Reap What You Sow: Agricultural Productivity, Structural Change and Urbanization

D. McGowan and Ch. Vasilakis

Discussion Paper 2015-19

Institut de Recherches Économiques et Sociales
de l'Université catholique de Louvain

Reap What You Sow: Agricultural Productivity, Structural Change and Urbanization

Danny McGowan and Chrysovalantis Vasilakis *†

November 2015

Abstract

This paper explores the effects of agricultural productivity shocks on structural change. We exploit the invention of hybrid corn seed as an exogenous source of variation in US agricultural productivity. The technology significantly increased land productivity in counties suited to producing corn. Using a difference-in-difference estimation strategy we show that the treatment group experienced structural change as economic activity became more concentrated in agriculture. Owing to the factor bias of the technology, agricultural labor demand increased leading labor to reallocate from manufacturing to agriculture. We also find the rate of urbanization significantly decreases in treated counties, consistent with structural change causing a decrease in living standards. The findings support recent economic theory that argues factor-biased productivity shocks in agriculture can differentially affect structural change and economic development.

Keywords: agriculture, productivity, structural change, urbanization
JEL Codes: D22, D24, L16, Q11

*We thank Mirko Draca, Frederic Docquier, Oded Galor, Paola Guiliano, Richard Kneller, Stelios Michalopoulos, William Pariente, Luigi Pascali, Nico Voigtlaender, and seminar participants at the University of Louvain and the University of Warwick for helpful comments and suggestions. We are grateful to Leah Boustan for sharing data. Samuel Bottom provided excellent research assistance.

†Danny McGowan (corresponding author): University of Nottingham, danny.mcgowan@nottingham.ac.uk. Tel: +44 (0)115 846 7420. Chrysovalantis Vasilakis: Bangor University and IRES, c.vasilakis@bangor.ac.uk.

1 Introduction

Across the developing world, a large share of the population live in rural areas and work in the agricultural sector. In these economies agriculture typically accounts for between 25% and 60% of GDP and employs roughly 40% of the workforce. As agriculture makes up such a large share of economic activity, agricultural productivity shocks may affect other sectors of the economy through general equilibrium effects, and provoke changes in economic development.

Early theoretical contributions to the development literature emphasize a positive link between productivity within agriculture and structural change. In these models, improvements in agricultural productivity release labor from rural areas leading to a more industrialized economy (Clark, 1940, Rosenstein-Rodan, 1943, Kuznets, 1957). However, this finding hinges on the closed-economy assumption. Matsuyama (1992) shows that relaxing this assumption can produce markedly different results. Using a two-sector model he shows that in an open economy a comparative advantage in agricultural production can lead to a slowdown in industrial growth and a greater concentration of employment in agriculture.

Building on these insights, Bustos et al. (2015) show that in an open economy the relationship between agricultural productivity and structural change depends on the factor bias of the productivity shock. Improvements in land productivity (output per acre) increase agricultural labor demand as a greater number of workers are required to process the additional output. In equilibrium, this bids up wages leading to factor reallocation as manufacturing workers transition to employment in agriculture. Consequently, economic activity becomes more concentrated in agriculture and the economy becomes less industrialized. On the other hand, improvements in agricultural labor productivity (output per worker) have the opposite effects and reduce agricultural labor requirements, leading to an expansion of industry as workers relocate to manufacturing.

We test these predictions using a hitherto unexplored natural experiment. The heart of our identification strategy is an exogenous shock to agricultural land productivity arising from the introduction of hybrid corn seeds in the United States (US) in 1935. Hybrid corn plants' genetic traits allow them to produce approximately 20% more output relative to traditional seed varieties with no differences in input requirements. The new technology therefore caused an increase in output per acre (land productivity). Importantly, the development and adoption of hybrid corn seeds is exogenous with respect to the key economic outcomes we are interested in. As our review of contemporary and historical accounts indicate, the technology was only invented through a series of fortunate accidents, and that adoption was due to farmers observing hybrids' superior performance on neighboring farms. The land-productivity shock was uncorrelated with difficult-to-observe determinants of structural change and economic development, satisfying the exogeneity criterion.

The second source of variation is cross-sectional and arises from differences in regions' suitability for cultivating corn, as determined by time-invariant geoclimatic conditions. Climactic conditions over the growing season, as well as local topography and geological characteristics restrict the areas within which corn can be grown.[1] This provides exogenous, cross-sectional variation that allows us to construct treatment and control groups.

Our empirical estimates support the predictions in Bustos et al. (2015). Using county-level data from the Censuses of Agriculture and Population, from 1930 to 1940, the difference-in-difference (DID) specifications compare changes between counties suited and unsuited to growing corn with similar pre-1935 characteristics.[2] Over time we find the productivity shock causes structural change as treated economies become relatively more agrarian. For example, following the shock labor reallocates from manufacturing to agriculture. In addition, the number of manufacturing establishments per capita fall whereas the number of farms per capita increase substantially. Diagnostic tests confirm that these findings are not driven by anticipation effects, and that the key identifying assumption of parallel trends holds.

Our second contribution is to provide evidence on how structural change affected the overall level of development. Owing to the historical context, data on GDP per capita is not available at the micro level. However, urbanization and economic development go hand-in-hand as an economy moves from a rural-agricultural base to an industrial service-based economy. The high degree of correlation between per capita income and urbanization has led economists to use urbanization as a proxy for economic development where income data is not available (see, for example, Acemoglu et al. (2001) and Nunn and Qian (2011)).

We find that counties that were suitable for corn cultivation experienced substantial decreases in urbanization after the land-productivity shock. The average treatment effect is equivalent to a 3% reduction in urbanization and is highly statistically significant. This evidence is consistent with a decline in living standards. There appear to be three reasons why structural change reduced the level of economic development. First, a stylized fact across regions and time is the commonly observed total factor productivity differential between agriculture and manufacturing (Restuccia et al., 2008). As labor reallocated from manufacturing to agriculture in treated economies a greater share of the workforce were employed in relatively lower productivity jobs. Second, the data show that the land-productivity shock caused an increase in unemployment. Whereas employment in agriculture increases, these gains are more than offset

[1] The suitability of an area to producing corn is not influenced by the type of technology farmers use. Rather suitability is a time-invariant characteristic provided by the FAO-GAEZ database which calculates suitability using agronomic algorithms.

[2] Counties are effectively small, open economies within the US which closely resembles the model set-up in Bustos et al. (2015).

by contractions in manufacturing employment levels. The net effect of these changes is a 3% increase in unemployment in the average treated county. This stems from the land-productivity shock increasing labor demand, forcing up wages in manufacturing, and leading manufacturing establishments to shut down. Finally, the expansion of the agricultural sector appears to come through the entry of small tenant farms which tend to be less productive.

A unique feature of the US labor market during our sample window is that it was effectively closed to foreign migrants due to legislation passed during the 1920s (Boustan et al., 2010). This allows us to pinpoint the reasons behind the reduction in the rate of urbanization. Our estimates show that the fall in urbanization is directly related to workers migrating from urban to rural areas within the same county. There is no evidence of significant changes to immigration from other counties. Our main interpretation of the technology's impact is that it led to a decline in urbanization and local migration within counties. Specifically, the land-productivity shock increases agricultural labor demand, leading workers that were previously employed in manufacturing to migrate to take less productive jobs in rural areas where agriculture is concentrated.

Our empirical methodology shares most of the advantages and disadvantages inherent in DID estimators. On the one hand, it allows us to control for both county and time-period fixed effects so that all time-invariant differences across counties, such as geography, preferences or institutions (to the extent that they change slowly over time), and secular changes over time - such as global improvements in health, sanitation, and technological advancements - are controlled for. On the other hand, the strategy relies on there being no other shocks occurring around the same time that hybrid corn seeds were adopted that are correlated with counties' suitability for corn production. In the baseline estimates we address this identification concern by directly controlling for time- and county-varying factors that might bias our estimates. For example, we take care to ensure our findings are not simply an artefact of differential trends in the severity of the Great Depression across space by controlling for state-level unemployment rates.[3] We also conduct a host of sensitivity checks to rule out alternative explanations. For example, our findings could reflect local variation in mechanization or the availability of other agricultural technologies, productivity improvements in other closely-related crops, environmental factors, or differences in governmental assistance through the New Deal program. We explore these alternative mechanisms but find little support for them in the data.

In subsequent analyses we test the model's predictions on how structural change and economic development respond to labor-productivity shocks in agriculture. These tests leverage the diffusion of tractors as a shock to labor productivity. Consistent with the theory's predictions, we find that increasing labor productivity leads to a more industrialized economy as

[3]The results are robust to using county-level unemployment rates.

employment shifts from agriculture to manufacturing. While we also obtain a positive effect on urbanization, the coefficient estimate is below conventional levels of significance.

These findings contribute to several existing literatures. First, our results add to the rapidly-evolving debate on the importance of agricultural productivity in provoking structural change (Gollin et al., 2002, Foster and Rosenzweig, 2004, Hornbeck and Keskin, 2014, Bustos et al., 2015, Eberhardt and Vollrath, 2015). They provide support for the Bustos et al. (2015) model and illustrate that the agricultural productivity-structural change relationship may be more nuanced than has been assumed in previous theoretical work.

Second, our findings also contribute to the understanding of the relationship between agricultural productivity and aggregate economic growth. As highlighted above, the agricultural sector tends to be less productive relative to manufacturing in most economies. This is one of the primary reasons for widespread differences in international productivity and cross-country living standards (Restuccia et al., 2008). Because of reverse causality and omitted variable bias, identifying the causal link between agricultural productivity and economic growth is difficult. Our findings provide evidence that, depending on the nature of the shock, increasing agricultural productivity can reduce economic development, as measured by urbanization rates.

Finally, our findings speak to the literature about the importance of agriculture for urbanization. Nunn and Qian (2011) present evidence that the adoption of the potato caused a significant increase in the rate of urbanization in Old World countries with suitable growing conditions. As potatoes increase labor productivity, the factor bias of the technology may explain the difference in the direction of the effect between their findings and ours.

The paper proceeds in the following order. Section 2 provides an overview of the data sets. Section 3 reports information on our economic laboratory and hybrid corn seed. We outline our identification strategy and present econometric results in Section 4. Section 5 deals with potential threats to identification. Finally, we draw conclusions in Section 6.

2 Data and Summary Statistics

We rely upon three data sets for our empirical analysis: agricultural productivity data; county-level information on labor statistics, population and urbanization; and crop suitability data.

2.1 Agricultural Productivity Data

We retrieve information on agricultural productivity from the National Agricultural Statistics Service (NASS). The NASS is the statistics branch of the US Department for Agriculture. Every year it conducts hundreds of surveys on issues relating to agricultural production, demographics, and the environment. As part of this mission the NASS administers a survey of field crop yields (output per acre) in each county. This provides annual information between 1930 and 1940 on land productivity for industry i in county c during year t. The industries in the data set are corn, barley, wheat, and soybeans.[4] For all industries, land productivity is measured as output per acre, in bushels. In addition, the NASS provides information on the number of acres planted for each industry in each county-year, and the final good price ($ per bushel) for each industry in each state-year.

