Research and Regional Development: Evidence From American Agriculture*

Shawn Kantor[†] Alexander Whalley[‡]

Abstract

While innovation is central for understanding development over time, less is known about how innovation affects development over space. In this paper we use the establishment of agricultural experiment stations to examine the effects of a permanent increase in local research on long-term regional development. Our analysis of county-level agricultural census data from 1870 to 2000 reveals that station establishment increased local land productivity over the medium term, with land one standard deviation closer to research receiving a 36% increase in productivity 20 years later. Research proximity effects disappear, however, after 30 to 50 years. Furthermore, while the effects of research proximity on extensive and intensive adjustments of producers do persist, they cannot account for the limited persistence of research proximity on land productivity. Instead, changes in the development and diffusion of basic science appear to have reduced the advantages of proximity.

*We thank Peter Berck, Alain de Janvry, Richard Hornbeck, Caroline Hoxby, Bill Kerr, Enrico Moretti, Petra Moser, Alan Olmstead, Paul Rhode, Mitchell Stevens, Bill Sundstrom, Sarah Turner, John Van Reenen, Romain Wacziarg, Fabian Waldinger, Bruce Weinberg, Heidi Williams, David Zilberman, and participants at the All-UC Economic History, Cliometric, NBER DAE, RCEF and Urban Economics Association conferences, Berkeley, Bocconi, LSE, Santa Clara, Stanford, SUNY-Binghamton and UCL seminars for helpful comments. We also thank Jeremy Atack, Price Fishback, Robert Margo, and Paul Rhode for sharing their data. We thank the National Science Foundation for financial support under grants: SES-0851788 and SBE-1158794. Chris Abrescy, Aly Anderson, Tony Hua, Michael Wall, and Marko Zivanovic provided excellent excellent research assistance.

[†]Rensselaer Polytechnic Institute & NBER.

[‡]University of California, Merced & NBER.

1 Introduction

The concentration of knowledge production in high income areas is striking. For instance, OECD countries account for 83% of the world's R&D expenditure and 98% of its patenting (OECD, 2004). Likewise, the five MSAs that produce 30% of US patents have per capita income 35% above the US average (USPTO, 2010). While knowledge production is central to understanding economic development over time, less is known about how knowledge production relates to development over space. This is especially true over the long term.

Two central challenges have limited progress. First, distinguishing knowledge spillovers from fundamentals is a tall order. Important problems arise because those who produce local spillovers also receive them (Lucas, 2009) and persistent unobserved factors can lead to the co-location of knowledge production and highly productive firms (Thompson and Fox-Kean, 2005). Second, information on producer productivity at the local level is typically unavailable over long time periods.[1] Because regional adjustments are often slow, historical settings allow us to see whether long term responses of producers amplify or reduce the long term effects of regional shocks (Davis and Wienstein 2002; Collins and Margo, 2007; Miguel and Roland, 2011; Hornbeck, 2012; Bleakley and Lin, 2012; Kline and Moretti, 2012).

In this paper we study the long term effects of agricultural research on regional development. The tacit nature of agricultural technology suggests an important role for proximity to research, as geographic proximity leads to the person to person interactions central to the diffusion of new technology.[2] We propose the establishment of federal

[1] Recent work measuring knowledge spillovers from university research to local firms has focused on a much shorter 10 to 20 year time frame (Hausman, 2012; and Kantor and Whalley, 2012). Work examining longer term knowledge spillovers has typically focused on the human capital intensive knowledge producing sector where adjustment likely occurs more quickly (Azoulay, Wang and Zivin, 2010; Waldinger, 2012; Moser, Voena, and Waldinger, 2012; Helmers and Overman, 2013).

[2] Frictions in the diffusion of agricultural technology are well documented. The speed with which new agricultural technologies are adopted depends on social contact (Munshi, 2004; Bandiera and Rasul, 2006; and Conley and Udry, 2010. Other studies demonstrate important roles for education (Foster and Rosenzweig, 1995), heterogeneous returns to technology (Suri, 2011) and the present-bias of farmers (Duflo, Kremer and Robinson, 2011) in determining agricultural technology adoption. Whether these frictions persist over longer time horizons is less known. Comin and Hobijn (2010) show the technology

agricultural experiment stations as a source of exogenous variation in the location of knowledge production. Because the stations were opened at pre-determined land grant college locations in response to nationwide concerns about agriculture, they create a positive shock to local research virtually independent of local economic conditions. Detailed data allow for an examination of research proximity effects from 1870 to the present, with an empirical analysis that compares changes in agricultural production between counties close and far from the experiment stations.

The establishment of the federal experiment stations in 1887 provided each state with physical and human capital necessary to generate the biological innovation that drove agricultural productivity growth (Olmstead and Rhode, 2008). Early research focused on finding new crop varieties, plant breeding, and animal husbandry. These innovations are readily apparent in the much higher crop yields achieved in station experiments relative to local farmers, as shown in Figure 1. Later, as basic research gained prominence, newly acquired knowledge of genetics led to important breakthroughs like hybrid corn (Griliches, 1957). While the aggregate returns to agricultural research spending are large – most estimates are over 40% (Alston et al., 2000) – no research to date has explored how proximity to research affected regional development in agriculture. Our paper seeks to fill this gap.

Our paper has three main findings. We first examine how land productivity responded to proximity to agricultural research. Our estimates indicate local spillover effects initially grew – peaking between 20-30 years after the experiment stations opened. Our peak estimates imply that being one standard deviation closer to research increased land productivity by 36 percent. While this estimate is economically significant, it suggests research proximity accounts for only a small fraction of the dispersion in land productivity across space. In our sample, the 90th - 10th percentile ratios of land productivity in 1880 is 4.5:1.[3] A one standard deviation in research proximity thus leads to productivity effects that account for less than 10 percent of this dispersion. Research proximity effects

across a range of sectors diffuses globally after 40 years.

[3]This ratios are more similar to modern developed economies than present day US firms. Syverson (2011) notes that within industry 90th - 10th percentile ratios for within industry TFP dispersion are close to 2 in the US. Hsieh and Klenow (2009) find within industry TFP 90-10 ratios of 5:1 in China and India.

do not persist indefinitely; they died out after the stations were open for 30 to 50 years.

Our land productivity results are robust to the inclusion of controls for distance to minor stations, urban markets, railways, and alternative measures of crop revenue. Furthermore, we show that geographic distance is central. Adding controls for time varying effects of distance between experiment station and far away counties in terms of land quality, agricultural technology, population literacy, and ethnicity does little to change our main findings.

We explore three opportunities to evaluate the credibility of our estimation approach. We first examine whether land productivity localization effects appear before the opening of the experiment stations. Second, we test whether distance to United States Department of Agriculture research facilities – that open much later than experiment stations – show a similar pattern of results during the period of experiment station opening. We lastly examine whether manufacturing output per worker also reveals an agricultural experiment station localization effect, where no direct effect would be expected. We find little evidence from any of these analyses that our identification strategy is violated.

There are two possibilities for the lack of persistence in research proximity effects. A first possibility is that the adjustments of producers to research proximity, that may take substantial time to occur, offset the short run benefits of proximity to research. For example, with prices for agricultural commodities set in a global market, productivity enhancing research in the United States stimulates entry into agriculture. As producers already farm the best land in an area, entry of land into agriculture will likely lower the average quality of land farmed. If extensive adjustment primarily takes place close to experiment stations, where the most productivity enhancing knowledge is received, changes in the composition of land quality could explain the lack of persistence. Of course, as the experiment stations were placed in areas where the majority of the land was already used in agriculture, the land that the experiment station research made marginally profitable may have been far from the stations.

Second, while the opening of experiment stations represented a permanent shock to local research infrastructure and funding for agricultural research permanently increased, changes in how basic science became commercialized were afoot. As the plant Plant Patent

Act enabling the patenting of new plant varieties was passed in 1930, private sector firms became increasingly involved in the commercialization of basic research. With private sector involvement the relevant locational factors not only included where knowledge was developed, but also where it was marketed to agricultural producers. Hybrid corn is a case in point. While the development of hybrid corn was based new knowledge of genetics discovered at the Connecticut experiment station in 1921, the technology diffusion curves famously presented in Griliches (1957) date the introduction of hybrid corn in a state from *commercial* introduction, beginning with Iowa in 1933. In addition, the introduction of a national agricultural extension program before 1920, that disseminated the innovations of experiment stations throughout the country, may have reduced the benefits of research proximity.

To understand limited persistence we examine the adjustments of producers to research proximity in our second set of results. We find adjustment responses on the extensive margin that decrease with proximity, and persist to the present day. These extensive adjustments may be expected as experiment stations were located in areas whether agriculture was already well established due to favourable climatic and agronomic conditions, leaving most land near the margin of profitability far from the stations. Further results indicate that research proximity led to adjustments on the intensive margin with local farmers allocating more farmland to crop production and employing more labor. The dynamics of the crop land and labor responses follow a similar patter to those for land productivity. However, they do persist until the present day, perhaps reflecting irreversibility of durable land investments.

Lastly, we examine total factor productivity and land value effects of proximity to research. We find that total factor productivity (TFP) effects take 20 years after the experiment stations opened to emerge, and have no persistence. As the TFP effects are less persistent than the the land productivity effects, slow producer adjustments cannot explain the limited persistence of the land productivity effects. Instead, the opening of the experiment stations had only a short term research proximity effect. This finding points to changes the location of commercialization of basic science or a highly effective agricultural extension program reducing the benefits of proximity to basic research. In addition we find little effect of research proximity on land values. As the land productivity localization effects are primarily due to increasingly intense use of costly inputs by local farmers, any value of proximity to basic research is small.

Our paper relates to a number of literatures. While regional productivity is central to determining spatial equilibrium, where highly mobile knowledge is produced matters little in many textbook models. More recently the implications of knowledge diffusion frictions across space have begun to be explored: Duranton (2007), Comin and Hobijn (2010), Kerr (2010), Keller and Yeaple (2012), Davis and Dingel (2012) and Comin, Dmitriev, and Rossi-Hansberg (2013). As even the largest effects of proximity to basic research we find are modest we suggest that other factors, such as where basic research is commercialized, may be more relevant over the long term than where basic research takes place.

Second, our paper speaks to the debate on internal and external drivers of firm productivity (Syverson, 2011). Recent contributions have demonstrated the importance of knowledge spillovers (Bloom, Schankerman and Van Reenen, 2012; Kantor and Whalley, 2012; Hausman, 2012), technology sourcing (Van Reenen, Harrison and Griffith, 2006) and agglomeration spillovers (Greenstone, Hornbeck and Moretti, 2010) for firm productivity. Our contribution is to show how the effects of one external driver of productivity – proximity to research – differs over the short and long term.

Lastly, we contribute to the literature on the role of reallocation in long term productivity growth. Economic historians have long debated the importance of changes in technology versus the quality of land farmed for long term trends in agricultural productivity growth (Johnson and Gustafon (1962); Olmstead and Rhode (2002)). We contribute to this literature by showing that location of reallocation can matter for the location of productivity growth over the long term.

Our paper unfolds as follows. The next section discusses the opening of the federal experiment stations, their research and discoveries. Section three presents our measurement framework. In section four we discuss the data and descriptive statistics. Section five presents our central results. Section six concludes.

2 Historical Background

This section describes the establishment, research and discoveries of the federal agricultural experiment stations in the United States.[4] Early experiment station research contributed to agricultural productivity through biological crop innovations (Olmstead and Rhode, 2008) and the application of emerging findings from chemistry (Rosenberg, 1971).[5] Summarizing the contributions of the stations, Rasmussen writes "The flowering of agricultural research in the last decades of the nineteenth century resulted in the establishment of principles that eventually affected every aspect of agricultural production and the welfare of the nation" (Rasmussen, 1962 p. 584). The high rates of return to agricultural research and development spending – many exceeding 40% – support Rasmussen's statement.[6]

American experiment stations were based on the model of the Moeckern plan in Saxony (Knoblauch et al., 1962).[7] After a trip to Germany and convincing local agriculturists of its value, Samuel Johnson founded the first American experiment station to receive state

[4]The discussion here focuses mainly on U.S. agriculture in the late 19th and early 20th century to guide our empirical analysis that examines the effects of station establishment in 1887. Excellent surveys of the role of public research in U.S. agriculture spanning both the 19th and 20th century include Huffman and Evenson, 2006 and Allston et al. (2010).

