is highly statistically significant. Finally, in Column 4 I include a fixed effect for each individual's hometown.

[11]The fact that the share in the same county as the college is less than one in 1940 reflects the failure of the fuzzy name matching procedure to link all individuals in the yearbooks to the census. This is likely due to typos in either the yearbook or the census records, individuals recording their names differently to the yearbook and census enumerators, etc. The fact that the rate is virtually identical between the agricultural and non-agricultural alumni gives confidence that the results are not driven by differential match rates across these two groups.

Because there is no uniform restriction on how individuals report their home towns in the college yearbooks, the same town may be recorded multiple ways, and even close neighboring areas will each have their fixed effect. These fixed effects therefore exhaust many degrees of freedom, so it is not surprising that the results are no longer statistically significant and are small in magnitude. While all these results are for linear probability models, results of logit models are similar. While all these coefficients are very small in magnitude and do explain only a small share of the variation in individuals' migration choices, they do paint a consistent picture in which migration is slightly less likely for individuals who have studied agriculture, even after including a battery of fixed effects.

In ongoing work, I match alumni to the 1940 census even if they do not live in the college county. This is more computationally intensive, since it requires searching for each alumnus record in the census data for multiple counties. Preliminary results (which will be available in future versions of this paper) show that agricultural alumni tend to live closer to the college county than do alumni from other majors, using either geodesic or latitudinal distance. Comparing within agricultural alumni, those who live further away from the college are more likely to leave agriculture, further supporting the conjecture that the human capital they acquired at college is less applicable further from the college. I stress that these results are preliminary.

V Conclusion

References

Alston, J. M. (2002, September). Spillovers. *Australian Journal of Agricultural and Resource Economics 46*(3), 315–346.

Alston, J. M. and P. G. Pardey (1996). *Making science pay: the economics of agricultural R&D policy*. Washington, DC: American Enterprise Institute Press.

Andrews, M. J. (2019a). Bar talk: informal social interactions, alcohol prohibition, and invention. Unpublished, National Bureau of Economic Research.

Andrews, M. J. (2019b). Historical appendix: How do institutions of higher education affect local invention? Evidence from the establishment of U.S. colleges. Unpublished, NBER.

Andrews, M. J. (2019c). How do institutions of higher education affect local invention? Evidence from the establishment of U.S. colleges. Unpublished, NBER.

Bandiera, O. and I. Rasul (2006, October). Social networks and technology adoption in northern Mozambique. *Economic Journal 116*(514), 869–902.

Beaman, L. and A. Dillon (Forthcoming). Diffusion of agricultural information within social networks: evidence on gender inequalities from Mali. *Journal of Development Economics*.

Burnham, J. F. (2006). Scopus database: a review. *Biomed Digital Libraries 3*(1).

Chatterji, A. K. and K. Fabrizio (2012, July-August). How do product users influence corporate invention? *Organizational Science 23*(4), 951–970.

Clark, J. A., J. H. Martin, and C. R. Ball (1922). Classification of American wheat varieties. Bulletin 1074, Washington.

Conley, T. G. and C. R. Udry (2010, March). Learning about a new technology: pineapple in Ghana. *American Economic Review 100*(1), 35–69.

Crabb, A. R. (1947). *The hybrid-corn makers: prophets of plenty*. New Brunswick, NJ: Rutgers University Press.

Diamond, J. (1997). *Guns, germs, and steel: the fates of human societies*. New York: W. W. Norton & Company.

Doyle, M., S.-B. Sung, and R. Amasino (2002). The genetic control of flowering time. *Annual Plant Reviews 6*, 33–60.

Feder, G., R. E. Just, and D. Zilberman (1985, January). Adoption of agricultural innovations in developing countries: a survey. *Economic Development and Cultural Change 33*(2), 255–298.

Foster, A. D. and M. R. Rosenzweig (2010). Microeconomics of technology adoption. *Annual Review of Economics 2*, 395–424.

Ganguli, I., J. Lin, and N. Reynolds (2019). The paper trail of knowledge spillovers: evidence from patent interferences. Unpublished, University of Massachusetts-Amherst.

Glaeser, E. L. and N. Hausman (2019). The spatial mismatch between innovation and joblessness. Unpublished, Harvard University.

Griliches, Z. (1957, October). Hybrid corn: an exploration in the economics of technological change. *Econometrica 25*(4), 501–522.

Griliches, Z. (1958, October). Research costs and social returns: hybrid corn and related innovations. *Journal of Political Economy 66*(5), 419–431.

Haines, M., P. Fishback, and P. Rhode (2018). United States agricultural data, 1840-2012

(ICPSR 35206). https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/35206/publications, accessed May 12, 2019.

Hanlon, W. W. (2015, January). Necessity is the mother of invention: input supplies and directed technical change. *Econometrica 83*(1), 67–100.

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

Juhász, R. (2018, November). Temporary protection and technology adoption: evidence from the Napoleonic Blockade. *American Economic Review 108*(11), 3339–3376.

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

Khan, B. Z. (2015, Winter). The impact of war on resource allocation: "creative destruction," patenting, and the American Civil War. *Journal of Interdisciplinary History 46*(3), 315–353.

Li, D., P. Azoulay, and B. N. Sampat (2017, 07 April). The applied value of public investments in biomedical research. *Science 356*(6333), 78–81.

