NBER WORKING PAPER SERIES

AGRARIAN ORIGINS OF INDIVIDUALISM AND COLLECTIVISM

Martin Fiszbein
Yeonha Jung
Dietrich Vollrath

Working Paper 29643
http://www.nber.org/papers/w29643

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2022, revised May 2026

We thank Sam Bazzi, Quoc-Anh Do, Paul Rhode, Cory Smith, David Weil, and seminar/conference participants at the World Bank, Brown University, University of Maryland - Agricultural and Resource Economics, University of Connecticut, University of California - San Diego, Universidad de San Andrés, Seoul National University, the Southern Economic Association, North Carolina State University, Tulane University, the Econometric Society's Australasian Meeting, the Virtual Economic History seminar, UMass - Boston, Universidad Alberto Hurtado and Universidad Andrés Bello for helpful comments and suggestions. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Agrarian Origins of Individualism and Collectivism
Martin Fiszbein, Yeonha Jung, and Dietrich Vollrath
NBER Working Paper No. 29643
January 2022, revised May 2026
JEL No. N51, O13, P16

ABSTRACT

This study examines the influence of agricultural labor intensity on individualism across U.S. counties. We measure historical labor intensity in agriculture by combining data on crop-specific labor requirements with county-level crop mix around 1900. Potential endogeneity in agricultural labor intensity is addressed using climate-induced variation in crop mix. The estimates indicate that a one standard deviation increase in labor intensity is associated with a reduction of 0.2-0.3 standard deviations in individualism (as captured by the share of children with infrequent names). We also document significant changes in individualism over time, in relation to within-county shifts in labor intensity due to mechanization and the boll weevil shock. Further evidence from contemporary online search queries and social media language suggests that historical labor intensity continues to influence geographic variation in individualism today. Lastly, our decomposition of labor intensity provides insights into the mechanisms underlying this cultural impact.

Martin Fiszbein
Boston University
Department of Economics
and NBER
fiszbein@bu.edu

Yeonha Jung
Sungkyunkwan University
yhjung0408@skku.edu

Dietrich Vollrath
University of Houston
devollrath@uh.edu

1 Introduction

Individualism is a salient dimension of cross-country cultural variation and a key determinant of economic and political organization (e.g. Heine, 2010; Cline and Williamson, 2017; Gorodnichenko and Roland, 2017, 2020; Henrich, 2020). In this paper, we examine the effects of labor intensity in agriculture on individualism, adding to a growing literature on this link between culture and production structure (see Talhelm et al., 2014; Talhelm and Oishi, 2018; Talhelm, 2020; Talhelm and English, 2020; Ang, 2019). We provide novel evidence using rich subnational variation from the United States, leveraging climate-induced variation in crop mix, and studying changes in response to major historical shocks to labor intensity. In addition, we study the mechanisms through which labor intensity affects individualism.

Our baseline analysis examines the impacts of agricultural labor intensity on individualism across U.S. counties, which provide extensive and granular cross-sectional variation in production patterns. U.S. agriculture included a wide range of crops with different levels of labor intensity, creating not only sharp regional differences but also substantive local variation. Our cross-sectional empirical approach is based on within-state variation and controls for a wide array of geo-climatic characteristics. Moreover, we introduce an instrumental variable strategy that isolates climate-induced variation in crop mix. The results establish significant positive effects of historical labor intensity on both contemporaneous and present-day measures of individualism.

We also show that historical *changes* in labor intensity led to *changes* in individualism in two empirical settings, each leveraging a different source of identifying variation in labor intensity: the mechanization of agriculture in the late 19th century and the boll weevil infestation in the early 20th century. Beyond confirming the cross-sectional results by accounting for unobserved county characteristics, these exercises contribute to our understanding of cultural formation and the time frame over which it occurs.

Labor intensity could have fostered non-individualistic norms through multiple mechanisms. It may have been driven by production processes requiring many operations or large teams of workers within tasks. These features often entailed interdependence among workers, promoting the formation of non-individualistic social norms that favor cooperation and coordination while reducing shirking and absenteeism. Another mechanism relates to the distribution of labor requirements over the growing season. Crops that were more often produced by stable, dedicated groups over the course of the growing season (as opposed to crops where labor was concentrated in short planting and harvesting windows) led to repeated interactions that would contribute to the development of cooperative norms. We lay out a simple conceptual background and empirical analysis to explore these mechanisms using detailed historical data on agricultural production processes.

Our empirical analysis is based on extensive data collection and construction. To measure labor intensity in agriculture at the county-level, we collect data on hours of work per acre for various crops from multiple historical sources that capture standard farming techniques in the 19th and early 20th century. We combine this information with Census data on the acreage devoted to each crop to construct a county-level index of agricultural labor intensity. Our measure of individualism comes from social psychology: the prevalence of infrequent children names, which reflect a desire to stand out, as opposed to common names, which reflect a desire to fit in (Twenge et al., 2010; Varnum and Kitayama, 2011; Bazzi et al., 2020; Knudsen, 2022). This measurement approach builds on the defining traits of individualism emphasized by Hofstede (1991) and Triandis (1995)—the view of the self as independent rather than interdependent, and the regulation of behavior by personal attitudes rather than social norms. We use Census microdata to compute these data at the county-level and for different groups at different points in time.

As part of our identification strategy, we exploit cross-sectional variation in the composition of agricultural production generated by climatic features. We construct an instrumental variable (IV) for agricultural labor intensity using data on climate-based potential yields for different crops from IIASA and FAO (2012). We estimate a fractional multinomial logit (FML) model of crop choice, in which the county-level shares of agricultural products are functions of the crop-specific potential yields. With the predicted shares for each crop, we compute an index of potential agricultural labor intensity, which we use as an IV for actual intensity.

The IV estimates indicate that a one standard deviation increase in labor intensity in 1900 is associated with a 0.27 standard deviation reduction in individualism during that period. Despite a smaller magnitude (0.19 reduction), the ordinary least squares (OLS) estimates also support the strong negative association between labor intensity and individualism. These results survive an array of robustness checks, including additional controls for agricultural properties, demographic characteristics, and socioeconomic conditions at the time. They also hold *within* different regions of the United States, such as the South and non-South, and for both white and black populations. Moreover, the impacts of agricultural labor intensity on culture persist today. Using Google Trends data, we find that counties with higher labor intensity in 1900 exhibit lower search interest in typical examples of individualistic words (e.g., "unique", "peculiar", "autonomous", "solo") relative to their antonyms (e.g., "common", "ordinary", "collaborative", "together"). This pattern is corroborated by a greater search interest in team sports relative to individual sports.

Next, we turn to examine the effect of *changes* in agricultural labor intensity on individualism by leveraging two historical shocks. The first shock stems from the mechanization of agricultural production, which differentially affected the labor requirements of various crops, leading to county-level variation in labor intensity depending on the original crop mix. Information on labor requirements by crop before and after mechanization is available from the *Hand and Machine Labor* Census report (Wright, 1899). Our estimates indicate that a one standard deviation increase in labor intensity from 1860 to 1900 is associated with a 0.31 standard deviation reduction in individualism during the same period. These first-differenced results are not affected by unobservable time-invariant characteristics of U.S. counties that might affect the cross-sectional estimates. We also confirm that the cultural shift driven by changes in labor intensity is robust to a range of time-varying controls potentially associated with agricultural mechanization. Our findings also hold when excluding the South from the sample.

The second historical shock draws on the boll weevil infestation, an insect that feeds on cotton and spread throughout the U.S. South between the 1890s and 1930s. This devastating pest had heterogeneous incidence across counties at different points in time. By forcing a change in crop mix, the weevil induced changes in labor intensity. These shocks led to shifts from cotton, a very labor intensive crop, to alternative crops with different degrees of labor intensity, depending on the agricultural suitability of each county. Our results show that the arrival of the boll weevil induced a 0.20 standard deviation increase in individualism across all counties in the South. Moreover, the effects of the boll weevil shock on individualism were larger where cotton production gave way to low labor intensity crops, such as wheat or rye.

Using more detailed crop-level data from the *Hand and Machine Labor* Census report (Wright, 1899), we examine the mechanisms linking agricultural labor intensity and individualism. We find evidence for two. The first is the number of distinct steps involved in production of a given crop (e.g. wheat only had five distinct operations, while tobacco had seventeen). More steps not only raised labor intensity but also, given their complementarities, made workers more dependent on one another, as in an "O-ring" production process (Kremer, 1993). This made communitarian norms that reduced shirking and absenteeism more valuable. The second involves the distinction between hours and workers. Crops that involved many hours from each individual worker over the course of a season (e.g. cotton) were associated with a persistent set of workers that provided more opportunities for interaction with the same group over time, leading to a more communitarian culture. In contrast, crops with more sporadic seasonal demands (e.g. wheat) could engage many workers, but each for only short periods of time, with less need for cooperation and fewer opportunities for repeated interaction.

Our paper adds to a large literature on the deep roots and persistence of cultural traits (e.g., Guiso et al., 2016; Nunn and Wantchekon, 2011; Spolaore and Wacziarg, 2013; Voigtländer and Voth, 2012; Bazzi et al., 2020), particularly echoing those that focus on historical conditions in agricultural production. Alesina et al. (2011) show that crops associated with plow use are linked to a lower status for women and higher fertility. Bardhan (2000) and Buggle (2020) suggest that the coordination of labor required by irrigation systems is associated with collectivist cultures, which is consistent with arguments that the nature of food production is a determinant of where cultures fall on the individualism-collectivism dimension (Barry et al., 1957; Mann, 1986; Wyer et al., 2013). Specific features of agricultural production have been linked to the evolution of time preferences (Galor and Özak, 2016), in-group favoritism (De Vliert, 2011), generalized trust (Buggle and Durante, 2021), and the origin of religious communities (Ager and Ciccone, 2018). We provide rigorous new evidence on the roots of individualism in agriculture.

Our analysis echoes other works in this literature that study both cultural persistence and cultural change, not viewing them as contradictory but rather as two aspects of cultural evolution (see Giuliano and Nunn, 2021). Examining cultural change allows us to overcome a common limitation of persistence studies based solely on cross-sectional variations. Combining evidence on persistence and change dispels the notion that persistence implies that culture is immutable. While our results suggest that culture can change rapidly (i.e. in decades) in response to relevant shocks, culture may also be persistent as long as temporal shocks are not systematically related to fundamental determinants of cultural background.

We also contribute to a burgeoning literature on individualism across the social sciences. This cultural dimension has attracted enormous interest from social psychologists, whose contributions framed it as a fundamental dimension of cross-cultural variation (e.g. Triandis, 1995; Heine, 2010). A number of contributions in economics and other social sciences have emphasized how individualism shapes crucial aspects of economic and political organization (e.g. Greif, 1994; Kyriacou, 2016; Cline and Williamson, 2017; Gorodnichenko and Roland, 2011, 2017, 2020). Our findings add to the emerging strand of research on the origins of individualism (e.g., Knudsen, 2022; Olsson and Paik, 2016; Hoang-Anh et al., 2019; Bazzi et al., 2020; Roland, 2020), and in particular to those focusing on the role of agricultural labor intensity (see Talhelm et al., 2014; Talhelm and Oishi, 2018; Talhelm, 2020; Talhelm and English, 2020). Considering the United States as historical case study and extrapolating micro-level evidence to higher levels of aggregation, our findings complement other explanations of American individualism—in particular, frontier settlement (Bazzi et al., 2020)—by suggesting that low-intensity cereal agriculture in the Northeast and Midwest may have also played a significant role.

Our results complement previous studies while offering new insights into the mechanisms underlying labor intensity's effects on cultural norms. In their analysis of rice versus wheat in China, Talhelm et al. (2014); Talhelm and Oishi (2018); Talhelm (2020); Talhelm and English (2020) underscore that rice cultivation required critical and binding labor exchange arrangements among farmers, as well as cooperation and coordination to build, maintain, and manage irrigation networks. In contrast, labor exchange arrangements were very limited in our context, with little to no production of paddy rice using extensive irrigation networks.[1] Our findings suggest that the links between labor intensity and collectivism are broader, likely involving additional channels—i.e., interdependencies unrelated to irrigation as well as the stability of work interactions throughout the agricultural season. Another suggestive implication is that, while the cooperation involved in large-scale irrigation may affect norms at the community-level (across kin networks or families), in our context, the influence likely operated largely within families or kin networks—and then plausibly spilled over to the broader community.

Finally, we add to the growing body of work showing how agricultural labor intensity shapes various economic and cultural outcomes. Several studies link variation in labor intensity to

[1] Rice was produced in the U.S. in this period, but was farmed mainly as a rain-fed grain similar to wheat or barley, or in limited areas as wet rice in natural floodplains.

economic development and structural change (see Vollrath, 2011; Eberhardt and Vollrath, 2018; Johnson and Vollrath, 2020). Studies on cultural outcomes have not been restricted to individualism: another key contribution is Fouka and Schläpfer (2020), which establishes a positive effect of agricultural labor intensity on work ethics, showing that this aspect of production shapes individuals' attitudes not only about interactions with others but also about their own work and leisure. Our paper makes a contribution to the measurement of labor intensity, complementing the use of U.S. Census data with various historical sources to validate the data on crop-specific labor requirements and to cover a wide array of crops, which may be useful in future studies.

The remainder of the paper is organized as follows. Section 2 explains how we measure agricultural labor intensity and individualism. Section 3 shows the cross-sectional results, and Section 4 explores the dynamic effects based on agricultural mechanization and the boll weevil shock. Section 5 provides further evidence on the underlying mechanisms. Section 6 concludes.

2 Measuring Labor Intensity and Individualism

2.1 Agricultural Labor Intensity

We measure labor intensity by the number of man-hours per acre used to produce a specific crop, based on data from historical sources described below. These crop-specific labor intensity measures are then combined with the actual acreage share of crops in a county (or the potential share based on geoclimatic conditions) to calculate county-level measures of agricultural labor intensity. In these county-level measures, labor intensity for each crop is fixed, and variation in overall labor intensity across counties results from variations in crop mix.

We consider four sources that measure crop-specific labor requirements from the late 19th and early 20th century, a period in which agriculture represented a major share of the national labor force. We combine the information from these different studies, which cover different regions and crops, to provide a comprehensive and cross-validated set of information on crop-specific labor intensities.

The small scale study *Labor requirements of crop production* by Cooper et al. (1916) was designed to yield detailed and accurate measures of agricultural labor use, production, and finances. It calculated man-hours per acre for 13 crops based on a comprehensive analysis of 45 farms from three counties in Minnesota between 1902 and 1912. Route men visited each farm on a regular basis, weighted products, and recorded labor use. Taking into account all farm-year observations, the study reports the average hours of labor per acre for each crop, including an allowance for general tasks pertaining to the entire farm operation. The figures were compared with those for other farms not included in the original study, finding deviations in man-hours per acre of only approximately 2%. While this source provides high quality data, it does not cover all relevant crops and its geographic coverage is narrow.

A complementary source with much broader geographic coverage is *Labor used for field crops* by Hecht and Vice (1954), published in the U.S. Department of Agriculture's *Statistical Bulletin*. It compiles labor requirements for 21 crops between 1910 and 1914 from the Works Progress Administration's National Research Project series, the Bureau of Agricultural Economics, and reports from state experiment stations. Man-hours per acre by average adult males are reported for each crop, considering the time for all the steps in the production process, as in hauling manure, plowing, planting, cultivating, harvesting, and hauling the crops. The time for farm maintenance and overhead work was calculated separately and allocated to total man-hours by crop, similar to Cooper et al. (1916). The U.S. Bureau of the Census (1960) provides an additional source of information for three major crops—wheat, corn, and cotton—from 1800 to 1900, based on publications by state and federal agencies. Finally, Reuss (1930) provides information on sugarcane production, which is not available from the other sources, based on a farm-level study.

The crop-specific labor requirements from our sources are displayed in Figure 1, with values expressed relative to wheat requirements from the same source, the only crop available in all sources (Appendix Table A3 reports the data in absolute terms). The dots are color-coded by source, and the blue bars indicate the average across sources for each crop. The figure reveals considerable variation in labor intensity across crops. There was a group of crops—barley, oats, hemp, hay, and flax—with labor requirements similar to those for wheat. Corn and rye required twice as much labor per acre as wheat, and potatoes more than three times as much. Man-hours for cotton were approximately 7.5 times those for wheat, while sugar beet and tobacco were even more labor intensive.

Differences across sources in the reported man-hours for given crops are substantive in some cases (e.g., rye and potatoes), but the differences are generally much smaller than those in average values across crops. Consistency across diverse sources suggests that our measures capture labor requirements that were relatively stable across locations and for different years within our period of interest. This is in line with the statement by Cooper et al. (1916) that man-hours per acre tend not to vary with yield or labor productivity, for which they offer empirical support in the setting of their study.

The comments above might be interpreted to mean that recorded man hours per acre reflect crop-specific labor requirements that are identical across locations. However, we need not make such an extreme assumption. We use the magnitude of man-hours per acre for each crop *relative* to wheat. We hold these relative values constant across counties, but implicitly allow the absolute amount of man-hours per acre to vary across counties. The relative amount of man-hours per acre of any crop relative to a baseline crop (e.g. wheat) will be constant across counties so long as the shape of the production function is similar across counties.

This argument is formalized in Appendix G, where we also connect our results to others that estimate historic crop-specific production functions (Fouka and Schläpfer, 2020). The intuitive logic is as follows: If farms within a county take crop prices, wages, and rental rates as given, then in choosing how to allocate land and labor across crops, the optimal man-hours per acre for any crop will be proportional to the wage/rent ratio of that county, scaled by the relative importance of labor and land in the production function (i.e. their elasticities) *of that crop*. Given the same wage/rent ratio across crops within a county, the only difference in man-hours per acre across crops will come from dissimilarities in the relative importance of labor and land in the crop-specific production functions. Our measure therefore captures variations in the elasticities of the production function for different crops.

In sum, by using the relative man-hours per acre of crops, our measure of labor intensity is not a proxy for absolute differences in productivity, prices, or the wage/rent ratio across counties. It is a measure of the relative labor intensity of crop production based on the shape of the underlying production functions.

While the conditions described above are sufficient for the accuracy of our county-level measure of agricultural labor intensity, the necessary conditions for the empirical analysis are indeed less stringent. For instance, labor requirements by crop could vary across regions without invalidating our analysis. Idiosyncratic differences in labor intensity across counties that are not associated with our variables of interest would create classical measurement errors, entailing attenuation bias but not other sources of bias. While regional differences in institutions and agricultural organizations could create non-classical measurement errors, this concern is alleviated by the fact that our results are based on within-state cross-county variation and the fact that they hold both within the North and the South. Furthermore, we introduce an IV strategy in Section 3.2 that hinges on climate-based variation in the crop mix.