2.2 County-Level Data

Historic versions of the US census provide county-level data on population, agriculture, manufacturing, and urbanization for years ending in "0". Although there are some changes in which variables are reported through time, most are available for 1920, 1930, and 1940. For each county we construct the variables: land value per acre (the ratio of agricultural land value to planted acres), capital per acre (the ratio of the value of agricultural machinery and equipment to planted acres), farm buildings value per acre (the ratio of the value of agricultural buildings excluding dwellings to planted acres), farms per capita, agricultural employment, agricultural labor intensity (agricultural workers per acre), manufacturing establishments per capita, average manufacturing wages, manufacturing employment, manufacturing labor intensity (workers per establishment), unemployment (ratio of persons totally unemployed to total population), partly unemployed (ratio of persons partly unemployed to total population), and urbanization (the ratio of the urban population to total population). We also construct a tractors per acre variable (the ratio of total tractors to planted acres) using state-level information reported in the Agricultural Census.

Data on net migration is taken from Winkler et al. (2013). This provides information on the net flow of migrants into each county over the past decade for the years 1930 and 1940.[5] An important feature of the US labor market during the sample period is that it was effectively closed to foreign migrants. As part of the move towards isolationism following WWI, laws enacted during the 1920s prohibited migration from abroad (Boustan et al., 2010). The net migration variable therefore exclusively captures internal migration between counties.

To capture business-cycle fluctuations we construct state-level unemployment rates. We follow the method used by Boustan et al. (2010), although the results are unchanged when we use the county-level unemployment rate data from the Census. We prefer to use the state-level data as it constructs unemployment rates using the labor force rather than the population.[6]

[4] These are the most important cereal crops during the sample period. Corn refers to yellow-kerneled field corn, or maize. The corn industry does not include sweet corn which is a different type of cereal crop.

[5] For example, the 1930 values indicate the net change in county c's population between 1921 and 1930.

[6] Boustan et al. (2010) create unemployment rates using individual microdata from the IPUMS database. This is a

2.3 Corn Suitability

Our empirical strategy revolves around difference-in-difference estimations that exploit the fact that some regions of the US were affected by the agricultural productivity shock whereas other similar regions were not. We assign counties to treatment and control groups based on their suitability to produce corn. To measure a county's suitability for corn production, we take advantage of an extremely rich micro-level data set: the Food and Agriculture Organization's (FAO) Global Agro-Ecological Zones (GAEZ) database. The primary goal of the GAEZ project is to inform farmers and government agencies about optimal crop choice in any given location on Earth. The GAEZ data set uses agronomic models and high resolution data on geographic characteristics such as soil, topography, elevation and, crucially, climatic conditions to predict the suitability of an area to growing various cereal crops (corn, wheat, barley). This ensures that all regions have a suitability index for each crop, regardless of whether it is actually planted. The GAEZ data set is also used by Costinot et al. (2015) in their study of climate change and Nunn and Qian (2011) to examine urbanization.

To approximate historical conditions as closely as possible, we follow Nunn and Qian (2011) and use a suitability variable that is constructed under the assumption that cultivation occurs under rain-fed conditions and under low input intensity. Doing so ensures that subsequent technological developments, such as irrigation technologies do not influence suitability, leading to a suitability index that closely approximates historical conditions. The results are, however, unchanged if alternative input intensities are used.

Importantly, for our purposes, the GAEZ reports suitability values for each state s regardless of whether crop k is actually grown in state s. The GAEZ data report a suitability index for each state ranging between 0 (not suitable) to 100 (very suitable). We define land to be suitable for cultivation if its state has a suitability index of at least 50 (good suitability). Counties that meet this criteria are assigned to the treatment group ($T_i = 1$) while the rest form the control group ($T_i = 0$).[7] Later we experiment with alternative suitability thresholds.

It is important to recognize that the FAO-GAEZ corn suitability index is not based on specific contemporary or historical technologies (e.g. traditional or hybrid corn seeds). Suitability is calculated using a number of seed varieties that intentionally span a wide range of geographic environments (boreal, temperate, tropical, sub-tropical). The key determinants of suit-

representative repeated cross-section survey of the US labor market in 1930 and 1940. An employed individual is defined as one holding a job in the private sector or a non-relief public sector job. Using this variable and information on the number of individuals reported as being unemployed, the unemployment rate is calculated as the ratio of the total number of unemployed to the labor force.

[7] We match the state-level suitability index data to counties within each state. For example, all counties in Illinois are assigned the suitability index value reported by the GAEZ for Illinois.

ability are local climatic and topographical characteristics as well as soil conditions that have not changed much over time. The suitability index is therefore invariant to whether farmers plant traditional or hybrid corn seeds, and both varieties require the same climatic and growing conditions. See section 6 in Fischer et al. (2012) for further details.

2.4 Summary Statistics

[Insert Table 1]

Table 1 provides a description of the variables we use in the empirical analysis. Panel A provides a summary of the NASS variables used to examine the effect of hybrid corn seed on land productivity. The panel shows that the average acre of land produces 2.7 (ln) bushels of output, although this masks some of the variation in yields across crops. For example, an acre of corn yields 2.68 bushels (ln) whereas mean wheat yields are somewhat lower at 1.97 bushels (ln). Information on the number of acres planted in each county are also reported along with price per bushel.[8]

Panel B tabulates summary statistics for the variables we use in the structural change and urbanization regressions. In the interests of parsimony we provide a description of how each variable is constructed in Appendix B. In addition to the information we derive from the US Census, we also rely on data sets provided by Fishback (2005), Skinner and Staiger (2007), and Hornbeck (2012) for covariates that we use in the robustness checks.

Finally, we report a summary of the GAEZ suitability index in Panel C. The average county has a suitability rating of 49 although given the standard deviation of 17, there is considerable variation across counties. The suitability index is time invariant and does not change across time. Moreover, the suitability index does not differ depending on whether traditional or hybrid seeds are used. We also report a summary of the treatment group indicator, T_i. This variable is equal to 1 if the GAEZ suitability index is greater than or equal to 50, 0 otherwise. The data show that 54% of observations belong to the treatment group.

3 Institutional Details

We begin by providing background information on the industry and technological breakthrough that underpin our empirical tests.

[8] The data in Table 1 Panel A are based on a sample that uses observations from the corn and wheat industries. In the empirical analysis we also experiment with different control groups (barley and soybeans). Summary statistics for these samples are reported in Appendix A1.

Corn is one of the largest sub-sectors of the US agriculture industry, and typically accounts for between 30% and 40% of agricultural output. It is a major cereal crop that is planted in Spring and harvested during Fall. Corn plants require evenly-distributed rainfall and temperatures between 10 and 20 degrees Celsius over the growing season. Plants are made up of a leafy stalk that produces ears containing kernels (seeds). The kernels may be used to produce food, livestock feed, future generations of corn seed or biofuel. The more kernels a plant produces, the higher its yield and the greater is land productivity.

[Insert Figure 1]

Until the 1930s farmers had access to only one type of corn seed. This traditional seed variety was produced by seed companies randomly mating two corn plants and using their kernels as seed for the next growing season.[9] Traditional seeds produced plants that yielded approximately 26 bushels per acre.[10] Figure 1 shows that between 1860 and 1935 corn yields per acre were essentially constant with some fluctuations around the mean due to weather shocks. Seed producers could to some extent manipulate corn traits by selecting the best plants and cross-breeding them. However, while selection can improve some qualitative traits, it proved ineffective in raising yields (Crow, 1998).[11]

Geneticists had long realized that inbred corn plants produced higher yields relative to traditional varieties. However, inbred plants' resistance to disease was often limited meaning that most never grew to maturity to make commercialization possible (Crow, 1998). During the 1930s researchers made significant advances in understanding how to combine the high-yield traits of inbred plants with the longevity of traditional varieties. This inbred variety is called hybrid corn.[12] The main reason behind hybrid plants' yield advantages is that they produce bigger ears containing more kernels. Hybrid corn seeds therefore increase land productivity (output per acre).

[Insert Figure 2]

[9]Typically plants were randomly mated through wind-borne pollination in fields owned by a seed company. Pollination involves pollen grains being blown from one plant's staemen to another plant's stigma. Following successful pollination a plant is fertilized and the reproduction process begins leading to the development of kernels. Seed producers then harvest the plants and store the kernels to be sold as seed the next year.

[10]A bushel of corn contains 8 US dry gallons of kernels.

[11]The reason for the lack of yield growth is that corn reproduces sexually each year. This process randomly selects half the genes from a given plant to propagate the next generation. Consequently, mass production of seeds with desirable genes (such as yield) can be lost in subsequent generations as high- and low-yielding plants are randomly mated. Selection of high yielding traditional plants was therefore not viable.

[12]Hybrid corn is a plant produced by mating two inbred corn plants. Experimental trials eventually found double-cross hybrids - crossing two inbred plants and crossing that hybrid with the hybrid of two other inbred plants - to reliably produce high yields.

Scientists first began to explore ways of increasing corn yields in the early 1900s (Schull, 1908, East, 1908, Schull, 1909). The reason behind academic interest in this topic was researchers' desire to find a solution to the historical stagnation of corn productivity (Crow, 1998). Crow (1998) reports that it took a long time for researchers to discover how to produce hybrid plants fit for commercialization and that this process proceeded on a trial-and-error basis. After numerous experimental trials, hybrid corn seeds became commercially available in 1935 (Skinner and Staiger, 2007).[13]

Farmers purchased corn seed from seed suppliers that typically had local shops or traveling sales men. From 1935 onwards suppliers offered farmers a choice of either the traditional or the hybrid technology. Figure 2 shows that adoption of the new technology was rapid. By 1940 approximately 30% of acres were planted using hybrid seed. However, the speed of adoption was also quicker in some states. For example, Griliches (1957) reports that hybrids accounted for 90% of planted acres in Iowa by 1940 with similar incidences in other Midwestern states. At the latest, hybrid corn seeds became available in all regions by 1937. Crow (1998) notes that the rapid adoption of hybrid corn seeds was mainly due to word of mouth between farmers and casual observation of crop yield performance on neighbors' farms.

3.1 The Effect of Hybrid Corn Seed on Land Productivity

Figure 1 provides the first indication that hybrid corn seeds improved agricultural land productivity. Historically, corn yields averaged 26 bushels per acre between 1860 and 1935. Post 1935 there is a clear break in the land productivity trend as corn yields jump sharply upward, followed by a steady rise through time. On average, land productivity was 22% higher in 1940 compared to 1934.

Although the descriptive evidence is suggestive, it neither accounts for possible unobserved heterogeneity nor pins down the average treatment effect of the technology. For this, we turn to regression analysis. To measure the productivity effects of the hybrid technology we use a difference-in-difference (DID) estimator that compares the evolution of land productivity within the corn industry to land productivity among closely-related cereal crops that are often grown in close proximity to corn and rely on very similar production methods. These features

[13]Planting hybrid and traditional varieties in close proximity has no effect on the productivity of either type of plant. This is because the plant's genetic make-up (that is, the number of kernels it produces) is determined solely by the traits in its seed. Instead, if a hybrid and traditional plant were to mate, it would only determine the genetic composition of the next generation of seeds. The current generation's genetic make-up would be unaffected as this depends on the seed from which the plant germinated. Given that farmers purchase seed from seed companies for one growing season at a time, proximity of the two varieties had no effect on the productivity of each variety during the current growing season.

obviate the confounding influence of climatic conditions, local economic shocks and mechanization that may affect productivity more generally. Unlike corn, there were no coinciding technological developments within the barley, soybean and wheat industries. These features allow us to establish a counterfactual for what land productivity within the corn sector would have been in the absence of technical change.