Manski, C. F. (1993). Identification of endogenous social effects: the reflection problem. *Review of Economic Studies 60*, 531–542.

Manson, S., J. Schroeder, D. V. Riper, and S. Ruggles (2018). IPUMS national historical geographic information system: version 13.0. http://doi.org/10.18128/D050.V13.0.

Maredia, M. K., R. Ward, and D. R. Byerlee (1996, August). Econometric estimation of a global spillover matrix for wheat varietal technology. *Agricultural Economics 14*(3), 159–173.

Moscona, J. (2019). Downstream effects of research incentives: evidence from agricultural innovation. Unpublished, MIT.

Moser, P. (2005, September). How do patent laws influence innovation? Evidence from nineteenth-century world's fairs. *American Economic Review 95*(4), 1214–1236.

Moser, P. and P. W. Rhode (2012). Did plant patents create the American rose? In J. Lerner and S. Stern (Eds.), *The Rate and Direction of Inventive Activity Revisited*, pp. 413–438. Chicago: University of Chicago Press.

Myers, K. A. and P. Rhode (2019). Hybrid corn adoption in response to 1930s drought stress. Unpublished.

Nelson, O. E. (1993, December). A notable triumvirate of maize geneticists. *Genetics 135*(4), 937–941.

Nelson, R. R. (1959, June). The simple economics of basic scientific research. *Journal of Political Economy 67*(3), 297–306.

Newman, C. L. (1899). The comparative yield of corn from seed of the same variety grown in different latitudes. Bulletin 59, Fayetteville, AR.

Olmstead, A. L. and P. W. Rhode (2008). *Creating abundance: biological innovation and American agricultural development*. Cambridge: Cambridge University Press.

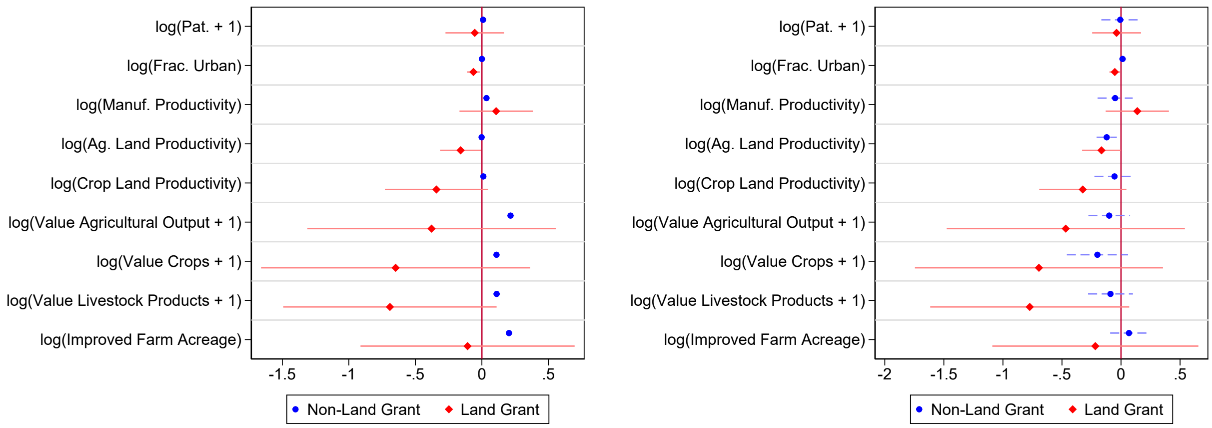
Scopus (2017, August). Scopus: content coverage guide. https://www.elsevier.com/__data/assets/pdf_file/0007/69451/0597-Scopus-Content-Coverage-Guide-US-LETTER-v4-HI-singles-no-ticks.pdf, accessed November 1, 2018.

Steckel, R. H. (1983, January). The economic foundations of east-west migration during the

19th century. *Explorations in Economic History 20*(1), 14–36.
Stokes, D. E. (1997). *Pasteur's quadrant: basic science and technological innovation.* Washington, DC: Brookings Institution Press.
Thomas, B. and D. Vince-Prue (1997). *Photoperiodism in plants: second edition.* San Diego: Academic Press.
University of Illinois (1889-1894). Field experiments with corn. Agricultural Experiment Station Bulletin 4, 8, 13, 20, 25, 31, Champaign, IL.
von Hippel, E. (1988). *The sources of innovation.* Oxford: Oxford University Press.
Wright, B. D. (2012, December). Grand missions of agricultural innovation. *Research Policy 41*(10), 1716–1728.