We use the crop-level measures of labor intensity to construct county-level measures of labor intensity in agricultural production by taking the weighted average of crop-level man hours with the acreage share of each crop in county (c) as weights:

$$\text{Intensity}_c = \sum_j \text{Crop share}_{j,c} \times \text{Crop intensity}_j \tag{1}$$

Crop share$_{j,c}$ is the acreage share of crop j in county c, which we draw from the Census of Agriculture digitized by Haines et al. (2005). Crop intensity$_j$ denotes the relative labor requirements of crop j. Therefore, Intensity$_c$ is the weighted average of crop-specific labor intensity using the acreage share of each crop as the weight.

Figure 2 shows the spatial distribution of Intensity$_c$ (Panel a) and its deviations from the state-level means (Panel b). While there was stark regional contrast, with the South characterized by much higher levels of labor intensity, there was also significant within-state variation everywhere, which is leveraged in our empirical approach.

Appendix Table A3 reports the data on crop-specific labor requirements from the four sources used in our measure of labor intensity as well as figures from another source, the *Thirteenth Annual Report of the Commissioner of Labor* (Wright, 1899), which considers production methods before and after mechanization in the 19th century. The figures for the latter method come from observations very close to 1900, our period of interest, but it is not clear that they are comparable to the figures from our other sources, so we leave them out of our baseline measure of labor intensity. We use data from the report in Section 4.1.[2]

2.2 Individualism

Our measures of individualism rely on insights from the social psychology literature. Twenge et al. (2010) argue that frequent names convey an intention to fit in, whereas infrequent names convey an intention to stand out. By analyzing the first names of American babies from 1880 to 2007, they document an increasing trend toward infrequent names, which they interpret as a growing interest in uniqueness and individualism. Varnum and Kitayama (2011) argue that the relative frequency of popular baby names decreases in the western regions of the U.S. and Canada, where an identical pattern is observed between regions where Europeans have settled (e.g., Australia, New Zealand, and the US) and European countries. Using Hofstede's individualism score, they argue that the differences in naming patterns reflect the individualistic culture of frontier settlement, a thesis that is studied with more granular data and different methods by Bazzi et al. (2020). Names-based measures of individualism have also been used in non-Western societies. The increasing prevalence of unique names is also observed in China (Cai et al., 2018) and Japan (Ogihara et al., 2015), coinciding with a rise in individualism scores. Knudsen (2022) provides a battery of validation tests showing a strong correlation between the distribution of names and standard measures of individualism-collectivism across and within countries.

Following these contributions, we measure county-level individualism in the historical period using the complete count census data. We compute the share of children aged 0-9 whose names are out of the top 10 most frequent names. This threshold is common in the social psychology literature, but different thresholds lead to analogous results. The frequency of names is calculated at the U.S. Census region level for each decade's birth cohort (Appendix C shows that the results are robust to defining name frequency at the county, state, or national levels).

We restrict the sample to native white and black children with native parents to eliminate the divergent naming patterns of recent immigrants. To avoid bias due to different naming patterns by race and compositional differences, we consider the frequency of names within the white and black population separately. Cook et al. (2014) document that distinctive naming patterns of African Americans are a long-standing cultural norm, that traces back at least to the late 19th century. For example, Appendix A.2 shows the top 10 most frequent names of whites and blacks at the national

[2]Appendix Figure A1 plots the crop-specific labor requirements from Wright (1899) for both production methods against the average of our baseline measures, in logs. While the cross-crop differences are broadly consistent, labor requirements for some crops are very different. For wheat, which is our baseline numeraire from the other sources, the machine method specifies extremely low requirements. This makes the values for other crops relative to wheat in the report very high. And since the report includes twelve crops, including these relative values in our baseline measure creates major shifts and changes in the ranking of crops that do not seem sensible. However, the results of our baseline analysis hold if we only use data from the report, and we use the data by itself later in specific sections of the paper.

level. The sample does not include other racial categories because of their small population sizes; in the 1900 Census, the proportion of population other than whites and blacks was 0.6%. While our baseline measures includes both blacks and whites (in each case considering the corresponding set of most frequent names), Section 3.3 provides comparative results when using only white or black children to measure individualism. The spatial patterns of the share of infrequent names computed in 1900 are illustrated in Figure 3.

3 Labor Intensity and Individualism Across U.S. Counties

This section examines the hypothesis that agricultural labor intensity fosters non-individualistic cultures. Using rich subnational data from the U.S. around 1900, we find a strong negative correlation between actual labor intensity and the share of infrequent names, a proxy for individualism. To address the possible presence of bias due to endogenous crop choice, we use a measure of potential labor intensity as an instrumental variable. We also conduct a number of additional checks to establish the robustness of the estimated effects and probe their heterogeneity by race and region. In Section 5, we provide empirical results that shed light on mechanisms of this relationship using detailed crop-level data.

Our empirical analysis relies on variation across counties controlling for state fixed effects, therefore comparing units within a given institutional context. The U.S. provides substantive identifying variation even when we do not rely on comparing crops at the extremes of the distribution of labor intensity (i.e. sugarcane in the South versus wheat in the North), but rather on comparisons of crops of intermediate labor intensity within each region. Regional institutional differences therefore do not drive our results, but they are relevant to the context and may have influenced the responsiveness of culture to agricultural labor intensity; in particular, slavery and later forms of labor coercion in the South, by extracting effort through external forces, may have decreased the scope for cultural mechanisms favoring cooperation. What is key from our perspective is that, insofar as we find a significant link between labor intensity and individualism with the same sign in both regions despite their institutional differences, this link would be of general relevance.

3.1 Estimating Equation

Our estimating equation is

$$y_{c,1910} = \alpha + \beta \text{Intensity}_{c,1900} + \gamma' \boldsymbol{X}_c + \mu_s + \epsilon_c \ , \tag{2}$$

where $y_{c,1910}$ denotes the proportion of infrequent names measured in year 1910 for county c, and $\text{Intensity}_{c,1900}$ is the index of agricultural labor intensity measured in 1900 .$\boldsymbol{X}_c$ is the vector of predetermined county characteristics that could have impacted both labor intensity and cultural orientation simultaneously, which consists of climatic (temperature and precipitation), ecological (agricultural land suitability, terrain elevation, and slope), and geographical (distance to major cities, distance to coastal line, distance to navigable river, latitude, and longitude) conditions. μ_s and ϵ_c represent the state fixed effects and error terms, respectively. To address the likelihood of spatial correlation, standard errors are clustered on 60mi-by-60mi grid squares following Bester et al. (2011). Appendix Table A1 provides summary statistics for all of our outcome variables and various controls.

While initial socioeconomic conditions could be another source of bias, they are not controlled in our baseline specification because they could be "bad controls," in the sense that socioeconomic conditions are co-determined with individualism, and thus their inclusion may mask the relationship between crop labor intensity and individualism. Nevertheless, Section 3.3 shows that the effects of labor intensity on individualism are robust to a large set of initial socioeconomic conditions, such as population density, other agricultural characteristics aside from labor intensity, and local demographic characteristics.

3.2 Instrumental Variable Strategy

Crop choice is endogenous. A natural concern about the OLS estimates is that individualism could influence the composition of production and thus labor intensity. For example, if certain historical conditions in a county precipitated an influx of collectivist migrants, the production of labor-intensive crops could increase as a result of selective migration. Besides reverse causality, there could also be omitted variables affecting both agricultural labor intensity and individualism. To address these concerns, we develop a measure of *potential* labor intensity based on climatic characteristics .

We begin with attainable yields of each crop in each county, which are computed at 0.5 arc-min resolution based on climatic conditions and crop-specific characteristics by the FAO-GAEZ. We consider the measures for rain-fed conditions and intermediate input levels, which match the environment of agricultural production in the late 19th and early 20th century in the US. Adopting the measures for the other scenarios does not affect our results.

Using the county-specific attainable yields of each crop, we estimate the potential share of crops following Fiszbein (2022), which adopts a conditional logit framework of McFadden (1973). The intuition follows a typical discrete choice model. Suppose a farmer in county c maximizes $\pi_{j,c} = \phi_j \mathbf{A}_c + u_{j,c}$ where $\pi_{j,c}$ is the profit from growing crop j and $\mathbf{A}_c$ is the vector of crop-specific attainable yields and $u_{j,c}$ captures all other factors affecting crop choice (e.g. cultural preferences). If $u_{j,c}$ follows the Type 1 extreme value distribution, then the profit-maximizing probability of growing crop j is derived as $e^{\phi_i' \mathbf{A}_c} / \left(1 + \sum_{j=1}^{I-1} e^{\phi_j' \mathbf{A}_c}\right)$. Based on a fractional multinomial logit (FML) framework (Papke and Wooldridge, 1996; Mullahy, 2015), we estimate the parameters $\hat{\phi}_i'$ with data on acreage shares of crops as the measure of probabilities, and the attainable yield data from the FAO-GAEZ as the vector of determinants in $\mathbf{A}_c$.[3] The estimated parameters of the FML model reflect the price and cost differentials among agricultural products as well as other factors that affect the link between physical productivity and profits for each crop.

Given the estimated values of $\hat{\phi}_i'$, we then estimate the potential acreage share—$\hat{\theta}_{j,c}$—for a given county as

$$\hat{\theta}_{j,c} = E[\theta_{j,c}|\mathbf{A}_c] = \frac{e^{\hat{\phi}_j'\mathbf{A}_c}}{1+\sum_{i=1}^{I-1} e^{\hat{\phi}_i'\mathbf{A}_c}} , \tag{3}$$

which is the fitted value for a county from the FML estimation. Thus, it strips out the non-agro-climatic variation present in the residuals.

Using the estimated potential acreage share, $\hat{\theta}_{j,c}$ of different crops by county, we construct the potential labor intensity of a county as

$$\text{Potential intensity}_c = \sum_j \hat{\theta}_{j,c} \times \text{Crop intensity}_j , \tag{4}$$

where $\hat{\theta}_{j,c}$ is the potential share of crop j in county c. Crop Intensity$_j$ is the relative crop-specific labor intensity from Table A3.

Figure 4 plots the actual labor intensity index in 1900 against the potential labor intensity index from the same year. There is a positive relationship, as expected, actual labor intensity varies around this potential, likely reflecting endogenous crop choices influenced by factors beyond inherent productivity. In the empirical analysis that follows, we use the measure of potential labor intensity as an instrument for the actual labor intensity.

3.3 Main Results

This section examines the contemporaneous relationship between agricultural labor intensity and individualism in 1900–1910 in the US. The estimates using the actual and potential labor intensity indicate that agricultural labor intensity had strong negative effects on individualism at the local level. The results are robust to including a battery of additional controls and to splitting the sample by race or region.

Table 1 shows the baseline relationship between agricultural labor intensity and individualism. The outcome variable is the share of children aged 0-9 with infrequent names in 1910, and the index of agricultural labor intensity is measured in 1900. The OLS estimates in Panel 1 show strong negative correlations between the two variables. The associations are robust to agricultural land suitability and other geo-climatic controls, implying that the effects of agricultural labor intensity are not arising due to general agricultural conditions. According to Column (4) with all controls, a one standard deviation increase in labor intensity in 1900 is associated with a 0.187 standard

[3] For the estimation of (3), attainable yields of 19 crops are used: cotton, oat, corn, wheat, barley, pasture legumes, potato, tobacco, rye, flax, alfalfa, dry pea, buckwheat, chickpea, bean, sugarcane, sorghum, sugar beet, and soybean.

deviation decrease in the share of infrequent names in 1910. The IV estimation, displayed in Panel 2, supports the findings of strong negative effects of labor intensity on individualism, with coefficients between 0.264 and 0.275. The effect size is comparable to the findings of Talhelm et al. (2014).[4]

The IV estimates are larger in absolute magnitude than the OLS coefficients, although not statistically different. While keeping in mind that the differences are not significant at conventional levels, it may be interesting to consider the factors that could explain them. The differences might result from attenuation bias or potential upward bias in the OLS estimates. It could also reflect a differentially higher effect of labor intensity on culture for counties with higher levels of cultural malleability. Given the close link between cultural and socioeconomic configurations, higher cultural malleability could also be associated with higher responsiveness of agricultural production patterns (in particular, actual labor intensity) to the geo-climatic conditions captured in the IV, which would likely imply stronger treatment effects of the potential labor intensity on individualism-collectivism (for a discussion of hardwired and fluid cultures, see Acemoglu and Robinson, 2025).

The findings are robust to considering alternative names-based measures of individualism. For our baseline measures, the frequency of names is calculated at the U.S. Census region level for each decade's birth cohort. This reflects the notion that there may be broad regional differences in naming patterns, while at the same time naming choices respond to cultural norms broader than purely local micro-communities. To address the potential concern that finer or broader geographic benchmarks for naming frequency may be appropriate, Appendix C shows that the results are robust to considering names-based measures at the county-, state-, and country-level.

Robustness. Given that our measure of potential labor intensity is constructed with exogenous measures of crop suitability, the IV strategy mitigates concerns about omitted variables. Nevertheless, some initial socioeconomic conditions might confound the interpretation of the results. Though such local characteristics could be in part results of labor intensity and thus "bad controls," robustness to their inclusion would support a causal interpretation of our findings.

Table 2 documents the OLS and IV estimates with various additional controls. Column (1) controls for overall population density, which could foster the production of labor-intensive crops

[4] Talhelm et al. (2014) report that a one-standard-deviation increase in the rice share is associated with a 0.17 standard deviation decrease in implicit individualism. To compare this with our result, we use two additional pieces of information. First, while paddy rice was not present in our context and thus we have no historical U.S. data on its labor requirements, we know from Talhelm and English (2020) that it "required twice the labor hours per hectare as crops like wheat, corn, and potatoes," so we assume it required 71.7 hours/acre per year—two times the average of requirements for corn (33.1 hours/acre per year), potatoes (60.2), and wheat (14.2) in our data. Second, thanks to information kindly shared by Thomas Talhelm through personal communication, we know that a one-standard-deviation increase in the rice share is equivalent to an increase of 0.31. This means that, compared to a county that only produces wheat, a county that devoted 31% of its land acreage to paddy rice would have, using the distribution of labor intensity observed in our context, an increase in intensity of 0.8 standard deviations. According to our estimates, this translates into a 0.15-0.22 standard deviation decrease in the individualism index (considering the range of effect sizes across OLS and IV specifications), which encompasses the 0.18 standard deviation decrease observed in the context of China.

and may also promote more collectivist behavior through channels unrelated to agricultural labor intensity. Column (2) controls for the urbanization rate to test whether local development status affected our findings. Columns (3)-(7) consider agricultural properties other than labor intensity: the proportion of employment in agriculture, average farm size, yield per farm acre, crop concentration, and soil heterogeneity. We control for agricultural employment share in consideration that the prevalence of agriculture alone could be a source of collectivism.[5] By including the average farm size and average farm yields, we assess whether scale economies or productivity of crop production confounds the effects of labor intensity. Crop concentration, measured by the Herfindahl–Hirschman index of crop acreage shares, is controlled to mitigate potential bias from homogeneous conditions of agricultural production. Similarly, we control for the soil heterogeneity index, a determinant of close- versus loose-knit social structures (Raz, 2025).

Given the lower labor intensity in western regions, our results could be affected by historical frontier experience, which is closely linked to the formation of individualistic cultures (Bazzi et al., 2020). In this respect, Column (8) controls for the index of total frontier experience. Columns (9)-(11) consider county characteristics related to ethnic composition. While the name-based measure of individualism is computed from native children with native parents, it does not rule out the possibility that immigrant grandparents influenced naming practices of their grandchildren. Moreover, if the location of the immigrant grandparents were associated with the crop mix across counties, it could bias our estimates. In this context, among native children aged 0 to 9 with native parents, we control for the proportion of native grandparents and English-speaking grandparents. We also include the diversity of birth countries in the county population to address concerns that ethnic diversity could affect naming patterns among the native population. Column (12) includes all additional controls. The results remain largely unchanged in all specifications.

Heterogeneity by Region and Race Heterogeneity across regions and races might challenge the generalization of the cultural effects of agricultural labor intensity. Differences between the South and non-South are a particular concern, given the distinctive nature of southern agriculture and its coercive institutions. To address this, we examine the relationship between agricultural labor intensity and individualism within the South and the rest of the country, and among whites and blacks separately.

Within-region effects can be heterogeneous for multiple reasons. The sets of crops grown in the South and elsewhere were different, so it is important to assess whether the link between agricultural labor intensity and individualism holds for dissimilar comparison sets. Another reason to examine possible differential effects in the South is the higher importance of agriculture: According to IPUMS microdata for 1900, the agricultural employment share was 64.7% in the South and 32.2% elsewhere.

[5] For example, Singelis et al. (1995) conduct survey of undergraduate students from diverse backgrounds and conclude that agricultural-rural cultures tend to be more collectivist.

The sharp historical contrasts in institutions and social stratification between regions are also likely sources of heterogeneity in cultural responses to agricultural patterns. Even after the abolition of slavery, the South continued to have various formal and informal institutional mechanisms of labor coercion and racial oppression. The old plantation system was replaced by alternative contractual structures, prominently including sharecropping and tenancy farming that maintained high degrees of labor control (see, e.g., Alston and Higgs, 1982; Ransom and Sutch, 2001). The history of slavery and continued racial discrimination in the U.S. also imply that the cultural effects of agricultural labor intensity could be different for blacks and whites.

Panel 1 of Table 3 splits the sample by region. Columns (1) and (2) present the estimates for the southern sample, while Columns (3) and (4) show the estimates for the rest of our observations. We find significant estimates with the same sign across all columns. The link between labor intensity and individualism thus holds across sharply different demographic structures and historical institutions in the southern and non-southern regions. The results are also consistent with a weaker relevance of cultural mechanisms in the South, where the history of direct labor coercion may have limited the differential fitness of cooperative social norms for different levels of labor intensity.[6] The key takeaway from our perspective is that, despite institutional differences, we find a significant link between labor intensity and individualism with the same sign in both regions, suggesting that this relationship would be of general relevance.

In Panel 2, we reproduce the estimates for blacks and whites separately, and find similar results for both groups. For those sub-samples, the measure of individualism is calculated based on naming patterns only within the respective racial group (e.g. the infrequent name share for blacks is calculated as the proportion of black children with names outside the 10 most common names for black children). While the OLS estimate is slightly larger for whites, the IV estimates of the effects of labor intensity on individualism are consistent across races. Thus, the regional differences in Panel 1 do not appear to be driven by racial composition.

Migration. The estimated impacts of labor intensity on culture may reflect lower individualism among the existing population as well as an inflow of less individualistic migrants. If the latter was the dominant factor, our results would reflect spatial sorting rather than a shift in local cultural attitudes. As a simple additional check, we consider individualism calculated separately for migrants and non-migrants. Appendix Table A10 shows that the estimated coefficients for migrant and non-migrant families are negative and significant across the OLS and IV specifications, with magnitudes similar to the baseline estimates. This suggests that, while selective migration may have contributed to the effects of higher labor intensity on culture, it likely complemented the effects on local cultural attitudes rather than providing an alternative explanation for our findings.

[6] While regional differences in the importance of agriculture could in principle influence the heterogeneity in link between labor intensity and culture, this factor would actually work against the observed patterns, given that the South employed a far larger percent of workers in agriculture in 1900.