[5]While the transition from man power to animal power to mechanical power in 19th and 20th centuries dominate many textbook accounts of productivity growth in agriculture more detailed studies show that biological innovation plays a central role. For example, Olmstead and Rhode (2008, p.61) present a particularly through account of productivity improvement in in U.S. wheat production and conclude that about half of the increase in wheat productivity from 1839 to 1909 can be attributed to biological innovation. Similarly, Johnson and Gustafon (1962) show that new varieties accounted for 60 percent of the increase in yields for U.S. grain production from 1928 to 1954 in the West. After the mid-1930's national agricultural productivity growth accelerated with innovations based on an improved understanding of genetics, such as the development of hyrid corn famously studied by Griliches (1957), taking center stage.

[6]The meta-analysis of rate of return estimates in Alston et al. (2000) includes 292 studies that reported a total of 1,852 estimates of rates of return to agricultural R& D, with an overall mean internal rate of return of 81.3%, with a mode of 40%, and a median of 44.3%.

[7]The Moeckern station, perhaps the first to use the term "Agricultural Experiment Station", appointed chemists to analyse farmer's soil and fertilizer, promote experimentation as a source of farm improvement, as well as, travel throughout the district to bring scientific advice to farmers.

funds at Yale in 1875 based on the Moeckern plan. Advocates for Federal experiment stations who saw the value of scientific agriculture began creating plans for stations in the 1880s.[8] Further support was provided by land grant colleges themselves as they struggled to live up to expectations on the original Morril Act. Yet limited land grant income alone and the still quite new Department of Agriculture meant that research activity was quite modest. Opposition emerged from states rights advocates and those concerned about the oversight of Federal stations finances. Rosenberg (1964) argues that "President Cleveland had grave doubts as to the constitutionality of the Hatch Act, the first direct cash grant-in-aid to individual states." (Rosenberg, 1964; p.3). Ultimately, the broader benefits to agriculture – petitions favouring passage came from 34 states (Harding, 1947; p.175)– carried the day and the Hatch Act passed on a voice vote with widespread support in 1887.

The Hatch Act charged the stations with acquiring and spreading practical information on subjects connected to agriculture and performing original scientific research. Each state was allocated $15,000 to support investigation and distribution of the results. Experiment stations were to send all reports to local newspapers and cover the costs of mailing to farmers requesting materials. Subsequent legislation in the form of the Adams Act of 1906 further strengthened the prominence of basic research and the role of the Office of Experiment Stations as overseer of a national agricultural research program (Ferleger 1990, 20).

Notable biological innovations were achieved quite soon after the stations were established. For example, the corn trials at the Illinois Station in 1888-1895 showed that soft corn and Flint varieties had only 64 to 71 percent of yield of the Leaming variety (Steckel, 1983). Some early Wheat trails showed similar yield differences amongst varieties. In trials in the Red River Valley region between 1892-1894 China Tea was found to have only 88 percent of the yields of the Blustem and Fife varieties (Olmstead and Rhode (2008), p.33). These early advances were possible because the Hatch Act led experiment stations to be established, hire staff and begin research quickly. By the end of 1888, 46 states

[8] Rosenberg (1971) writes "The political needs of the station scientists guaranteed that the educated, adequately capitalized farmer would be their natural ally in the achievement of power. Indeed, the larger the scale of an enterprise the more likely it was in general to find experiment station scientists relevant." Rosenberg (1971, p.18)

already had stations established employing 369 staff nationwide (True, 1888).[9]

The testing of new varieties continued to be a major focus of crop research and a cornerstone of the biological innovations produced by the stations. New crop varieties better suited to climatic conditions were often found. In addition, as the productivity of planted varieties declined relatively quickly, new varieties were constantly required for farmers to simply keep productivity from falling (Olmstead and Rhode, 2002).[10] While some farmers adopted varieties shown to be superior, quickly the differences in average yields between experiment station trials and farmers in the same county we display in Figure 1 suggest that many did not.[11] While some farmers initially derided experiment station techniques as 'book farming' that lacked real world applicability, the value of new innovations was soon more broadly appreciated: "Years ago doubts of the value of experiment station work were commonly expressed among farmers... No one now doubts that such discoveries amply justified the trouble and expense involved." (Indiana Farmer's Guide, 1922).[12]

Other innovations in agriculture, such as the use of appropriate crop rotations and commercial fertilizer were built from applying new knowledge of chemistry to agricultural practices.[13] Indeed, the majority of scientists involved with experiment stations, such

[9] Prior to the Hatch Act, 12 states had already established their own experiment station and a number of others reported such work without having a formal structure governing such work (Carstensen 1960, 13). We refer to the states with pre-Hatch stations as the non-Hatch states in our empirical analysis.

[10] Crop varieties depreciate because pests and fungis adapt to infest available varieties. As a result many varieties are only is use for a short period of time. For example, Olmstead and Rhode (2008, p.32) note that in 1919 roughly 80 percent of the U.S. wheat acreage consisted of varieties that had not existed in North America before 1873.

[11] An important caveat to comparing yields from experiment trails to those of farmers in the field is that experiment station trails frequently dropped varieties that appeared to have low yields. For this reason we plot the average yield reported in each trial rather than the maximum yield to represent the state of knowledge of experiment station scientists at the time.

[12] Other contemporary accounts also note the value of experiment station research to farmers. "Intelligent farmers long since ceased to scoff at the work of the agricultural experiment stations, and the unintelligent ones may revise their opinions..." (Christian Advocate, 1901) "If we know our business we ought to keep in touch with our experiment station, not carelessly glance at the bulletins and toss them in the waste basket, but study them until we have the ideas learned right at our fingertips, using them in everyday business." (Michigan Farmer, 1904)

[13] The importance of chemistry for agriculture was well known. As Rosenberg writes: "Indeed, to many

as Samuel W. Johnson at Yale and Wilbur Olin Atwater the first director of the Office of Experiment Stations, were German trained chemists.[14] Much like with crop varieties experiment station trials with fertilizers and crop rotations showed notable effects. For example, including legumes in a wheat crop rotation rather than continuous wheat planting was shown to roughly double wheat yields in a trial in Kansas (Elwood et al., 1939, p.74).

Many important agricultural innovations came from other sources. Immigrant farmers often brought their native seeds with them which was sometimes well suited to the United States climate and soil. The introduction of Turkey hard red winter wheat of Russian origin in Kansas in 1873 (Quisenberry and Reitz, 1974), and the hardy alfalfa brought by German settlers to Minnesota in 1858 (Kletzer, 1957) are just some examples.[15] Plant breeders with another source of innovation with varieties such as Fultz wheat in 1862 or Robert Reid's Yellow Dent corn achieving prominence (Olmstead and Rhode, 2008; p. 29). Similarly, the United States Department of Agriculture's research program led to breakthroughs such as the use of airplanes to deliver insecticide and to the development of disease resistant plants (Harding, 1947).

A last element of the modern agricultural research infrastructure was the development of the agricultural extension program. Again, a number of states had already begun experimenting with small scale extension programs, such as offering short courses for farmers off campus (Smith and Wilson, 1930: p.31) before the federal policy was enacted. Seamen Knapp is generally credited with developing the first large scale co-operative demonstration program at the Porter Farm in Texas in 1903. Under Knapp's direction large scale co-operative demonstration extension spread throughout the south to combat the boll weevil. The contours of the debate for funding a federal Extension program under the Smith-Lever Act of 1914 were similar to those of the Hatch and Adam's acts. The Smith-Lever Act was rolled out across the nation quickly, partly to support the war effort, with over 2600 counties having extension agents by 1918 (Smith, 1930).

Americans agricultural science *was* chemistry..." Rosenberg (1971, p. 3, emphasis in the original).

[14] In 1930 the first year disciplinary breakdowns for station scientists are available, 40 appointments were in agricultural chemistry, 10 appointments were in bacteriology and three were in biology (Huffman and Evenson, 2006, p.66).

[15] Ethnic communities continue to be an important source of technology transfer. See Kerr(2008) for modern evidence.

3 University Research and Agricultural Production

3.1 Measurement Framework

Our empirical strategy exploits the timing of federal experiment station establishment dates within a flexible event-study framework. Formally, we first estimate the equation,

$$(1) \qquad Y_{it} = \theta_i + \gamma_t + \gamma_{Cit} + \pi_{1870} D_{1870} \times Distance_i + \sum_{j=1890}^{2000} \pi_j D_j \times Distance_i + \epsilon_{it}.$$

Here Y_{it} is the agricultural outcome in county i in years t=1870,...,2000. θ_i is a set of county fixed effects, which absorbs time-invariant differences in observable and unobservable characteristics and allows consistent estimation of π_{1870} and π_j even in the presence of fixed differences across locations. D_{1870} and $D_{1890}, ..., D_{2000}$ are dummy variables that take a value of one in the year indicated and a value of zero otherwise. γ_t is a set of year fixed effects that flexibly control for national time series variation in agricultural outcomes. γ_{C_it} is a set of year fixed effects or county characteristics in 1880-by year fixed effects that may be included depending on the specification. $Distance_i$ is the linear distance in miles between the experiment station and county i bases on the geographic coordinates of each county seat. To ease comparisons to other non-geographic distances, we normalise all the distance variables to have a mean of zero and a standard deviation of 1.

An advantage of the less-restrictive event-study specification we use is that it describes the dynamics of the experiment station effects. Our parameters of interest are $\pi_{1890}, ..., \pi_{2000}$ that measure how the relationship between distance to the experiment station and the agricultural outcome differs from the reference year 1880. If experiment stations generate local spillovers to production that persist indefinitely, then we would expect that for all years $\pi_j < 0$. However, if local experiment station effects only persisted until say 1960 then we would expect that $\pi_j < 0$ for $j < 1960$ and that $\pi_j = 0$ for $j >= 1960$. Our specification allows us to estimate how long knowledge spillovers take to emerge, and whether they grow or persist without imposing any specific dynamic structure on the effects of interest.

Our empirical approach utilizes variation in university research that is plausibly exogenous to local agriculture. As noted above, the passage of the Hatch Act in 1887 provides variation in the location of university agricultural research. As the experiment stations were opened in already chosen locations (counties with land grant colleges) the combination of timing and location of the Hatch Act spending provides a compelling source of exogenous variation in the location of university research. Our central identifying assumption is that changes in university research at land grant colleges are unrelated to changes in unobserved determinants of local agricultural development.

While it is not possible to test our identification assumption directly the use of a flexible event-study specification allows us to gauge its plausibility. Estimates of π_{1870}, π_{1890}, ..., π_{2000}, allow a visual and statistical evaluation of the evolution of pre-treatment unobservables that are correlated with distance to the experiment station. More specifically, estimates of π_{1870} also allow an explicit test of whether any "effects" preceded the opening of the experiment stations – an important falsification test.

After presenting the fully flexible event-study estimates, we summarize the magnitudes and joint statistical significance for the years 1940-2000 in a specification that replaces those individual year exposure dummies with a year group for the last periods by estimating,

$$Y_{it} = \theta_i + \gamma_{Cit} + \pi_{1870} D_{1870} \times Distance_i + \sum_{j=1890}^{2000} \pi_j \times Distance_i +$$

$$\sum_{j=1890}^{1930} \pi_j D_j \times Distance_i + \pi_{1940-2000} D_{1940-2000} \times Distance_i + \epsilon_{it}. \quad (2)$$

Here D_{1870} and $D_{1890}, ..., D_{1930}$ are again dummy variables that take a value of one in the year indicated and a value of zero otherwise, where as $D_{1940-2000}$ is a dummy variable that takes a value of one for the years 1940 to 2000 and zero otherwise. This specification allows us to study the dynamic effects of station opening for the 50 years following opening in a flexible way, while summarizing the magnitudes and joint statistical significance of any persistence beyond that period with the parameter $\pi_{1940-2000}$. We choose to present this specification for the majority of our reported results as the parsimony it provides for a more straightforward presentation of alternative specifications, while retaining the flexible structure of the dynamics for the first fifty years after the stations opened.