Graphs

Figure 1: Geodesic Distance

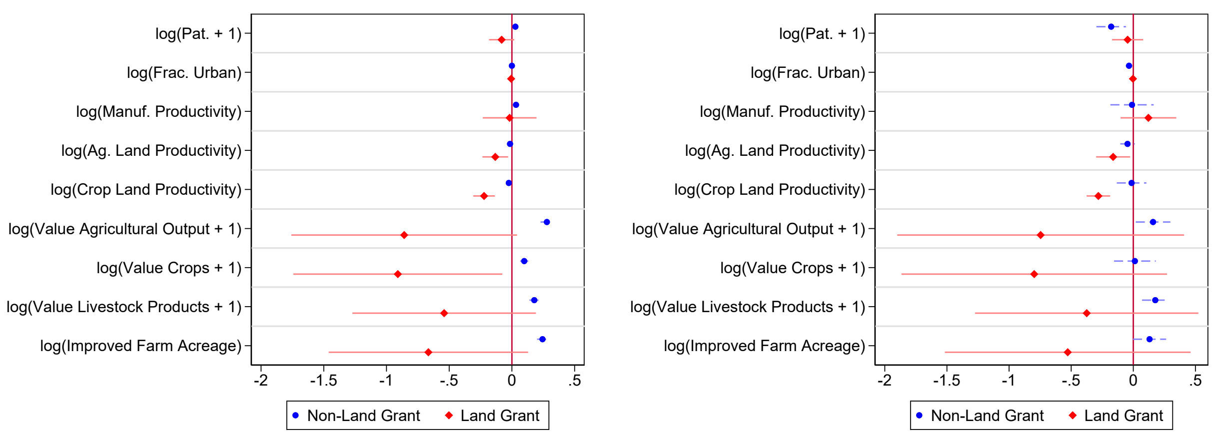


(a) All Counties

(b) Runner-Up Counties

Notes: Estimates of the change in the elasticity of the dependent variable with distance to the college county after establishment of the college, using geodesic distance. The red diamonds show these results for the sample of land grant colleges. The blue circles show these results for the sample of non-land grant colleges. 90% confidence intervals are shown.

Figure 2: Latitudinal Distance

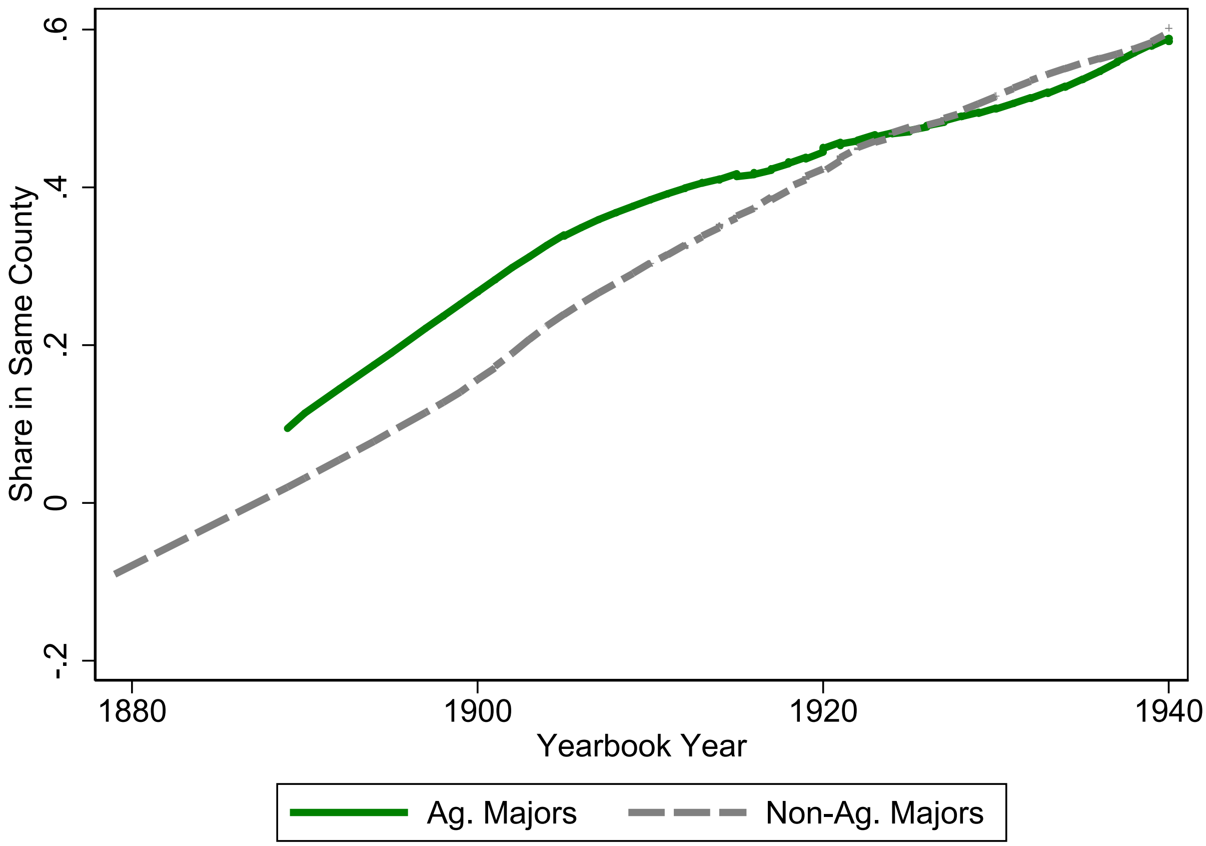


(a) All Counties

(b) Runner-Up Counties

Notes: Estimates of the change in the elasticity of the dependent variable with distance to the college county after establishment of the college, using latitudinal distance. The red diamonds show these results for the sample of land grant colleges. The blue circles show these results for the sample of non-land grant colleges. 90% confidence intervals are shown.

Figure 3: Share of Agriculture and Non-Agriculture Alumni Living in the Same County as their Alma Mater



Notes: The Lowess-smoothed share of agriculture alumni (green solid line) and non-agriculture alumni (gray dashed line) living in the same county as their alma mater as of the 1940 census. The x-axis is the year of each college yearbook.