3.4 Long-Term Effects of Historical Labor Intensity on Individualism

This section investigates whether the historical labor intensity in agriculture left an imprint in culture over the long term. Since Census data on names are not available beyond 1940, analysis on contemporary individualism requires alternative measures. To this end, we adapt two approaches from social psychology to examine subnational variation in cultural attitudes using online search interests. First, we track online search volumes for individualistic versus communitarian terms and those for individual versus team sports using Google Trends data. Second, we use the contemporary individualism–collectivism index from Havaldar et al. (2024), constructed using knowledge-guided lexical models applied to Twitter data.

Google Trends data For the analysis based on online search interests, we rely on Google Trends data, which are available across metropolitan regions defined by Designated Metro Areas (DMA). Since DMAs are larger than counties, we harmonize the county-level data with the DMA boundaries as detailed in Appendix J. Google Trends provides relative search volumes ranging from 0 to 100, indicating proportions relative to the highest search queries among the specified terms for the designated period and regions. The search volumes are computed from 2004 (when the data become available) to the present. Our analysis reports standardized coefficients to facilitate interpretation.

Our first approach captures the relative search frequency of individualistic versus communitarian terms. We consider pairs of words (e.g., "unique" versus "common") to capture relative interest, ensuring that our measures are not conflated with unrelated overall differences in search intensities. We seek for terms that strongly connote either an individualistic or collectivist orientation, representing a clear association with defining features of each. Drawing on previous studies that track word usage, we exclude terms that encode multiple meanings or interpretations (e.g., "independence."). We also exclude terms with low search frequencies, as their cross-sectional variation tends to be limited and likely noisier.

Based on these criteria, we consider four pairs of terms associated with core features of individualism and collectivism. We begin with *"unique"* and its antonym *"common,"* which are respectively linked to individualism's emphasis on standing out and collectivism's emphasis on fitting in (see, e.g., Twenge et al., 2012; Greenfield, 2013). The second pair, *"peculiar"* versus *"ordinary,"* captures the same defining contrast between valuing distinctiveness and embracing conformity. The remaining pairs—*"autonomous"* versus *"collaborative"* and *"solo"* versus *"together"*—have been used in prior text-based analyses of individualism (e.g., Oyserman and Lee, 2008; Twenge et al., 2012), and capture another key dimension of individualistic cultures: independence and autonomy (Oyserman et al., 2002; Santos et al., 2017). Our analysis considers each of the four pairs separately (the ratio of search intensities for the individualistic and the communitarian term), and the first principal component of the four measures.

Our second approach uses search interests in individual sports versus team sports, which complements the above measure along analogous dimensions of individualism and collectivism.

Given their emphasis on cooperation and shared success, team sports likely reflect a community-oriented mindset. Their prevalence may therefore be used to capture cultural attitudes along the individualism-collectivism spectrum (see House et al., 2004). Consistent with this idea, empirical studies show that individuals with a stronger collectivist orientation are more likely to value communal aspects of team sports, such as group attachment, community pride, and collective identity (Armstrong, 2002; Wann et al., 2008; Han et al., 2016).

We consider three popular team sports (football, basketball, and baseball) and four individual sports (golf, tennis, boxing, and wrestling) that exhibit considerable volume and variation in online search intensity across metropolitan areas. While it makes sense to track relative interest in individual versus team sports, there is no one-to-one correspondence connecting each individual sport with a team sport. Thus, we measure the search volume of each individual sport relative to the combined interest in the three team sports. Examining different sports mitigates concerns that sports preferences may be correlated with socioeconomic characteristics (e.g. tennis and golf being high-income "country club sports"). We measure the interest in each of the four individual sports (relative to team sports) separately, also the first principal components of those four measures.

Table 4 presents our estimates of the long-term cultural effects of historical labor intensity. Panel 1 displays the estimated effects on the four measures of relative search volumes for word pairs capturing individualism versus collectivism, as well as on the first principal component of these measures. The coefficients are consistently negative across all specifications, indicating that higher labor intensity in 1900 predicts lower search interest in individualistic terms. Using the first principal component as the outcome variable, the results indicate that a one standard deviation increase in labor intensity corresponds to a 0.49–0.74 standard deviation decrease in individualism. Similarly, Panel 2 shows that metro regions with higher historical labor intensity exhibit lower search interest in individual sports relative to team sports, as well as the first principal component. The coefficients are negative in all cases and statistically significant for three of the five OLS estimations and all of the IV estimations.[7] In both panels, the results are stronger for the first principal component than other measures, suggesting that the negative estimates largely reflect shared characteristics across terms and across sports, supporting these measures as meaningful proxies of cultural attitudes in the individualism-collectivism dimension.

Twitter data Despite consistent results across different settings, analyses based on Google Trends outcomes may raise two concerns. First, because words with broader connotations or insufficient metro-level variation must be excluded, the set of search terms is necessarily limited. Second, even if the selected terms accurately capture individualism–collectivism, this relationship may not be reliably reflected in their search interests.

To address such concerns, we conduct a complementary analysis using the index of

[7] The weaker associations in the case of golf and tennis could reflect their broader connotations as country club sports, often associated with wealthier individuals and sometimes perceived as part of a broader package for college admissions.

individualism and collectivism proposed by Havaldar et al. (2024), constructed from a large corpus of geolocated Tweets. The methodology begins with a small set of expert-curated "seed words" for individualism and collectivism. This list is then expanded using word embeddings to identify and add semantically similar terms, which are weighted by relevance. Next, the expanded list is filtered to remove any term that does not positively correlate with the other words in its category, ensuring the internal consistency. The index score for a given geographical unit is calculated by summing the weighted frequencies of these words in tweets originating from that location. The data include 2,042 counties with sufficient tweet observations to compute the individualism–collectivism scores, with 1,862 counties included in our sample.

Table 5 presents the results. Given that the individualism and collectivism scores are constructed as distinct indices, we show the results using each score as the outcome variable separately. Columns (1) and (2) show the OLS and IV estimates for the individualism score. Consistent with previous evidence, higher historical labor intensity in agriculture predicts lower individualism in tweet patterns. Columns (3) and (4) report the corresponding estimates for the collectivism score, indicating a significant positive influence on collectivist orientation. In addition to the metro-level evidence based on Google Trends data, these results confirm the persistent effects of historical labor intensity on contemporary individualism-collectivism.

4 Cultural Change: Responses to Major Shocks in Labor Intensity

Even if individualism shaped by fundamental factors exhibits long-term persistence, major shocks could lead to cultural change over shorter periods. Having established the long term effects of historical labor intensity, we now document that changes in agricultural labor intensity had significant effects on individualism over relatively short periods of time. Taken together, our findings suggest that cultural persistence and change are interconnected aspects of the same process.

Analysis of cultural change also strengthens our identification. While the cross-sectional evidence supports the cultural effects of agricultural labor intensity, this relationship might be influenced by unobserved county characteristics. Adding to the IV strategy, we are able to further address this concern by examining the link between *changes* in labor intensity and *changes* in individualism, which allows us to account for unobserved cross-sectional cultural determinants.

Specifically, we consider two major shocks to agricultural production patterns: the mechanization of the late 19th century and boll weevil infestation in the early 20th century, and examine the responses in local individualism as proxied by name-based measures at different points in time. The process of mechanization in agriculture entailed heterogeneous changes in production methods across crops, creating cross-county variation in labor intensity depending on crop mix. The boll weevil induced changes in labor intensity depending on crop mix, combined with the spatial and temporal patterns of its spread across the U.S. south. We use the resulting changes in labor intensity to study whether they induced changes in individualism, controlling for county fixed effects to absorb time-invariant unobserved heterogeneity.

4.1 Mechanization

In the late 19th century, American agriculture experienced sweeping technological change and rapid productivity growth (Kendrick, 1961; Evenson, 1978; Olmstead and Rhode, 2008). Part of this process involved mechanization and declines in labor inputs per acre, with differences across crops. We measure changes in crop-level labor intensity using information from the Thirteenth Annual Report of the Commissioner of Labor (Wright, 1899), a source of extremely detailed production data used by Atack et al. (2019) in their study of mechanization in manufacturing. For the agricultural sector, the report provides detailed information on the operations and total man-hours necessary to plant, grow, and harvest an acre for many specific crops, before ("hand method") and after mechanization ("machine method"), based on data from representative farms.[8]

Mechanization reduced the total labor requirements for most crops, but the changes were heterogeneous. For example, under the "hand method" corn production was slightly more labor intensive than cotton production (182.7 versus 167.8 man-hours per acre), but under the "machine method," cotton production required 3 times more labor inputs than corn production (78.7 versus 27.5 man-hours per acre). In the case of tobacco, the total man-hours per acre even increased. Figure 5 displays the labor requirements for the 12 crops included in the report under the pre- and post-mechanization methods.[9]

Based on the man-hours per acre under the hand and machine methods, we can construct two different indices of agricultural labor intensity. Following the same approach as in Section 2, we use Equation 1 to compute county-level labor intensity (Intensity_c), now with the crop-specific labor requirements (Crop intensity_j) from Wright (1899), first for hand methods and then for machine methods.

Figure 6 presents county-level measures of agricultural labor intensity calculated for 1860 under the hand method and for 1900 under the machine-method. The indices display broadly similar spatial distributions but also differences. For example, the high levels of labor intensity in the cotton belt stand out more in 1900 than in 1860, as the decline in labor requirements was smaller for cotton farming than for other crops.

We use the two measures of labor intensity for each county c at different points in time: $\text{Intensity}_{c,1860}$ computed with the hand labor requirements and $\text{Intensity}_{c,1900}$ with the machine

[8] While the report used the terms throughout (even in its title), the use of hands and machines was not exclusive to each method. The hand method was defined as "*the primitive method of production which was in vogue before the general use of automatic or power machines*," while the machine method corresponds to the most advanced production methods utilizing machinery in 1895–1896.

[9] The data on labor requirements from Wright (1899) are not included in our baseline measure of labor intensity because, as explained in Section 2.1, combining these data with the figures with other sources raises issues of comparability. Such issues do not arise in this section because we only use data from Wright (1899). However, our baseline cross-sectional results hold even if we use a measure of labor intensity based only on data from the report. See Appendix D.

labor requirements.[10] Our estimating equation is

$$y_{c,t+10} = \alpha + \beta \text{Intensity}_{c,t} + \gamma X_{c,t} + \theta_c + \theta_{s,t} + \lambda \text{Intensity}_{c,1860} \times I[\text{Year} = 1900] + \epsilon_{c,t} \ , \tag{5}$$

where $y_{c,t+10}$ denotes the share of infrequent names in county c in year $t+10$, and $\text{Intensity}_{c,t}$ is the index of labor intensity based on the hand and machine methods for 1860 and 1900, respectively.[11] $X_{c,t}$ is a vector of time-varying county characteristics across three categories. First, to account for the confounding effects of demographic changes, we control for population density, fertility rate, and the share of migrants. Second, considering the impact of mechanization on structural change and local development, we include the proportion of farmland, manufacturing population ratio, and urban population ratio. Finally, to capture the potential impact of mechanization on agricultural conditions, we incorporate average farm size, yield per acre, and the value of farm machinery per acre. θ_c indicates county fixed effects absorbing unobserved time-invariant local characteristics, and $\theta_{s,t}$ denotes state-year fixed effects.

Despite an extensive set of controls, concerns about cultural trends may still remain. Mean reversion could be a particular issue, as counties initially high in labor intensity and low in individualism might naturally have experienced declines in labor intensity and increases in individualism over time, converging to the average. To address this possibility, we control for the interaction between the index of labor intensity in 1860 and the dummy variable for 1900 ($\text{Intensity}_{c,1860} \times I[\text{Year} = 1900]$), which captures any temporal trends in individualism potentially linked to initial labor intensity.

The results are displayed in Table 6. Changes in labor intensity are strongly (negatively) associated with changes in the prevalence of individualism. According to the OLS estimate in Column (1), a one standard deviation reduction in labor intensity through mechanization was associated with a 0.33 standard deviation increase in the share of infrequent names. Column (2) mitigates concerns about cultural trends in relation to the initial labor intensity, and Column (3) confirms the robustness of our findings to the time-varying county characteristics related to different aspects of mechanization. Using potential labor intensity as an IV, Columns (4) to (6) strengthen the causal interpretation of the results. Potential labor intensity is based not only on the crop-specific attainable yields, which are time-invariant measures, but also on crop-specific market conditions that change between one period and another. Thus, potential labor intensity can be constructed separately in 1860 and 1900, and we use them to instrument actual labor intensity in each period. Consistent with the OLS regressions, the IV estimates indicate that changes in labor intensity due to mechanization negatively impacted individualism. The results are also robust to

[10]We intersect the 1860 and 1900 county shapefiles, and the area of each intersection is used as weight to harmonize the variables measured in 1860 to the 1900 county boundaries. Appendix J illustrates more details about the harmonization process with a different example.

[11]As described in Section 2.2, the share of infrequent names in $t+10$ measures the extent of individualism between t and $t+10$.

cultural trends and time-varying county controls.

Given the estimation period (1860–1900), the validity of our findings outside the South could be an issue, as fundamental changes in Southern agriculture—such as the abolition of slavery—could confound the relationship between shifts in labor intensity and individualism. Moreover, as shown in Figure 5, the effects of mechanization on labor requirements vary more widely across Southern plantation crops, potentially threatening the generalizability of the results. To address these concerns, Table 7 presents the estimation results excluding southern counties. The estimates are not significantly different: the negative association between changes in labor intensity and individualism remains consistent, even in the absence of southern counties. Similar to the baseline estimation in Section 3.3, the IV estimates from the non-South sample are larger in magnitude, which may suggest that coercive institutions in the South could have weakened the cultural effects of changes in labor intensity.

4.2 The Boll Weevil Shocks

This section assesses the effects of changes in labor intensity by leveraging the exogenous shocks of boll weevil infestations, which induced major changes in crop choice across counties in the South. The boll weevil entered Texas from Mexico in 1892 and spread throughout the U.S. South affecting many counties at different points over the following decades. Given the exclusive feeding of this insect on cotton, the infestations were devastating for the production of cotton but not for other crops (Coakley et al., 1969). Within 5 years, the boll weevil reduced cotton production by 50%; in response to the shocks, farmers did not take land out of agriculture, but rather shifted to other crops (Lange et al., 2009).

Cotton was one of the most labor-intensive crops in that period. According to the labor requirements described in Section 2, cotton production in the early 20th century required 7.63 times more man-hours than wheat production. Thus, the boll weevil's arrival and resulting shift to crops other than cotton usually lowered labor intensity. We therefore start by examining whether counties hit by the boll weevil experienced an increase in individualism. Moreover, the shift away from cotton induced by boll weevil shocks would be larger in places where farmers switched to low intensity crops. We conduct a second exercise based on this intuition.

Our first estimating equation is

$$y_{c,t+10} = \beta \text{ Boll Weevil}_{c,t} + \gamma' \boldsymbol{X}_{c,t} + \theta_{s,t} + \theta_c + \epsilon_{c,t} \,, \tag{6}$$

where Boll Weevil$_{c,t}$ is a dummy variable indicating the presence of the weevil in county c one year before t, when decisions on crop mix for period t were made. For each county, the dummy takes a value of 1 in the year of boll weevil's arrival and thereafter, based on data from Lange et al. (2009). The equation is estimated for $t = \{1900, 1910, 1920\}$. The sample is a balanced panel of 560 southern counties in which the boll weevil was present at least once during these time periods. For consistency, we exclude counties that were not used for the baseline estimation in Section 3. We include state-year fixed effects, $\theta_{s,t}$, absorbing any state-level responses to the boll weevil, and county fixed effects, θ_c, which capture time-invariant county characteristics. $\boldsymbol{X}_{c,t}$ is a vector of time-varying controls potentially associated with the boll weevil shock and cotton farming. We first control for characteristics specific to cotton farmers, including the share of African Americans, average literacy rate, and the share of tenant farms.[12] Second, recognizing broader impacts of the weevil, we include controls for demographic characteristics (proportion of migrants and fertility rate), sectoral labor reallocation (manufacturing employment share), and local development (population density and urban population ratio).

Our hypothesis suggests that the effects of the negative shocks on cotton would be larger in locations where farmers shifted more to less labor intensive crops. To examine the heterogeneous effects of the boll weevil shock depending on the labor-intensity of the other local crops, we estimate the following equation:

$$\begin{aligned} y_{c,t} = \beta \text{ Boll Weevil}_{c,t} + \lambda \text{ Low-Intensity Crops}_{c,t} + \eta \text{ (Boll Weevil}_{c,t} \times \text{ Low-Intensity Crops}_{c,t}) \\ + \gamma' \boldsymbol{X}_{c,t} + \theta_{s,t} + \theta_c + \epsilon_{c,t} \ , \end{aligned} \tag{7}$$

where Low-Intensity Crops$_{c,t}$ is the acreage share of crops with labor requirements below those for corn, calculated after excluding cotton acreage.[13] The coefficient of interest here is the one on the interaction term (Boll Weevil$_{c,t}$ $\times$ Low-Intensity Crops$_{c,t}$).

Table 8 shows the estimates from Equations 6 and 7. As shown in Columns (1) and (2), the arrival of the boll weevil increased the share of infrequent names by 0.21-0.23 standard deviations. This is consistent with the prediction that shifts away from cotton, presumably toward less labor-intensive crops, would have increased the prevalence of individualism. Supporting this view, Columns (3) and (4) indicate significant positive effects of boll weevil shock on individualism that were stronger in locations with low labor-intensity among crops other than cotton. A one standard deviation increase in the share of low-intensity crops implies that a boll weevil shock led to an additional 0.09-0.11 standard deviation increase in individualism. These results are robust to the additional county controls, mitigating concerns about alternative explanations.

Pre-trends might be a concern in interpreting our results. For example, given the eastward spread of its dissemination, the positive coefficients of the weevil could reflect pre-existing regional differences in the growth of individualism. To address this concern, we conduct placebo tests using the share of infrequent names measured in the earlier period (1870, 1880, and 1900) in place of

[12] Cotton plantations in the postbellum South were divided into small tenant farms, which were mainly managed by black farm families who had lower levels of education than average farm laborers (Aiken, 1998; Ransom and Sutch, 2001; Jung, 2020).

[13] We use corn as a threshold considering that its labor requirements fall in the middle range of the crops we considered (See Figure 1). Moreover, Lange et al. (2009) suggest that a major consequence of the boll weevil shock was the shift from cotton to corn production, implying that corn would be an appropriate benchmark for the change in crop production following the boll weevil shock.

the 1900, 1910, and 1920 measures.[14] If the findings in Table 8 resulted from pre-existing trends rather than from the causal impact of the boll weevil shock, the results from the placebo tests would demonstrate a consistent pattern. However, the placebo estimates differ from the original estimates. As shown in Table 9, the coefficients of the boll weevil shock are not positive, and their statistical significance is low. Treatment heterogeneity exhibits a similar pattern; the coefficients of the boll weevil shock and its interaction with the share of less intensive crops are both close to zero.