To explore the sensitivity of our results we add a range of covariates to the model in equation (2) to capture other time trends. We explore whether our results are robust to state times year fixed effects as controls. We examine how robust the results are to political economy concerns by adding county vote share in 1880 for president Cleveland, who initially advocated states rights over federal experiment stations, times year fixed effects. We also explore whether adding distance to markets or transportation affects our results. We add a range of year times various distance interactions. We further examine whether our results are consistent with an effect of agricultural research driven by geography alone by separating the effects of East-West versus North-South distance to the experiment station, and including time effects that depend on an array of non-geographic distances. We explain further details of these specifications in the results section below.

A few other estimation details are worth noting. All our analysis is based on the balanced panel of counties that report agricultural output every year. Each county is weighted by its total land area in 1880 so that estimates capture the effects for a typical acre of land rather than county. In all specifications, to address the possibility of persistent autocorrelation in outcomes within a county, we cluster the standard errors at the county level.

4 Data and Descriptive Statistics

Agricultural, Population and Geographic Data. The primary data we use to estimate the impact of university agricultural-related research on local agricultural development comes from the Agricultural Census (Haines, 2005). We obtain county-level data on agricultural outcomes from 1870 to 2000 from this source. We use county-level data on crop revenue, farm value, farm acreage, improved acres, and equipment value for the full sample period. Our main outcome of interest is total crop revenue per acre of farmland. This variable is collected in most, but not all, years in the agricultural census data we use. For the years 1870, 1880, 1890 and 1900 total farm revenue, but not crop revenue, is reported. Our crop revenue measure for these years is interpolated using the reported level of total farm revenue times the average fraction of revenue from crops from the county in the years where both total and crop revenue are reported.[16] Starting in 1880 we also

[16]The years where both are reported are from 1950 to 2000. The average fraction of revenue from crops

have data on acreage planted with wheat and corn, and fertilizer expenditure, allowing us to calculate crop specific yields.

We merge into our dataset variables reported at the county level from the decennial Population Census and the Fishback, Horrace and Kantor (2005, 2006) (FHK) county geography database. We measure time-invariant land suitability to capture agricultural technology as of 1880. To construct our soil suitability measure we take the prediction from a regression of all 12 county level soil quality measures, and all 13 county level climate, elevation, and geographic coordinate measures that FHK report on crop revenue per acre in the 1880 cross section. Further details of this procedure as well as the variable definitions are in the data appendix.

We also create a measure to capture the social distance between the experiment station county population and others in the state. Our measure of ethnic relatedness is based on the genetic distance measure used in recent work by Spolaore and Wacziarg (2009). This simple measure captures how far back in time the populations of two countries split off from the same initial population, so that the population of China would be more genetically distant from England than from Japan, for example. We utilize a similar approach to Spolaore and Wacziarg (2009) here based on the place of birth of the population of each county in 1880. We use the 100% sample of the 1880 census to construct a measure of the ethnic distance of the population of the reference county form the population of the experiment station county in the same state.[17] To do so we first create country of origin shares in each county as well as the experiment station county. We then use country of

in our sample is 37.6%. As our fraction of revenue for crops is based on observations from the end of our sample long term changes in crop revenue share could affect our measure. To address this concern we also consider another crop revenue measure that is the summation of the quantity of each of four crops times their national price. We use corn, wheat, oats and barley to create this second measure of crop revenue as national prices for these crops are available in each year of our sample from the historical statistics and these crops account for a large share of planted acreage in our sample states.

[17]We also explored measuring the genetic distance between station staff and a counties population by typical places of birth for station scientists. However, the name based place of birth for station scientists reveals that the vast majority are U.S. born, and so there is little variation across states. In addition, as the empirical evidence on learning frictions in agricultural technology adoption indicates that farmers learn from observing other farmers genetic differences between the county populations are likely to capture the relevant diffusion friction.

origin shares from the census data and the genetic distance between each country of origin to measure the weighted genetic distance from county i to experiment station county e as $F_{ie}^{ST} = \sum_{j=1}^{n}\sum_{k=1}^{n}(s_{j,e} \times s_{k,i} \times d_{jk})$ for birthplaces $j, k = 1, ...n$ with genetic distances between them of $d_j k$ (Spolaore and Wacziarg, 2009) to create a measure of the ethnic distance between each county in the experiment station county. We then have a measure of the relatedness or social connections in place between the experiment station county and each other county in the state.

Locational Data. The locations of the federal experiments stations are obtained from the Office of Experiment Stations (1910; p.300). While nearly all of the states in our sample have a single federal experiment station located at their land grant college there are a couple of exceptions. Connecticut, Missouri, and New York had two stations in different locations in operation in 1910.[18] For these states we denote one station as primary and the other as secondary. We denote the land grant college station as primary in these states and denote the other station as a minor station.[19] We use these locations to compute the distances from the primary and secondary experiment stations for each county. We specify the distance to the experiment station in model (1) and (2) as the distance to the primary station and add year time distance to secondary station as additional controls as a sensitivity analysis.

We next merge in data on the locations of USDA research stations, railways, and large urban markets to calculate distances to these for each county. We measure the presence of a USDA station by the location of research staff in 1929, the first year such information is available, from United States Department of Agriculture (1929). We measure the presence of a railroad in 1880 referenced against the NHGIS boundary file for 1880.[20] A county is coded as having railroad access if the county boundary is crossed by or touches a railroad.

[18]New Jersey also has two stations reported, however as they are both in the land grant college county the existence of multiple stations does not affect how we measure station location.

[19]The primary and minor stations for these states are as follows. Connecticut: Storrs (Primary) and New Haven (Minor). Missouri: Columbia (Primary) and Mountain Grove (Minor). New York: Geneva (Primary) and Ithaca (Minor). Most states that have multiple stations are in the south (i.e. Louisiana) or far west (i.e. California and Montana) and are not in our sample.

[20]The railroad mapping itself is based upon the digitized image from the Library of Congress of "Colton's railroad map of the United States" published in 1882 in the manner described in Atack, Bateman, Haines and Margo (2010). We thank Jeremy Atack for making this data available to us.

A county is coded as a large market if is in the top 5 percent in the 1880 distribution of county percentage urban. For every county we then compute distance to the nearest county with a USDA station, railroad, and urban market.

Sample Selection and Descriptive Statistics. To create our balanced panel of counties we impose the following sample restrictions. We first keep states east of the continental divide so that we are working with counties that were relatively settled in 1870 and did not experience large boundary changes over our period. We drop southern states (as defined in Olmstead and Rhode (2002) who use the regional definitions from Parker and Klein (1966)) because attributing Hatch Act funding or agricultural extension to locations is less clear in the south as some Historically Black Colleges received funding that is difficult to document. This sample of 20 states in the Northeast and Midwest has 1277 counties based on their boundaries in 1880 (Horan and Hargis, 1995).[21] We drop the 44 counties with large boundaries changes during our sample period based on Horan and Hargis (1995). Lastly, we drop counties that do not report crop revenue every year. Our final analysis sample contains a balanced panel of 1063 counties based on consistent county boundaries defined as of 1880.

Table 1 provides a first look at summary statistics of relevant measures. We report the 1880 summary statistics for the full sample and then stratify the counties based on whether their geographic distance to the experiment station was below- or above-median. As shown in Panel A, agricultural outcomes in below-median distance counties were quite different from their counterparts in counties that were further away. The comparisons indicate many differences exist with the closer counties having higher levels of crop revenue and land value per acre. These differences can also be seen visually in the map of crop revenue per acre in 1880 displayed in Figure 2. Similarly, farms further from the station have lower input intensities, fraction of county acres in a farm and smaller rural populations. In panel C we see that land closer to experiment stations is more suitable for agriculture suggesting that the differences in outcome before the experiment station opened could be partly explained by more favourable conditions near the experiment stations. Lastly, we see that counties close by the experiment station have characteristics that may allow them to absorb experiment station discoveries at lower cost, with higher

[21]The 20 states contained in our sample are: CT, IA, IL, IN, KS, MA, ME, MI, MN, MO, NE, NH, NJ, NY, OH, PA, RI, TX, VT, and WI.

levels of literacy rates and closer ethnicity to the experiment station population.

5 Results

Land Productivity. We first examine how crop revenue per farm acre responds to research proximity with the fully-flexible model in (1). We report these results in Table 2. Column (1) of the table reports the year-by-year distance interaction estimates under the baseline model. To capture underlying trends the model in column (1) of Table 2 includes year fixed effects interacted with woodland above 4% in 1880 to capture the diffusion of bared wire at this time (Hornbeck, 2010), and year fixed effects interacted with the fraction of land in farms in 1880, as well as, year fixed effects interacted total population in 1880, to capture secular trends in settlement patterns. In column (2) we exclude the interactions of initial county characteristic variables to examine a more parsimonious specification. Lastly, in column (3) we estimate the baseline model for the 10 states without any permanent experiment stations before the passage of the Hatch Act. As the opening of experiment stations in these states is completely determined by federal policy, the station openings in these states are most likely to satisfy the exclusion restriction.

The first result to note in Table 2 is that none of the 1870 times distance interactions is statistically significant. Moreover, the lack of statistical significant in the models is not driven by imprecision in the estimates as the sign of the point estimates differs across the specifications, and magnitude of the point estimates are quire small. We continue to use the model in column (1) as our baseline model throughout our analysis as the pre-trend point estimates in Table 2 are very small. These first results in Table 2 indicate that the opening of federal experiment stations did not follow changes in local agricultural production; an important falsification test for our specification.

The central results in Table 2 concern the dynamics of the research proximity effects. Across all three specifications the results are very similar. Localization effects emerge quite quickly with a distance effect first emerging and becoming statistically significant in some specifications in 1890. The distance effects grow over time until 1910 – more than 20 years after the experiment stations first opened. The effects of distance then decline and become statistically insignificant by 1920 or 1950 depending on the specification. There is little evidence of a statistically significant localization effect longer than 60 years after the experiment stations opened. Figure 3 visually depicts the year-by-year interactions for all three models along with the 95th percentile confidence interval for the estimates in the first baseline model.

The magnitude of these proximity effects, while economically significant, are not especially large. Our peak estimates in column (1) imply that that being one standard deviation closer to research increases land productivity by 36 percent. In our sample, the 90th - 10th percentile ratios of land productivity in 1880 is 4.5:1.[22] A one standard deviation in research proximity leads to productivity effects that account for less then 10 percent of this dispersion. And these are the *peak* effects. This finding suggests that research proximity accounts for a relatively small fraction land productivity across space.

Land Productivity: Alternative Specifications. We next examine the robustness of these results in Table 3 by fitting the more parsimonious equation (2). This allows us to examine a range of possible concerns with our approach thus far. We first report estimates from our baseline model for the more parsimonious equation (2) in column (1) for comparison purposes. We then present models where we further includes state-year interactions (column (2)) and percent voting Democrat in 1880-year interactions (column (3)).[23] Our results change little. Similarly, the estimates in column (4) indicate that using distance to the nearest station, regardless of whether it was in the same state or not, does little to alter the results.

While we might consider agricultural markets as largely integrated at this time, if they were not experiment station opening could affect local commodity prices and our revenue based land productivity measure, even if productivity remained unchanged. The

[22]This ratios are more similar to modern developed economies than present day US firms. Syverson (2011) notes that within industry 90th - 10th percentile ratios for within industry TFP dispersion are close to 2 in the US. Hsieh and Klenow (2009) find within industry TFP 90-10 ratios of 5:1 in China and India.

[23]One important issue in the 1880 Presidential election was railroad regulation, often pushed for by agricultural interests at the time to lower transportation prices. We thus use vote share in 1880 to measure voters agriculture related policy preferences before the Hatch Act passed to address potential policy endogeneity.

direction of any bias depends on the importance of local fiscal versus productivity effects of experiment stations. On the one hand, if the opening of the station was primarily a local fiscal stimulus we may expect local prices to increase. This would lead farm revenue to run counter to any productivity effects, and lead us to underestimate the productivity effects of the experiment station. On the other hand if experiment stations primarily increase productivity, then local agricultural prices may be expected to fall. In this case our estimates would overestimate the productivity effects of experiment stations. To address this issue we construct an alternative measure of crop productivity that is based on county level quantities of the three crops reported consistently throughout our sample (corn, wheat and barley) times their national prices which are unaffected by local shocks. When we estimate our model using this measure (in column (5)) we obtain similar but slightly larger estimates, perhaps indicating a small short term fiscal impact of experiment stations.[24]

Similarly, the results in columns (6), (7) and (8) of Table 3 show that controlling for distance to minor stations, railways or urban markets does little to change the results. As minor stations did produce a modest amount of research and market access can be an important driver of regional growth (i.e. Redding and Strum, 2008) we see these as important robustness analyses.