Tables

Table 1: Effect of Land Grant Colleges on New Wheat Introductions

	Num. Varieties	log(Varieties + 1)	ihs(Varieties)	Share Acreage Varieties	Av. Share Acreage Varieties
CollegeCounty * PostCollege	0.00722	0.00356	0.00458	-0.00000121	-0.00000315
	(0.00493)	(0.00212)	(0.00274)	(0.00000380)	(0.00000266)
PostCollege	-0.000917	-0.00123	-0.00154	-0.00000363	-0.00000438
	(0.00178)	(0.00122)	(0.00154)	(0.00000428)	(0.00000300)
Num. Counties × Years	9235	9235	9235	9235	9235
Adj. r-Sqr.	0.000307	0.00503	0.00501	0.000911	0.00156

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Results of a differences-in-differences regression for the differences between land grant college counties and runner-up counties before and after the college is established. The dependent variable in Column 1 is the number of new wheat varieties, the dependent variable in Column 2 is $\log(NewWheatVarieties + 1)$, the dependent variable in Column 3 is $ihs(NewWheatVarieties) = \log(NewWheatVarieties + (NewWheatVarieties^2 + 1)^{\frac{1}{2}})$, the dependent variable in Column 4 is the number of new wheat varieties multiplied by the share of total national acreage of the new variety, and the dependent variable in Column 5 is the average share of total national acreage in each new variety. Standard errors are clustered at the county level.

Table 2: Share of Wheat Varieties from Land Grant Research

	Share Varieties		Share Acres	
	Land Grant Research	Land Grant Counties	Land Grant Research	Land Grant Counties
Post Morrill Act	0.303	0.355	0.097	0.131
Post Hatch Act	0.347	0.389	0.113	0.166

Notes: Columns 1 and 2 list the share of new wheat varieties introduced since the passage of the Morrill Act in 1862 (row 1) and the passage of the Hatch Act in 1887 (row 2). Column 1 shows the share of varieties introduced as a result of land grant college research. Column 2 includes any varieties introduced in land grant college counties, regardless of whether they were the result of programmatic research. Columns 3 and 4 do the same but weight each variety by acreage planted.

Table 3: Logit Results for Research on the College County Crop

	Common Crop	Common Crop	Latin Crop	Latin Crop
Constant	1.620***	1.580***	1.746***	1.690***
	(0.158)	(0.155)	(0.165)	(0.162)
Early Period		3.165		2.959
		(2.471)		(1.893)
Num. College × Years	448	448	486	486
log-Likelihood	-297.9	-296.5	-318.7	-316.9

Exponentiated coefficients; Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Logit regressions in which the dependent variable is the probability of a paper mentioning the lead crop in the land grant college county at the time the college was established. The sample is of all papers in which either the lead crop in the college county or the lead crop in the rest of the state is mentioned in the paper title. Columns 1 and 2 use crops' common names (e.g., wheat, corn, etc.) as the dependent variable, while Columns 3 and 4 includes scientific or Latin crop names. Columns 2 and 4 include a dummy variable for years in the first half of the sample, from 1960 to 1985.

Table 4: Multinomial Logit Results for Research on the College County Crop or Lead Crop in the Rest of the State

	Common Crop	Common Crop	Latin Crop	Latin Crop
College_Crop				
Constant	0.0156***	0.0162***	0.0174***	0.0179***
	(0.000945)	(0.00100)	(0.00100)	(0.00105)
Early Period		0.477*		0.652
		(0.154)		(0.174)
RunnerUp_Crop				
Constant	0.00963***	0.0103***	0.00999***	0.0106***
	(0.000740)	(0.000794)	(0.000755)	(0.000807)
Early Period		0.151**		0.220**
		(0.107)		(0.128)
College - RunnerUp Const. = 0: χ^2-Stat.	24.60	21.65	34.94	30.07
College - RunnerUp Const. = 0: p-value	0.000000706	0.00000328	3.40e-09	4.16e-08
Num. College × Years	18201	18201	18201	18201
log-Likelihood	-2399.9	-2389.6	-2559.0	-2551.8

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Multinomial logit regressions in which the dependent variable is the probability that a paper mentions the lead crop in the land grant college county at the time the college was established (the "College_Crop"), the probability that a paper mentions the lead crop in the rest of the state at the time the college was established (the "Statewide_Crop"), and the probability that a paper mentions neither (the omitted category). The sample is of all papers in agricultural journals in which at least one author has an affiliation with a land grant college. Columns 1 and 2 use crops' common names (e.g., wheat, corn, etc.) as the dependent variable, while Columns 3 and 4 includes scientific or Latin crop names. Columns 2 and 4 include a dummy variable for years in the first half of the sample, from 1960 to 1985.

Table 5: Probability that Agriculture Alumni Live in Same County as their Alma Maters

	(1)	(2)	(3)	(4)
Ag. Major	0.0106	0.0156**	0.0183***	0.00663
	(0.00561)	(0.00527)	(0.00536)	(0.00634)
Constant	0.133***	0.132***	0.131***	0.130***
	(0.00218)	(0.00198)	(0.00201)	(0.00209)
Num. Alumni	28753	28753	27987	25441
Adj. R-sqr.	0.0000895	0.189	0.190	0.207

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Linear probability results for the probability that agriculture alumni live in the same county as their alma mater, relative to non-agriculture alumni. Column 1 shows the unconditional correlation. Column 2 includes college and yearbook year fixed effects. Column 3 includes college, yearbook year, and home state fixed effects. Column 4 includes college, yearbook year, and home town fixed effects.