Using the NASS productivity data we estimate the equation

$$yield_{ict} = \alpha_c + \beta_1 Corn_{ic} + \beta_2 Post_t + \beta_3 Corn_{ic} * Post_t + X'_{ict}\delta + \gamma_t + \varepsilon_{ict}, \quad (1)$$

where $yield_{ict}$ is output per acre in industry i of county c at time t; $Corn_{ic}$ is a dummy variable equal to 1 if the observation is from the corn industry, 0 otherwise; $Post_t$ is a dummy equal to 1 for the years 1935 to 1940, 0 otherwise; X_{ict} is a vector of control variables; α_c and γ_t are county and year fixed effects respectively; ε_{ict} is the error term. In line with Bertrand, Duflo and Mullianathan (2004) we cluster the standard errors at the county level.[14]

[Insert Table 2]

We report the estimation results in Table 2. The results in column 1 of the table are estimated using a simple specification that excludes the control variables, and uses wheat as the counterfactual.[15] Wheat is a natural control group as it is also a widely-grown cereal crop in the regions where corn is planted. Unlike corn, however, there were no technological developments within the wheat sector at the same time hybrid corn seeds became available. The coefficient on the corn variable indicates that within the average county corn productivity is approximately 49% higher relative to productivity in the wheat industry. Notably, the post dummy is statistically insignificant. This implies there was no change in land productivity between the pre- and post-treatment periods within the control group, reinforcing the view that there was no technical change within the wheat sector. But this pattern is not mirrored within the corn industry. Rather the coefficient on the $Corn - Post$ interaction term is both positive and highly statistically significant. The economic magnitude of the effect is also large: equivalent to a 17% productivity gain. Hybrid corn seeds therefore caused a significant increase in land productivity.

In Table 2 column 2 we append the original model with some control variables to rule out potential confounding influences. To ensure that our results do not stem from changes in

[14]This specification uses a balanced panel to ensure the results are not driven by counties entering or exiting the data set. That is, it includes only observations from counties that grow both crops in all years between 1930 and 1940. The results are unchanged when we use an unbalanced panel that includes observations of counties that begin producing either crop or drop out of the sample by stopping production.

[15]In this specification, the corn dummy variable captures the industry fixed effect.

economies of scale we include an acres planted variable. The coefficient estimate is statistically insignificant. We find a negative and statistically significant association between final goods price and yields. The negative physical productivity-price correlation is consistent with equilibria where producers move along the demand curve in response to demand shocks (Foster et al., 2008). Finally, we include the unemployment rate variable to ensure our results are not simply capturing farmers' responses to business cycles. We find that yields negatively covary with the business cycle. Despite including the control variables our key finding remains robust: hybrid corn seed significantly increased productivity within the corn sector.

Next, we augment the regression equation with controls for land value and capital per acre to rule out the confounding effect of improvements to land and capital inputs. The estimates provided in column 3 show that our main finding is preserved. This is also the case when we include county-industry fixed effects in column 4. County-industry effects purge any time invariant, county-specific unobservables that differentially affect productivity within the corn industry relative to wheat productivity (for example, altitude or soil acidity).

Could the observed sensitivity of productivity to technology adoption reflect pre-treatment trends or spurious forces? One way to test the exogeneity assumption is to create placebo shocks as in Bertrand et al. (2004). We therefore generate a dummy $Placebo_t$ (equals 1 for 1934, 0 otherwise) and interact this with the corn dummy. Since hybrid corn seeds were not introduced until 1935 we know that the null of zero effect on the corn-placebo interaction is true. Alternatively, if our previous results were capturing pre-treatment trends in corn productivity we would expect the $Corn - Placebo$ coefficient to be similar in magnitude and statistical significance to the $Corn - Post$ coefficient. This is not the case. Instead the $Corn - Placebo$ interaction coefficient is statistically insignificant in Table 2 column 5. Given there were no significant differences in productivity between the treatment and control groups pre-treatment, it appears that the parallel trends assumption holds.

In the remainder of Table 2 we carry out further validation checks to ensure our findings are not driven by the choice of counterfactual. The results in columns 6 and 7 continue to show a positive and statistically significant effect of the new technology on corn productivity when we use barley and soybeans as the control group, respectively.[16]

Our results could be contested on the grounds that changes to input usage coincided with

[16]Whereas barley also belongs to the cereal crop family, soybeans do not. An advantage of using soybeans as the control group is that its genetic composition is substantially different from corn, which allows us to be more certain that technological developments within the control group do not bias the average treatment effect. Moreover, soybeans are often grown in the same regions as corn. The obvious disadvantage is of course that a non-cereal crop control group is somewhat genetically removed from corn. This presumably explains the considerably larger average treatment effect when we use soybeans as the control group. An advantage of using wheat or barley as the control group is that neither is planted in rotation with corn.

the introduction of hybrid corn seed such that the productivity gains are driven by more intensive input use rather than the new technology. While the estimates above go some way to ruling out this possibility by controlling for capital usage and land quality, they no not capture changes to other types of inputs. Although the NASS do not provide county-level input expenditure data, Fernandez-Cornejo et al. (1999) report annual values for seed, fertilizer, and lime expenditure within the corn industry over 1910 to 1950. Our strategy to rule out the effect of inputs is to use a type of structural break procedure.

[Insert Table 3]

Specifically, we examine whether there exists a structural break in each of these series using the $Post_t$ dummy to examine if there are systematic differences in input usage between the periods. We estimate the equation

$$m_t = \alpha + Post_t + \varepsilon_t, \tag{2}$$

where m_t is annual expenditure on input m; $Post_t$ is equal to 0 for the years 1910 to 1934, 1 for 1935 to 1950; and ε_t is the error term. The results in Table 3 display no evidence of an increase in input usage. There are no significant differences in the amount of seed, fertilizer and lime used between periods, and there is actually a trend towards using fewer inputs such as pesticides.

Further evidence that it was the genetic traits of hybrids and not input usage that increased corn productivity can be found in Russell (1974). Using hybrid and traditional corn seeds from 1935 he grew the plants in controlled laboratory conditions, allowing him to isolate the yield effect of the hybrid technology. He finds that the hybrids produced plants with at least 60% higher yields.

An additional concern is that following the adoption of hybrid corn seeds farmers reduced the number of corn acres planted such that output and labor demand remained unchanged. We test this hypothesis by re-estimating equation (1) using acreage planted and output as the dependent variable. The results in column 1 of Appendix Table A2 show that there was no significant difference in the number of acres planted between the treatment and control groups through time. Corn farmers did not therefore reduce acreage planted but kept acreage constant. However, the average treatment effect in column 2 reveals a significant increase in corn output. Hence, the primary effect of hybrid corn seed was to increase output through higher yields.

Our final robustness check addresses the concern that our findings are driven by omitted agricultural technologies. The productivity regressions capture to some extent improvements and diffusion of other technology shocks by conditioning on capital per acre. However, given our time period coincides with the mechanization of agriculture, we re-run the analysis and directly control for tractor usage. The estimation results reported in column 3 of Appendix Table A2 show that controlling for tractor technologies has little impact on our main conclusions. Rather hybrid corn seeds remain a positive and significant determinant of land productivity.

4 Empirical Analysis

In this section, we proceed in two stages. We first outline our econometric strategy to assess how the land-productivity shock impacted counties' labor markets, structural change, and urbanization. We then present the estimation results and discuss the findings.

4.1 Identification Strategy

Our empirical methodology exploits two sources of exogenous variation. Cross-time variation in land productivity comes through the invention and adoption of hybrid corn seed whereas cross-sectional variation in the effect of the shock arises from differences in counties' suitability to produce corn.

[Insert Figure 3]

The GAEZ data in Figure 3 illustrate that only some regions are suited to growing corn. Most western states are unsuitable because of too high temperatures, lack of rainfall, or both. We assign counties to the treatment (control) group based on whether the suitability index is above (below) 50: equivalent to good suitability.[17] Suitability is a time invariant characteristic that does not depend on what type of seed is used. As outlined in Section 2, both the traditional and hybrid seed varieties require the same growing conditions (that is, temperatures between 10 and 20 degrees Celsius and evenly distributed rainfall over the growing season).

For our estimation strategy to generate credible inferences requires that the suitability index accurately predicts where corn is grown. Otherwise, any statistically significant relationships may be purely coincidental. To examine this issue we use a probit model to estimate the equation

$$y_{ct} = \alpha + \beta SI_c + X_{ct}'\delta + \gamma_t + \varepsilon_{ct}, \tag{3}$$

where y_{ct} is a dummy variable equal to 1 if corn is planted in county c during year t, 0 otherwise; SI_c is the GAEZ suitability index in county c; X_{ct} is a vector of controls; γ_t are year dummies; and ε_{ct} is the error term. The marginal effects from this test are provided in Table

[17] The treatment group includes counties in Georgia, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maryland, Michigan, Minnesota, Missouri, Nebraska, New York, Ohio, Oklahoma, Pennsylvania, South Dakota, Tennessee, and Wisconsin.

4. The unconditional results in column 1 show that a 1% increase in the suitability index raise the probability of corn being grown in a county by 2.9%. This effect is significant at the 1% level. The coefficient magnitude is unchanged in column 2 when we include control variables in the model. This suggests that: 1) the GAEZ suitability indicator does a good job at predicting where corn is grown, and 2) because the suitability coefficient estimate is invariant to including observable characteristics, suitability is uncorrelated with omitted variables contained in the error term, and is plausibly exogenous.

[Insert Table 4]

A concern is the extent to which the FAO-GAEZ suitability measure is an accurate indicator of suitability during the 1930s and 1940s given it is constructed by researchers to inform contemporary individuals and governments. The construction of the suitability measure does not give any obvious cause for concern. In fact, the suitability index should be a good proxy for historical conditions because they are primarily based on climatic characteristics such as temperature, humidity, length of days, sunlight, topography and rainfall that have not changed significantly since the period of our study. Moreover, in constructing our suitability measure, we intentionally used the FAO-GAEZ suitability measure that assumes rain-fed conditions and low input usage to avoid measurement error from changes over time in irrigation intensity and technologies. Nunn and Qian (2011) also follow this approach.

To evaluate the consequences of land productivity on structural change and urbanization we estimate the following reduced-form equation

$$y_{it} = \alpha_i + \beta T_i * Post_t + X_{it}'\delta + \gamma_t + \varepsilon_{it}, \tag{4}$$

where y_{it} is the outcome of interest in county i at time t (structural change or urbanization); T_i is a dummy variable equal to 1 if the county is in the treatment group (that is, the GAEZ suitability index is 50 or above), 0 otherwise (GAEZ suitability index less than 50); $Post_t$ is a dummy variable equal to 1 in the year 1940, 0 in 1930; X_{it} is a vector of control variables; α_i and γ_t are county and year fixed effects, respectively; and ε_{it} is the error term. Because T_i is a time invariant county specific characteristic, it is captured by the county fixed effects. Similarly, because our estimations use data from two years (1930 and 1940) the $Post_t$ dummy is captured by the year effects. Following Bertrand et al. (2004) we cluster the standard errors at the county level, although we note that because the data set contains only two time periods for each county, the efficiency of the estimates is unlikely to be affected by serial correlation in the dependent variable.[18]

[18]In Section 5 we test the robustness of our findings to more restrictive definitions of corn suitability. The results are unchanged in these tests.