5 Mechanisms Linking Labor Intensity and Individualism

The combined results from Sections 3.3 to 4.2 substantiate the link between agricultural labor intensity and individualism. However, they do not explain how or why this link operated. This section analyzes mechanisms through which labor intensity would have influenced individualism.

5.1 Conceptual Background and Empirical Approach

Labor intensity is higher in production processes that involve many operations, large team sizes, many hours per worker, or a combination of these elements. The first two elements (the number of operations and their average team size) may be seen as two dimensions of production scale, while the third one (hours per individual worker) depends on how persistent the workforce is over the growing season. Each of these dimensions of the production process may influence individualism in different ways, as discussed below. We can study the relevance of these various channels empirically in our context by leveraging detailed historical data from the *Hand and Machine Labor* Census report (Wright, 1899).

The report offers detailed information, including not only data on hours per acre by crop, but also breakdowns of the number of operations involved in growing each crop (e.g. plowing, planting, cultivating, etc.), the number of workers involved in each operation, and the total hours those workers were employed. Appendix Table A4 provides an example of the details included in the report for a specific crop. With this crop-level information and county-level data on crop mix, we construct a county-level index for each element, allowing us to explore which aspects of crop production drive the overall relationship between labor intensity and individualism.

(i) Production operations and the role of interdependence. The first potential driver of labor intensity is the number of operations involved in the production process, each of which requires labor. The number of operations is associated with the complexity of the production process. An operation refers to a specific task within the cultivation, harvesting, or post-harvest handling of a crop, classified sequentially to represent each distinct stage of the production process. For example,

[14] Due to the unavailability of the 1890 census, we are not able to maintain a consistent interval. However, placebo estimates using the measures from the 1860, 1880, and 1900 are not significantly different. In this case, we need to exclude black children from the analysis due to the lack of names of Enslaved in the 1860 census. Alternatively, including both black and white children in 1880 and 1900, while using only white children in 1860, yields similar results.

wheat production involves five steps: breaking the soil, planting seeds, covering the seeds with soil, harvesting and bagging, and transporting to storage. In contrast, cotton had ten operations, maize had fifteen, and tobacco seventeen. The average for all crops is just over ten operations.

Agricultural production is characterized by strong complementaries, as each operation relies on the ones coming before it, and the final product requires that all operations are completed successfully. In this sense, agriculture is an "O-ring" type of production process (Kremer, 1993). The more operations a given crop requires, the greater the risk of failure due to a lack of skills or effort at any stage of the process. A high number of operations thus entails greater dependence among workers on the work of others at different steps. The organization of production for complex crops would then favor the adoption of communitarian norms that discourage shirking and absenteeism.

This explanation rests on the premise that the number of operations and the number of distinct workers are positively related, meaning that more complex production processes typically involve more people who rely on one another. The argument does not require that each operation be performed by a different worker; rather, it only requires that crops involving more operations tend to involve more distinct workers on average, increasing the scope for interdependence across workers. Data from Wright (1899) in Appendix I provide evidence that this was the case, and the number of distinct workers used over a growing season increased in line with the number of operations involved.[15]

(ii) Workers per operation and the role of team size. The second element that may drive labor intensity is the average number of workers engaged in the production process at each different step, which captures the notion of team size. Operations with large scale may involve stronger needs for coordination among individual workers, which may favor non-individualistic norms. Even with limited needs for coordination, larger teams may be a result of higher elasticity of output with respect to labor. This would imply larger proportional reductions in output from non-marginal reductions in labor input due to shirking or absenteeism, favoring communitarian norms to limit such output losses.

Production processes with many workers per operation would also tend to induce high population density in or around the area of agricultural production. Density has been associated with stronger social connections (Chay and Munshi, 2015), so this could be a mechanism through

[15] Historical accounts suggest the assumption is plausible as well. Isern (1986) describes the Plains wheat belt, where harvests were carried out by itinerant crews and custom threshing outfits with specialized roles (e.g., engineer, separator man), hired separately from those who plowed or seeded. Cotton also illustrates clear labor segmentation: post-harvest ginning was often carried out by independent ginners, while field labor markets distinguished spring chopping and hoeing from fall picking (U.S. Bureau of the Census, 1911). Similarly, Bennett (2014) notes that tobacco production involved skill-specific tasks (e.g., topping, suckering, priming, grading) divided among hired workers. Corn followed a comparable pattern: short-term huskers were paid per bushel, while custom shellers traveled farm-to-farm, providing post-harvest services (Myrick, 1911). Taken together, these historical accounts suggest that operations requiring heterogeneous skills and inputs were typically handled by distinct labor. Even if this relationship did not hold for every operation in the Commissioners report, it is reasonable to assume that a greater number of operations correlates with a higher degree of labor division.

which labor intensity influences county-level population density (see Table 2), this does not conclusively rule out the potential relevance of this mechanism. With the data on the number of workers per operation, we can assess the importance of team size more directly.

(iii) Hours per worker and repeated interactions. Another element that may drive labor intensity is hours per worker. Some crops may employ the same set of individuals throughout the entire growing season across all or many operations, leading to high total hours for each, while others cycle through many individuals, leading to low hours per worker even if the total hours required for production are large. Hours per worker can thus indicate whether a dedicated workforce was needed over the growing season, which would be driven in part by the seasonality of the operations involved with specific crops.

The grains and hay that were prominent on northern farms required concentrated labor at very specific times of planting and harvest, but little labor between those steps. These crops required many workers, but each worker was only needed for a small number of hours, and there was less need for a dedicated or stable workforce over the growing season. Compared to those grains, crops such as maize and cotton had more consistent and less seasonal demand for labor, involving more attention throughout the growing season (e.g. multiple rounds of cultivation). Greater flexibility in planting and harvesting these crops also reduced the need for concentrated labor effort at those stages compared to northern grain crops, which led to the distinction between "few-day" northern grain crops and "multiple-day" southern staples (Earle, 1978; Reid, 1979). In these cases the hours of each individual worker would be relatively high.

For multiple-day crops with high hours per worker, it made sense to have a dedicated group of workers (family and/or hired labor) throughout the year. This in turn would imply more stable and consistent interactions, plausibly promoting communitarian norms by having that small number of workers interact through a repeated game. In contrast, few-day crops did not require the same need for dedicated workers, typically relying on occasional hired-in workers. Few-day crops with low hours per worker thus created limited opportunities for workers to build a collective identity among workers, whereas multiple-day crops created many such opportunities.[16]

We assess the empirical significance of these different dimensions of labor intensity using detailed crop-level production stage information from Wright (1899). For each crop j, labor intensity can be expressed as a function of the three key elements as follows:

$$\text{Crop intensity}_j \equiv \frac{\text{Hours}_j}{\text{Acre}_j} = \frac{\text{Operations}_j}{\text{Acre}_j} \times \frac{\text{Workers}_j}{\text{Operation}_j} \times \frac{\text{Hours}_j}{\text{Worker}_j}. \tag{8}$$

Figure 7 illustrates the cross-crop variation we leverage in this analysis, plotting total labor hours

[16] Production scale and the need for dedicated workforces have been emphasized in theories linking the history of slavery in the U.S. South due to its specialization in labor intense crops. Earle (1978) and Reid (1979) argue that slavery was not profitable in crops with sporadic labor needs (e.g. wheat), while Engerman and Sokoloff (1997, 2002) argue that slavery was favored by the presence of scale economies and the benefit of persistent teams of workers. These ideas are consistent with our analysis, with the difference that they highlight labor intensity's effects on institutions while our paper studies its effects on culture.

per acre, the number of operations, the number of workers per operation, and labor hours per worker at the crop level. In all cases, crops are ordered by total labor hours per acre, providing some visual sense of the associations across variables. Labor intensity is highly correlated with the three elements, with simple correlation coefficients in the three of 0.66, 0.46, and 0.76, respectively. The correlation of overall labor intensity with the number of workers per operation (0.46) is driven by the two crops at the upper tail of the distribution (sugarcane and tobacco). The correlations of labor intensity with operations and hours per worker are higher (0.66 and 0.76, respectively) and there is a more consistent alignment with total labor intensity throughout the rankings. Given these strong but not perfect correlations, we have substantive variation to identify the impacts of the different components of labor intensity.

The figure also suggests that differences in labor intensity across crops may be driven by different components. Among high-intensity crops, cotton required relatively small teams, but its moderate yet consistent labor demands throughout the year resulted in high hours per worker. Sugarcane's high labor intensity, in contrast, was primarily rooted in large team sizes. Sweet potato, despite not being among "plantation crops," exhibits high labor intensity due to both numerous operations and sustained labor requirements throughout the year, despite small team sizes. Potatoes, with similar or slightly larger team sizes than sweet potatoes, required fewer operations, leading to lower overall labor intensity. Among cereals—all of which had below-median labor intensity—barley, wheat, and oats had similar levels of hours per worker, but oats required larger teams and many more operations, resulting in greater overall labor intensity.

Combining these crop-level differences with county-level variation in crop mix, as in our baseline analysis, we explore mechanisms using the following specification:

$$\begin{aligned} y_{c,1910} = \alpha + \beta_1 \text{Operations}_{c,1900} + \beta_2 \text{Workers per operation}_{c,1900} + \beta_3 \text{Hours per worker}_{c,1900} \\ + \gamma' \boldsymbol{X}_c + \mu_s + \epsilon_c \ , \end{aligned} \tag{9}$$

where each element of labor intensity is constructed as the county-level weighted average of that element, with weights based on each crop's acreage in the county's farmland, as in our main regressions for overall labor intensity. We also keep the same set of controls as in our baseline specification, $\boldsymbol{X}_c$, as well as state fixed effects μ_s.

We adopt this linear specification for the county-level regressions as a coarse approach to exploring the contributions of the different components of labor intensity, noting that this is not a precise aggregation of the crop-level expression above. In addition, a specific caveat applies to the construction of the county-level measure of hours per worker, as it could involve some measurement error in the presence of agricultural diversification and crop-rotation schemes. Nevertheless, our approach would remain informative unless such measurement errors are systematically linked to unobserved determinants of local cultural traits, which is unlikely.[17]

[17] While taking the weighted average across crops has a straightforward interpretation for the number of operations or

5.2 Results and Discussion

Table 10 shows the estimation results using each element of labor intensity as the outcome variable. For reference purposes, Columns (1) and (3) report the OLS and IV regressions of individualism on our overall measure of labor intensity, labor hours per acre. As data on specific components of labor intensity is available only from the Commissioner's report, these results rely exclusively on that source, leading to slight differences in the point estimates compared to those in Table 1. Nevertheless, the same significant relationship remains with a similar effect size.

To address which element appears most important to the overall relationship, Columns (2) and (4) show the results of horse-race regressions using the county-level aggregate of the three drivers of labor intensity. In the IV specification, we simultaneously use the potential indices of the three labor intensity components as instruments, which are constructed following the same approach as for potential labor intensity. The first-stage results are displayed in Appendix H.

In the OLS estimates, only hours per worker remains significant. In the IV specification, we get a more complex story. Again the point estimate for hours per worker (-0.271) is significant, indicating a strong role for this dimension of labor intensity. Additionally, operations per acre becomes statistically significant and exhibits an even larger effect (-0.366).

These results suggest that both the complexity of production (as proxied by operations per acre) and the need for dedicated workers (as proxied by hours per worker) drove the overall relationship of labor intensity and individualism. Crops requiring multiple operations spread out over a long growing season would have been most conducive to non-individualistic norms. Such crops relied more on a group of dedicated workers who had regular interactions with one another, and whose success depended on each step being performed correctly. These environments would reasonably foster norms emphasizing group success and mutual reliance. In contrast, crops requiring few steps and only sporadic labor—even if involving many workers—had less incentive to form communitarian norms. It was not necessary or profitable to involve the same individuals repeatedly during the year, and those individuals would not see their rewards as tied to the efforts of others. Individualistic norms would be more likely to take root in such an environment.

In addition, the small and insignificant coefficients for workers per operation suggest that the influence of labor intensity on individualism did not operate through the scale of work teams at any given production step. This aligns with qualitative evidence (Earle, 1978; Reid, 1979), which discusses how northern crops like wheat and barley required large—yet very short-term—work

for team size, linear aggregation could be problematic for hours per worker and capturing the need for dedicated workforces. For example, if two crops had seasonal patterns that mirrored one another, but both had low hours per worker, aggregating might instead show high hours per worker in that county if workers shifted between crops during the year. This introduces measurement error in the county-level measure of hours per worker. However, the measure is likely to be meaningful for several reasons. In practice, labor requirements over the year tend to overlap across crops. Moreover, diversification, crop-rotation, and mixed farming schemes were relatively limited, and thus the overall need for a dedicated workforce was strongly correlated with the needs of local dominant crop. Specialization was even higher at the farm level than at the county level, further mitigating the concerns—especially if farms were, at least in part, the relevant level for cultural formation.

crews during brief planting and harvesting windows. Although these crops mobilized substantial labor forces, the groups were transitory in nature, comprised of workers brought together for only short periods of time.

These findings are consistent with the role of large-scale irrigation as a driver of collectivism in paddy-rice regions of China (see Talhelm et al., 2014; Talhelm and Oishi, 2018; Talhelm, 2020; Talhelm and English, 2020), while pointing to additional mechanisms operating across a range of crops. Even in a context where large-scale irrigation networks and labor exchange arrangements were relatively uncommon, we observe a strong effect of labor intensity on cultural norms.[18] Moreover, we empirically link collectivist norms to specific dimensions of labor intensity—interdependencies unrelated to irrigation, per se, but rather to complexity in general and the need for a dedicated workforce over the growing season.

Our results shed light on the socioeconomic scale at which cultural formation took place. In settings with large-scale irrigation projects—where cooperation and binding labor exchange arrangements were essential—interactions shaping norms occurred primarily at the community-level, spanning across families and kinship networks. In our context, while hired labor may have contributed to community-level cultural formation, the influence of labor intensity likely operated largely within families or kin-networks, before potentially spilling over to the rest of the community through cultural transmission. While beyond the scope of this paper, we see the mechanisms facilitating the diffusion of cultural traits across society and the channels for persistence over time as key areas for future studies.

6 Conclusion

This paper studies the effects of agricultural labor intensity on individualism, leveraging rich historical variation across U.S. counties and over time through an array of complementary approaches, and advancing our understanding of the underlying mechanisms. The links between agricultural production—the dominant sector in societies at early stages of development—and individualism—a key dimension of cross-cultural variation—is of core interest for the literature on how economic and social relations influence cultural formation. We contribute to this key area by offering new insights and raising fresh questions in multiple ways.

The vast range of crops in the U.S., combined with the sharp variation in their composition both across and within regions, make this a well-suited context for establishing the broader relevance of the link between agricultural labor intensity and culture, as advanced by the seminal work of Talhelm et al. (2014) in a more narrowly focused comparison between rice and wheat in China.

[18] Large-scale irrigation projects existed but were infrequent in the context of our study. One instance is mentioned by Arrington and May (1975), who notes that the settlement patterns of the Mormons around the Great Salt Lake were designed to foster "social contact needed to plan and execute cooperative construction of canals and other irrigated works" necessary for crop production in an arid climate. While it is quite plausible that large-scale irrigation networks had similar cultural effects in the U.S. as in southeast Asia, conducting a meaningful empirical analysis would be hard due to their limited presence in this period.

Our cross-sectional analysis provides compelling results based on climate-induced variation in crop mix, controlling for a wide array of geo-climatic characteristics and documenting robustness across different regions and subpopulations. Beyond the strong effects on contemporaneous individualism, we further demonstrate long-term effects of historical labor intensity on contemporary cultures. To this end, we develop two additional measures of individualism, adapted from social psychology for application to online search data, which may be of interest to future studies.

We also go beyond cross-sectional differences by examining changes in response to major historical shocks to labor intensity, induced by mechanization and the boll weevil. This is a key aspect of our contribution, both in terms of identification and conceptual insight. A potential concern with studies relying on cross-sectional variation in deeply-rooted factors across geographical units is that the regressors of interest may be correlated with time-invariant unobserved characteristics. Our analysis mitigates this concern by leveraging time variation. Moreover, these findings provide insight into the coexistence of cultural persistence and change. Taken together, our results—drawing on variation across counties and over time—demonstrate that agricultural conditions have powerful, enduring effects on cultural formation, but that they do not represent an unchanging destiny, as cultural attitudes remain malleable when shocks are large enough.

Finally, our paper advances the understanding of the mechanisms through which labor intensity affects individualism, laying out a simple conceptual background and empirical approach. Relying on detailed historical data on the production process for different crops, we explore the relevance of (i) interdependence across workers, as captured by the number of complementary operations required in production, (ii) team size, captured by the number of workers per operation, and (iii) the use of a dedicated workforce, captured by hours per worker, which entailed more stable interactions among that group. We find evidence supporting the relevance of interdependence and dedicated workforces in driving the relationship of labor intensity and non-individualistic norms.

We make a novel contribution to the literature on the links between agricultural labor intensity and individualism by examining mechanisms—a key area of advancement in recent research on cultural formation. Our findings complement previous studies on rice versus wheat in China (Talhelm et al., 2014; Talhelm and Oishi, 2018; Talhelm, 2020; Talhelm and English, 2020), which highlight the role of labor exchange arrangements among farmers, a necessity under irrigation-based rice cultivation. In the U.S. context, where labor exchange arrangements were limited, our results point to the relevance of additional channels. We provide exploratory but rich evidence on the mechanisms underlying the inception of non-individualistic norms in response to high labor intensity. While this represents a novel contribution, it also underscores the need for further studies on the mechanisms driving the diffusion and persistence of individualism.

References

Acemoglu, D. and Robinson, J. A. (2025). Culture, institutions, and social equilibria: A framework. *Journal of Economic Literature*, 63(2):637–692.

Ager, P. and Ciccone, A. (2018). Agricultural risk and the spread of religious communities. *Journal of the European Economic Association*, 16(4):1021–1068.

Aiken, C. S. (1998). *The Cotton Plantation South since the Civil War*. Johns Hopkins University Press.

Alesina, A., Giuliano, P., and Nunn, N. (2011). Fertility and the plough. *The American Economic Review*, 101(3):499–503.

Alston, L. J. and Higgs, R. (1982). Contractual mix in southern agriculture since the civil war: Facts, hypotheses, and tests. *The Journal of Economic History*, 42(2):327–353.

Ang, J. B. (2019). Agricultural legacy and individualistic culture. *Journal of Economic Growth*, 24(4):397–425.

Armstrong, K. L. (2002). Race and sport consumption motivations: A preliminary investigation of a black consumers' sport motivation scale. *Journal of Sport Behavior*, 25(4):309.

Arrington, L. J. and May, D. (1975). "a different mode of life": Irrigation and society in nineteenth-century utah. *Agricultural History*, 49(1):3–20.