A The Importance of Proximity Using Alternative Distance Measures

In this section, I repeat the exercise in Section II using a number of alternative distance measures. In Figures 1 and 2, I show how agricultural productivity and output decline with logged geodesic and latitudinal distance, respectively. But the effect of distance may exhibit crucial nonlinearities that complicate the interpretation of the results. To test for this, I instead estimate alternative specifications that divide counties into those that are "near" versus "far" from the college. I start by dividing counties by the median distance between the college and all counties in the state, again using both geodesic and latitudinal distance; results are similar using alternative arbitrary distance cutoffs. In Figures 4-6 I plot these cutoffs for several land grant colleges in the sample. One thing these maps make clear is that, when changing the distance measure from geodesic to latitudinal, the sample of counties that are considered either "near" or "far" can change dramatically. This is especially the case when restricting attention to the runner-up counties.[12] This fact makes the robustness of the overall conclusions even more remarkable.

I now estimate

$$\begin{aligned} Dep.Var._{it} = &\beta_1 NearCollegeCounty_i \times PostCollege_{it} + \beta_2 PostCollege_{it} + \\ &+ County_i + College_i + Year_t + \epsilon_{it}, \end{aligned} \tag{6}$$

where $NearCollegeCounty_i$ is equal to one if county i is less than the median distance from the college. This is now a differences-in-differences specification with discrete treatment groups. In contrast to the results in Section II, here a positive coefficient indicates that the dependent variable increases in counties *closer* to the college after the college is established.

Figures 7 and 8 plot the results. Regardless of whether these distance cutoffs are determined by geodesic or latitudinal distance, I find results consistent to those in Section II: counties near the college see an increase in agricultural productivity and output and virtu-

[12] Indeed, there are several cases in which all of the runner-up counties are either considered near or far in particular regressions. In these cases, there is no variation for a given college and so the in these specification the effect of distance is not identified off of these land grant colleges.

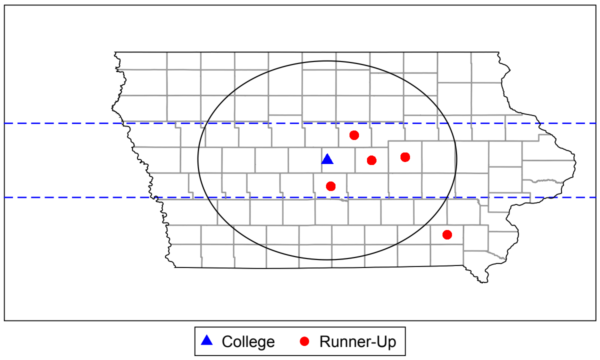
ally no change to patenting, urbanization, or manufacturing productivity.[13] Non-land grant colleges see smaller changes in magnitude for agricultural variables and similar small changes for the non-agricultural variables.

Figure 9 plots results using another distance measure: "adjacency." I consider two counties to be adjacent if they share a border. This measure may capture important aspects of the agricultural environment when, for example, geographical features like rivers or mountain ranges that may affect the diffusion of agricultural technology also determine borders, as well as when county-level political decisions affect diffusion directly. Results when using adjacency are very similar for the full sample; when restricting attention to the runner-up counties, the proximity to the non-land grant college also appear to have similar effects on agricultural outcomes.

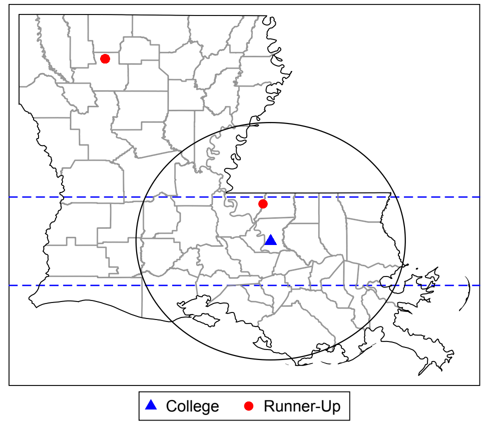
Finally, I attempt to measure the extent to which counties are suited to the same types of agriculture, regardless of geographic distance. I record whether counties were growing the same lead crop in the last census prior to the establishment of the college, as in Section III. Because there is typically little variation when the college county grows the same lead crop as the rest of the state (and because the counties growing other crops in those cases are possibly different in crucial ways, such as having very little agricultural activity overall which can skew the lead crop measure), I restrict attention to the cases in which the college county grows a different lead crop from the rest of the state. Note that this leaves a very small sample size. Results are plotted in Figure 10. The results here are the noisiest and most difficult to interpret out of all the distance measures, but there still appears to be no effect for non-agricultural dependent variables, and the coefficient on proximity for agricultural dependent variables is smaller for the non-land grant colleges.

[13]The only exception is that proximity to land grant colleges appears to increase manufacturing productivity when using the median of latitudinal distance as the cutoff.