Central to the internal validity of our DID design is the exogeneity of the land-productivity shock. Crow (1998) cites academic curiosity among agricultural researchers in solving the long-standing stagnation of corn productivity as the main reason behind research into developing new seed varieties. The discovery process was haphazard and followed a trial-and-error approach at agricultural research stations and universities as geneticists experimented with different ways to produce reliable hybrids (Crow, 1998). After its invention, hybrid corn seed was made commercially available by seed companies that were set up during the 1920s and 1930s (Crow, 1998). Seed companies quickly developed hybrid seeds that could be sold throughout the US (Griliches, 1957). As a result, hybrid corn became available between 1935 and 1937 (Zuber and Robinson, 1941, Skinner and Staiger, 2007, Olmstead and Rhode, 2008). Farmers' choice to adopt hybrid seeds appears to have been driven by word of mouth and their observations of hybrids' growth performance on neighbors' farms (Crow, 1998, Sutch, 2011).

The reasons behind the invention and adoption of hybrid corn seeds are unrelated to structural change and urbanization. It is therefore unlikely that endogeneity bias arises through a simultaneous relationship in equation (4). This would require that hybrids were invented to rebalance economies towards agriculture which is implausible. In addition, the invention of hybrids is due to luck and academic curiosity, and is therefore uncorrelated with unobservable determinants of structural change and urbanization contained within the error term in equation (4). Estimates of β therefore capture the causal effect of land productivity on the outcomes of interest.

Estimating the causal effect of land productivity shocks also relies on establishing an implied counterfactual. That is, what would have happened to structural change and urbanization within the treatment group had hybrid corn seed not been invented. We therefore test the key identifying assumption underlying our empirical model: the parallel trends assumption. This assumes that in the absence of hybrid corn seeds, the dependent variables of interest would have evolved in similar fashion within the treatment and control groups. To test this assumption we compare the evolution of the key dependent variables during the pre-treatment period by estimating the equation

$$y_{it} = \alpha_i + \beta T_i * D1930_t + \gamma_t + \varepsilon_{it}, \tag{5}$$

where $D1930_t$ is a dummy variable equal to 1 if an observation is from 1930, 0 otherwise. We estimate the model using data from 1920 and 1930. Owing to gaps in the questions asked in the Census, not all variables are available before 1930 (e.g. agricultural employment), or are only available in 1910 rather than 1920 (e.g. urbanization).

[Insert Table 5] [Insert Table 6] [Insert Table 7]

The results of this test are provided in Table 5. For each variable the coefficient estimates support the parallel trends assumption. In all columns of the table the interaction coefficient is statistically insignificant. This implies that economic conditions within the treatment and control groups evolved similarly through time and that there was no anticipation of the productivity shock: structural change and urbanization rates do not change in anticipation of hybrid corn seeds. We can therefore be confident that the control group constitutes a reliable counterfactual for conditions within the treatment group in the absence of the hybrid technology.

A related question is, to what extent do the treatment and control groups resemble each other? Establishing a valid counterfactual also rests on the extent to which economic conditions are similar across the groups prior to treatment. We therefore undertake a series of t-tests to examine whether there exist significant differences in the key dependent variables and control variables between the treatment and control groups in 1930. These results are reported in Table 6. From the table it is clear that there are no significant differences in the levels of agricultural and manufacturing employment between the treatment and control groups. There is also no evidence of significant differences in the prevalence of the agricultural and manufacturing sectors: both the number of farms and manufacturing establishments per capita are similar. Moreover, business cycle conditions appear similar given there are no significant differences in state unemployment rates.

From these diagnostic tests we conclude that: 1) there were no anticipation effects, 2) the control group serves as a good approximation of the treatment group in the untreated state, and 3) the treatment and control groups closely resemble one another.

Finally, one may question to what extent a productivity shock in the corn sector raised land productivity in the broader agricultural industry. The evidence in Table 7 suggests a large effect. Within the treatment group corn acreage covers almost 9% of the average county's total land area and accounts for 35% of all planted acres. It therefore seems reasonable to conclude land productivity in the wider agricultural sector was affected by the introduction of hybrid corn seeds. From Table 7 it is also apparent that almost no corn production takes place within the control group. This reinforces the view that the control group was unaffected by the shock which aids construction of the implied counterfactual.

4.2 Results and Discussion

We explore the consequences of the land productivity shock in three stages. We begin by documenting the labor market effects and then proceed to examine the impact on structural change. Finally, we test whether the shock triggered changes in the rate of urbanization.

4.2.1 Labor Market Effects

[Insert Table 8]

Table 8 reports the results of estimations in which we test the labor market effects of the productivity shock. Column 1 presents the results for agricultural employment. The estimated coefficient for the T_i-$Post_t$ interaction term is 0.0972. This indicates that the land-productivity shock increased agricultural employment by approximately 97 workers within the treatment group post treatment. Since the standard deviation of agricultural employment is 0.69, this is a sizeable response. Compared to the average pre-treatment level of agricultural employment, a 0.0972 increase in employment represents an increase of around 10%. The effect is also statistically significant at the 1% level.

In column 2 we examine whether the employment effects of the technology are driven by more intensive labor usage within agriculture. The average treatment effect shows that labor intensity increased by 37% within the treatment group post treatment. The coefficient is precisely estimated and is significant at the 1% level. One effect of the technology was therefore to increase agricultural demand for labor as more workers were required to deal with the greater output produced by hybrid plants compared to traditional varieties.

Columns 3 and 4 examine how the productivity shock affected the manufacturing sector. In column 3 we find a high sensitivity of manufacturing employment to the shock. On average, manufacturing employment contracts by 169 workers in treated counties. This is equivalent to an 8.4% reduction relative to the pre-treatment mean. Next, we examine the impact on manufacturing labor intensity. The T_i-$Post_t$ interaction coefficient is negative but not significantly different from zero, indicating that the number of workers per firm did not change.

Given that in the employment regressions the average treatment effect is estimated to be larger for manufacturing than agriculture, it appears that factor reallocation was imperfect. Potential explanations are differences in skill requirements between the sectors or geographic immobility due to agriculture being concentrated in rural areas and industry within cities. We should therefore be able to document an increase in unemployment within the treatment group post 1935. Indeed this is the case in columns 5 and 6 of Table 8. The share of the population reported as being totally and partly unemployed increased by around 3%. Hence, while the productivity shock led to an expansion of the agricultural sector this did not outweigh the contraction in manufacturing employment.

Among the control variables we find that some are economically important. For example, manufacturing value added is negatively correlated with employment within agriculture although the coefficient is only marginally significant. Labor intensity in both sectors is significantly related to unemployment rates although the direction of the effect is positive for agriculture and negative for manufacturing. To the extent that manufacturing value added proxies productivity within the manufacturing sector, these results are consistent with increasing productivity reducing manufacturing labor requirements. We also find that the size of the agriculture sector, proxied by farms per capita, is significantly associated with all the dependent variables in Table 8.

4.2.2 Structural Change

The labor market regressions indicate that the improvement in land productivity caused by hybrid corn seeds led to factor reallocation from manufacturing to agriculture. This is consistent with structural change as treated economies rebalance towards agriculture. We now dig deeper to provide further evidence of structural change.

[Insert Table 9]

The decrease in employment within the manufacturing sector does not come about through firms shedding workers, which would be the case if manufacturing labor intensity decreased. This suggests that the contraction in manufacturing employment is driven by firms closing down and exiting. We find evidence consistent with this hypothesis in column 1 of Table 9. The point estimate on the interaction term shows a significant reduction in the number of establishments per capita. The magnitude of this effect corresponds to a 41% reduction, indicating net exit. However, the average number of establishments per capita is only 0.045 meaning that the absolute size of the effect is smaller and equates to a 3% reduction in the number of establishments out of an average population of 25 per county.

Bustos et al. (2015) hypothesize that the effects of land productivity on structural change are transmitted through wage effects. Specifically, higher land productivity increases agricultural demand for labor which, due to the market clearing condition, bids up wages in manufacturing, increases firms' break-even costs, leading some to exit. The findings in Table 9 column 2 support this view. Specifically, we observe a significant 5% increase in manufacturing wages in counties suited to growing corn post 1935. This is consistent with manufacturing labor supply decreasing as workers take jobs in agriculture.[19] Together these findings suggest that the land productivity shock triggered an increase in wages that caused some manufacturing firms to shut down and lay off employees.

The results in column 3 of Table 9 provide insights into how the size of the agricultural sector was affected by the productivity shock. In sharp contrast to manufacturing, we find that the number of farms per capita increased in treated counties. The magnitude of this effect corresponds to a 12% increase. One explanation for this behavior could be that at the higher land

[19]Unfortunately, the Census does not contain comparable wage data for agriculture.

productivity level, the expected profits of entering the sector increased leading new producers to enter (Melitz, 2003). Together with our previous labor market findings, this suggests that hybrid corn seeds triggered structural change as counties suited to growing corn experienced de-industrialization and became more agrarian. The factor bias of the productivity shock appears to be important in explaining this result because it altered demand and supply of labor in the two sectors. Our findings therefore support the predictions made by Bustos et al. (2015).

In the remainder of Table 9 we try to gain deeper insights into how the productivity shock affected the agricultural sector more generally. We find that the increase in the number of farms was primarily due to the entry of small tenant farms. In column 4 and 5 the average treatment effect indicates a 0.6 and 1.1 percentage point decrease in farms operated by the full or part owner. This is almost completely offset by a 1.9 percentage point increase in tenant farming found in column 6. Hence, the entry of new firms came through existing landowners renting land to tenant farmers.[20]

The evidence in column 7 of Table 9 indicates that the average size of farms decreased in counties affected by the productivity shock. The interaction coefficient shows that average farm size (measured in acres (ln)) fell by 8%. This is consistent with the increase in tenant farming (leading to subdivision of existing farms) observed above.

Previous evidence reported by Helpman et al. (2008) shows that firm size is strongly correlated with total factor productivity. The reduction in average farm size is therefore consistent with a decrease in productivity within the agricultural sector more generally. Further evidence that the entry of small farms led to lower agricultural productivity is reported in column 3 of Appendix Table A3. We find that the greater the share of tenant farms within a county the lower is yield per acre within the corn industry. This effect is highly statistically significant. Possible explanations for this negative correlation are that tenants operate small farms that are unable to take advantage of economies of scale or that they are inexperienced farmers that are not fully aware of best operating practices.

Together this evidence suggests that the introduction of hybrid corn seeds led to the entry of relatively small, less productive firms into the corn industry.