In the last two columns of Table 3 we examine the effects of research proximity for the yields of two main crops: corn and wheat. While examining the effects of research proximity on crop yields is a more direct way to measure productivity spillovers, there are two challenges in this context. First, while the agricultural census recorded information on each farm's output of corn and wheat in 1870, the acreage planted for each crop was not recorded. It is thus not possible to calculate the crop yield for each crop in 1870, making the examination of prior trends in crop specific productivity impossible. A second challenge is that agricultural crops are geographically specialized in certain areas based on the suitability of the soil and climatic conditions. We thus examine the responses of each crop within its region of production using the Parker and Klein (1966)'s corn states for our corn productivity analysis and grain states for wheat analysis.[25]

[24] See Hornbeck and Keskin (2012) for evidence that spillovers from the Ogallala Aquifer primarily had short term impacts on regional agriculture.

[25] We do this so that we are focusing on counties where the crop is likely to be planted each year and

The effects of research proximity on crop productivity in columns (9) and (10) are relatively modest. The effects are stronger and similar to crop revenue per acre for corn, while appearing later for wheat. A central difference between the crop yields and the revenue per farm acre results is likely due to how farm land is utilized, and how land use responds to research proximity. We return to this issue later.

Land Productivity: Future USDA Research and Manufacturing Comparisons. One possible concern with our estimation approach thus far is that perhaps local trends in unobservable determinants of agricultural productivity drove the opening of the experiment stations at land grant colleges. For example, if local producers who were experiencing productivity gains were able to determine the locational choice of research facilities our identification strategy may be threatened. To evaluate the importance of this concern we conduct two additional analyses.

We first examine whether each county's distance to research stations opened by the USDA much later display similar dynamics around the experiment station opening in 1887.[26] The results for this analysis are reported in the first three columns of Table 4 where the distance included in model (1) is now the distance to the nearest USDA station open in 1929. Comfortingly we find little evidence that the distance to the future USDA station mattered where the experiment station effects are prominent. We see some distance proximity effects for the USDA facilities in the 1950's and 1960's. This would be consistent with our main results showing experiment station effects peak 20-30 years after the stations opened, if most USDA stations opened in the 1920s. Of course as we do not know the precise date each USDA station opened and USDA facility openings could be endogenously related to agricultural outcomes (i.e. in responses to plant disease outbreaks) we cannot be more definitive. In any case, the results showing little dynamics

the experiment stations are likely to specialize in research related to that crop for time-invariant reasons. We follow Olmstead and Rhode (2008) in using the Parker and Klein (1966) crop region classifications. The corn states in our sample are: IA, IL, IN, MO, and OH. The grain states in our sample are: KS, NE and TX.

[26]We measure USDA station location as of 1929, the earliest year we have been able to find a comprehensive source. While the dates that the facilities opened is not recorded the legislative history of USDA research suggests most facilities would have opened after the Bureau of Plant Industry was established in 1901. Most opening probably occurred 10 to 20 years after this as the USDA began to play a larger role in inspection and conservation.

near the dates of experiment station opening are robust across all three columns. We display them visually in Figure 4.

We also examine whether manufacturing productivity respond to experiment station opening in an similar fashion to agricultural land productivity. As manufacturing productivity should not be directly affected by the agricultural innovations from the experiment stations we should expect little response if the results thus far are due to research.[27] We report the result of fitting equation (1) with manufacturing output per worker as the outcome variable in columns (4)-(6) of Table 4. We find statistically significant negative point estimates early in our sample period. Importantly, the pattern of point estimates and statistical significant show little effect of experiment station opening, suggesting that the prior results cannot be accounted for by changes by correlated changes in agricultural research. Instead, the manufacturing results indicate very stable distance effects that do not change when the experiment station opening, consistent with an educational effect that diminishes over time. These results are displayed visually in Figure 5.

Land Productivity: Geographic versus Other Distances. The results thus far demonstrate that the productivity of agricultural land geographically close to experiment stations is enhanced when the stations open. However as producers in counties close and far from experiment stations differ along a number of dimensions that can be related to regional productivity growth, we might wonder if these other differences account for the research proximity effect. For example, recent research has indicated central roles for education (Glaeser and Mare, 2001; Glaeser and Gottlieb, 2009; Gennaioli, La Porta, Lopez-de-Silanes, and Shleifer, 2012), and ethnic distance to technology (Spolare and Wacziarg, 2009) in explaining development. To address this question we examine whether our research proximity results are robust to controlling for time-varying relationships between other distances to the experiment station county and land productivity.

We present the results of our analysis in Table 5. In columns (1) and (2) we first consider the possibility that the effects of agricultural research differ across space. As agricultural innovations often have local applicability they may be difficult to transfer

[27]Of course linkages between agriculture and manufacturing could lead to indirect effects. The role of structural transformation away from agriculture in the United States is related increasing urbanization (Micheals, Rauch and Redding, 2012), and regional income convergence (Caselli and Colman, 2001).

to further away conditions profitably, making them especially valuable for local farmers. Alternatively, policymakers could value innovations that make distant marginal land more productive, and push experiment station scientist in that direction. To examine whether heterogeneous effects of agricultural innovation can explain our results we add year times distance in terms of the suitability of land and the initial level of productivity interactions to the model. We find that adding time varying controls for these distances does little to alter our central results.

We next examine whether differences between experiment station counties and other counties in terms of their knowledge absorption costs can explain the results we find. To do so we add distances between the experiment station county and each county in terms of literacy rates and ethnic background. In columns (3) and (4) of Table 5 we find that adding these controls does little to alter the geographic based research proximity results above.

Agricultural technology is generally more transferable to areas with similar latitudes than those with similar longitudes (Diamond, 1997). This suggests a natural test for research proximity effects. If the results thus far are driven by agricultural research from the experiment stations then east-west distance to the experiment station should be more important that north-south distance to experiment station. The results in Table 5 column (5) are consistent with this expectation, and provide more direct evidence of a research proximity based effect.

Extensive and Intensive Adjustment Responses. Adjustments made by agricultural producers to the location of knowledge production may be crucial for understanding the longer term localization effects of university research. Persistent extensive margin responses could account for our results. If nearby producers respond to persistent productivity spillovers from research proximity by bringing lower quality less productive land into production over a longer time horizon this could counteract any productivity spillovers. However, as areas nearby experiment stations have extensive agricultural land use before experiment stations opened, land made marginally productive by experiment station innovation may not be near the stations.

We first examine these extensive margin responses in Table 6 column (1). Our results reveal that entry of land into agriculture occurs far from experiment stations. That these extensive response also happen quickly and persist suggests that responses on the extensive margin are unlikely to explain the lack of persistence in the effect of research proximity on land productivity.

We next turn examining intensive adjustment responses across a range of margins. The results in columns (2)-(5) reveal that producers to adjust to research proximity on some intensive margins. We find research proximity increases both the intensity of land used for crops as well as the size of the rural population, and that both responses persist to the present day. These results echo Foster and Rosenzweig (1996) who find that new agricultural technologies lead to the accumulation of complementary inputs, in their case education. We find little effects of research proximity on the intensity of farm equipment or fertilizer use per acre in columns (4) and (5) however, suggesting these inputs were less complementary to the biological innovations produced by the experiment stations.

Total Factor Productivity and Land Values. How important are adjustments of producers in explaining the effects of research proximity on land productivity? In particular can they account for the lack of persistence in effects of proximity to a permanent research facility? To examine this issue we follow Hornbeck, Moretti and Greenstone (2010) and examine the total factor productivity responses. While the use of county-level production data on a limited number of inputs renders the analysis necessarily more crude than a study of individual firm productivity with modern data, we view it as worthwhile to understand how meaningful the adjustments are. Even more so, as many estimates of knowledge spillovers are based on single factor productivity measure – like our land productivity estimates above– it is important to know how much of these responses are due to input responses and how much are total factor productivity spillovers.

We estimate the effects of proximity to research on total factor productivity in Table 7. The results in column (1) differ from the land productivity results above in two key ways. First, TFP effects take more time to manifest than the land productivity effects. This suggests that producers adjustments are highly responsive to research proximity. A second feature of the TFP effects is that the magnitude of the 1910 point estimate is less than half of the land productivity results above.

Most centrally our TFP results show no more persistence than our land productivity results above. Thus the adjustments of producers to research proximity, while persistent, do not account for the lack of persistent land productivity effects. Our results are more consistent with research proximity producing a temporary local TFP shock. That the commercialization of basic research no longer was co-located with where the basic science was discovered and the introduction of the extension program seems to have reduced the impact of proximity to basic research.

One potential caveat with our TFP estimates is that they do not address the endogeneity of input choices by producers. We address this issue by examining how robust our estimates are to the use of methods using the adjustment of investment (Olley and Pakes, 1996) to control for the endogeneity of input choices. To do so we add fourth degree polynomials of log capital and log investment, and their interactions, for both equipment and land. These results are reported in column (2), and are very similar to our baseline estimates. We also follow Levinsohn and Petrin (2003) in using material inputs to address endogeneity concerns. Here we use fertilizer as the relevant material input by adding fourth degree polynomials of log equipment, log cropland, and log fertilizer, as well as their interactions to the model in column (3). Again the results are very similar to our baseline estimates.

Our last analysis examines whether land values respond to the production of university research. As some input responses may be unobserved the use of land values allows us to further assess whether our small and temporary TFP results are robust to concerns about unobserved inputs. Of course, as land supply is relatively elastic (due to significant amounts of public land being available for claim) we might expect land value effects to be smaller than any TFP effects. We present the land value effect estimates in the last column of Table 7. The results reveal little land value impacts. This finding further strengthens the case for small TFP effects of research proximity that do not persist.

6 Conclusion

This paper uses the establishment of federal agricultural experiment stations in 1887 to provide new evidence on the effects of university research on long-term regional development. Our analysis of county-level agricultural census data from 1870 to 2000 reveals that station establishment increased local land productivity for 20-40 years, but research proximity effects do not persist to the present day. Permanent changes the location of knowledge production do not have permanent effects of the geography of economic activity, and contrast with recent evidence on the persistence economic effects of cultural traits and environmental shocks.

Our analysis also highlights the role of producer adjustments to knowledge production. We find that producers engage in meaningful adjustments in response to research proximity that persist to the present day. However, these persistent adjustment do not account for the lack of persistence in the effects of proximity to research. Indeed TFP effects of research proximity are notably smaller than those for land productivity, and display no further persistence.

Many argue that policies to create knowledge hubs are central to regional economic development (Saxenian, 1994). Our results echo Durranton (2011) in casting doubt upon this view. Two features of our context are worth pointing out, however. First, the historical record and our TFP estimates indicate the agricultural experiment stations produced only transitory local productivity effects. Research facilities that produce permanent local productivity shocks could well have longer term impacts on regional development. Second, widely dispersed agricultural production may not capture effects for spatially concentrated sectors, if strong agglomeration effects are necessary for persistent local knowledge spillovers to emerge. Care should be taken in transporting our results to other settings with much different innovation processes or stronger agglomeration economies. We view examining these issues in other contexts as an important direction for future work.

7 References

Azoulay, Pierre, Joshua S. Graff Zivin and Jialan Wang (2010) "Superstar Extinction," *Quarterly Journal of Economics*, 125(2): 549-589.

Alston, J.M., C. Chan-Kang, M.C. Marra, P.G. Pardey, and T.J. Wyatt (2000), "A Meta Analysis of Rates of Return to Agricultural R & D: Ex Pede Herculem?" Washington D.C., IFPRI Research Report No. 113.

Alston, J.M., M.A. Andersen, J.S. James and P.G. Pardey. (2010) *Persistence Pays: U.S. Agricultural Productivity Growth and the Benefits from Public R D Spending*, New York: Springer.

Atack, Jeremy, Fred Bateman, Michael Haines and Robert Margo (2010) "Did Railroads Induce or Follow Economic Growth? Urbanization and Population Growth in the American Midwest, 1850-1860," *Social Science History*, 34(2): 171-97.