Figure 4: Maps of College and Runner-Up Locations along with Distance and Latitude Cutoffs



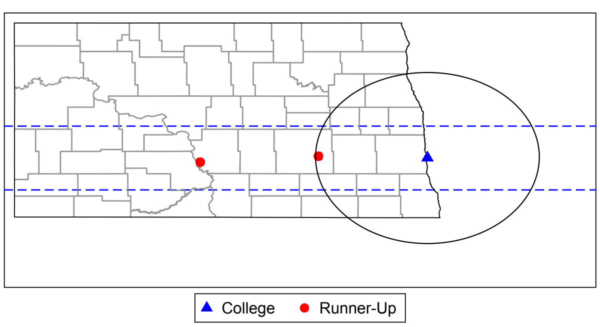
(a) Iowa State University



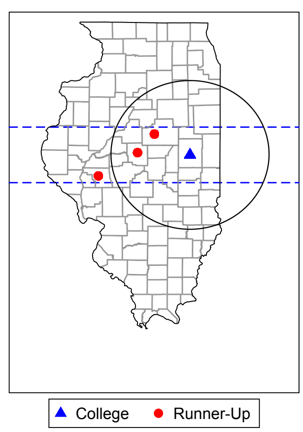
(b) Louisiana State University

Notes: The blue triangle indicates the location of the college site. The red circles indicate the locations of the runner-up sites. The black circle indicates areas within 150 km of the college site. The blue dashed lines indicate areas that are plus or minus one degree of latitude of the college site.

Figure 5: Maps of College and Runner-Up Locations in which the Runners-up Vary by Distance but not Latitude



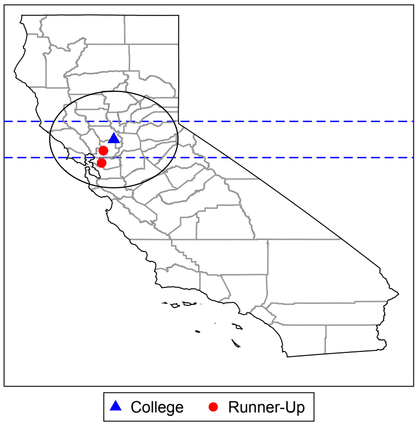
(a) North Dakota State University



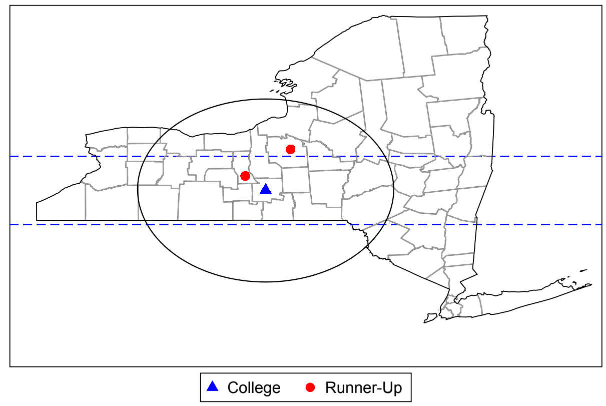
(b) University of Illinois Urbana-Champaign

Notes: The blue triangle indicates the location of the college site. The red circles indicate the locations of the runner-up sites. The black circle indicates areas within 150 km of the college site. The blue dashed lines indicate areas that are plus or minus one degree of latitude of the college site.

Figure 6: Maps of College and Runner-Up Locations in which the Runners-up Vary by Latitude but not Distance



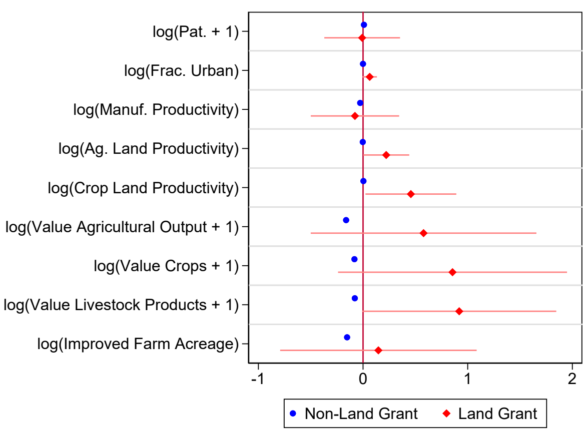
(a) University of California Davis



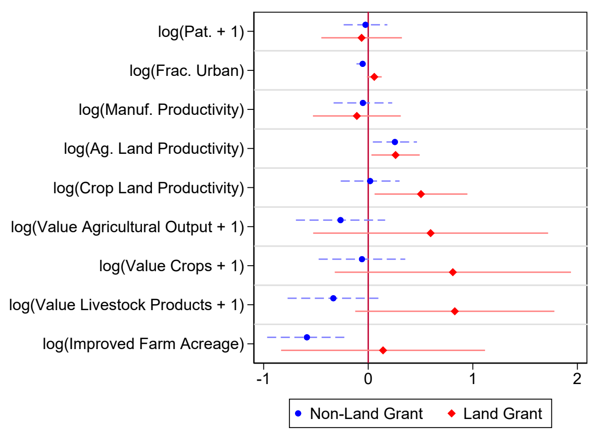
(b) Cornell University

Notes: The blue triangle indicates the location of the college site. The red circles indicate the locations of the runner-up sites. The black circle indicates areas within 150 km of the college site. The blue dashed lines indicate areas that are plus or minus one degree of latitude of the college site.