Atack, J. (2016). Historical geographic information systems (gis) database of steamboat-navigated rivers during the nineteenth century in the united states. Technical report, Vanderbilt University.

Atack, J., Margo, R. A., and Rhode, P. W. (2019). Automation of manufacturing in the late nineteenth century: The hand and machine labor study. *Journal of Economic Perspectives*, 33(2):51–70.

Bardhan, P. (2000). Irrigation and cooperation: An empirical analysis of 48 irrigation communities in south india. *Economic Development and cultural change*, 48(4):847–865.

Barry, H. I., Bacon, M. K., and Child, I. L. (1957). A cross-cultural survey of some sex differences in socialization. *The Journal of Abnormal and Social Psychology*, 55(3):327–332.

Bazzi, S., Fiszbein, M., and Gebresilasse, M. (2020). Frontier culture: The roots and persistence of "rugged individualism" in the united states. *Econometrica*, 88(6):2329–2368.

Bennett, E. P. (2014). *When tobacco was king: families, farm labor, and federal policy in the Piedmont*. University Press of Florida.

Bester, C. A., Conley, T. G., and Hansen, C. B. (2011). Inference with dependent data using cluster covariance estimators. *Journal of Econometrics*, 165(2):137–151.

Buggle, J. C. (2020). Growing collectivism: Irrigation, group conformity and technological divergence. *Journal of Economic Growth*, 25(2):147–193.

Buggle, J. C. and Durante, R. (2021). Climate risk, cooperation and the co-evolution of culture and institutions. *The Economic Journal*, 131(637):1947–1987.

Cai, H., Zou, X., Feng, Y., Liu, Y., and Jing, Y. (2018). Increasing need for uniqueness in contemporary china: Empirical evidence. *Frontiers in Psychology*, 9:554.

Chay, K. and Munshi, K. (2015). Black networks after emancipation: evidence from reconstruction and the great migration. *Unpublished working paper*.

Cline, B. N. and Williamson, C. R. (2017). Individualism, democracy, and contract enforcement. *Journal of Corporate Finance*, 46:284–306.

Coakley, J. M., Maxwell, F. G., and Jenkins, J. N. (1969). Influence of feeding, oviposition, and egg and larval development of the boll weevil on abscission of cotton squares. *Journal of economic entomology*, 62(1):244–245.

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

Cooper, T. P., Boss, A., and Peck, F. W. (1916). *Labor requirements of crop production*, volume 157. University Farm.

De Vliert, E. V. (2011). Climato-economic origins of variation in ingroup favoritism. *Journal of Cross-Cultural Psychology*, 42(3):494–515.

Earle, C. V. (1978). A staple interpretation of slavery and free labor. *Geographical Review*, 68(1):51–65.

Eberhardt, M. and Vollrath, D. (2018). The effect of agricultural technology on the speed of development. *World Development*, 109:483–496.

Engerman, S. L. and Sokoloff, K. L. (1997). Factor endowments, institutions, and differential paths of growth among new world economies. *How Latin America Fell Behind*, pages 260–304.

Engerman, S. L. and Sokoloff, K. L. (2002). Factor endowments, inequality, and paths of development among new world economics. Technical report, National Bureau of Economic Research.

Evenson, R. E. (1978). A century of productivity change in us agriculture: an analysis of the role of invention, research and extension. Technical report, Center Discussion Paper, No. 296, Yale University, Economic Growth Center.

Fiszbein, M. (2022). Agricultural diversity, structural change, and long-run development: Evidence from the united states. *American Economic Journal: Macroeconomics*, 14(2):1–43.

Fouka, V. and Schläpfer, A. (2020). Agricultural returns to labour and the origins of work ethics. *The Economic Journal*, 130(628):1081–1113.

Galor, O. and Özak, Ö. (2016). The agricultural origins of time preference. *American Economic Review*, 106(10):3064–3103.

Giuliano, P. and Nunn, N. (2021). Understanding cultural persistence and change. *The Review of Economic Studies*, 88(4):1541–1581.

Gorodnichenko, Y. and Roland, G. (2011). Which dimensions of culture matter for long-run growth? *American Economic Review*, 101(3):492–98.

Gorodnichenko, Y. and Roland, G. (2017). Culture, Institutions, and the Wealth of Nations. *The Review of Economics and Statistics*, 99(3):402–416.

Gorodnichenko, Y. and Roland, G. (2020). Culture, institutions and democratization. *Public choice*, pages 1–31.

Greenfield, P. M. (2013). The changing psychology of culture from 1800 through 2000. *Psychological science*, 24(9):1722–1731.

Greif, A. (1994). Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies. *Journal of Political Economy*, 102(5):912–950.

Guiso, L., Sapienza, P., and Zingales, L. (2016). Long-term persistence. *Journal of the European Economic Association*, 14(6):1401–1436.

Haines, M. R. et al. (2005). Historical, demographic, economic, and social data: The united states, 1790-2000. *Computerized Data sets from the Inter-University Consortium for Political and Social Research (ICPSR). ICPSR*, (2896).

Han, D., Mahony, D. F., and Greenwell, T. C. (2016). A comparative analysis of cultural value orientations for understanding sport fan motivations. *International Journal of Sports Marketing and Sponsorship*, 17(3):260–276.

Havaldar, S., Giorgi, S., Rai, S., Talhelm, T., Guntuku, S. C., and Ungar, L. (2024). Building knowledge-guided lexica to model cultural variation. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 211–226.

Hecht, R. W. and Vice, K. R. (1954). *Labor used for field crops*. Number 144. US Department of Agriculture.

Heine, S. J. (2010). Cultural psychology. In Fiske, S. T., Gilbert, D. T., and Lindzey, G., editors, *Handbook of Social Psychology*, pages 254–266.

Henrich, J. (2020). *The WEIRDest People in the World: How the West Became Psychologically Peculiar and Particularly Propserous*.

Hoang-Anh, H., Martinsson, P., and Olsson, O. (2019). The origins of cultural divergence: Evidence from a developing country. EfD Discussion paper 19-8, Environment for Development, University of Gothenburg.

Hofstede, G. (1991). *Cultures and Organizations: Software of the Mind*. McGraw-Hill.

House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., and Gupta, V. (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Sage publications.

IIASA and FAO (2012). Global agro-ecological zones (gaez v3. 0). Technical report, IIASA, Laxenburg, Austria and FAO, Rome, Italy.

Isern, T. D. (1986). Folklife of the threshing outfit. *South Dakota History*, 16(1).

Johnson, T. R. and Vollrath, D. (2020). The role of land in temperate and tropical agriculture. *Economica*, 87(348):901–937.

Jung, Y. (2020). The long reach of cotton in the US south: Tenant farming, mechanization, and low-skill manufacturing. *Journal of Development Economics*, 143:102432.

Kendrick, J. W. (1961). *Productivity trends in the United States.* Princeton: Princeton Univ. Press.

Knudsen, A. S. B. (2022). Those who stayed: Selection and cultural change in the age of mass migration. *Department of Economics, University of Copenhagen*.

Kremer, M. (1993). The O-Ring theory of economic development. *The Quarterly Journal of Economics*, 108(3):551–575.

Kyriacou, A. P. (2016). Individualism–collectivism, governance and economic development. *European Journal of Political Economy*, 42:91–104.

Lange, F., Olmstead, A. L., and Rhode, P. W. (2009). The impact of the boll weevil, 1892–1932. *The Journal of Economic History*, 69(3):685–718.

Mann, M. (1986). *A history of power from the beginning*, volume 1 of *The sources of social power*. Cambridge University Press.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In Zarembka, P., editor, *Frontiers in Econometrics*, chapter 4, pages 105–142. Academic Press.

Mullahy, J. (2015). Multivariate fractional regression estimation of econometric share models. *Journal of econometric methods*, 4(1):71–100.

Myrick, H., editor (1911). *The Book of Corn: A Complete Treatise Upon the Culture, Marketing and Uses of Maize in America and Elsewhere*. Orange Judd Company, New York and Chicago.

Nunn, N. and Wantchekon, L. (2011). The slave trade and the origins of mistrust in africa. *American Economic Review*, 101(7):3221–52.

Ogihara, Y., Fujita, H., Tominaga, H., Ishigaki, S., Kashimoto, T., Takahashi, A., Toyohara, K., and Uchida, Y. (2015). Are common names becoming less common? the rise in uniqueness and individualism in japan. *Frontiers in psychology*, 6:1490.

Olmstead, A. L. and Rhode, P. W. (2008). *Creating Abundance*. Cambridge University Press.

Olsson, O. and Paik, C. (2016). Long-run cultural divergence: Evidence from the neolithic revolution. *Journal of Development Economics*, 122:197–213.

Oyserman, D., Coon, H. M., and Kemmelmeier, M. (2002). Rethinking individualism and collectivism: evaluation of theoretical assumptions and meta-analyses. *Psychological bulletin*, 128(1):3.

Oyserman, D. and Lee, S. W. (2008). Does culture influence what and how we think? effects of priming individualism and collectivism. *Psychological bulletin*, 134(2):311.

Papke, L. E. and Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of applied econometrics*, 11(6):619–632.

Ransom, R. L. and Sutch, R. (2001). *One kind of freedom: The economic consequences of emancipation*. Cambridge University Press.

Raz, I. T. (2025). Soil heterogeneity, social learning, and the formation of close-knit communities. *Journal of Political Economy*, 133(8):000–000.

Reid, J. D. (1979). The evaluation and implications of southern tenancy. *Agricultural History*, 53(1):153–169.

Reuss, G. (1930). An economic study of the factors affecting farm organization and power utilization of sugar cane farms, 1929. Technical report, Louisiana State University and Agricultural and Mechanical College, Agricultural Experiment Station.

Roland, G. (2020). The deep historical roots of modern culture: A comparative perspective. *Journal of Comparative Economics*, 48(3):483–508.

Ruggles, S., Flood, S., Sobek, M., Backman, D., Chen, A., Cooper, G., Richards, S., Rogers, R., and Schouweiler, M. (2020). IPUMS USA: Version 10.0 [dataset]. DOI: 10.18128/d010.

Santos, H. C., Varnum, M. E., and Grossmann, I. (2017). Global increases in individualism. *Psychological science*, 28(9):1228–1239.

Singelis, T. M., Triandis, H. C., Bhawuk, D. P., and Gelfand, M. J. (1995). Horizontal and vertical dimensions of individualism and collectivism: A theoretical and measurement refinement. *Cross-cultural research*, 29(3):240–275.

Spolaore, E. and Wacziarg, R. (2013). How deep are the roots of economic development? *Journal of Economic Literature*, 51(2):325–369.

Talhelm, T. (2020). Emerging evidence of cultural differences linked to rice versus wheat agriculture. *Current Opinion in Psychology*, 32:81–88.

Talhelm, T. and English, A. S. (2020). Historically rice-farming societies have tighter social norms in china and worldwide. *Proceedings of the National Academy of Sciences*, 117(33):19816–19824.

Talhelm, T. and Oishi, S. (2018). How rice farming shaped culture in southern china. In Uskul, A. and Oishi, S., editors, *Socioeconomic Environment and Human Psychology*, pages 53–76. Oxford University Press.

Talhelm, T., Zhang, X., Oishi, S., Shimin, C., Duan, D., Lan, X., and Kitayama, S. (2014). Large-scale psychological differences within china explained by rice versus wheat agriculture. *Science*, 344(6184):603–608.

Triandis, H. (1995). *Individualism & Collectivism*. Westview Press, Boulder, CO.

Twenge, J. M., Abebe, E. M., and Campbell, W. K. (2010). Fitting in or standing out: Trends in american parents' choices for children's names, 1880–2007. *Social Psychological and Personality Science*, 1(1):19–25.

Twenge, J. M., Campbell, W. K., and Gentile, B. (2012). Increases in individualistic words and phrases in american books, 1960–2008. *PloS one*, 7(7):e40181.

U.S. Bureau of the Census (1911). Cotton production and statistics of cottonseed products: 1910. Bulletin 111, Department of Commerce and Labor, Washington, D.C.

U.S. Bureau of the Census (1960). *Historical statistics of the United States, colonial times to 1957*.

Varnum, M. E. and Kitayama, S. (2011). What's in a name? popular names are less common on frontiers. *Psychological Science*, 22(2):176–183.

Voigtländer, N. and Voth, H. (2012). Persecution perpetuated: the medieval origins of anti-semitic violence in nazi germany. *The Quarterly Journal of Economics*, 127(3):1339–1392.

Vollrath, D. (2011). The agricultural basis of comparative development. *Journal of Economic Growth*, 16:343–370.

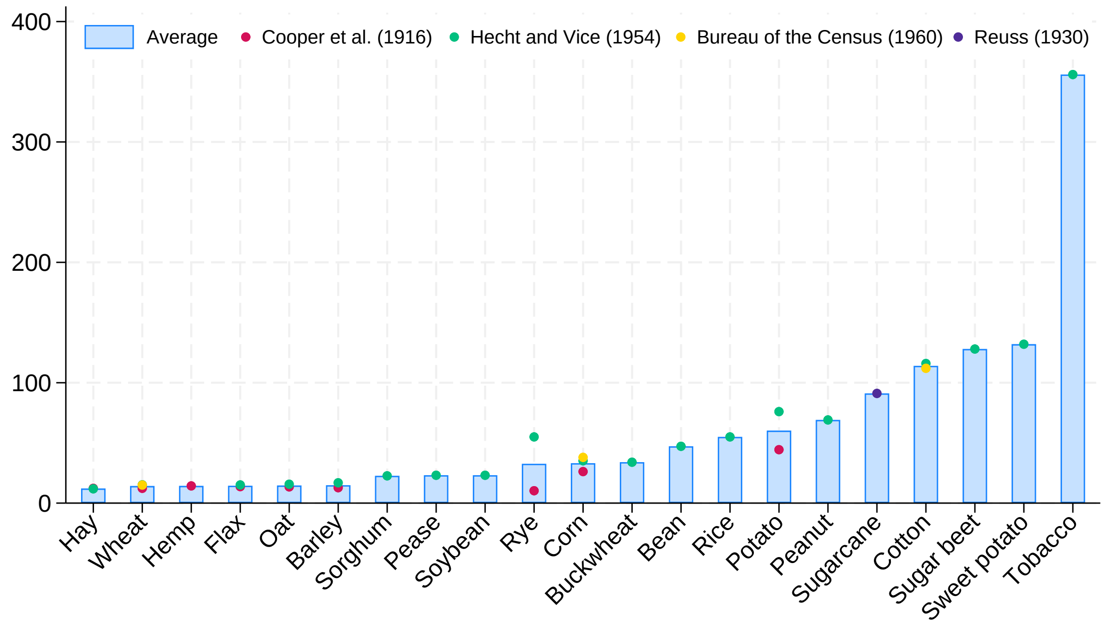
Wann, D. L., Grieve, F. G., Zapalac, R. K., and Pease, D. G. (2008). Motivational profiles of sport fans of different sports. *Sport Marketing Quarterly*, 17(1):6.

Wright, C. D. (1899). *Hand and machine labor*. US Government Printing Office.

Wyer, R. S., yue Chiu, C., and yi Hong, Y. (2013). *Understanding culture: theory, research, and application*. Psychology Press.

Figures

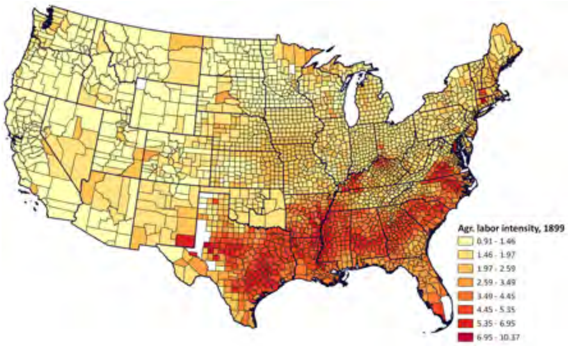
Figure 1: Crop-specific Labor Intensities from Different Data Sources



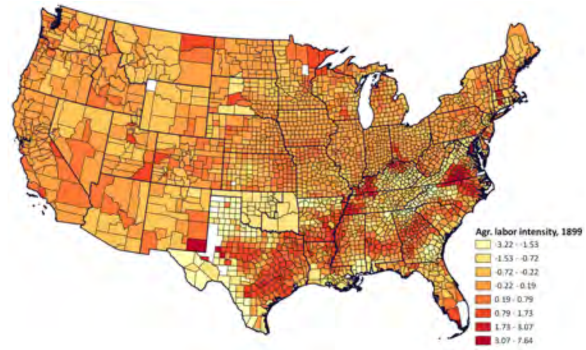
Notes: The dots show the man-hours per acre to produce each crop from each individual source, and the bar denotes their average values across sources.

Figure 2: Agricultural Labor Intensity, 1900

(a) Raw figures



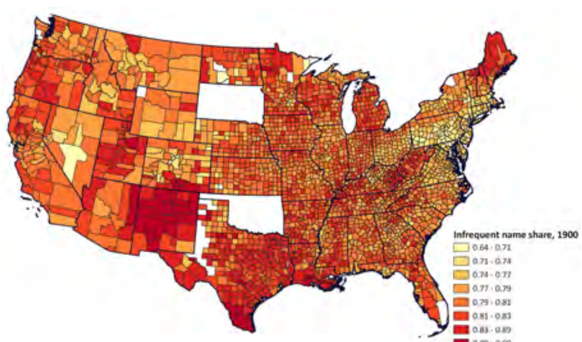
(b) Deviations from state-level means



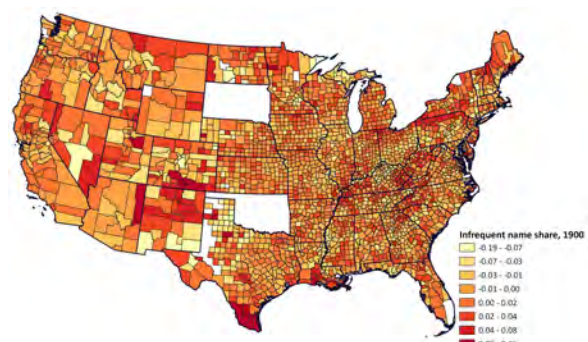
Notes: The map shows the index of agricultural labor intensity in 1900 estimated from Equation 1. Panel (b) displays the difference between the index of agricultural labor intensity and its state-average. White areas are counties not included in the sample because of data availability.