Bandiera, Oriana and Imran Rasul (2006) "Social Networks and Technology Adoption in Northern Mozambique," *Economic Journal*, 116(514): 869-902.

Bleakley, Hoyt and Jeffrey Lin (2012) "Portage and Path Dependence," *Quarterly Journal of Economics*, 127(2): 587-644.

Bloom, Nicholas, Mark Schankerman and John Van Reenen (2012) "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, forthcoming.

Caselli, Francesco and Wilbur John Coleman (2001) "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation," *Journal of Political Economy*, 109(3), pp. 584-616.

Carstensen, Vernon (1960) "The Genesis of an Agricultural Experiment Station, *Agricultural History*, 34 (Jan), pp. 13-20.

Collins, William J. and Margo, Robert A. (2007) "The Economic Aftermath of the 1960s Riots in American Cities: Evidence from Property Values," *The Journal of Economic History*, 67(4), pp. 849-883.

Comin, Diego and Bart Hobijn (2010) "An Exploration of Technology Diffusion," *American Economic Review*, 100(5): 2031-59.

Comin, Diego Mikhail Dmitriev and Esteban Rossi-Hansberg (2013) "The Spatial Diffusion of Technology," *Princeton University*, Working Paper.

Conley, Timothy G. and Christopher R. Udry (2010) "Learning about a New Technology: Pineapple in Ghana," *American Economic Review*, 100(1): 35-69.

Davis, Donald R. and David E. Weinstein (2002) "Bones, Bombs, and Break Points: The Geography of Economic Activity," *American Economic Review*, 92(5): 1269-1289.

Davis, Donald R. and Jonathan I. Dingel (2012) "A Spatial Knowledge Economy," *National Bureau of Economic Research*, Working Paper no. 18188.

Duflo, Esther, Michael Kremer, and Jonathan Robinson (2011) "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." *American Economic Review*, 101(6): 2350-90.

Duranton, Gilles (2007) "Urban Evolutions: The Fast, the Slow, and the Still," *American Economic Review*, 97(1), pp. 197-221.

Duranton, Gilles (2011) "California Dreamin: The Feeble Case for Cluster Policies," *Review of Economic Analysis*, 3(1): 3-45.

Elwood, Robert B., Lloyd E. Arnold D Clarence Schmutz and Eugene G. Mckibben (1939) "Changes in Technology and Labor Requirements in Crop Production: Wheat and Oats," Works Progress Administration, National Research Project, Report A-10, Philadelphia, PA

Fishback, Price V., Horrace, William C. and Shawn Kantor (2005) "Did New Deal Grant Programs Stimulate Local Economies? A Study of Federal Grants and Retail Sales During the Great Depression," *Journal of Economic History*, 65(1): 36-71.

Fishback, Price V., Horrace, William C. and Shawn Kantor (2006) "The impact of New Deal expenditures on mobility during the Great Depression," *Explorations in Economic History*, 43(2): 179-222.

Foster, Andrew D and Rosenzweig, Mark R. (1995) "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*, 103(6): 1176-1209.

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez-de-Silanes and Andrei Shleifer (2011) "Human Capital and Regional Development," *Working Paper* 17158, National Bureau of Economic Research.

Glaeser, Edward L. and David C. Mare (2001) "Cities and Skills," *Journal of Labor Economics*, 19(2):316-342.

Glaeser, Edward L. and Joshua D. Gottlieb (2009) "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States," *Journal of Economic Literature*, 47(4): 983-1028.

Greenstone, Michael, Richard Hornbeck and Enrico Moretti (2010) "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," *Journal of Political Economy*, 118(3): 536-598.

Griliches, Z. (1957) "Hybrid corn: An exploration in the economics of technological change." *Econometrica*, 25, pp. 501-522.

Haines, Michael R. (2010) "Historical, Demographic, Economic, and Social Data: The United States, 1790-2002," ICPSR02896-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-05-21.

Harding, T. Swann (1947) *Two Blades of Grass: A History of Scientific Development in the U.S. Department of Agriculture*, University of Oklahoma Press, Norman: OK.

Hausman, Naomi (2012) "University Innovation, Local Economic Growth, and Entrepreneurship," US Census Bureau Center for Economic Studies Paper No. CES-WP- 12-10.

Helmers, Christian and Henry Overman (2013) "My precious! The location and diffusion of scientific research: evidence from the Synchrotron Diamond Light Source", *Universidad Carlos III de Madrid*, Working Paper.

Horan, Patrick M., and Peggy G. Hargis (1995) "County Longitudinal Template, 1840-1990" [Computer file]. ICPSR06576-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

Hornbeck, Richard (2010) "Barbed Wire: Property Rights and Agricultural Development," *Quarterly Journal of Economics*, 125(2): 767-810.

Hornbeck, Richard (2012) “The Enduring Impact of the American Dust Bowl: Short- and Long-Run Adjustments to Environmental Catastrophe,” *American Economic Review*, 102(4): 1477-1507.

Hornbeck, Richard and Pinar Keskin (2012) “Does Agriculture Generate Local Economic Spillovers? Short-run and Long-run Evidence from the Ogallala Aquifer, National Bureau of Economic Research, Working Paper no. 18416.

Hsieh, Chang-Tai and Peter J. Klenow (2009) “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics*, 124(4), pp. 1403-1448.

Huffman, W.E. and R.E. Evenson (2006) *Science for Agriculture: A Long Term Perspective*, Blackwell Publishing, Ames:IA.

Johnson, D. Gale and Robert L. Gustafson (1962)*Grain Yields and American Food Supply: An Analysis of Yield Changes and Possibilities*, Chicago: University of Chicago Press.

Kantor, Shawn and Alexander Whalley (2012) “Knowledge Spillovers from Research Universities: Evidence from Endowment Value Shocks,” *Review of Economics and Statistics*, forthcoming.

Keller, Wolfgang and Stephen R. Yeaple (2012) “The Gravity of Knowledge,” *American Economic Review*, forthcoming.

Kerr, William R. (2010) “Breakthrough Inventions and Migrating Clusters of Innovation,” *Journal of Urban Economics*, 67(1): 46-60.

Kletzer, Frank (1957) “The History of Grimm Alfalfa,” University of Minnesota Extension.

Kline, Patrick and Enrico Moretti (2012) “Local Economic Development, Agglomeration Economies and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority,” *Working Paper*, University of California - Berkeley.

Knoblach, H.C, E.M. Law, and W.P. Meyer (1962) *State Experiment Stations: A History of Research Policy and Procedure*, United States Department of Agriculture, Miscellaneous Publication 904; Washington, DC.

Levinsohn, James and Amil Petrin (2003) "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70(2), pp. 317-342.

Lucas, Robert E. (2009) "Ideas and Growth," *Economica*, 76 (2009), 119.

Michaels, Guy, Ferdinand Rauch and Stephen J. Redding (2012) "Urbanization and Structural Transformation," *Quarterly Journal of Economics*, 127(2), pp. 535-586.

Miguel, Edward and Roland, Gerard (2011) "The Long-run Impact of Bombing Vietnam," *Journal of Development Economics*, 6(1): 1-15.

Moser, Petra, Alessandra Voena, and Fabian Waldinger (2012) "German-Jewish Emigres and U.S. Invention," *Working Paper*, Stanford University.

Munshi, Kaivan (2004) "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution," *Journal of Development Economics*, 73(1): 185-213.

Olley, Steven G., and Ariel Pakes (1996) "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64(6), pp. 1263-98.

Olmstead, Alan L. and Paul W. Rhode (2002) "The Red Queen and the Hard Reds: Productivity Growth in American Wheat, 1800 1940," *The Journal of Economic History*, 62(4), pp. 929-966.

Olmstead, Alan L. and Paul W. Rhode (2008) *Creating Abundance: Biological Innovation and American Agricultural Development*, Cambridge University Press, New York, NY.

Organization for Economic Cooperation and Development (2004) *Science and Technology Indicator Scoreboard*.

Parker, William N. and Judith L.V. Klein (1966) "Productivity Growth In Grain Production in the United States, 1840-60 and 1900-10," in *Output, Employment and Productivity in the United States After 1800*, 20: 532-80. New York: Columbia University Press for the National Bureau of Economic Research.

Quisenberry, K. S. and L. P. Reitz (1974) "Turkey Wheat: The Cornerstone of an Empire," *Agricultural History*, 48(1), (Jan., 1974), pp. 98-110.

Rasmussen, Wayne D. (1962) "The Impact of Technological Change on American Agriculture, 1862-1962," *Journal of Economic History*, 22(4), pp. 578-591.

Redding,Stephen J. and Daniel M. Sturm (2008) "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review*, 98(5): 1766-97.

Rosenberg, Charles E. (1964) "The Adams Act: Politics and the Cause of Scientific Research," *Agricultural History*, 38 (Jan), pp. 3-12.

Rosenberg, Charles E. (1971) "Science, Technology, and Economic Growth: The Case of the Agricultural Experiment Station Scientist, 1875-1914," *Agricultural History*, 45(1), pp. 1-20.

Saxenian, AnnaLee (1994) *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*, Cambridge, MA: Harvard University Press.

Smith, Clarence Beaman, and Meredith Chester Wilson (1930) *The Agricultural Extension System of the United States*, John Wily and Sons, New York: NY.

Spolaore, Enrico and Romain Wacziarg (2009) "The Diffusion of Development," *Quarterly Journal of Economics*, 124(2): 469-529.

Steckel, Richard (1983) "The Economic Foundations of East-West Migration During the Nineteenth Century," *Explorations in Economic History*, 20(1), pp 14-36.

Sunding, and David Zilberman (2001) "The Agricultural Innovation Process: Research and Technology Adoption in a Changing Agricultural Sector," *Handbook of Agricultural Economics*, Volume 1, Edited by B. Gardner and G. Rausser, pp. 207-261.

Suri, T. (2011) "Selection and Comparative Advantage in Technology Adoption", *Econometrica*, 79(1): 159-209.

Synder, Thomas D. (1993) *120 Years of American Education: A Statistical Portrait*, U.S. Department of Education, National Center for Education Statistics, Washington DC.

Syverson, Chad (2011) "What Determines Productivity?," *Journal of Economic Literature*, 49(2), pp. 326-65.

Thompson, Peter and Melanie Fox-Kean (2005) “Patent Citations and the Geography of Knowledge Spillovers: A Reassessment,” *American Economic Review*, 95(1), pp. 450-460.

True, A.C. (1888) ”Origin and Development of the Agricultural Experiment Stations in the United States,”*Annual Report*, U.S. Department of Agriculture, Washington, DC, pp. 541-547.

United States Department of Agriculture (1929) “List of Technical Workers in the Department of Agriculture and Outline of Department Functions,” Miscellaneous Publication N. 63, Washington, DC.

United States Office of Experiment Stations (1910) *Annual Report of the Office of Experiment Stations, June 30, 1910. Washington*, U.S. Department of Agriculture. UNT Digital Library. http://digital.library.unt.edu/ark:/67531/metadc5000/. Accessed October 17, 2012.

United States Trademark and Patent Office (2010) “Calendar Year Patent Statistics: By Metropolitan Area ”, accessed at: http://www.uspto.gov/web/offices/ac/ido/oeip/taf/reports.htm

Van Reenen, John, Rupert Harrison and Rachel Griffith (2006) “How Special Is the Special Relationship? Using the Impact of U.S. R& D Spillovers on U.K. Firms as a Test of Technology Sourcing,” *American Economic Review*, 96(5), pp. 1859-1875.

Waldinger, Fabian (2012) “Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany,” *Review of Economic Studies*, 79(2), pp.838-861.

Wright, Brian D. (2012) “Grand missions of agricultural innovation,” *Research Policy*, 41(10), December, pp. 1716-1728.