4.2.3 Urbanization

An important question is how structural change affected the level of economic development in treated counties. Considering the productivity gap between agriculture and manufacturing

[20] Consistent with greater demand for agricultural land, in Appendix Table A3 we find evidence that the productivity shock caused a significant increase in land value per acre. In addition, the value of farm building values also significantly increases through time within the treatment group.

through time (Restuccia et al., 2008), and the fact that the land-productivity shock caused an increase in unemployment within the treatment group, it seems likely that living standards fell.

Unfortunately, answering such questions is difficult as the Census does not provide GDP or income data. However, previous historical studies have found urbanization to strongly correlate with per capita income levels and have accordingly used it to proxy economic development (Acemoglu et al., 2001, Nunn and Qian, 2011). Using county-level urbanization rates we are therefore able to gain some insights into the development effects of structural change.

[Insert Table 10]

Column 1 in Table 10 reports the results when the urbanization rate is used as the dependent variable in equation (4). The interaction coefficient is negative and statistically significant at the 1% level. Economically, the effect size is equal to a 0.59 percentage point reduction in the share of population living in urban areas. This is equivalent to a 3% reduction in urbanization relative to the treatment group's pre-treatment mean. The transition towards a more agrarian economy appears to have caused a moderate decline in development and living standards.

De-urbanization also confirms our earlier findings on labor reallocation between sectors. In a world where agriculture is concentrated in rural areas, manufacturing is concentrated in cites, and agents are mobile across locations and choose between agriculture and non-agriculture as sectors of employment, we would expect agricultural productivity shocks to lead to changes in urbanization. Land productivity shocks increase labor demand in agriculture leading workers to migrate to rural areas.

A concern is that changes in urbanization and factor reallocation stem from broader migratory patterns because each county represents a small, open economy within the US. We therefore conduct a validation check in column 2 of Table 10 by examining how net migration was affected by the land productivity shock. The interaction coefficient point estimate is highly statistically significant but economically small. On average, net migration increased by approximately 1 individual in treated counties because of the productivity shock. It therefore appears that not many agents moved between counties in search of employment. Rather migration from urban to rural areas and reallocation of labor between manufacturing and agriculture took place within counties.

5 Threats to Identification

So far the results show that after the introduction of hybrid corn seeds, counties suited to corn cultivation experienced a shift to a more agrarian economy and de-urbanization. In our analysis, we control for a baseline set of county characteristics and a set of county and year fixed effects. We now consider a host of additional factors that might have affected structural change or urbanization. Perhaps the most severe threat to identification in our setting are shocks that systematically coincide with the invention of hybrid corn and correlate with corn suitability. It is not entirely obvious what these factors might be, although other productivity shocks in agriculture and new technologies appear to be the most likely candidates. We therefore conduct a series of robustness tests to affirm the validity of our key results. Owing to the large number of dependent variables used above, we focus attention on the key outcomes: agricultural and manufacturing employment, farms and establishments per capita, and urbanization.

5.1 Productivity Change in Related Industries

[Insert Table 11]

A key element of our research design is that improvements in land productivity are driven by developments within the corn industry. One danger is that the productivity gain is due to changes in related industries such that our inferences are spurious. We therefore append equation (4) with controls for wheat and barley yield in the county. These are the most closely related field crops to corn and are also ubiquitous throughout the regions where corn is grown. While we find some significant associations between wheat and barley yields in the estimation results reported in Table 11, including the additional controls has no effect on our key inferences. Rather the interaction coefficient remains broadly unchanged in terms of statistical significance and economic magnitude.

5.2 Environmental Factors

[Insert Table 12]

Previous research by Hornbeck (2012) has shown that the Dust Bowl conditions between 1935 and 1938 reduced agricultural productivity and triggered a shift from agriculture to manufacturing in affected areas.[21] If corn tends to be grown in areas that were less affected by the Dust Bowl, our results might be driven by the control group reacting to the environmental catastrophe, thereby biasing the counterfactual, and generating spuriously large average treatment effects. We control for the potential confounding effect of the Dust Bowl using county-specific erosion levels reported by Hornbeck (2012). As the dust storms began in 1935 we set

[21] The Dust Bowl was triggered by severe droughts between 1934 and 1936 which resulted in crop failure across the American plains. This loss of ground cover made farmland susceptible to wind erosion. Dust storms between 1935 and 1938 eroded enormous quantities of topsoil (Hornbeck, 2012). The harsh conditions reduced agricultural production and productivity by restricting the land available for farming.

the erosion variable to 0 for observations from 1930. The estimation results reported in Table 12 show that controlling for counties' exposure to the Dust Bowl has no effect on our conclusions about the effect of hybrid corn seeds. We also observe that more eroded counties have significantly less (more) agricultural (manufacturing) employment suggesting that economic activity reallocated towards manufacturing in counties affected by the Dust Bowl.

5.3 Mechanization

[Insert Table 13]

Our next set of controls captures the increasing mechanization of agriculture during the sample period. Tractors first became available during the 1920s and are often credited with revolutionizing agricultural productivity. While the rate of adoption was relatively slow (Manuelli and Seshadri, 2014), the new technology became more prevalent during the 1930s.

Although the increasing mechanization of agriculture coincides with the invention of hybrid corn seed, it seems unlikely that the diffusion of tractors biases our previous findings. This is because of the factor bias of the new technology: tractors increase labor productivity, rather than land productivity. Following the predictions in Bustos et al. (2015), we would expect tractors to have the opposite effects on the labor market variables and structural change compared to hybrid corn seed. That is, increasing the incidence of tractors would result in labor reallocating from agriculture to manufacturing leading to a more industrialized economy.

This is precisely what we find in Table 13. Following the adoption of tractors agricultural employment falls, presumably because tractors reduce labor requirements. A one per cent increase in the number of tractors per acre is associated with a 1.42% reduction in agricultural employment. On the other hand, increasing tractor usage is positively and significantly related to manufacturing employment. In addition, we find that tractors are associated with more farms per capita but are negatively though statistically insignificantly related to manufacturing establishments per capita. This suggests some heterogeneity in the way the manufacturing sector adjusts to factor-biased productivity change in agriculture. Whereas land-productivity shocks lead manufacturing establishments to close down, labor-productivity shocks lead existing establishments to hire additional workers. Likewise, the results show that increasing the incidence of tractors leads to more but smaller farms (in employment terms). The final column in Table 13 shows that increasing tractor use is positively correlated with urbanization, although the coefficient estimate falls below conventional levels of statistical significance.

Controlling for tractor usage has no quantitative or qualitative effects on our conclusions about how hybrid corn seeds affected structural change or urbanization. In all columns of Table 13 the T_i-$Post_t$ interaction coefficient is robust to the inclusion of the tractors per acre variable.

The economic magnitude of the effect is similar to before and the coefficient estimates remain statistically significant.

5.4 New Deal Programs

[Insert Table 14]

Part of the government's response to the Great Depression aimed to stimulate the economy by providing subsidies to agriculture and increasing certain crop and animal prices. The Agricultural Adjustment Act of 1933 (AAA) paid farmers to reduce crop acreage and the government purchased and destroyed livestock to bid up prices. Although these measures were subsequently declared unconstitutional, efforts to support agriculture by limiting crop acreage were included in subsequent legislation (Hornbeck, 2012).

The land productivity gains we observe in corn production may therefore reflect a transition to using less marginal land. Alternatively, subsidy payments may allow credit-constrained producers to adopt other productivity-enhancing technologies, such as tractors. To rule out these concerns we use information from Fishback (2005) on government support provided through the AAA and related programs to each county. This provides a time-varying, county-specific value of government support to agriculture. The results are reported in Table 14. Our findings are unaffected by the inclusion of this new covariate in the estimating equation.

5.5 Spatial effects

The final plausible source of contamination of the estimates is through reactions within the control group. For example, if the introduction of hybrid corn seeds trigger a change in behavior within the control group the average treatment effects may either be biased or spurious as they are driven by developments within the control group rather than the reaction of treated counties to the labor-productivity shock. This seems unlikely given corn accounts for very few planted acres within the control group. Nevertheless, we test for spillover effects using a Monte Carlo procedure.

To implement this test we use only observations from the control group. We randomly assign 54% of counties to the placebo treatment group and the remainder to the control condition.[22] Next, we estimate equation

$$y_{it} = \alpha + \beta Placebo_i * Post_t + X_{it}'\delta + \gamma_i + \gamma_t + \varepsilon_{it}, \tag{6}$$

where $Placebo_i$ is equal to 1 if a county is in the placebo treatment group, 0 otherwise; and save the p-value on the coefficient β. We repeat this process 500 times for each dependent variable.

[22]This values is based on the fact that in our whole sample 54% of counties are in the treatment group.

Intuitively, what this procedure does is check whether there are any differential trends in y_{it} within a control-group county following treatment. If spillovers are not important we know that the null of zero effect on the coefficient β is correct and we would only reject the null if we make a type-1 error. Hence, if the control group is unaffected by the shock, the rejection rates should be around 10% at the 10% significance level, 5% at the 5% significance level etc.[23] On the other hand, if the control group is systematically affected by the shock we would expect to observe substantially higher rejection rates. In Appendix Table A4 we find rejection rates that are broadly in line with type-1 errors. It therefore seems unlikely that our findings arise due to spillovers on the control group.

A related concern is that the standard errors are serially correlated across space. This would be the case if a given observation is correlated with observations in other locations. To deal with this issue we use a spatial error model (SEM) to estimate equation (1). The key difference between the previous specification is that the error term is now given by

$$\varepsilon_{it} = \lambda W \varepsilon_{it} + u_{it},$$

where W is the spatial weight matrix and the parameter λ is a coefficient on the spatially correlated errors, and u_{it} is the error term. The main coefficient of interest in the model is λ which indicates whether the errors are spatially correlated. The results in Appendix Table A5 show that our findings are robust to the change in estimation strategy. Moreover, there is no evidence that the residuals are spatially correlated.

5.6 Suitability Index

So far, our analyses have assumed that a county is suited to growing corn if the GAEZ index is greater than or equal to 50. Our findings could be contested on the grounds that this threshold is inappropriate or that it captures other forces. To rule out this possibility, we examine the robustness of our key inferences to using a higher suitability threshold. In Appendix Table A6 we designate the treatment group as counties with GAEZ index values of 70 or above, 0 otherwise. This provides a tougher test as a smaller number of counties comprise the treatment group. Despite this change, our inferences remain unaltered: post treatment the treatment group experience a significant increase in agricultural employment and farms per capita whereas manufacturing employment, establishments per capita and urbanization decrease.

[23]The rejection rates will not be perfectly equal to the significance level because we run 500 replications rather than an infinite number.

6 Conclusions

A rapidly expanding body of literature has begun to examine the role of agriculture in structural change (Foster and Rosenzweig, 2004, Nunn and Qian, 2011, Hornbeck and Keskin, 2014, Bustos et al., 2015, Eberhardt and Vollrath, 2015). This study tests theoretical predictions emanating from a new theory about how factor-biased productivity shocks in agriculture affect structural change by Bustos et al. (2015). Our empirical tests leverage the introduction of hybrid corn seed as an exogenous source of variation in US land productivity. This technology produced larger corn plants leading to a 17% increase in land productivity (output per acre).