Figure 3: Individualism as Proxied by Shares of Infrequent Names, 1900

(a) Raw figures

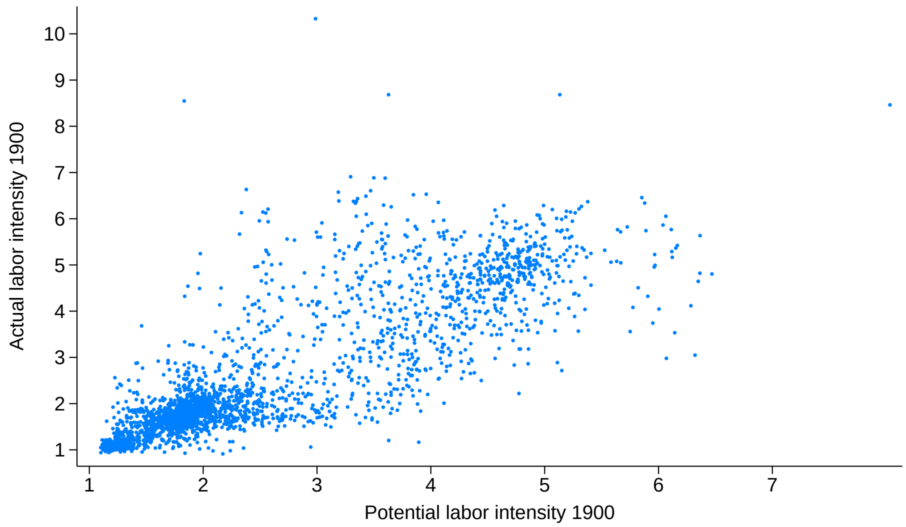


(b) Deviations from state-level means



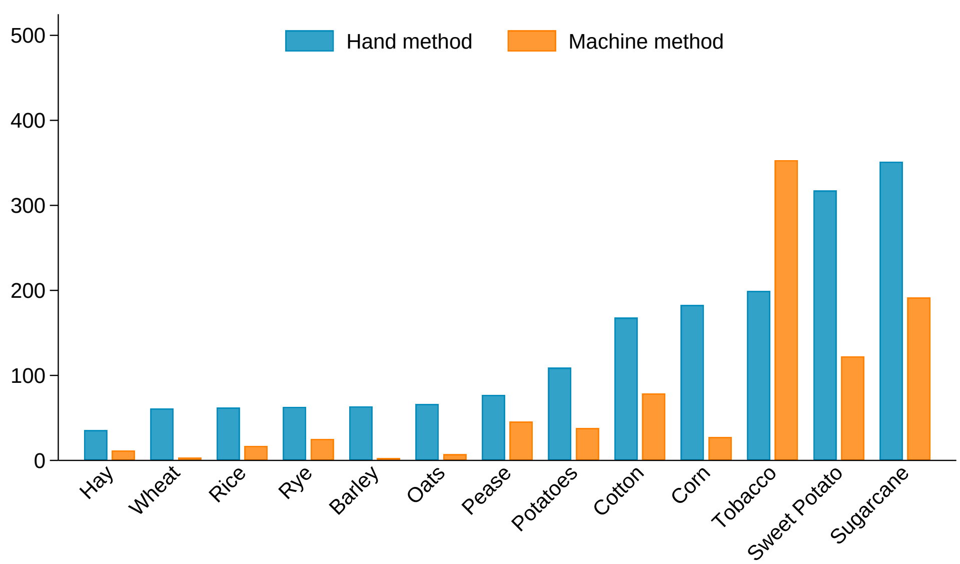
Notes: The map shows the share of infrequent names in 1900. Panel (b) displays the difference between the index value and its state-average. White areas are counties not included in the samples because of data availability.

Figure 4: Actual and Potential Labor Intensity



Notes: The figure shows a scatter plot of the actual and potential labor intensity in 1900. The potential labor intensity is estimated from Equations 3 and 4, while the actual labor intensity is from Equation 1.

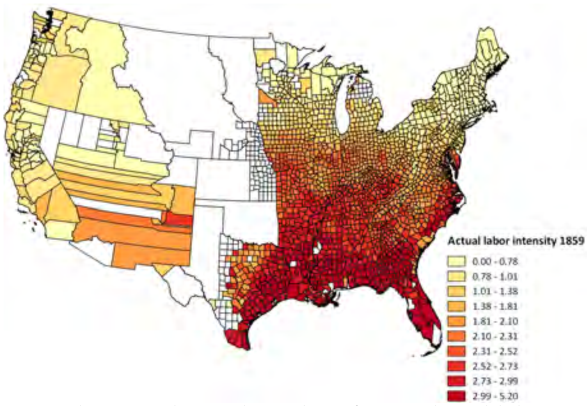
Figure 5: Labor Requirements Before and After Mechanization



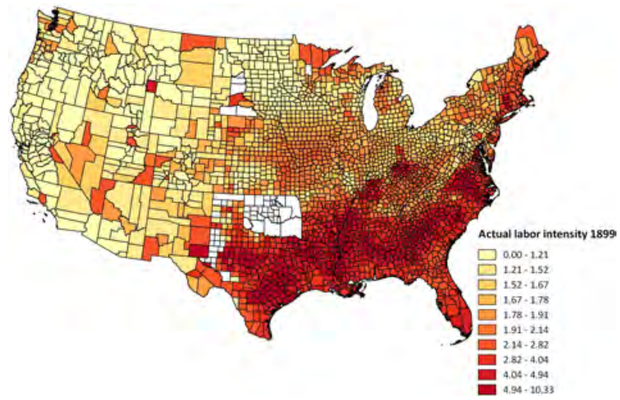
Notes: The bars denote the total man-hours necessary to produce an acre of each crop. The man-hours under the hand method are based on observations between 1828 and 1872. The machine method reflects production technologies between 1893 and 1896. Data are obtained from Wright (1899).

Figure 6: Agricultural Labor Intensity in 1860 and 1900

Labor intensity 1860, Hand-Method

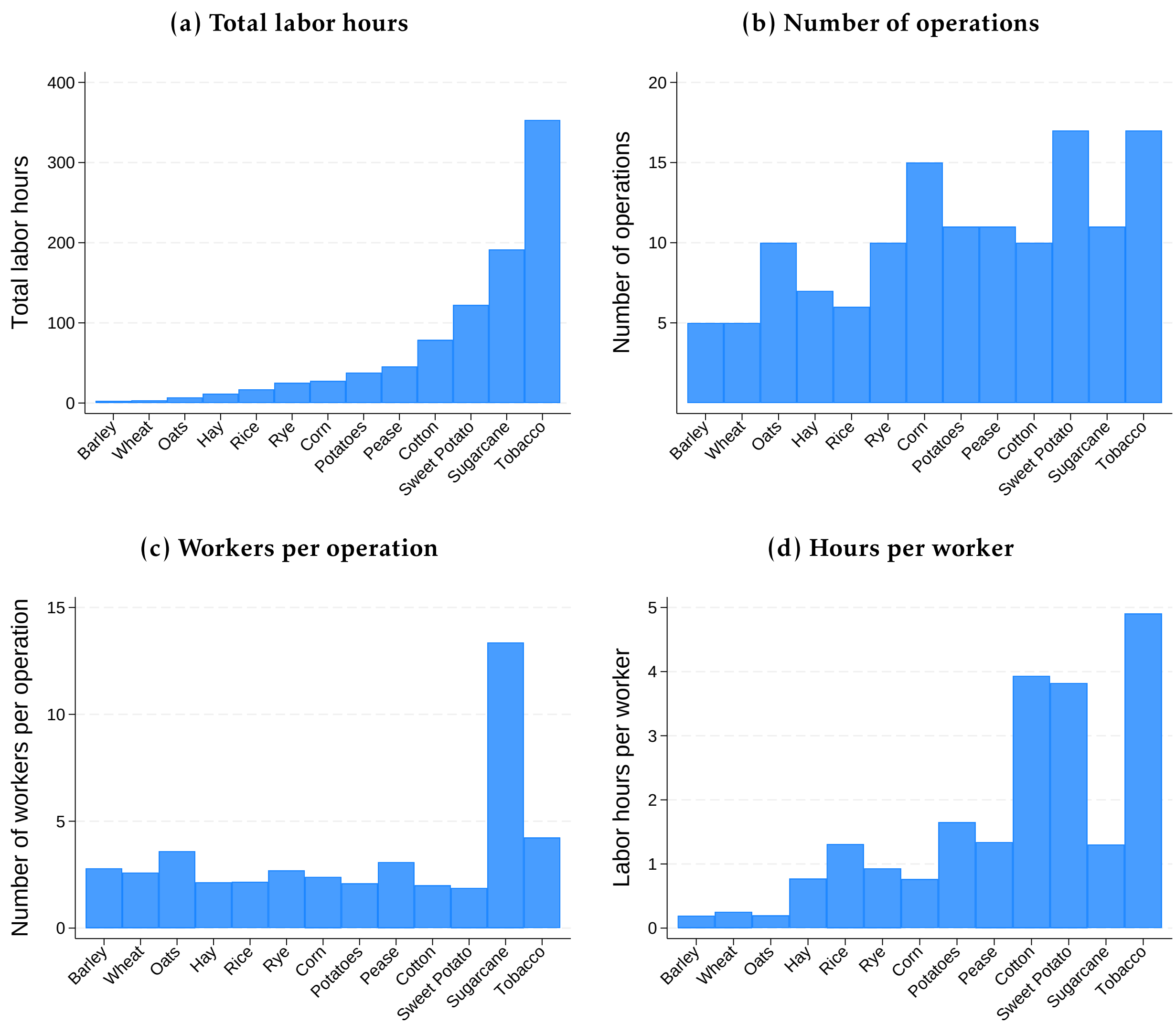


Labor intensity 1900, Machine-Method



Notes: The map shows the index of agricultural labor intensity in 1860 and in 1900 estimated from Equation 1. White areas are counties not included in the samples because of data availability.

Figure 7: Labor Intensity and its Drivers (Operations, Workers per Operation, Hours per Worker)



Notes: Based on data from the *Hand and Machine Labor* Census Report Wright (1899), each figure illustrates (a) total labor hours per acre, (b) the number of unique operations, (c) the number of workers per operation, and (d) the number of labor hours per worker. Crops on the X-axis are ordered by total labor hours per acre.

Tables

Table 1: Agricultural Labor Intensity and Individualism

Dep. var: Infrequent name share 1910	(1)	(2)	(3)	(4)
	Panel 1: OLS estimates			
Laboor intensity 1900	-0.217***	-0.212***	-0.198***	-0.194***
	(0.037)	(0.038)	(0.034)	(0.034)
R-squared	0.55	0.55	0.60	0.60
Observations	2732	2732	2732	2732
	Panel 2: IV estimates			
Laboor intensity 1900	-0.274***	-0.275***	-0.276***	-0.285***
	(0.072)	(0.073)	(0.068)	(0.067)
F-stat	133.64	128.18	81.04	78.47
Observations	2732	2732	2732	2732
Agricultural land suitability	N	Y	N	Y
Geo-climatic controls	N	N	Y	Y
State fixed effects	Y	Y	Y	Y

Notes: The table shows standardized coefficients. Robust standard errors clustered at 60mi×60mi grid cells are shown in parentheses. The share of infrequent names is computed for counties whose number of relevant observations (individuals) is no less than 30. The geo-climatic controls are: temperature, precipitation, terrain elevation and slope, distance to major cities, distance to coast, distance to navigable river, latitude, and longitude.

Table 2: Robustness to Additional Controls

Dep. var: Infrequent name share 1910	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
						Panel 1: OLS Estimation						
Labour intensity 1900	-0.198***	-0.195***	-0.237***	-0.193***	-0.166***	-0.192***	-0.190***	-0.158***	-0.185***	-0.206***	-0.177***	-0.164***
	(0.034)	(0.034)	(0.036)	(0.034)	(0.033)	(0.035)	(0.034)	(0.034)	(0.034)	(0.035)	(0.034)	(0.034)
R-squared	0.60	0.66	0.67	0.60	0.62	0.60	0.59	0.60	0.60	0.60	0.61	0.69
Observations	2717	2732	2726	2732	2732	2731	2690	2690	2707	2730	2732	2649
						Panel 2: IV Estimation						
Labour intensity 1900	-0.297***	-0.263***	-0.311***	-0.284***	-0.318***	-0.282***	-0.286***	-0.252***	-0.274***	-0.315***	-0.235***	-0.242***
	(0.066)	(0.065)	(0.065)	(0.067)	(0.067)	(0.068)	(0.066)	(0.067)	(0.068)	(0.067)	(0.067)	(0.073)
F-stat	77.74	78.59	79.89	77.85	84.70	77.13	77.88	79.74	73.44	76.42	75.88	67.86
Observations	2717	2732	2726	2732	2732	2731	2690	2690	2707	2730	2732	2649
Population density	√											√
Urban population ratio		√										√
Agricultural employment share			√									√
Average farm size				√								√
Average farm yields					√							√
Crop concentration						√						√
Soil heterogeneity							√					√
Frontier experience								√				√
English-speaking grandparents									√			√
Native grandparents										√		√
Birthplace diversity											√	√
Agricultural land suitability	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Geo-climatic controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Notes: The table shows standardized coefficients with robust standard errors clustered at 60mi×60mi grid cells in parentheses. The share of infrequent names is computed for counties whose number of relevant observations (individuals) is no less than 30. The definition and data sources of the additional controls are provided in Appendix A.2. The geo-climatic controls are: temperature, precipitation, terrain elevation and slope, distance to major cities, distance to coast, distance to navigable river, latitude, and longitude.

Table 3: Heterogeneity by Region and Race

Dep. var: Infrequent name share 1910	(1)	(2)	(3)	(4)
	Panel 1: Heterogeneity by Region			
	South		Non-South	
	OLS	IV	OLS	IV
Labour intensity 1900	-0.108***	-0.207**	-0.551***	-0.840*
	(0.039)	(0.089)	(0.133)	(0.436)
F-stat		69.97		35.00
R-squared	0.55		0.64	
Observations	1215	1215	1517	1517
	Panel 2: Heterogeneity by Race			
	Whites		Blacks	
	OLS	IV	OLS	IV
Labour intensity 1900	-0.193***	-0.259***	-0.110***	-0.268***
	(0.035)	(0.067)	(0.041)	(0.099)
F-stat		77.50		29.60
R-squared	0.59		0.49	
Observations	2693	2693	1188	1188
Agricultural land suitability	Y	Y	Y	Y
Geo-climatic controls	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y

Notes: The table shows standardized coefficients. Robust standard errors clustered at 60mi×60mi grid cells are shown in parentheses. The share of infrequent names is computed for counties whose number of relevant observations (individuals) is no less than 30. The geo-climatic controls are: temperature, precipitation, terrain elevation and slope, distance to major cities, distance to coast, distance to navigable river, latitude, and longitude.

Table 4: Long-Run Effects: Metro-Level Evidence

Dep. var.: Relative search interest based on Google Trends data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel 1. Individualistic versus Collectivistic Terms										
	"Unique" vs. "Common"		"Peculiar" vs. "Ordinary"		"Autonomous" vs. "Collaborative"		"Solo" vs. "Together"		First Principal Component	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Labour intensity 1900	-0.282**	-0.573***	-0.196**	-0.254**	-0.290***	-0.281**	-0.281**	-0.488**	-0.493***	-0.741***
	(0.129)	(0.167)	(0.094)	(0.117)	(0.095)	(0.123)	(0.140)	(0.193)	(0.129)	(0.159)
F-stat		106.83		105.29		106.83		106.83		105.29
R-squared	0.51		0.37		0.49		0.48		0.61	
Observations	200	200	197	197	200	200	200	200	197	197
Panel 2. Individual Sports versus Team Sports (Baseball, Basketball, and Football)										
	Golf		Tennis		Boxing		Wrestling		First Principal Component	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Labour intensity 1900	-0.157	-0.331**	-0.114	-0.200*	-0.368***	-0.554***	-0.146**	-0.265***	-0.537**	-0.922***
	(0.113)	(0.136)	(0.090)	(0.121)	(0.135)	(0.125)	(0.073)	(0.089)	(0.210)	(0.241)
F-stat		106.87		106.87		106.87		106.87		106.87
R-squared	0.54		0.59		0.70		0.72		0.68	
Observations	201	201	201	201	201	201	201	201	201	201
Agricultural land suitability	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Geo-climatic controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Division fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Notes: The table shows standardized coefficients with robust standard errors in parentheses. The measure of labor intensity and other county-level variables are harmonized with the DMA boundaries. In Panel 1, Columns 1-8 use the search volume of each individualistic word relative to its corresponding collectivist word as the outcome variable. In Panel 2, Columns 1-8 use the search volume of each individual sport relative to three major team sports: baseball, basketball, and football. In both panels, Columns 9 and 10 use the first principal component of the variables in Columns 1-8 as the outcome variable.

Table 5: Long-Run Effects: County-Level Evidence

Dep. var.: Twitter-based individualism and collectivism scores (Havaldar et al., 2024)				
	(1)	(2)	(3)	(4)
	Individualism		Collectivism	
	OLS	IV	OLS	IV
Labour intensity 1900	-0.257***	-0.424***	0.223***	0.437***
	(0.042)	(0.100)	(0.044)	(0.115)
F-stat		52.50		52.50
R-squared	0.45		0.33	
Observations	1862	1862	1862	1862
Agricultural land suitability	Y	Y	Y	Y
Geo-climatic controls	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y

Notes: The table shows standardized coefficients, with robust standard errors clustered at 60mi×60mi grid cells in parentheses. The dependent variables are Twitter–based indices of individualism-collectivism obtained from Havaldar et al. (2024).

Table 6: Mechanization, Labor Intensity, and Changes in Individualism

Dep. var: Infrequent name share	(1)	(2)	(3)	(4)	(5)	(6)
	OLS			IV		
Labour intensity	-0.177***	-0.187***	-0.200***	-0.342***	-0.379***	-0.479***
	(0.037)	(0.038)	(0.038)	(0.095)	(0.094)	(0.110)
F-stat				58.71	57.08	44.34
R-squared	0.89	0.89	0.90			
Observations	3592	3592	3362	3592	3592	3142
County fixed effects	Y	Y	Y	Y	Y	Y
State-year fixed effects	Y	Y	Y	Y	Y	Y
Cultural trends	N	Y	Y	N	Y	Y
Additional county characteristics	N	N	Y	N	N	Y

Notes: The table shows standardized coefficients with robust standard errors clustered at 60mi×60mi grid cells in parentheses. The share of non top-10 names is computed for counties whose number of relevant observations (individuals) is no less than 30. Additional county characteristics include the proportion of farmland, average farm size, yield per acre, value of farm machinery per acre, population density, urban population ratio, manufacturing population ratio, fertility rate, and the share of migrants.

Table 7: Mechanization, Labor Intensity and Changes in Individualism: Excluding the South

Dep. var: Infrequent name share	(1)	(2)	(3)	(4)	(5)	(6)
	OLS			IV		
Labor intensity	-0.138**	-0.162***	-0.158**	-0.586**	-0.613**	-0.571**
	(0.057)	(0.062)	(0.074)	(0.278)	(0.273)	(0.247)
F-stat				42.21	30.32	29.36
R-squared	0.90	0.91	0.92			
Observations	1916	1916	1810	1916	1916	1710
County fixed effects	Y	Y	Y	Y	Y	Y
State-year fixed effects	Y	Y	Y	Y	Y	Y
Cultural trends	N	Y	Y	N	Y	Y
Additional county characteristics	N	N	Y	N	N	Y

Notes: The table shows standardized coefficients with robust standard errors clustered at 60mi×60mi grid cells in parentheses. The share of non top-10 names is computed for counties whose number of relevant observations (individuals) is no less than 30. Additional county characteristics include the proportion of farmland, average farm size, yield per acre, value of farm machinery per acre, population density, urban population ratio, manufacturing population ratio, fertility rate, and the share of migrants.