TABLE 1: Descriptive Statistics, 1880

	Full Sample	Distance from Experiment Station: Below Median	High Median	t-stat [p-value] (2)-(3)
	(1)	(2)	(3)	(4)
Panel A: Crop Production and Farm Values				
Crop Revenue Per Farm Acre	4.15	4.94	3.51	-5.69
	(2.68)	(2.36)	(2.76)	[0.000]
Corn Produced Per Corn Acre	28.24	32.47	24.78	-6.82
	(11.17)	(8.02)	(12.14)	[0.000]
Wheat Produced Per Wheat Acre	12.22	13.62	10.98	-6.04
	(4.68)	(4.26)	(4.68)	[0.000]
Farm Value Per Farm Acre	20.68	25.87	16.51	-6.55
	(18.26)	(18.28)	(17.14)	[0.000]
Panel B: Agricultural Inputs				
Cropland Acre Per Farm Acre	0.56	0.64	0.49	-6.30
	(0.23)	(0.18)	(0.25)	[0.000]
Farm Equipment Value per Farm Acre	0.86	1.05	0.71	-6.33
	(0.60)	(0.54)	(0.60)	[0.000]
Fertilizer Expenditure Per Farm Acre	0.03	0.04	0.02	-3.43
	(0.08)	(0.10)	(0.06)	[0.001]
Rural Population Per Farm Acre	0.07	0.07	0.08	0.71
	(0.08)	(0.04)	(0.10)	[0.475]
Farm Acre Per County Acre	0.60	0.71	0.51	-5.68
	(0.31)	(0.25)	(0.33)	[0.000]
Panel C: Other Characteristics				
Agricultural Land Suitability	1.12	1.44	0.87	-4.64
	(0.84)	(0.46)	(0.98)	[0.000]
Any Literacy Rate	0.84	0.90	0.80	-4.33
	(0.18)	(0.12)	(0.20)	[0.000]
Ethnic Distance to US Population	0.70	-0.12	1.35	2.31
	(2.89)	(0.06)	(3.76)	[0.021]
County Observations	1063	532	531	

Notes: Authors' calculations with 1880 county data unless indicated otherwise, as described in the text. The unit of observation is county. The main entries in column (1) present the mean of the selected variables for all counties. The main entries in column (2) present the mean of the selected variables for counties below median distance from the experiment station. The main entries in column (3) present the mean of the selected variables for counties above median distance from the experiment station. The standard deviations of the selected variable are presented in parenthesis in columns (1)-(3). The main entries in column (4) present the test statistics for a test of differences in means between column (2) and (3), with the p-value of the test presented in square brackets. All monetary values are expressed in 1880 $.

TABLE 2: Research Proximity and Land Productivity

Dependent Variable=	Log(Crop Revenue Per Farm Acre)		
Specification:	Baseline	No Initial Conditions	Baseline: Hatch
	(1)	(2)	(3)
1870 × Station Distance	0.01	0.06	-0.02
	(0.05)	(0.06)	(0.06)
1890 × Station Distance	-0.07*	-0.04	-0.07*
	(0.04)	(0.04)	(0.04)
1900 × Station Distance	-0.24***	-0.20**	-0.28***
	(0.09)	(0.09)	(0.10)
1910 × Station Distance	-0.36***	-0.27**	-0.40***
	(0.11)	(0.12)	(0.10)
1920 × Station Distance	-0.26**	-0.14	-0.29**
	(0.12)	(0.13)	(0.13)
1930 × Station Distance	-0.20*	-0.07	-0.20
	(0.11)	(0.12)	(0.12)
1940 × Station Distance	-0.20*	-0.12	-0.22*
	(0.09)	(0.12)	(0.11)
1950 × Station Distance	-0.05	0.01	-0.03
	(0.12)	(0.11)	(0.14)
1960 × Station Distance	-0.05	0.00	-0.03
	(0.10)	(0.09)	(0.12)
1970 × Station Distance	-0.13	-0.06	-0.10
	(0.10)	(0.10)	(0.12)
1980 × Station Distance	-0.13	-0.08	-0.10
	(0.10)	(0.11)	(0.14)
1990 × Station Distance	-0.13	-0.08	-0.09
	(0.10)	(0.10)	(0.12)
2000 × Station Distance	-0.17*	-0.12	-0.14
	(0.10)	(0.10)	(0.12)
Year FEs	yes	yes	yes
County FEs	yes	yes	yes
1880 IC × Year FEs	yes	no	yes
Sample States	All	All	Hatch
R^2	0.92	0.91	0.92
County Observations	1063	1063	723

Notes: Authors' calculations with county data from 1870 to 2000 as described in the text. The unit of observation is a county-year. Each column reports the results from one regression, weighted by 1880 county land area. The main entries in the first row of columns (1)-(3) report estimates of β_t from equation (1) in the text. Standard errors clustered at the county level reported in parentheses. The excluded year interaction is 1880. The models in all columns include additional controls for year and county fixed effects. The model in column (1) and (3) also includes controls for: 1880 woodland fraction × year fixed effects, 1880 population × year fixed effects, and 1880 fraction of land in farms × year fixed effects. * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

TABLE 3: Research Proximity and Land Productivity: Alternative Specifications

Dependent Variable=	Log(Crop Revenue Per Farm Acre)								Log (Corn Output Per Corn Acre)	Log (Wheat Output Per Wheat Acre)
						Add Year-Distance To *x* Interaction:				
Specification:	Baseline	State × Year	1880 PE × Year	Closest Station	National Price Revenue	Minor Station	Railway	Urban Center	Baseline	Baseline
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1870 × Station Distance	0.01	-0.06	-0.04	0.03	0.00	-0.02	0.00	0.06		
	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)		
1890 × Station Distance	-0.07*	-0.07**	-0.07**	-0.05	-0.07	-0.08***	-0.07*	-0.09**	0.01	-0.11*
	(0.04)	(0.03)	(0.03)	(0.04)	(0.07)	(0.03)	(0.04)	(0.04)	(0.01)	(0.06)
1900 × Station Distance	-0.24***	-0.26***	-0.26***	-0.22**	-0.47***	-0.27***	-0.24***	-0.22***	-0.02*	-0.09
	(0.09)	(0.10)	(0.09)	(0.10)	(0.15)	(0.09)	(0.09)	(0.08)	(0.01)	(0.07)
1910 × Station Distance	-0.36***	-0.35***	-0.36***	-0.36***	-0.45***	-0.33***	-0.36***	-0.35***	-0.04***	0.03
	(0.11)	(0.10)	(0.10)	(0.10)	(0.14)	(0.10)	(0.10)	(0.09)	(0.01)	(0.04)
1920 × Station Distance	-0.26**	-0.28**	-0.27**	-0.25**	-0.20	-0.27**	-0.26**	-0.25**	-0.05***	-0.06
	(0.12)	(0.12)	(0.12)	(0.12)	(0.17)	(0.11)	(0.11)	(0.11)	(0.02)	(0.05)
1930 × Station Distance	-0.20*	-0.20**	-0.21**	-0.20*	-0.45*	-0.24**	-0.21*	-0.19*	-0.02	-0.13***
	(0.11)	(0.10)	(0.10)	(0.12)	(0.24)	(0.10)	(0.11)	(0.11)	(0.02)	(0.04)
[1940-2000] × Station Distance	-0.12	-0.10	-0.14	-0.11	-0.31*	-0.18**	-0.13	-0.06	-0.01	-0.04*
	(0.10)	(0.10)	(0.10)	(0.11)	(0.19)	(0.08)	(0.10)	(0.11)	(0.01)	(0.02)
Year FEs	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
County Fes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
1880 IC × Year FEs	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Sample States	all	all	all	all	all	all	all	all	corn	grain
R^2	0.92	0.93	0.92	0.92	0.87	0.92	0.92	0.92	0.93	0.76
County Observations	1063	1063	1063	1063	1063	1063	1063	1063	493	219

Notes: Authors' calculations with county data from 1870 to 2000 as described in the text. The unit of observation is a county-year. Each column reports the results from one regression, weighted by 1880 county land area. The main entries in the first row of columns (1)-(10) report estimates of β_t from equation (2) in the text. Standard errors clustered at the county level reported in parentheses. The excluded year interaction is 1880. The models in all columns include additional controls for year and county fixed effects, 1880 woodland fraction × year fixed effects, 1880 population × year fixed effects, and 1880 fraction of land in farms × year fixed effects. The model in column (3) also includes state × year fixed effects. The model in column (4) also includes 1880 democratic vote share × year fixed effects. The model in column (4) uses a station distance measure based on the nearest station regardless of state boundaries. The model in column (5) defines crop revenue based on crop quantity times national prices. The model in column (6) includes distance to secondary station × year fixed effects. The model in column (7) includes distance to an 1880 railway × year fixed effects. The model in column (9) includes distance to an 1880 urban center × year fixed effects. The sample in column (9) is the Parker-Klein (1966) Corn States: IA, IL, IN, MO, and OH. The sample in column (10) is the Parker-Klein (1966) Grain States: KS, NE and TX. * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

TABLE 4: Research Proximity and Land Productivity: Falsification

Dependent Variable=	Log(Crop Revenue Per Farm Acre)			Log(Manufacturing Revenue Per Worker)		
Distance to Station Type:	USDA (1929) Station			Agricultural Experiment Station		
Specification:	Baseline	No Initial Conditions	Baseline: Hatch	Baseline	No Initial Conditions	Baseline: Hatch
	(1)	(2)	(3)	(4)	(5)	(6)
1870 × Station Distance	-0.08*	-0.01	-0.13**	-0.07***	-0.10***	-0.09***
	(0.05)	(0.04)	(0.06)	(0.02)	(0.02)	(0.02)
1890 × Station Distance	-0.07*	-0.05	-0.12**	-0.06**	-0.07**	-0.09**
	(0.04)	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)
1900 × Station Distance	-0.03	-0.05	-0.07	-0.03	-0.01	-0.06**
	(0.09)	(0.08)	(0.12)	(0.02)	(0.03)	(0.02)
1910 × Station Distance	-0.05	-0.03	-0.12	-0.05*	-0.04	-0.07**
	(0.13)	(0.15)	(0.18)	(0.03)	(0.03)	(0.03)
1920 × Station Distance	-0.09	-0.01	-0.17	-0.06*	-0.05	-0.07**
	(0.12)	(0.11)	(0.16)	(0.03)	(0.03)	(0.03)
1930 × Station Distance	-0.14	-0.03	-0.22	-0.06*	-0.03	-0.05
	(0.11)	(0.11)	(0.15)	(0.04)	(0.03)	(0.04)
1940 × Station Distance	-0.11	-0.06	-0.20	0.01	0.03	0.00
	(0.11)	(0.11)	(0.15)	(0.02)	(0.02)	(0.03)
1950 × Station Distance	-0.16*	-0.10	-0.24*	-0.01	-0.01	-0.04
	(0.09)	(0.06)	(0.13)	(0.02)	(0.02)	(0.03)
1960 × Station Distance	-0.17**	-0.10*	-0.22*	-0.03	-0.03	-0.05*
	(0.09)	(0.06)	(0.12)	(0.03)	(0.03)	(0.03)
1970 × Station Distance	-0.15	-0.09	-0.19	-0.03	-0.04	-0.06*
	(0.09)	(0.06)	(0.12)	(0.03)	(0.03)	(0.03)
1980 × Station Distance	-0.13	-0.09	-0.18	-0.03	-0.03	-0.07**
	(0.10)	(0.07)	(0.14)	(0.03)	(0.03)	(0.03)
1990 × Station Distance	-0.15	-0.10	-0.19	-0.05	-0.06**	-0.08***
	(0.09)	(0.07)	(0.12)	(0.03)	(0.03)	(0.03)
2000 × Station Distance	-0.14	-0.11	-0.20	-0.04	-0.05**	-0.08***
	(0.10)	(0.08)	(0.13)	(0.03)	(0.03)	(0.02)
Year FEs	yes	yes	yes	yes	yes	yes
County FEs	yes	yes	yes	yes	yes	yes
1880 IC × Year FEs	yes	no	yes	yes	no	yes
Sample States	all	all	hatch	all	all	hatch
R^2	0.92	0.91	0.91	0.96	0.96	0.96
County Observations	1063	1063	723	670	670	413

Notes: Authors' calculations with county data from 1870 to 2000 as described in the text. The unit of observation is a county-year. Each column reports the results from one regression, weighted by 1880 county land area. The main entries in the first row of columns (1)-(6) report estimates of β_t from equation (1) in the text. Standard errors clustered at the county level reported in parentheses. The excluded year interaction is 1880. The models in all columns include additional controls for year and county fixed effects. The model in column (1), (3),(4) and (6) also includes controls for: 1880 woodland fraction × year fixed effects, 1880 population × year fixed effects, and 1880 fraction of land in farms × year fixed effects. The sample in columns (4)-(6) is the balanced panel of counties with manufacturing revenue reported every year. * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