Figure 7: Geodesic Distance, Closer vs. Farther than Median Distance



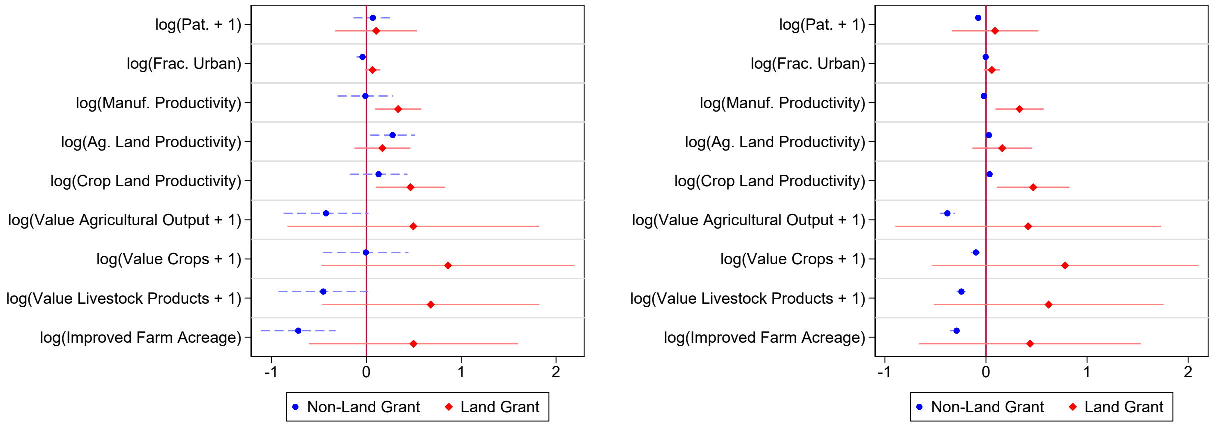
(a) All Counties



(b) Runner-Up Counties

Notes: Differences-in-differences estimates showing the difference between near counties and far counties before and after the establishment of a new colleges, using geodesic distance. The blue circles show these results for the sample of non-land grant colleges. The red diamonds show these results for the sample of land grant colleges. The solid lines show 90% confidence intervals.

Figure 8: Latitudinal Distance, Closer vs. Farther than Median Distance

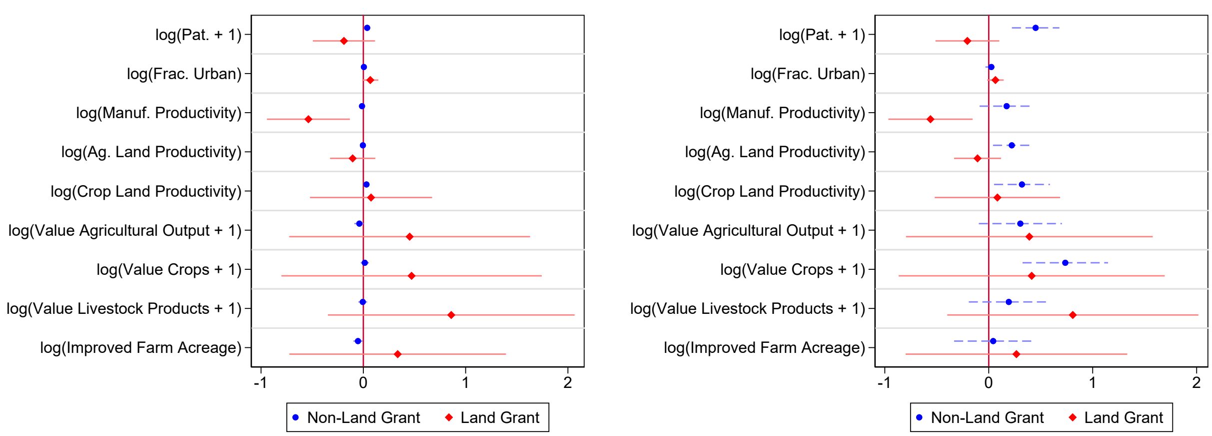


(a) All Counties

(b) Runner-Up Counties

Notes: Differences-in-differences estimates showing the difference between near counties and far counties before and after the establishment of a new colleges, using longitudinal distance. The blue circles show these results for the sample of non-land grant colleges. The red diamonds show these results for the sample of land grant colleges. The solid lines show 90% confidence intervals.