Consistent with the model's predictions we find that counties suited to cultivating corn experienced structural change as their economies became more agrarian. Specifically, the increase in land productivity triggered an increase in demand for labor in the agricultural sector, causing workers to move from manufacturing to agriculture. We also find a decline in the number of manufacturing establishments per capita in treated counties, whereas the number of farms per capita rise. Namely, the direction of structural change is from agriculture to manufacturing.

Furthermore, we are able to shed light on the overall effect of structural change on the level of economic development. Since urbanization rates provide reasonable proxies for economic development, changes in urbanization provide insights into the overall standard of living. Following the introduction of hybrid corn seeds, treated economies experienced a 3% reduction in the share of the population living in urban areas. This is consistent with a reduction in economic development. This effect stems from how the new technology affected the labor market and the structure of treated economies. As economic activity became more concentrated in agriculture, which tends to be less productive relative to the manufacturing sector, living standards fell. We also provide evidence that the land-productivity shock caused unemployment: while it created new jobs in agriculture these gains were offset by relatively larger declines in manufacturing employment. Moreover, many of the new firms created within the agricultural industry following the shock are small and owned by tenants. This suggests that a greater share of agricultural output was produced by relatively unproductive firms.

Together this evidence indicates that the factor-bias of agricultural productivity shocks is an important determinant of structural change and economic development.

References

Acemoglu, D., Johnson, S., and Robinson, J. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5):1369–1401.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust difference-in-difference estimates? *Quarterly Journal of Economics*, 119(1):249–275.

Boustan, L., Fishback, P., and Kantor, S. (2010). The effect of internal migration on local labor markets: American cities during the great depression. *Journal of Labor Economics*, 28(4):719–746.

Bustos, P., Caprettini, B., and Ponticelli, J. (2015). Agricultural productivity and structural transformation. evidence from brazil. CREI WOrking Paper.

Clark, C. (1940). *The Conditions of Economic Progress*. London: MacMillan and Co.

Costinot, A., Donaldson, D., and Smith, C. (2015). Evolving comparative advantage and the impact of climate change in agricultural markets: Evidence from 1.7 million fields around the earth. *Journal of Political Economy*.

Crow, J. (1998). 90 years ago: The beginning of hybrid maise. *Genetics*, 148:923–928.

East, E. (1908). Inbreeding in corn. Connecticut Agricultural Experimental Station.

Eberhardt, M. and Vollrath, D. (2015). Agricultural technology, structural change, and growth. *CSAE Working Paper WPS/2014-21*.

Fernandez-Cornejo, J., Caswell, M., and Klotz-Ingram, C. (1999). Seeds of change: From hybrids to genetically modified foods. *Choices*, pages 18–22.

Fischer, G., Nachtergaele, F., Prieler, S., Teixeira, E., Toth, G., van Velthuizen, H., Verelst, L., and Wiberg, D. (2012). Global agro-ecological zones model documentation. *IIASA Documentation Paper*.

Fishback, P. (2005). The new deal and retail sales. *Journal of Economic History*, 65(1):36–71.

Foster, A. and Rosenzweig, M. (2004). Agricultural productivity growth, rural economic diversity, and economic reforms: India 1970-2000. *Economic Development and Cultural Change*, 52(3):509–542.

Foster, L., Haltiwanger, J., and Syverson, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability. *American Economic Review*, 98(1):394–425.

Gollin, D., Parente, S., and Rogerson, R. (2002). The role of agriculture in development. *American Economic Review Papers and Proceedings*, 92(2):160–164.

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

Helpman, E., Melitz, M., and Rubenstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *Quarterly Journal of Economics*, 123(2):441–487.

Hornbeck, R. (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, R. and Keskin, P. (2014). The historically evolving impact of the ogallala aquifer: Agricultural adaptation to groundwater and drought. *American Economic Journal: Applied Economics*, 6(1):190–219.

Kuznets, S. (1957). Quantitative aspects of the economic growth of nations ii. industrial distribution of national product and labor force. *Economic Development and Cultural Change*, 5:1–111.

Manuelli, R. and Seshadri, A. (2014). Frictionless technology diffusion: The case of tractors. *American Economic Review*, 104(4):1368–1391.

Matsuyama, K. (1992). Agricultural productivity, comparative advantage, and economic growth. *Journal of Economic Theory*, 58(2):317–334.

Melitz, M. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

Nunn, N. and Qian, N. (2011). The potato's contribution to population and urbanization: Evidence from a historical experiment. *Quarterly Journal of Economics*, 126(2):593–650.

Olmstead, A. and Rhode, P. (2008). *Creating Abundance: Biological Innovation and American Agricultural Development*. New York: Cambridge University Press.

Restuccia, D., Yang, D., and Zhu, X. (2008). Agriculture and aggregate productivity: A quantitative cross-country analysis. *Journal of Monetary Economics*, 55(2):234–250.

Rosenstein-Rodan, P. (1943). Problems of industrialisation of eastern and south-eastern europe. *Economic Journal*, 53:202–211.

Russell, W. (1974). Comparative performance for maize hybrids representing different eras of maize breeding. Proceedings of the 29th Annual Corn Sorghum Research Conference.

Schull, G. (1908). The composition of a field of maize. *American Breeders Association*, 4:296–301.

Schull, G. (1909). A pure line method of corn breeding. *American Breeders Association*, 5:51–59.

Skinner, J. and Staiger, D. (2007). *Technology Adoption from Hybrid Corn to Beta-Blockers*. NBER: University of Chicago Press.

Sutch, R. (2011). *The Economics of Climate Change: Adaptations Past and Present*. Chicago: University of Chicago Press.

Zuber, M. and Robinson, J. (1941). The 1940 iowa corn yield test. *Iowa Agricultural Experiment Station Bulletin*, 19:519–593.

Tables

Table 1: Summary statistics

Variable	Obs	Mean	Std. Dev.	Level	Source
A: NASS Agricultural Variables					
Yield per acre	4,071	2.80	1.02	County-industry	NASS
Acres planted	4,071	10.72	1.21	County-industry	NASS
Price per bushel	4,071	-0.60	0.36	State-industry	NASS
B: County-Level Variables					
Agricultural employment	5,960	0.49	0.69	County	Census
Agricultural labor intensity	5,960	-0.123	1.42	County	Authors' calculations
Manufacturing employment	5,960	2.14	8.68	County	Census
Manufacturing labor intensity	5,960	25.96	34.16	County	Authors' calculations
Unemployed	5,960	0.47	0.52	County	Census
Part unemployed	5,960	0.44	0.51	County	Census
Population density	5,960	0.63	1.67	County	Authors' calculations
Net migration	5,921	3.89	0.09	County	Winkler et al. (2013)
Fertility	5,960	-0.209	0.91	County	Authors' calculations
Male-female ratio	5,960	1.07	0.11	County	Authors' calculations
Manufacturing wages	5,960	28.73	49.03	County	Census
Establishments per capita	5,960	-7.98	2.84	County	Authors' calculations
Farms per capita	5,960	8.35	10.52	County	Authors' calculations
Land value per acre	5,960	4.82	21.03	County	Authors' calculations
Farm buildings value per acre	5,960	1.84	11.50	County	Authors' calculations
Capital per acre	5,960	1.45	7.56	County	Authors' calculations
Urbanization	5,960	0.21	0.24	County	Authors' calculations
State unemployment rate	5,960	-3.23	0.36	State	IPUMS
Erosion	5,960	0.09	0.23	County	Hornbeck (2012)
AAA assistance	5,960	0.32	0.59	County	Fishback (2005)
Tractors per acre	5,960	8.45	10.92	State	Census
C: GAEZ Data					
Suitability	5,960	48.92	16.73	State	FAO-GAEZ
T_i	5,960	0.54	0.50	County	Authors' calculations

Table 2: Productivity Effects of Hybrid Seeds

	1	2	3	4	5	6	7
Control group			Wheat			Barley	Soybeans
Corn	0.4914***	-0.2100***	-0.2061***				
	(18.97)	(-4.13)	(-3.94)				
Post	0.0297	-0.0673	0.0597	0.0601	0.0444	-0.0382	-0.2731**
	(0.68)	(-0.97)	(0.50)	(0.50)	(0.38)	(-0.31)	(-2.27)
Corn * Post	0.1721***	0.1172***	0.1068**	0.1065**	0.1245***	0.1302***	0.4908***
	(4.89)	(3.01)	(2.58)	(2.55)	(2.89)	(2.72)	(10.07)
Acres planted		0.0051	0.0051	0.0040	0.0044	-0.0043	-0.0025
		(0.54)	(0.54)	(0.40)	(0.44)	(-0.49)	(-0.19)
Price		-2.0037***	-2.0085***	-2.0074***	-2.0123***	-1.7511***	-1.6675***
		(-17.82)	(-17.93)	(-17.81)	(-18.02)	(-15.59)	(-13.84)
Unemployment rate		0.4550***	0.3099**	0.3095**	0.3104**	0.2807*	0.3987**
		(4.06)	(1.99)	(1.98)	(1.98)	(1.84)	(2.53)
Land value			0.0576*	0.0576*	0.0576*	0.0350	0.1087***
			(1.92)	(1.91)	(1.91)	(1.20)	(3.59)
Capital			-0.0000	-0.0000	-0.0000	-0.0001	0.0000
			(-0.58)	(-0.58)	(-0.58)	(-1.20)	(0.31)
Placebo					-0.1345**		
					(-2.47)		
Corn * Placebo					0.0908		
					(1.24)		
County FE	√	√	√				
Year FE	√	√	√	√	√	√	√
County * industry FE				√	√	√	√
N	4,071	4,071	4,071	4,071	4,071	4,051	4,760
R^2	0.80	0.83	0.83	0.83	0.83	0.81	0.80

Notes: The dependent variable in all regressions is land productivity, measured in bushels per acre. The standard errors are clustered at the county level and the corresponding t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Table 3: Input Usage

	1	2	3	4
Dependent variable:	Seed	Fertilizer	Lime	Pesticides
Post	0.2750	0.1015	0.5327	-0.3587**
	(1.02)	(0.34)	(1.45)	(-2.67)
N	41	41	41	41
R^2	0.75	0.45	0.82	0.95

Notes: Robust standard errors are used to calculate heteroskedasticity robust t-statistics reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Table 4: Suitability and Corn Production

	1	2
SI	0.0290***	0.0290***
	(19.58)	(17.26)
State unemployment rate		-0.0857***
		(-2.98)
Manufacturing estabs.		-0.0084***
		(-2.95)
Manufacturing VA		-0.0003
		(-0.15)
Total farms		0.0278***
		(2.71)
Year FE	√	√
N	5,969	5,969

Notes: This table reports probit estimates of equation 3. Coefficient estimates are marginal effects. The dependent variable is equal to 1 if the number of corn bushels produced in the county is greater than zero, 0 otherwise. Heteroskedasticity robust t-statistics are reported in parentheses. *** and ** indicate statistical significance at the 1% and 5% levels, respectively.