Table 8: The Boll Weevil Shocks and Changes in Individualism

Dep. var: Infrequent name share	(1)	(2)	(3)	(4)
Boll weevil	0.208**	0.227***	0.185**	0.210***
	(0.084)	(0.073)	(0.087)	(0.076)
Boll weevil × Less intensive crops			0.108**	0.094**
			(0.043)	(0.042)
Less intensive crops			-0.056	-0.025
			(0.045)	(0.049)
R-squared	0.06	0.14	0.07	0.14
Observations	1680	1680	1680	1680
Counties	560	560	560	560
County fixed effects	Y	Y	Y	Y
State-year fixed effects	Y	Y	Y	Y
Additional county characteristics	N	Y	N	Y

Notes: Except for the dummy variable of boll weevil presence, all the dependent and explanatory variables are standardized. Robust standard errors clustered at 60mi×60mi grid cells are shown in parentheses. The sample is balanced.

Table 9: The Boll Weevil Shocks and Changes in Individualism: Placebo Tests

Dep. var: Infrequent name share	(1)	(2)	(3)	(4)
Boll weevil	-0.023	-0.038	-0.010	-0.029
	(0.106)	(0.111)	(0.108)	(0.112)
Boll weevil × Less intensive crops			-0.060	-0.046
			(0.050)	(0.050)
Less intensive crops			0.040	0.033
			(0.057)	(0.060)
R-squared	0.16	0.17	0.16	0.17
Observations	1670	1670	1670	1670
County fixed effects	Y	Y	Y	Y
State-year fixed effects	Y	Y	Y	Y
Additional county characteristics	N	Y	N	Y

Notes: The placebo regressions are identical to those in Table 8, except replace outcome variables in 1900, 1910, and 1920 with measures from 1870, 1880, and 1900, respectively. They thus estimate whether the boll weevil shock was predictive of changes in individualism going *backwards* in time. Except for the dummy variable of boll weevil presence, all the dependent and explanatory variables are standardized. Robust standard errors clustered at 60mi×60mi grid cells are shown in parentheses. The sample is balanced.

Table 10: Mechanisms Linking Agricultural Labor Intensity and Individualism

Dep. var: Infrequent name share	(1)	(2)	(3)	(4)
	OLS		IV	
Labour intensity	-0.153***		-0.291***	
	(0.034)		(0.067)	
Operations per acre		-0.032		-0.367***
		(0.032)		(0.106)
Workers per operation		0.017		-0.042
		(0.020)		(0.026)
Hours per worker		-0.239***		-0.293***
		(0.037)		(0.078)
F-stat			82.08	26.15
R-squared	0.60	0.60		
Observations	2732	2732	2732	2732
Agricultural land suitability	Y	Y	Y	Y
Geo-climatic controls	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y

Notes: The table shows standardized coefficients. Robust standard errors clustered at 60mi×60mi grid cells are shown in parentheses. The share of infrequent names is computed for counties whose number of relevant observations (individuals) is no less than 30. The geo-climatic controls are: temperature, precipitation, terrain elevation and slope, distance to major cities, distance to coast, distance to navigable river, latitude, and longitude.

Appendices

A Data Appendix

A.1 Summary Statistics

Table A1: Summary Statistics

	Obs.	Mean	Std. Dev.	Min.	Max.
	Panel 1: county-level variables				
	Obs.	Mean	Std. Dev.	Min.	Max.
Individualism measure					
Share of infrequent names 1910	2732	0.81	0.04	0.59	0.99
Twitter-based individualism score	1862	0.16	0.02	0.09	0.36
Labor intensity measures					
Actual labor intensity 1900	2732	2.58	1.44	0.91	10.33
Potential labor intensity 1900	2732	2.58	1.21	1.10	8.03
Initial local conditions					
Land suitability for agriculture	2732	0.60	0.26	0.00	1.00
Share of farmland	2732	0.39	0.28	0.00	1.03
Log average farm size	2732	5.04	0.93	1.67	13.02
Average farm yield	2732	2.28	0.58	-0.22	6.80
Herfindahl index of crop concentration	2731	0.45	0.14	0.18	1.00
Share of agricultural employment	2726	0.63	0.22	0.01	0.97
Log population density	2717	3.09	1.44	-3.25	10.39
Urban population ratio	2732	0.14	0.22	0.00	1.00
Share of elderly (aged 65 or over)	2726	0.06	0.02	0.01	0.23
Share of females	2726	0.48	0.03	0.26	0.57
	Panel 2: Metro-level variables				
	Obs.	Mean	Std. Dev.	Min.	Max.
Google Trends outcomes					
Relative search volume: Unique	205	25.74	2.94	17.00	35.00
Relative search volume: team sports	206	86.30	5.41	67.00	98.00
Labor intensity measures					
Actual labor intensity 1900	206	2.50	1.28	0.95	5.73
Potential labor intensity 1900	206	2.57	1.22	1.19	5.97

Notes: The table displays descriptive statistics for the outcome and control variables used in the analysis (without standardization).

A.2 Variables

Demographic and Socioeconomic Variables

Infrequent names of children: The complete-count census files including names are available through NBER by special arrangement with IPUMS. Using the complete-count censuses, we compute the share of infrequent names in three steps. First, we restrict the sample to include only native white and black children aged 0 to 9 whose parents are also native. Second, we count the total frequency of each first name by sex and race at the region level to identify the top-10 most frequent first names. Third, we calculate the county-level proportion of children whose first names are out of the top-10 in their corresponding regions. The top-10 most frequent names were different in each decade. For illustration, Table A2 shows the top-10 most frequent names of white and black children at the national level in 1910 and 1940.

Table A2: Top-10 Most Frequent Names at the National Level

	1910				1940			
	White		Black		White		Black	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
1	William	Mary	John	Mary	Robert	Mary	James	Mary
2	John	Ruth	James	Annie	John	Betty	John	Dorothy
3	James	Helen	Willie	Mattie	William	Dorothy	Robert	Annie
4	George	Mildred	William	Rosa	James	Ruth	Willie	Betty
5	Robert	Dorothy	George	Willie	Charles	Helen	William	Shirley
6	Charles	Margaret	Robert	Emma	Richard	Margaret	Charles	Willie
7	Joseph	Ethel	Henry	Sarah	Donald	Virginia	George	Helen
8	Edward	Edna	Joseph	Bessie	George	Mildred	Henry	Barbara
9	Walter	Elizabeth	Thomas	Hattie	Edward	Elizabeth	Joseph	Louise
10	Thomas	Alice	Frank	Ethel	Thomas	Doris	Joe	Ruth

Notes: The table shows the top-10 most frequent names at the national level in descending order of popularity.

Share of farmland: Total farm acreage divided by county area. The data are from Haines et al. (2005).

Average farm size: Average farm size in acres. The data are from Haines et al. (2005).

Average farm yield: Total value of farm output divided by total farm acres. The data are from Haines et al. (2005).

Crop concentration: Herfindahl–Hirschman Index of crop acreage shares. The data are from Haines et al. (2005).

Urban population ratio: The proportion of county residents in urban places above 2,500 population. The data are from Haines et al. (2005).

Share of blacks: The black population divided by the total population. The data are from Haines et al. (2005).

Share of agricultural employment: The proportion of agricultural labor among workers aged 18 to 70. Industrial classification follows the IND1950 variable in the Census from IPUMS-USA (Ruggles et al., 2020).

Share of elderly: The proportion of the population aged 60 and over, computed from Ruggles et al. (2020)

Share of females: The proportion of females computed from Ruggles et al. (2020)

Climatic, Geographic, and Ecological Variables

Temperature/Precipitation: County-level average of annual temperature/precipitation from IIASA and FAO (2012).

Terrain elevation: County-level average of the median elevation at 0.5 arc-min resolution from IIASA and FAO (2012).

Terrain slope: County-level average of the terrain slope index at 0.5 arc-min resolution from IIASA and FAO (2012).

Latitude/Longitude: Latitudinal/Longitudinal distance from the equator, calculated from the centroid of each county using shapefiles from Ruggles et al. (2020).

Distance to major cities: Minimum distance to major cities in 1880 (New York, Philadelphia, Chicago, Boston, St. Louis, Baltimore), calculated from the centroid of each county using shapefiles from Ruggles et al. (2020).

Distance to coastal line: Minimum distance to the nearest coastline, calculated from the centroid of each county using shapefiles from Ruggles et al. (2020) and raster files from Natural Earth.

Distance to navigable rivers: Minimum distance to the nearest navigable rivers, calculated from the centroid of each county using shapefiles from Ruggles et al. (2020) and Atack (2016).

Google Trends Outcomes

Google Trends provides a relative search volume, that is a search proportion normalized to the highest search queries across the entered terms during the designated period. The relative search volume is available from 2004 to the present and takes values between 0 and 100.

Relative search interests in words: We use the pairs of words that are considered to represent individualistic and collectivist values. The relative search volume (RSV) of each word from 2004 to the present is taken as an outcome variable.

Relative search interests in sports: For each individual sport, we use its RSV in comparison to three major team sports—football, baseball, and basketball.

B Agricultural Labor Requirements

Table A3 shows the raw labor intensity (hours/acre) for each crop by source. Column (5) averages the values across the first four sources (or uses the value of the single source), and is the basis for our baseline measure of labor intensity used in Section 3. Columns (6) and (7) show the "hand" and "machine" measures, respectively, from Wright (1899) used in Section 4.1.

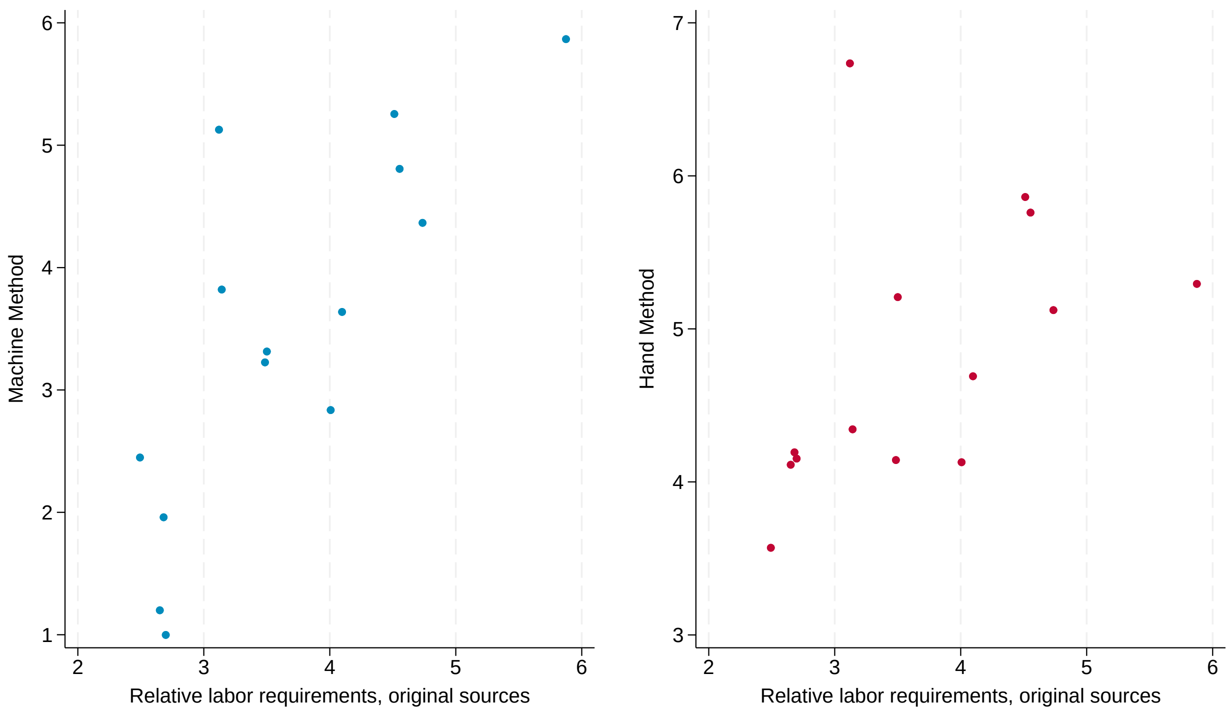
Table A3: Labor Requirements from Different Data Sources

	(1) Cooper et al. (1916)	(2) Hecht and Vice (1954)	(3) Census Bureau (1960)	(4) Reuss (1930)	(5) Average Columns (1)-(4)	(6) Wright (1899) methods based on Hand	(7) Wright (1899) methods based on Machine
Barley	12.8	16.9	-	-	14.9	63.6	2.7
Bean	-	47.2	-	-	47.2	-	-
Buckwheat	-	34.0	-	-	34.0	-	-
Corn	26.2	35.2	38.0	-	33.1	182.7	27.5
Cotton	-	116.0	112.0	-	114.0	167.8	78.7
Flax	13.7	15.1	-	-	14.4	-	-
Hay	12.3	11.9	-	-	12.1	35.5	11.6
Hemp	14.3	-	-	-	14.3	-	-
Oat	13.5	15.7	-	-	14.6	66.3	7.1
Peanut	-	69.1	-	-	69.1	-	-
Peas	-	23.2	-	-	23.2	77.0	45.6
Potato	44.4	76.0	-	-	60.2	109.0	38.0
Rice	-	55.0	-	-	55.0	62.1	17.0
Rye	10.3	-	-	-	10.3	63.0	25.2
Sorghum	-	22.7	-	-	22.7	-	-
Soybean	-	23.2	-	-	23.2	-	-
Sugar beet	-	128.0	-	-	128.0	-	-
Sugarcane	-	-	-	91.1	91.1	351.35	191.55
Sweet Potato	-	132.0	-	-	132.0	317.3	122.34
Tobacco	-	356.0	-	-	356.0	199.2	353.2
Wheat	12.3	15.2	15.0	15.9	14.2	61.1	3.32

Notes: The table shows the required man-hours per acre for crops from each data source.

Figure A1 compares the values of labor intensity between our baseline cross-sectional estimates—Column (5) of Table A3—to the "hand" and "machine" labor intensities of Wright (1899). While the sources are not identical, both capture similar information.

Figure A1: Labor requirements under hand and machine methods: comparison with the original sources



Notes: The figures plot the required man-hours per acre at the crop-level. The x-axis shows the values from the original sources used in Section 2. The y-axis shows the values from Wright (1899) under the machine(left panel) and hand (right panel) methods.

Finally, to illustrate how Wright (1899) defines the steps used in production with both "hand" and "machine" methods, Table A4 reproduces the original data for wheat on man-hours per step. In the original source the number of workers and hours per worker were listed for each step; the table combines them into a total number of man-hours for the entire step.

Table A4: Required Man-Hours per Acre for Producing Wheat, by Operation

Hand-Method		Machine-Method	
Operations	Man-Hours	Operations	Man-Hours
Breaking ground	6h 40m	Breaking ground	1h
Sowing seed	1h 15m	Sowing seed	15m
Pulverizing	2h 30m	Pulverizing	12m
Reaping, binding, and shocking	20h	Reaping, thrashing, and sacking	1h
Hauling sheaves to barn	4h	Hauling to granary	52m
Thrashing and stacking	13h 20m		
Winnowing	10h		
Gathering up and sacking	3h 20m		
Total	61h 5m	Total	3h 19m

Notes: Machine-method is the most advanced production method utilizing machinery in 1895-1896. Hand-Method implies traditional methods in 1829-30 before the general use of machinery.

C Geographic Levels of Name Frequencies

Our baseline measure of individualism is constructed using name frequency calculated at the census region level. To address concerns about the appropriateness of the census region as a reference unit, Table A5 reproduces the baseline estimates using the share of infrequent names with name frequency calculated at the county, state, and national levels. In all cases, the results indicate a significant negative relationship with agricultural labor intensity. Among the different estimates, those using county as the reference group raise the possibility of size-dependent bias, since in counties with a small number of children, the share of children with "common" names can be mechanically inflated; for example, in a county with only 10 female white children, the share of common names equals one for female white children by definition. Nevertheless, the estimates are robust and stable across different reference units, suggesting that the estimated effects of labor intensity on individualism are driven by broad tendencies of individuals to stand out from both their local community and the wider population.

Table A5: Robustness to the Geographic Scale of Name Frequencies

Dep. var.: Infrequent name share 1910						
	(1)	(2)	(3)	(4)	(5)	(6)
	County		State		Country	
	OLS	IV	OLS	IV	OLS	IV
Labour intensity 1900	-0.105***	-0.143**	-0.181***	-0.330***	-0.195***	-0.338***
	(0.035)	(0.072)	(0.037)	(0.071)	(0.040)	(0.078)
F-stat		80.23		80.23		80.23
R-squared	0.44		0.59		0.60	
Observations	2773	2773	2773	2773	2773	2773
Agricultural land suitability	Y	Y	Y	Y	Y	Y
Geo-climatic controls	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y

Notes: The table reports estimates using the share of infrequent names, with name frequency calculated at the county (Columns 1 and 2), state (Columns 3 and 4), and country (Columns 5 and 6) levels, respectively.

D Alternative Measure of Labor Intensity: Robustness

Our analysis of the mechanism relies on the crop-level characteristics based on Wright (1899). To validate this approach, it needs to be verified that the measures of labor intensity from different sources yield consistent implications. In this regard, Tables A6 and A7 replicate the robustness checks in the main text, but using the index based on Wright (1899). The results are not significantly different from those using the baseline index. The negative link between labor intensity and individualism holds across the additional controls, while the estimates by region and race also remain consistent. Although the OLS estimate for the South appears to be smaller with a relatively higher standard error, the IV estimate strongly supports our findings. It should also be noted that the alternative index of labor intensity could have exacerbated attenuation bias in the South-specific OLS estimate. Wright (1899) reports a significantly lower value of labor requirements for tobacco compared to the values from the original data sources, which may introduce a measurement error and lead the OLS estimate for the South sample toward zero.