Figure 9: Adjacent vs. Non-Adjacent County

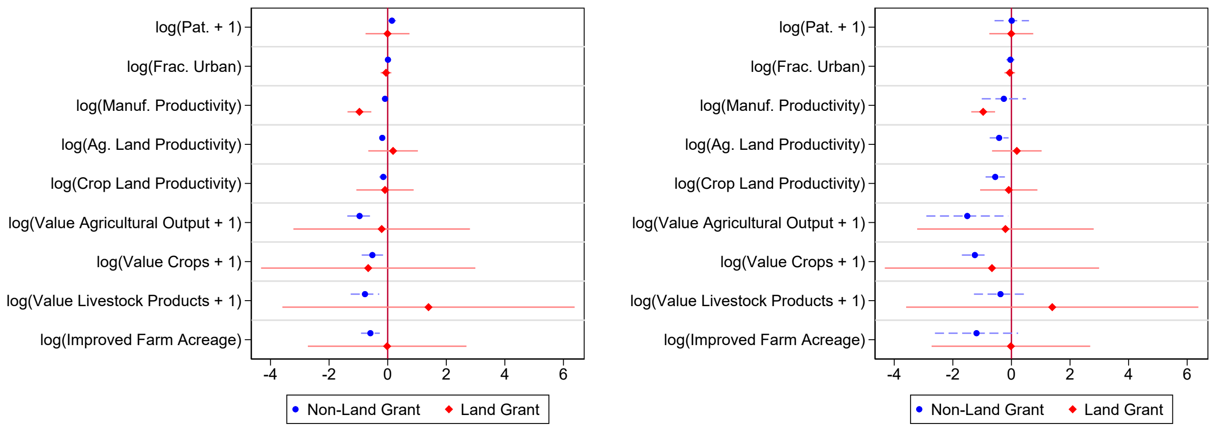


(a) All Counties

(b) Runner-Up Counties

Notes: Differences-in-differences estimates showing the difference between near counties and far counties before and after the establishment of a new colleges, using adjacency to determine nearness. The blue circles show these results for the sample of non-land grant colleges. The red diamonds show these results for the sample of land grant colleges. The solid lines show 90% confidence intervals.

Figure 10: County Grows Same vs. Different Lead Crop than the Land Grant College



(a) All Counties　　**(b)** Runner-Up Counties

Notes: Differences-in-differences estimates showing the difference between near counties and far counties before and after the establishment of a new colleges, using the lead crops to determine nearness. The blue circles show these results for the sample of non-land grant colleges. The red diamonds show these results for the sample of land grant colleges. The solid lines show 90% confidence intervals.

B Constructing Histories of Wheat Varieties and the Geographic Extent of Cultivation

The following description of the construction of wheat varietal histories and the determination of where these varieties are grown is reproduced from Clark et al. (1922, p. 15-17):

> The history of the origin of varieties can not be neglected in a complete classification, as many varieties are scarcely or not at all distinguishable from similar or closely related varieties and differ only in their origin or qualities. In this study much attention has been given to the history of varieties, and to many readers it probably will be the most interesting and valuable part of the classification. The compiling of these histories has been a long and arduous task. It has required a review of the literature on wheat varieties written during a period of more than 200 years. The sources of this information are varied. Definite information is readily available on the origin of only a comparatively few varieties. Introductions of foreign varieties have been recorded in recent years by the Office of Foreign Seed and Plant Introduction. Frequent reference is made to the

accession numbers and published inventories of that office. Many bulletins of the State agricultural experiment stations have contained valuable information on the origin of domestic varieties. Agricultural papers have been reviewed and much information as to the origin of varieties has been obtained from that source. There is still much to lean concerning the origin of our cultivated varieties. The origin of many probably has never been recorded, but fof some for which the origin has not been determined there probably is a recorded history somewhere. Reference is always given to the published sources of the histories that have been obtained...

To determine the commercially cultivated varieties of wheat in the United States and the extent of their distribution, a wheat varietal survey was made in cooperation with the Bureau of Crop Estimates. The first survey was made in 1917, when questionnaires were sent to one or two correspondents in each of the wheat producing counties of the various States. The incomplete returns from this survey were very interesting and contained so much valuable information from the counties reporting that it was decided in 1918 to send questionnaires to several correspondents in all counties not previously reporting, in order to have a more complete record. The replies were received and tabulated. They showed the varieties grown in the localities of the county where the correspondents lived, but it was soon determined that all of the varieties grown in each county were not included and that one or two reports from each county did not give an accurate estimate of the proportionate distribution of the different varieties. It was finally decided in 1919 to determine rather accurately the percentage each variety formed of the total wheat crop of each county. A new schedule was printed for this survey of about 70,000 were mailed. To the more important wheat-growing counties as many as 30 to 40 questionnaires were sent, fewer being sent to counties less important in wheat production. From the survey about 40,000 returns were received. About 19,000 of these gave definite information, and these results have been tabulated. In addition to the names of varieties grown and the percentage each formed of the total wheat crop, the questionnaires contained tabular spaces for descriptions of varieties. From these descriptions the correct naming of the variety was checked... The reports were edited before being tabulated and thus many recognizable mistakes were corrected. The summary of these reports revealed a large number of new names used for old varieties and also brought to light several wheats distinctly different from any of the varieties previously obtained. More than a thousand letters were written to the correspondents, requesting samples and additional information. A considerable number of additional varieties were obtained in this way.

C Cases in Which the College and State Grow Different Lead Crops

Table 6: Colleges with Different Lead Crops in College County and the Rest of the State

	College	Last Census	College Crop	Statewide Crop
1	North Dakota State University	1880	Wheat	Oat
2	New Mexico State University	1880	Wheat	Corn
3	Oregon State University	1860	Oat	Wheat
4	University of Maine	1860	Oat	Corn
5	University of Nevada	1880	Wheat	Barley
6	Virginia Polytechnic Institute	1870	Corn	Wheat

Notes: List of cases in which the land grant college county grows a different lead crop from the other counties in the state in the last census before the establishment of the land grant college. Column 1 lists the name of the college, Column 2 the year of the last census before the college is established, Column 3 the lead crop in the college county, and Column 4 the lead crop in the rest of the state.