TABLE 5: Research Proximity and Land Productivity: Alternative Distance Comparisons

Dependent Variable=	Log(Crop Revenue Per Farm Acre)				
Distance to Ag Station County One:	Miles	Miles	Miles	Miles	Miles EW
Distance to Ag Station County Two:	Land Suitability	1880 Technology	1880 Literacy	1880 Ethnic	Miles NS
Specification:	Baseline	Baseline	Baseline	Baseline	Baseline
	(1)	(2)	(3)	(4)	(5)
1870 × Station Distance One	-0.09*	0.01	0.01	-0.03	-0.01
	(0.06)	(0.04)	(0.05)	(0.06)	(0.04)
1890 × Station Distance One	-0.06***	-0.06*	-0.07*	-0.05**	-0.07**
	(0.03)	(0.04)	(0.04)	(0.02)	(0.04)
1900 × Station Distance One	-0.22***	-0.24***	-0.24***	-0.22**	-0.26***
	(0.08)	(0.09)	(0.09)	(0.10)	(0.06)
1910 × Station Distance One	-0.33***	-0.35***	-0.36***	-0.27***	-0.35***
	(0.08)	(0.10)	(0.11)	(0.10)	(0.08)
1920 × Station Distance One	-0.27***	-0.25**	-0.26**	-0.23**	-0.28***
	(0.10)	(0.11)	(0.12)	(0.11)	(0.08)
1930 × Station Distance One	-0.21**	-0.20*	-0.20*	-0.23***	-0.24***
	(0.09)	(0.11)	(0.11)	(0.07)	(0.06)
[1940-2000] × Station Distance One	-0.14	-0.12	-0.13	-0.22***	-0.15***
	(0.09)	(0.10)	(0.10)	(0.07)	(0.05)
1870 × Station Distance Two	-0.23***	-0.19***	0.12**	0.04	0.05
	(0.05)	(0.05)	(0.06)	(0.03)	(0.05)
1890 × Station Distance Two	0.02	-0.08***	0.06*	-0.02	0.03
	(0.04)	(0.02)	(0.03)	(0.05)	(0.04)
1900 × Station Distance Two	0.06	-0.14***	0.09**	-0.02	0.05
	(0.04)	(0.04)	(0.04)	(0.07)	(0.04)
1910 × Station Distance Two	0.06	-0.17***	0.14***	-0.10	0.06
	(0.08)	(0.05)	(0.04)	(0.10)	(0.06)
1920 × Station Distance Two	-0.04	-0.21***	0.20***	-0.03	0.13**
	(0.08)	(0.05)	(0.05)	(0.10)	(0.06)
1930 × Station Distance Two	-0.03	-0.20***	0.12**	0.02	0.17*
	(0.08)	(0.05)	(0.04)	(0.12)	(0.09)
[1940-2000] × Station Distance Two	-0.04	-0.24***	0.04	0.11	0.11
	(0.05)	(0.05)	(0.03)	(0.10)	(0.12)
Year FEs	yes	yes	yes	yes	yes
County FEs	yes	yes	yes	yes	yes
1880 IC × Year FEs	yes	yes	yes	yes	yes
R^2	0.92	0.92	0.92	0.93	0.92
County Observations	1063	1063	1063	1063	1063

Notes: Authors' calculations with county data from 1870 to 2000 as described in the text. The unit of observation is a county-year. Each column reports the results from one regression, weighted by 1880 county land area. The main entries in the first row of columns (1)-(5) report estimates of β_t from an extended version of equation (1) in the text, with the two sets of distance-year interactions indicated. Standard errors clustered at the county level reported in parentheses. The excluded year interaction is 1880. The models in all columns include additional controls for year and county fixed effects, 1880 woodland fraction × year fixed effects, 1880 population × year fixed effects, and 1880 fraction of land in farms × year fixed effects. The sample in columns (4)-(6) is the balanced panel of counties with manufacturing revenue reported every year. * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

TABLE 6: Research Proximity and Producer Adjustments

Dependent Variable=	Log (Farmland Acres Per County Acre)	Log (Cropland Acres per Farm Acre)	Log (Rural Population per Farm Acre)	Log(Farm Equipment per Farm Acre)	Log(Fertilizer Per Farm Acre)
Specification:	Baseline	Baseline	Baseline	Baseline	Baseline
	(1)	(2)	(3)	(4)	(5)
1870 × Station Distance	-0.19*	-0.01	0.19	0.15***	
	(0.11)	(0.04)	(0.16)	(0.04)	
1890 × Station Distance	0.02	-0.11**	-0.03	-0.01	-0.02
	(0.06)	(0.04)	(0.06)	(0.04)	(0.15)
1900 × Station Distance	0.29***	-0.44***	-0.27***	-0.18**	-0.54*
	(0.09)	(0.10)	(0.09)	(0.08)	(0.28)
1910 × Station Distance	0.31***	-0.34***	-0.23***	-0.09	-0.35
	(0.09)	(0.10)	(0.10)	(0.10)	(0.27)
1920 × Station Distance	0.32***	-0.30***	-0.23**	-0.07	-0.27
	(0.09)	(0.11)	(0.10)	(0.12)	(0.16)
1930 × Station Distance	0.36***	-0.27***	-0.24***	-0.05	-0.12
	(0.09)	(0.11)	(0.09)	(0.11)	(0.15)
[1940-2000] × Station Distance	0.41**	-0.22**	-0.30***	-0.04	-0.18
	(0.09)	(0.10)	(0.10)	(0.11)	(0.21)
Year FEs	yes	yes	yes	yes	yes
County FEs	yes	yes	yes	yes	yes
1880 IC × Year FEs	yes	yes	yes	yes	yes
R^2	0.82	0.86	0.81	0.97	0.89
County Observations	1063	1063	1063	1063	1063

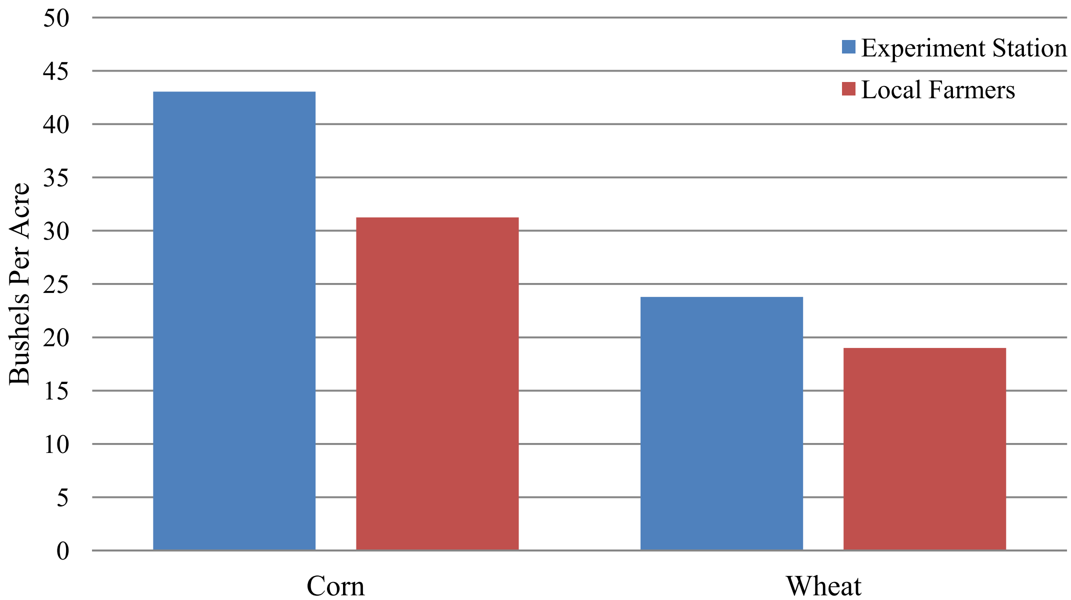
Notes: Authors' calculations with county data from 1870 to 2000 as described in the text. The unit of observation is a county-year. Each column reports the results from one regression, weighted by 1880 county land area. The main entries in the first row of columns (1)-(5) report estimates of β_t from equation (1) in the text for the dependent variable indicated. Standard errors clustered at the county level reported in parentheses. The excluded year interaction is 1880. The models in all columns include additional controls for year and county fixed effects, 1880 woodland fraction × year fixed effects, 1880 population × year fixed effects, and 1880 fraction of land in farms × year fixed effects. * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

TABLE 7: Research Proximity, Total Factor Productivity and Land Values

Dependent Variable=	Log (Crop Revenue)			Log(Farm Value Per Acre)
Specification:	Baseline	Investment-Capital Interactions	Fertilizer-Capital Interactions	Baseline
	(1)	(2)	(3)	(4)
1870 × Station Distance	-0.04		-0.02	0.01
	(0.09)		(0.06)	(0.02)
1890 × Station Distance	0.02	0.00	0.02	0.01
	(0.03)	(0.02)	(0.03)	(0.03)
1900 × Station Distance	0.02	-0.01	0.02	-0.08
	(0.07)	(0.03)	(0.05)	(0.09)
1910 × Station Distance	-0.14***	-0.16***	-0.13***	-0.01
	(0.05)	(0.03)	(0.04)	(0.11)
1920 × Station Distance	-0.06	-0.09*	-0.05	0.00
	(0.06)	(0.05)	(0.05)	(0.13)
1930 × Station Distance	-0.03	0.00	-0.01	0.02
	(0.05)	(0.04)	(0.03)	(0.12)
[1940-2000]	0.02	0.02	0.03	0.00
× Station Distance	(0.05)	(0.03)	(0.03)	(0.11)
Input Controls	yes	yes	yes	no
Year FEs	yes	yes	yes	yes
County FEs	yes	yes	yes	yes
1880 IC × Year FEs	yes	yes	yes	yes
R^2	0.95	0.97	0.95	0.95
County Observations	1063	1063	1063	1063

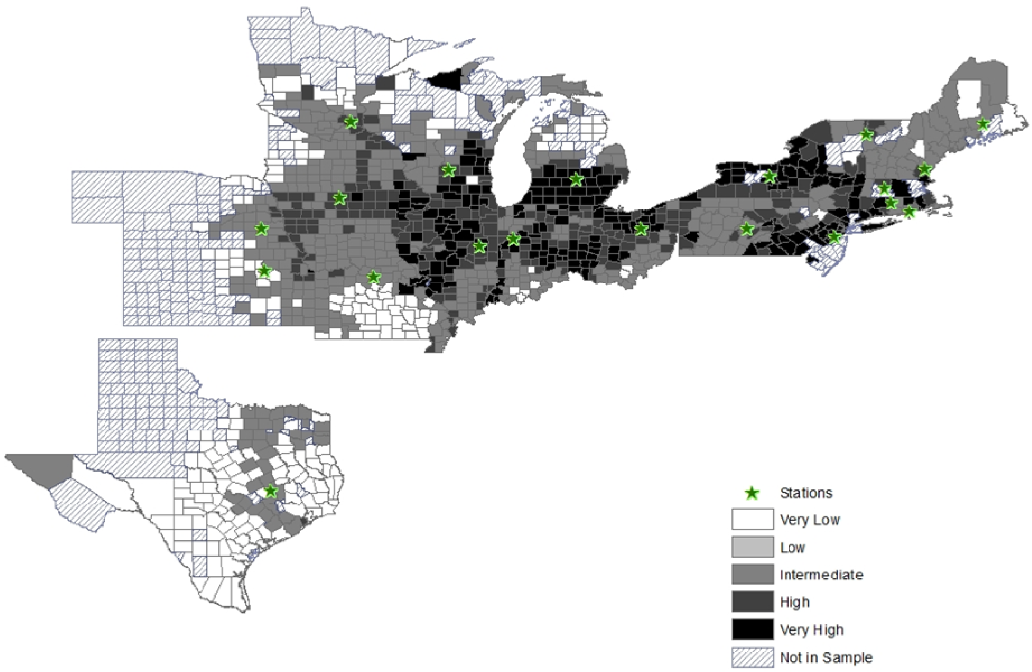
Notes: Authors' calculations with county data from 1870 to 2000 as described in the text. The unit of observation is a county-year. Each column reports the results from one regression, weighted by 1880 county land area. The main entries in the first row of columns (1)-(4) report estimates of β_t from equation (2) in the text. Standard errors clustered at the county level reported in parentheses. The excluded year is 1880. The models in all columns include additional controls for year and county fixed effects, 1880 woodland fraction × year fixed effects, 1880 population × year fixed effects, and 1880 fraction of land in farms × year fixed effects. The models in columns (1)-(3) control for Cropland Acres, Rural Population, Farm Equipment, Fertilizer Expenditure, and Farmland Acres. The model in column (2) also control for forth order polynomials in equipment per acre, cropland per acre, Δ equipment per acre, and Δ cropland per acre fully interacted. The model in column (3) also controls for forth order polynomials in equipment per acre, cropland per acre, and fertilizer per acre fully interacted. * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