Table 5: Parallel Trends Test

	1	2	3	4
Dependent variable	Farms per capita	Manufacturing employment	Establishments per capita	Urbanization
T_i * $D1930_t$	0.0064	0.0064	-0.1824	-0.0031
	(0.35)	(0.06)	(-1.40)	(-0.76)
County FE	x	x	x	x
Year FE	x	x	x	x
N	6,163	6,170	6,170	6,054
R^2	0.04	0.01	0.60	0.18

Notes: The standard errors are clustered at the county level and the corresponding t-statistics are reported in parentheses. $D1930_t$ is a dummy variable equal to 1 for the year 1930, 0 otherwise. All regressions use data for 1920 and 1930 except for the results reported in column 4 which use data for 1910 and 1930. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Table 6: Comparability of the Treatment and Control Groups

Variable	Difference	Std. Error	t-statistic
Agricultural employment	-0.1556	0.1233	-1.26
Manufacturing employment	2.3437	2.2324	1.05
Farms per capita	-0.1497	0.1776	-0.84
Establishments per capita	-0.2613	0.4779	-0.55
State unemployment rate	0.1599	0.1280	1.25
Urbanization	0.0362	0.0463	0.78
Manufacturing VA	-0.0333	0.3890	-0.09

Notes: This table reports the mean difference between the level of variable y in the treatment and control group and the associated standard error and t-statistic. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

Table 7: Size of the Corn Industry

	Land share	Acreage share
T_i=1	8.86	35.30
T_i=0	0.61	2.50
Difference	8.25***	32.80***

Notes: Both land share and acreage share are measured in percent. T_i is the treatment group indicator. $T_i = 1$ denote counties belonging to the treatment group (that is, they have GAEZ suitability index values of 50 or above). 0 denotes counties in the control group (GAEZ suitability index values below 50). *** indicates statistical significance at the 1% level.

Table 8: Labor Market Effects

	1	2	3	4	5	6
	Agriculture		Manufacturing			
Dependent variable	Employment	Labor intensity	Employment	Labor intensity	Totally unemployed	Part unemployed
T_i * $Post_t$	0.0972***	0.3770***	-0.1694***	-0.2268	0.0338***	0.0326***
	(4.87)	(11.93)	(-2.87)	(-0.25)	(6.76)	(6.59)
State unemployment rate	0.0659*	0.7340***	-0.0319	-11.4043***		
	(1.71)	(10.53)	(-0.27)	(-5.90)		
Establishments per capita	0.0022	0.0237***	0.0090	1.8274***	-0.0000	0.0002
	(1.00)	(2.74)	(1.31)	(8.87)	(-0.03)	(0.10)
Manufacturing VA	-0.0035*	0.0069**	-0.0040	-0.3756**	-0.0016**	-0.0017**
	(-1.87)	(2.02)	(-0.72)	(-2.48)	(-2.32)	(-2.34)
Farms per capita	-0.5965***	-0.3931***	-0.3719***	3.9031***	0.0245**	0.0253**
	(-32.14)	(-8.94)	(-3.97)	(5.14)	(2.14)	(1.97)
County FE	x	x	x	x	x	x
Year FE	x	x	x	x	x	x
N	5,960	5,960	5,960	5,960	5,960	5,960
R^2	0.84	0.88	0.05	0.10	0.04	0.04

Notes:The standard errors are clustered at the county level and the corresponding *t*-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Table 9: Structural Change

	1	2	3	4	5	6	7
	Establishments per capita	Manufacturing wages	Farms per capita	Full owner	Part owner	Tenant	Farm size
$T_i * Post_t$	-0.4077***	2.4465**	0.1183***	-0.0060**	-0.0110***	0.0191***	-0.0783***
	(-3.75)	(2.12)	(4.23)	(-2.52)	(-7.94)	(8.11)	(-7.39)
State unemployment rate	0.5345**	-2.1229	0.2410***	-0.0393***	-0.0047	0.0597***	-0.0695**
	(1.97)	(-0.82)	(3.61)	(-6.57)	(-1.35)	(10.99)	(-1.97)
Manufacturing VA	-0.0149	1.7236***	0.0230***	0.0003	0.0000	-0.0001	-0.0005
	(-1.07)	(6.98)	(8.31)	(1.27)	(0.26)	(-0.50)	(-0.47)
Establishments per capita		1.5398***	-0.0829***	-0.0022***	0.0010***	0.0008*	0.0122***
		(11.16)	(-18.61)	(-4.88)	(3.37)	(1.89)	(5.17)
Farms per capita	-1.3094***	-5.7486***					
	(-15.91)	(-6.36)					
County FE	x	x	x	x	x	x	x
Year FE	x	x	x	x	x	x	x
N	5,960	5,960	5,960	5,960	5,960	5,960	5,960
R^2	0.19	0.21	0.99	0.06	0.04	0.10	0.99

Notes: The standard errors are clustered at the county level and the corresponding *t*-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Table 10: Urbanization Effects

	1 Urbanization	2 Net migration
T_i * $Post_t$	-0.0059***	1.1169***
	(-2.58)	(5.23)
State unemployment rate	0.0076	2.5006**
	(1.53)	(2.30)
Manufacturing VA	0.0006**	-0.0642**
	(2.34)	(-1.98)
Establishments per capita	0.0006	-0.0883
	(1.06)	(-1.13)
Farms per capita	-0.0050***	1.1561*
	(-3.18)	(1.90)
County FE	x	x
Year FE	x	x
N	5,960	5,960
R^2	0.08	0.03

Notes: The standard errors are clustered at the county level and the corresponding t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Table 11: Other Agricultural Productivity Shocks

	1	2	3	4	5
Dependent variable:	Agricultural employment	Manufacturing employment	Farms per capita	Establishments per capita	Urbanization
T_i * $Post_t$	0.1067***	-0.1387**	0.1245***	-0.5105***	-0.0069***
	(4.90)	(-2.18)	(4.15)	(-4.41)	(-2.75)
Wheat yield	0.0031	0.0161**	0.0072*	-0.0585***	-0.0007**
	(0.96)	(2.50)	(1.72)	(-3.75)	(-2.01)
Barley yield	0.0022	-0.0123***	-0.0142***	0.0528***	0.0009**
	(0.81)	(-2.68)	(-3.40)	(2.91)	(2.21)
Control variables	x	x	x	x	x
County FE	x	x	x	x	x
Year FE	x	x	x	x	x
N	5,960	5,960	5,960	5,960	5,960
R^2	0.84	0.05	0.99	0.19	0.09

Notes: Both wheat and barley yields are measured in bushels per acre (in natural logarithms). The standard errors are clustered at the county level and the corresponding t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Table 12: Dust Bowl Conditions

	1	2	3	4	5
Dependent variable:	Agricultural employment	Manufacturing employment	Farms per capita	Establishments per capita	Urbanization
T_i * $Post_t$	0.1116***	-0.2069***	0.1074***	-0.3946***	-0.0059***
	(5.64)	(-3.43)	(3.85)	(-3.67)	(-2.61)
Erosion	-0.2121***	0.5519***	0.1537***	-0.1904	0.0005
	(-7.70)	(5.32)	(3.74)	(-1.03)	(0.16)
Control variables	x	x	x	x	x
County FE	x	x	x	x	x
Year FE	x	x	x	x	x
N	5,960	5,960	5,960	5,960	5,960
R^2	0.84	0.05	0.99	0.19	0.08

Notes: The standard errors are clustered at the county level and the corresponding *t*-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Table 13: Mechanization Tests

	1	2	3	4	5
Dependent variable:	Agricultural employment	Manufacturing employment	Farms per capita	Establishments per capita	Urbanization
T_i * $Post_t$	0.0908***	-0.1635***	0.1356***	-0.4127***	-0.0057**
	(4.55)	(-2.77)	(4.95)	(-3.83)	(-2.50)
Tractors per acre	-0.0142***	0.0133**	0.0537***	-0.0114	0.0005
	(-5.64)	(2.37)	(12.70)	(-0.53)	(1.15)
Control variables	x	x	x	x	x
County FE	x	x	x	x	x
Year FE	x	x	x	x	x
N	5,960	5,960	5,960	5,960	5,960
R^2	0.84	0.05	0.99	0.19	0.08

Notes: The standard errors are clustered at the county level and the corresponding *t*-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

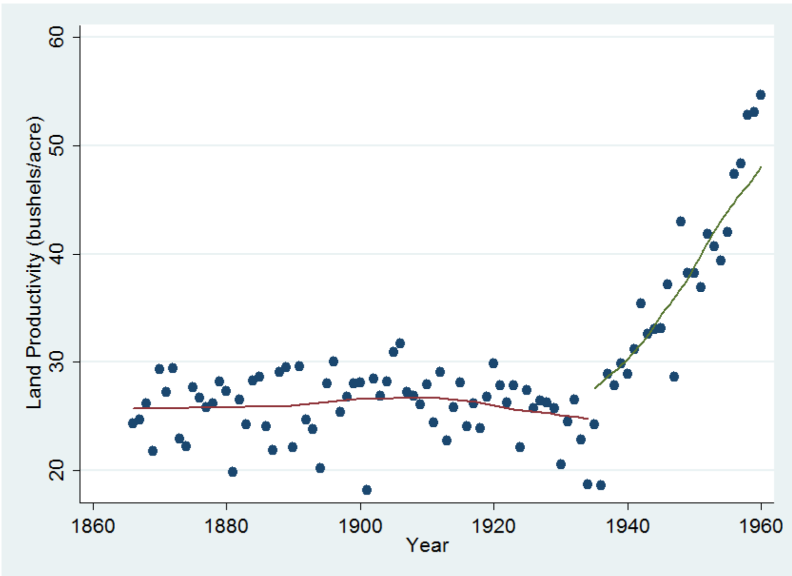
Table 14: New Deal Assistance

	1	2	3	4	5
Dependent variable:	Agricultural employment	Manufacturing employment	Farms per capita	Establishments per capita	Urbanization
$T_i * Post_t$	0.1424***	-0.2028***	0.1004***	-0.4516***	-0.0074***
	(7.65)	(-3.29)	(3.60)	(-4.16)	(-3.28)
AAA	-0.3091***	0.2284***	0.1136***	0.3150***	0.0102***
	(-13.71)	(4.76)	(5.87)	(4.02)	(5.68)
Control variables	x	x	x	x	x
County FE	x	x	x	x	x
Year FE	x	x	x	x	x
N	5,960	5,960	5,960	5,960	5,960
R^2	0.87	0.05	0.99	0.19	0.09

Notes: The standard errors are clustered at the county level and the corresponding t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