Table A6: Robustness to Additional Controls, Using Alternative Measure of Labor Intensity

Dep. var: Infrequent name share 1910	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
						Panel 1: OLS Estimation						
Labor intensity 1900	-0.157***	-0.149***	-0.168***	-0.153***	-0.112***	-0.151***	-0.156***	-0.124***	-0.146***	-0.162***	-0.142***	-0.107***
	(0.034)	(0.034)	(0.037)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.033)	(0.034)	(0.032)	(0.035)
R-squared	0.60	0.66	0.66	0.60	0.61	0.60	0.59	0.60	0.60	0.60	0.61	0.69
Observations	2717	2732	2726	2732	2732	2731	2690	2690	2707	2730	2732	2649
						Panel 2: IV Estimation						
Labor intensity 1900	-0.344***	-0.304***	-0.359***	-0.331***	-0.363***	-0.331***	-0.331***	-0.294***	-0.317***	-0.363***	-0.271***	-0.275***
	(0.078)	(0.078)	(0.080)	(0.080)	(0.080)	(0.082)	(0.078)	(0.080)	(0.080)	(0.081)	(0.078)	(0.086)
F-stat	53.43	54.09	54.43	52.40	62.48	51.24	53.99	55.45	49.98	51.90	51.74	51.91
Observations	2717	2732	2726	2732	2732	2731	2690	2690	2707	2730	2732	2649
Population density	√											√
Urban population ratio		√										√
Agricultural employment share			√									√
Average farm size				√								√
Average farm yields					√							√
Crop concentration						√						√
Soil heterogeneity							√					√
Frontier experience								√				√
English-speaking grandparents									√			√
Native grandparents										√		√
Birthplace diversity											√	√
Agricultural land suitability	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Geo-climatic controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Notes: The table shows the robustness to additional controls using the index of labor intensity based on Wright (1899). The definition and data sources of the additional controls are provided in Appendix A.2. Robust standard errors clustered at 60mi×60mi grid cells are in parentheses. The geo-climatic controls are: temperature, precipitation, terrain elevation and slope, distance to major cities, distance to coast, distance to navigable river, latitude, and longitude.

Table A7: **Heterogeneity by Region and Race, Using the Alternative Measure of Labor Intensity**

Dep. var: Infrequent name share 1910				
	(1)	(2)	(3)	(4)
	Panel 1: Heterogeneity by Region			
	South		Non-South	
	OLS	IV	OLS	IV
Labor intensity 1900	-0.062*	-0.204***	-0.483***	-1.601***
	(0.037)	(0.078)	(0.143)	(0.572)
F-stat		58.60		15.14
R-squared	0.55		0.64	
Observations	1215	1215	1517	1517
	Panel 2: Heterogeneity by Race			
	Whites		Blacks	
	OLS	IV	OLS	IV
Labor intensity 1900	-0.152***	-0.279***	-0.076**	-0.233***
	(0.034)	(0.067)	(0.035)	(0.085)
F-stat		80.84		39.04
R-squared	0.59		0.49	
Observations	2693	2693	1188	1188
Agricultural land suitability	Y	Y	Y	Y
Geo-climatic controls	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y

Notes: The table shows the race- and region-specific estimates using the index of labor intensity based on Wright (1899). Robust standard errors clustered at 60mi×60mi grid cells are shown in parentheses. The geo-climatic controls are: temperature, precipitation, terrain elevation and slope, distance to major cities, distance to coast, distance to navigable river, latitude, and longitude.

E Major Shocks in Labor Intensity: Economic Outcomes

We use mechanization and the boll weevil infestation as an empirical setting to examine the dynamic relationship between agricultural labor intensity and individualism. Given the significance of these two shocks, this section assesses their impacts on the local economy, which may influence interpretations of their cultural effects.

Table A8 presents the results of the mechanization shock. Columns (1) and (2) indicate that counties with a greater decline in agricultural labor intensity due to mechanization experienced a larger increase in manufacturing employment. This suggests that reduction in agricultural labor requirements, driven by changes in production technologies, led to a reallocation of labor across sectors within the localities. However, Columns (3) to (6) indicate that this shift did not translate into broader local development, as proxied by urbanization rates and population density.[19] Combined with the robustness checks accounting for time-varying county characteristics in Table 6, these suggest that the link between mechanization, agricultural labor intensity, and individualism is not attributable to changes in local development status.

Table A8: Mechanization, Labor Intensity, and Local Economic Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var:	Manufacturing employment		Urbanization		Population density	
	OLS	IV	OLS	IV	OLS	IV
Labour intensity	-0.068**	-0.186**	0.009	-0.096	0.089	0.072*
	(0.030)	(0.077)	(0.025)	(0.060)	(0.076)	(0.043)
F-stat		49.33		57.08		57.08
R-squared	0.00		0.00		0.00	
Observations	3190	3190	3592	3592	3592	3592
County fixed effects	Y	Y	Y	Y	Y	Y
State-year fixed effects	Y	Y	Y	Y	Y	Y
Cultural trends	Y	Y	Y	Y	Y	Y

Notes: The table shows standardized coefficients with robust standard errors clustered at 60mi×60mi grid cells in parentheses.

The estimates from the boll-weevil shock offer a similar pattern. As shown in Table A9, while the weevil led to a decline in cotton production, there was no significant impact on other local economic characteristics. This is consistent with the observation that the primary consequence of the boll-weevil shock was a shift of farm labor from cotton to non-cotton crops, rather than labor reallocation across sectors or regions (Lange et al., 2009).

[19] For instance, the labor shift from agriculture could have increased the supply of low-skilled workers in manufacturing, but without corresponding improvements in productivity.

Table A9: The Boll Weevil Shocks and Changes in Local Economic Characteristics

	(1)	(2)	(3)	(4)
Dep. var:	Cotton production	Manufacturing employment	Urbanization	Population density
Boll weevil	-0.066**	0.010	0.008	-0.055
	(0.032)	(0.007)	(0.009)	(0.055)
R-squared	0.33	0.34	0.21	0.36
Observations	1680	1680	1680	1680
County fixed effects	Y	Y	Y	Y
State-year fixed effects	Y	Y	Y	Y

Notes: Robust standard errors clustered at 60mi×60mi grid cells are shown in parentheses. The sample is balanced.

F Migration

The impact of higher labor intensity on culture may involve two components: lower individualism among the existing population and an inflow of less individualistic migrants. If the latter were the dominant factor, our results would reflect spatial sorting rather than a shift in local cultural attitudes. As a simple additional check for our findings, we consider here the name-based measure of individualism calculated separately for migrants and non-migrants.

Specifically, using linked Census records, we separately calculate the share of infrequent names among children whose parents resided in the same county 10 years earlier and among those whose parents migrated into the county during the past decade. As shown in Table A10, the results are consistent with our baseline findings. Both the coefficients for migrant and non-migrant families are negative and significant across the OLS and IV specifications, with magnitudes similar to the baseline estimates. This suggests that, while selective migration may have contributed to the effects of higher labor intensity on culture, it acted as a complementary mechanism, reinforcing the effects on local cultural attitudes rather than providing an alternative explanation for our findings.

Table A10: Agricultural Labor Intensity and Individualism by Migration Status

Dep. var: Infrequent name share 1910				
	(1)	(2)	(3)	(4)
	OLS		IV	
	Migrants	Non-Migrants	Migrants	Non-Migrants
Labor intensity 1900	-0.233***	-0.202***	-0.220**	-0.271***
	(0.049)	(0.040)	(0.087)	(0.082)
F-stat			78.41	78.41
R-squared	0.37	0.45		
Observations	2725	2725	2725	2725
Agricultural land suitability	Y	Y	Y	Y
Geo-climatic controls	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y

Notes: The table shows standardized coefficients. Robust standard errors clustered at 60mi×60mi grid cells are shown in parentheses. The share of infrequent names is calculated for children whose parents lived in the same county in the previous census decade.

G Man-hours per acre and Production Elasticities

Here, we provide a more formal justification for the assumption that the relative man-hours per acre between crops are comparable across counties.

In county c, let there be two crops denoted by a and b. Each crop has a Cobb-Douglas production function of the form

$$Y_i = A_i X_i^{\alpha_i} L_i^{1-\alpha_i} \tag{10}$$

where $i \in (a, b)$, X_i is land used in production and L_i is labor. A_i is total factor productivity. The price of each crop is p_i, and it is assumed that both the land rental market and labor market within the county operate efficiently, so that the value marginal product of land is equated across crops, and the value marginal product of labor is equalized across crops. That implies

$$\alpha_a \frac{p_a Y_a}{X_a} = \alpha_b \frac{p_b Y_b}{X_b} \tag{11}$$

and

$$(1-\alpha_a)\frac{p_a Y_a}{L_a} = (1-\alpha_b)\frac{p_b Y_b}{L_b}. \tag{12}$$

Write the labor market condition as

$$(1-\alpha_a)\frac{p_a Y_a}{X_a}\frac{X_a}{L_a} = (1-\alpha_b)\frac{p_b Y_b}{X_b}\frac{X_b}{L_b}, \tag{13}$$

and then combine this with the land market condition to arrive at

$$\frac{L_a/X_a}{L_b/X_b} = \frac{(1-\alpha_a)/\alpha_a}{(1-\alpha_b)/\alpha_b}. \tag{14}$$

The relative labor/land ratio of the two crops depends only on the coefficients of the production functions, which is a standard result.

Our assumption that man-hours per acre (the labor/land ratio) of each crop *relative* to a baseline crop (wheat) are the same for all counties thus amounts to assuming that the coefficients for production functions of crops are the same in all counties. Implicitly, we examine whether the shape of the production function (α_a and α_b) determines the level of individualism across counties.

It is also worth pointing out what our empirical approach is *not* assuming. We do not assume that the relative prices of crops are identical across counties, that total factor productivity of each crop is identical across counties, or that the aggregate labor/land ratio is identical across counties. These aspects may well differ across counties, thus accounting for differences in the ratio of labor being used across crops (L_a/L_b) or the ratio of land used across crops (X_a/X_b). Differences in productivity, prices, or endowments can also account for differences in observed yields (Y_a/X_a or Y_b/X_b) and labor productivity (Y_a/L_a or Y_b/L_b). Differences across counties in labor ratios, land ratios, yields, or labor productivity do not invalidate our approach.

What would create an issue for our approach is if the elasticities in the production function are themselves functions of labor/land ratios. Then the relative labor/land ratios we employ may only be appropriate for the sites from which they were observed, but are not universal, and hence we are mismeasuring labor intensity for counties. One argument against this is that the labor/land ratios obtained from farm-specific studies (Cooper et al., 1916) are quite similar to the aggregate labor/land ratios we cite (Hecht and Vice, 1954) from the same period. Nor are our results sensitive to controls for population density and other measures of the county aggregate labor/land ratio.

This section also shows how our approach is similar to the approach used by Fouka and Schläpfer (2020), who estimate the values of $(1-\alpha_i)$ for various crops using a land-market condition similar to that described above along with information on prices. They then relate the values of $(1-\alpha_i)$ to the differences in work ethics in modern times. In essence, we are pursuing a similar strategy, only using the observed man-hours per acre in place of the estimated values of $(1-\alpha_i)$ to measure labor intensity. One argument in favor of our approach is that we do not have to assume that the relative price of crops was identical across counties, which is something Fouka and Schläpfer (2020) require in their approach.

H First Stage Results for Labor Intensity Components

To examine the mechanisms linking labor intensity and culture, Table 10 shows estimation results incorporating three measures of labor intensity components within a single regression. In the IV specification, we simultaneously use the potential indices of these components as instruments. This approach might raise a concern about whether the first stage relevance of the IV reflects crop-level variations in labor intensity components or is instead driven by county-level variation in crop mix. To address this, Table A11 reports the first stage estimates, which confirm that each measure of labor intensity component specifically correlates with its corresponding potential index.

Table A11: First Stage Regressions Using Labor Intensity Components

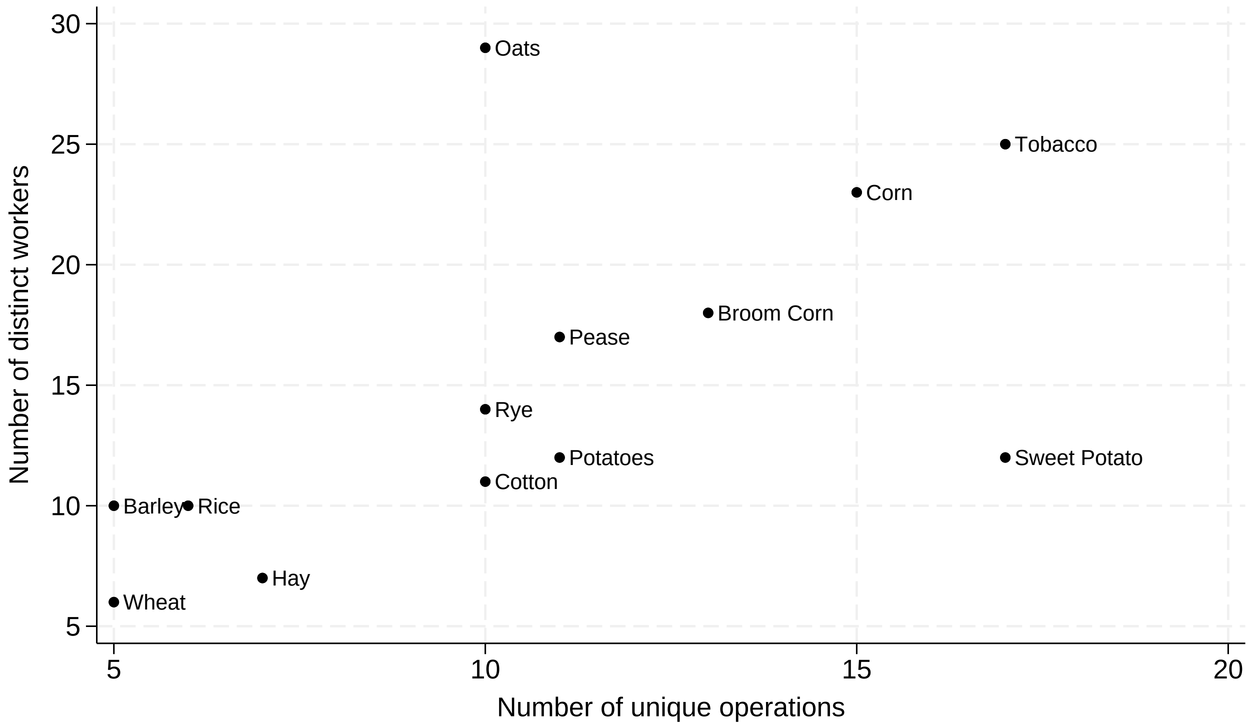
	(1)	(2)	(3)
Dep. var: actual index of	Operations per acre	Hours per worker	Workers per operation
Potential index of			
Operations per acre	0.519***	0.101**	-0.038
	(0.058)	(0.042)	(0.046)
Hours per worker	0.054	0.542***	0.143***
	(0.040)	(0.065)	(0.042)
Workers per operation	-0.016	-0.065*	1.057***
	(0.027)	(0.034)	(0.071)
R-squared	0.72	0.83	0.66
Observations	2732	2732	2732
Agricultural land suitability	Y	Y	Y
Geo-climatic controls	Y	Y	Y
State fixed effects	Y	Y	Y

Notes: The table shows standardized coefficients. Robust standard errors clustered at 60mi×60mi grid cells are shown in parentheses.

I Comparison of total versus distinct workers

Wright (1899) contains information both on the total number of workers used to produce a given crop (i.e. the sum of the number of workers involved in each operation involved with that crop) as well as information on the number of *distinct* workers used to produce a given crop, which eliminates any double counting if the same individual worked on several operations. Across crops, the correlation of those two measures is 0.96, and 0.70 (significant to less than 1%) even if we exclude sugarcane, which is an outlier in terms of both total workers and distinct workers.

Figure A2: Number of distinct workers versus number of operations, by crop



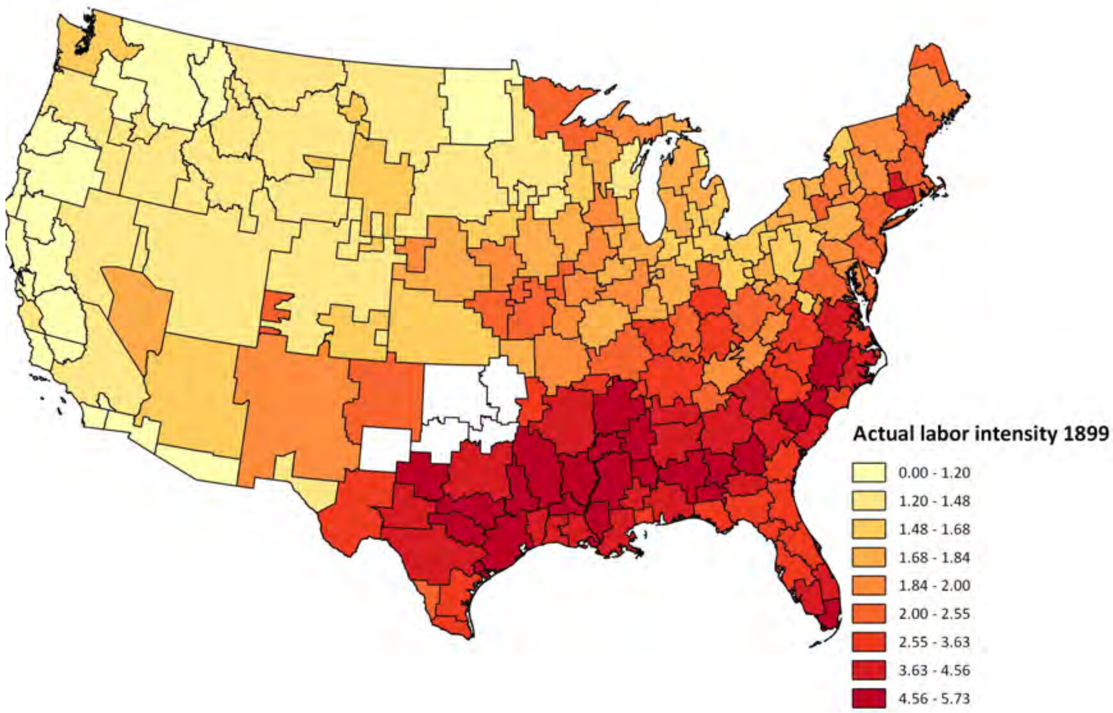
Notes: The figures plot the number of distinct workers engaged in crop production over a growing season against the number of unique operations that crop involves. Data are from Wright (1899), and refer the "Machine" techniques. Sugarcane is excluded, as it is an extreme outlier in terms of number of distinct workers (around 150) compared to other crops.

Moreover, there is a clear relationship between the number of distinct workers and the number of operations involved in a crop. Figure A2 plots the two measures against one another and the positive relationship is clear. For the data in the figure the correlation is 0.601, significant at less than 5%. This excludes sugarcane, which is an extreme outlier in terms of distinct workers (147).

J Agricultural Labor Intensity Harmonized with DMA Boundaries

For the analyses based on Google Trends data in Section 3.4, we need to harmonize the county-level explanatory variables with DMA boundaries. Harmonization is conducted in three steps. First, we intersect the shapefiles of the county and DMA borders. Second, county-level observations are assigned to fragments within each county, relying on the assumption that the county-level data are evenly distributed over county areas. Finally, the fragment-level data are weighted by the proportion of the county's area to the area of the corresponding DMA and aggregated into the DMA level.

Figure A3: Agricultural Labor Intensity Harmonized with DMA Boundaries



Notes: The map shows agricultural labor intensity in 1900 harmonized with the DMA boundaries. The sample excludes DMAs if their area of 30% or more overlaps counties whose labor intensity in 1900 is not available.