FIGURE 1: Corn and Wheat Crop Yields: Early Station Experiments and Local Farmers Compared



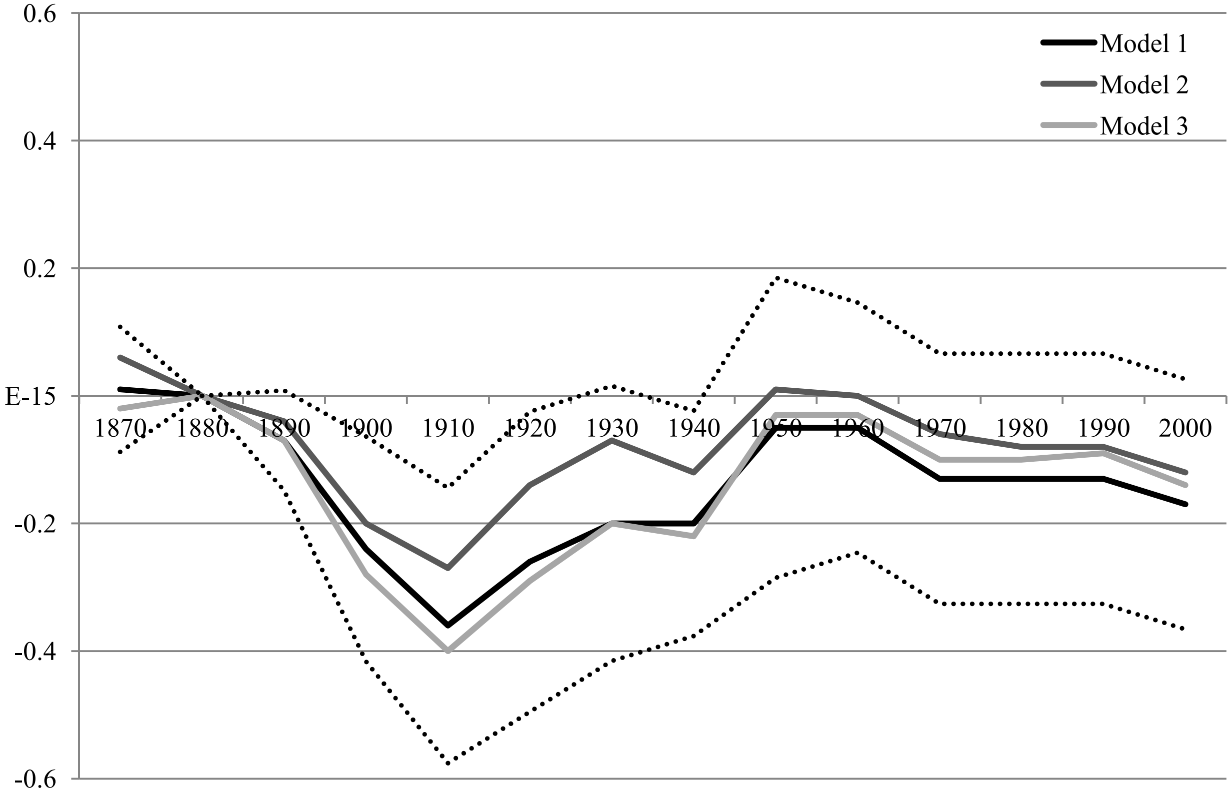
Notes: Authors' tabulation of experiment station reported crop yields from crop trials and agricultural census data for the primary experiment station county from 1889, 1899, 1909, and 1919.

FIGURE 2: Station Locations and Crop Revenue Per Farm Acre, 1880



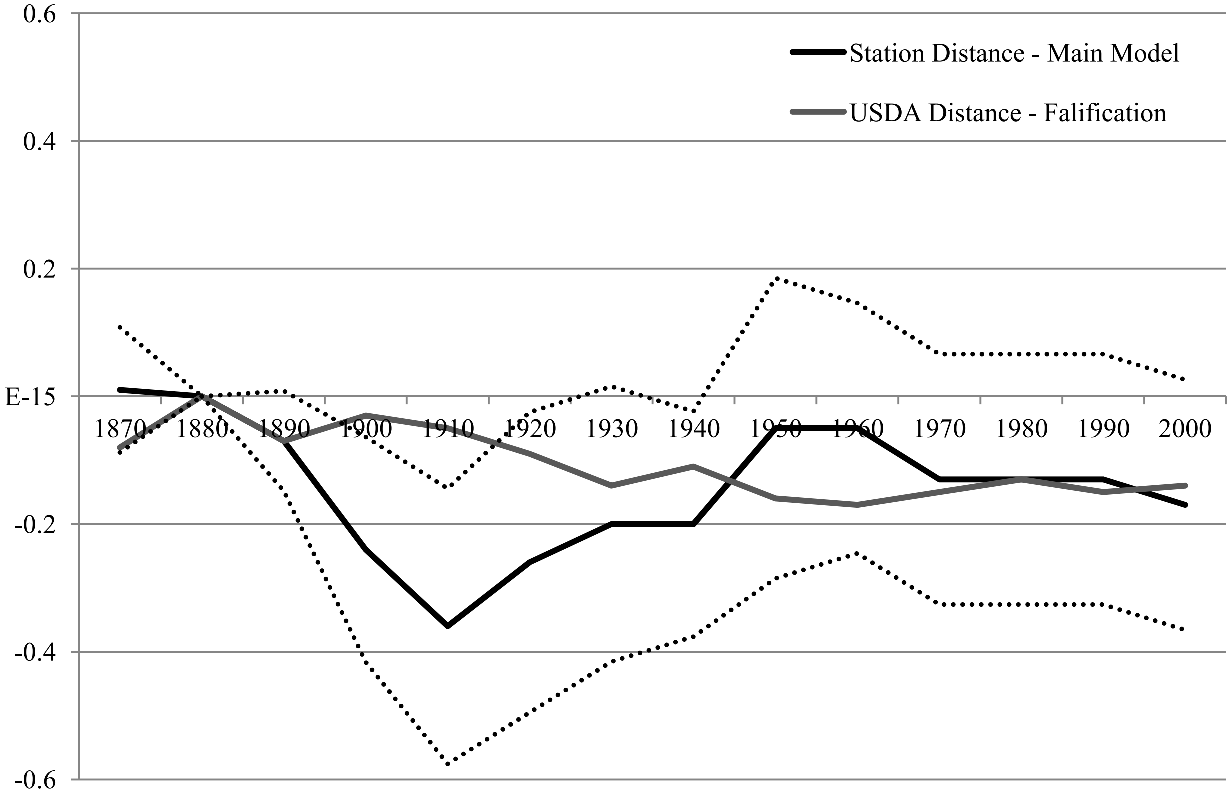
Notes: The map reflects station locations and crop revenue per acre in 1880 for the 1880 jurisdictional boundaries of each county. Station locations are for the primary federal experiment station in each state. The map reflects the balanced panel of counties that report crop revenue in each year and do not have larger boundary changes as described in the text.

FIGURE 3: Research Proximity and Crop Revenue per Farm Acre



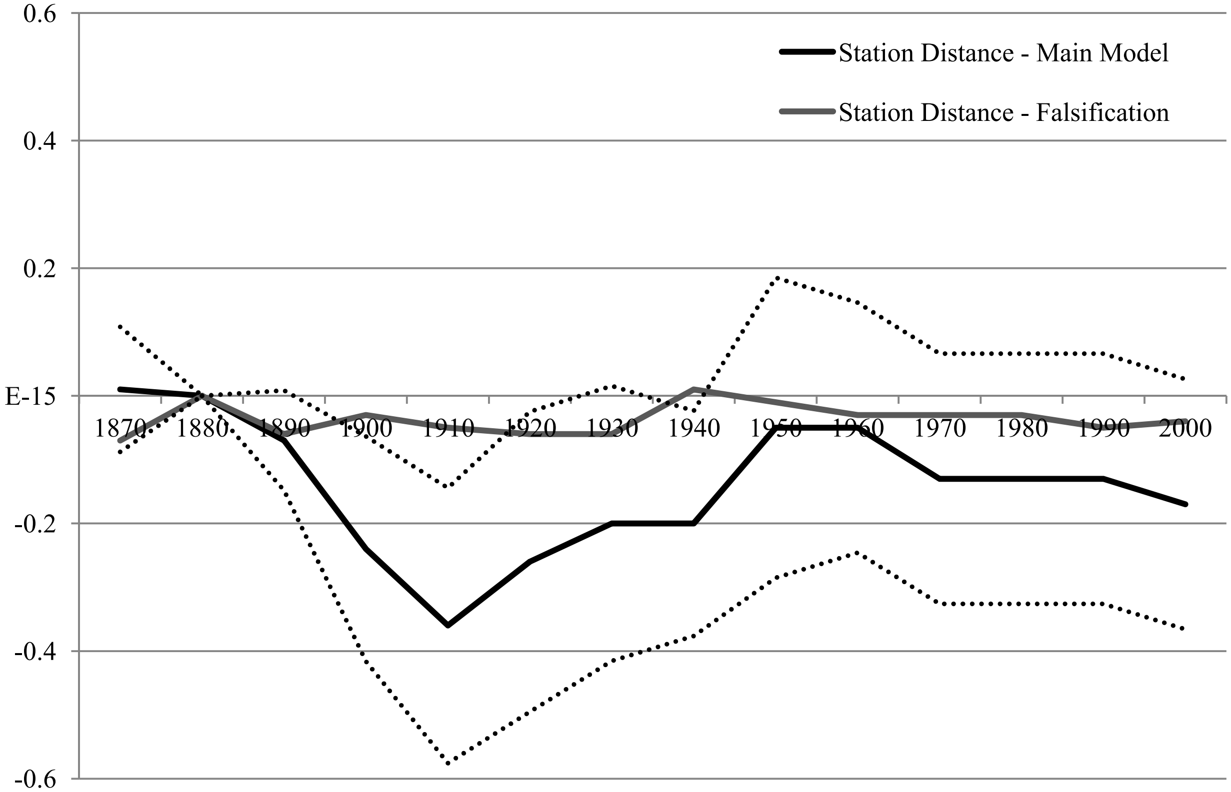
Notes: Authors' calculation from fitting equation (2). Each solid line plots the year-by-year distance interaction estimates from a different specification. Model 1 plots estimates from Table 2 column (1). Model 2 plots estimates from Table 2 column (2). Model 3 plots estimates from Table 2 column (3). The dotted lines depict 95% confidence intervals from fitting Model 1.

FIGURE 4: Research Proximity and Crop Revenue per Farm Acre: USDA Research Center Falsification



Notes: Authors' calculation from fitting equation (2). Each solid line plots the year-by-year distance interaction estimates from a different specification. *Station Distance – Main Model* plots the estimates from Table 2 column (1). *USDA Distance – Falsification* plots estimates from Table 4 column (1). The dotted lines depict 95% confidence intervals from fitting *Station Distance – Main Model*.

FIGURE 5: Research Proximity and Crop Revenue Per Acre: Manufacturing Falsification



Notes: Authors' calculation from fitting equation (2). Each solid line plots the year-by-year distance interaction estimates from a different specification. *Station Distance – Main Model* plots the estimates from Table 2 column (1). *Station Distance – Falsification* plots estimates from Table 4 column (4). The dotted lines depict 95% confidence intervals from fitting *Station Distance – Main Model*.