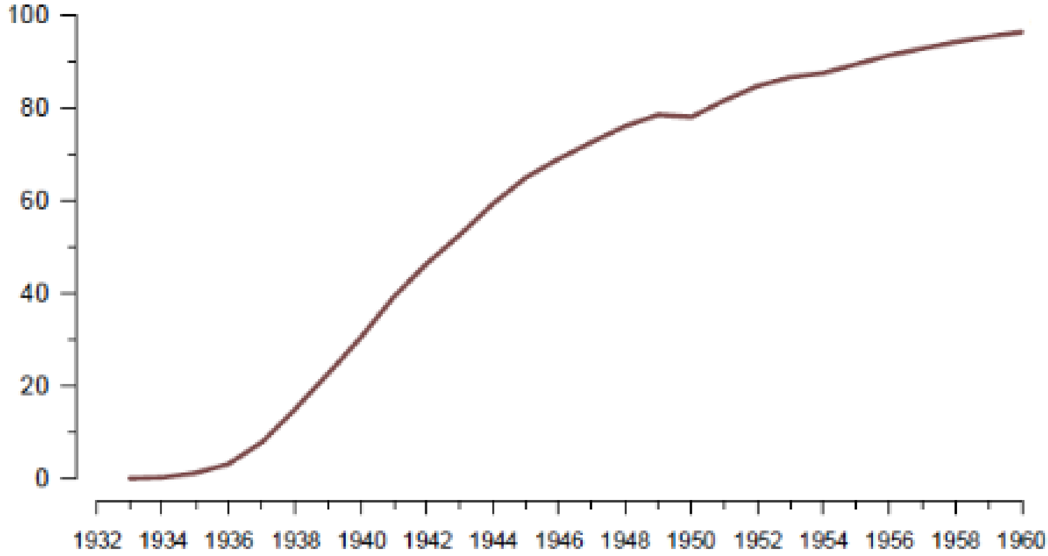
Figures

Figure 1: Historic Corn Productivity



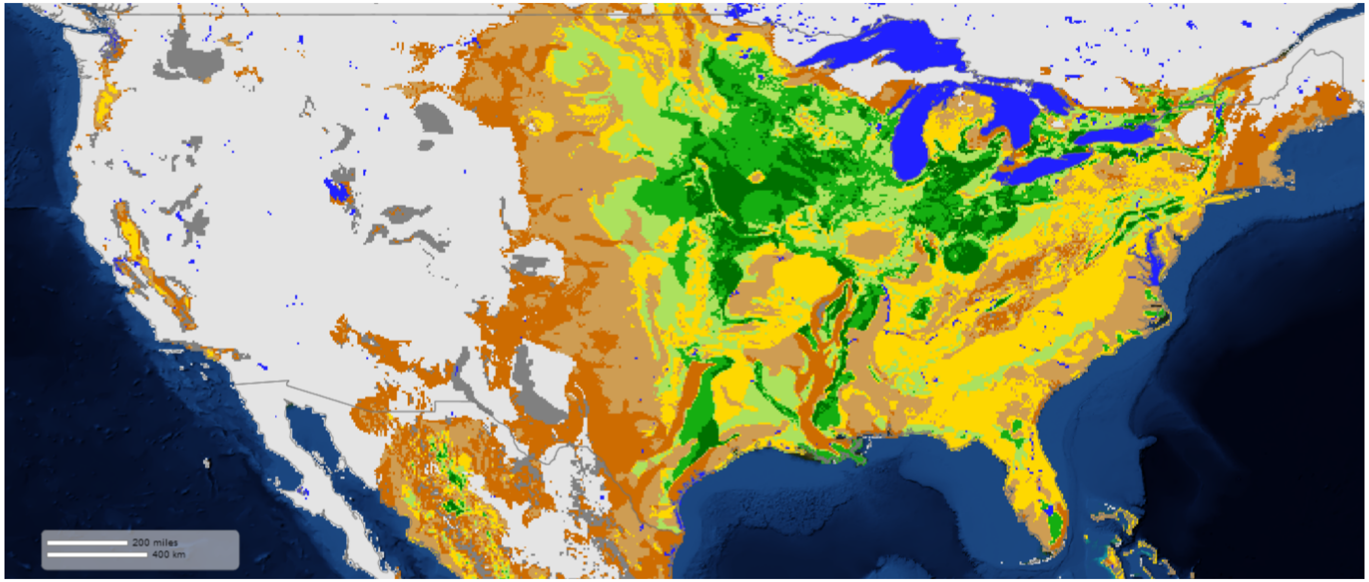
Notes: This figure plots annual yield per acre in the US corn industry between 1866 and 1960. Local polynomial regression functions are fitted to the data before and after 1935 in order to illustrate trends in the data.

Figure 2: Percentage of Corn Acres Planted with Hybrid Seeds



Notes: This figure plots the annual percentage of corn acres planted using hybrid corn seed between 1932 and 1960. Data is taken from the USDA (1962) *Agricultural Statistics*, Table 46, page 41.

Figure 3: Suitability to Corn Production



Notes: Green areas are most suited to growing corn. White areas are not suited to growing corn. Data are taken from the FAO-GAEZ database.

Appendix

Appendix A: Summary Statistics

Table A1: Additional Summary Statistics

Variable	Obs	Mean	Std. Dev.	Level	Source
A: NASS Agricultural Variables - Barley Counterfactual					
Yield per acre	4,051	2.82	0.97	County-industry	NASS
Acres planted	4,051	10.72	1.21	County-industry	NASS
Price per bushel	4,051	-0.71	0.41	State-industry	NASS
B: NASS Agricultural Variables - Soybeans Counterfactual					
Yield per acre	4,760	2.88	0.89	County-industry	NASS
Acres planted	4,760	10.87	1.10	County-industry	NASS
Price per bushel	4,760	-0.51	0.40	State-industry	NASS

Table A1 reports summary statistics for the variables used in the productivity regressions when barley (Panel A) and soybeans (Panel B) are used as the control group.

Appendix B: Variable Description

Agricultural employment: total number of workers employed in agriculture (in thousands).
Agricultural labor intensity: agricultural workers per acre (ln).
Manufacturing employment: total number of workers employed in manufacturing (in thousands).
Manufacturing labor intensity: number of workers per manufacturing establishment.
Unemployed: ratio of totally unemployed workers to the population (ln).
Part unemployed: ratio of partly unemployed workers to the population (ln).
Population density: ratio of total population to square miles.
Net migration: ratio of net migrants to total population (ln).
Fertility: ratio of children aged 5 or less to total population (ln).
Male-female ratio: ratio of men to women.
Manufacturing wages: ratio of total manufacturing wages to manufacturing employment.
Establishments per capita: ratio of manufacturing establishments to total population (ln).
Farms per capita: ratio of number of farms to total population.
Land value per acre: ratio of total value of agricultural land to acres planted.
Farm buildings per acre: ratio of total value of farm buildings to acres planted.
Capital per acre: ratio of machinery and equipment value to acres planted.
Urbanization: ratio of urban dwellers to total population.
State unemployment rate: see text for description.
Erosion: percentage of land eroded by dust storms between 1935 and 1938.
AAA assistance: total value of support provided through the Agricultural Adjustment Act of 1933 and related programs (in millions).
Tractors per acre: the tractors variable is taken from the US Agricultural Census. The Census reports the total number of tractors in each state for 1930 and 1940. We then divide this number by the total number of planted acres and take the natural logarithm.
Farm size: mean farm size in acres (ln).
Full-owner share: share of farms that are fully owned by an operator.
Part-owner share: share of farms that are part owned by an operator.
Tenant share: share of farms that are operated by a tenant.

Appendix C: Productivity Robustness Tests

Table A2: Robustness Testing

	1	2	3
Dependent variable:	Acres planted	Output	Yield
Post	-0.1154	0.1916	0.1641
	(-0.78)	(0.33)	(1.58)
Corn * Post	0.0762	1.9707***	0.1077**
	(0.62)	(4.29)	(2.57)
Acres planted			-0.0159
			(-1.55)
Price			-2.0613***
			(-17.89)
State unemployment rate			0.3312**
			(2.09)
Land value			0.0659**
			(2.29)
Capital			-0.0000
			(-0.47)
Tractors			-0.0058***
			(-4.75)
County-industry FE	x	x	x
Year FE	x	x	x
N	4,071	4,071	4,071
R^2	0.66	0.81	0.83

Notes: The standard errors are clustered at the county level and the corresponding t-statistics are reported in parentheses. *** and ** indicate statistical significance at the 1% and 5% levels, respectively.

Appendix D: Additional Structural Change Tests

Table A3: Structural Change Tests

	Land value	Building value	Yield
T_i * $Post_t$	4.6979***	2.5176***	
	(4.88)	(4.99)	
State unemployment rate	-1.6064	-1.1512*	-0.4561***
	(-1.16)	(-1.87)	(-4.45)
Manufacturing VA	0.1532***	0.0659**	0.0162*
	(2.96)	(2.35)	(1.73)
Establishments per capita	-0.0406	-0.0375	-0.0192
	(-0.60)	(-1.29)	(-1.39)
Farms per capita	3.0962***	1.4714***	
	(8.86)	(11.21)	
Tenant farm share			-5.7229***
			(-6.76)
County-industry FE	x	x	x
Year FE	x	x	x
N	5,960	5,960	5,960
R^2	0.10	0.05	0.10

Notes: The standard errors are clustered at the county level and the corresponding t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Appendix E: Spillover Effects

Table A4: Spillover Analysis

Dependent variable:	Agricultural employment	Manufacturing employment	Farms per capita	Establishments per capita	Urbanization
Rejection rate at the 10% level	10.6%	12.2%	9.6%	9%	9.8%
Rejection rate at the 5% level	6.2%	6.2%	4.4%	3.8%	5.6%
Rejection rate at the 1% level	1%	0.4%	1.2%	0.2%	1.8%

Notes: All rejection rates are calculated using two-tailed tests.

Table A5: Spatial Econometric Results

	1	2	3	4	5
Dependent variable:	Agricultural employment	Manufacturing employment	Farms per capita	Establishments per capita	Urbanization
T_i * $Post_t$	0.8838*	-0.4822***	0.7034*	-0.0560***	-0.0113***
	(1.88)	(-2.85)	(12.23)	(-1.78)	(-8.96)
λ	0.9899	0.9000	0.9800	0.9890	0.9800
	(0.10)	(0.10)	(0.12)	(0.10)	(0.10)
Control variables	x	x	x	x	x
County FE	x	x	x	x	x
Year FE	x	x	x	x	x
Spatial FE	x	x	x	x	x
N	5,960	5,960	5,960	5,960	5,960
R^2	0.56	0.78	0.98	0.77	0.64

Notes: λ is the spatial autocorrelation parameter. The standard errors are clustered at the county level and the corresponding t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Appendix F: GAEZ Threshold Robustness Tests

Table A6: Alternative GAEZ Thresholds

Dependent variable:	Agricultural employment	Manufacturing employment	Farms per capita	Establishments per capita	Urbanization
$T70_i * Post_t$	0.1707***	-0.4322**	0.1849***	-0.5141***	-0.0057**
	(9.14)	(-2.14)	(5.78)	(-4.59)	(-2.30)
Control variables	x	x	x	x	x
County FE	x	x	x	x	x
Year FE	x	x	x	x	x
N	5,960	5,960	5,960	5,960	5,960
R^2	0.84	0.05	0.99	0.18	0.08

Notes: $T70_i$ is a dummy variable equal to 1 if the GAEZ corn suitability indicator is greater than or equal to 70, 0 otherwise. The control variables are the state unemployment rate, manufacturing value added, establishments per capita, and farms per capita. We drop farms per capita and establishments per capita from the control variables in column 3 and 4, respectively. The standard errors are clustered at the county level and the corresponding t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Institut de Recherches Économiques et Sociales
Université catholique de Louvain

Place Montesquieu, 3
1348 Louvain-la-Neuve, Belgique

ISSN 1379-244X D/2015/3082/19