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

We thank Sam Bazzi, Paul Rhode, and conference participants at the Southern Economic Association 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
JEL No. N51,O13,P16

ABSTRACT

We study the influence of agricultural labor intensity on individualism across U.S. counties. To measure historical labor intensity in agriculture we combine data on crop-specific labor requirements and county-specific crop mix around 1900. To address endogeneity we exploit climate-induced variation in crop mix. Our estimates indicate that an increase of one standard deviation in labor intensity is associated with a reduction of 0.2-0.4 standard deviations in individualism (as captured by the share of children with infrequent names). We further document consistent patterns using within-county changes in labor intensity over time due to both mechanization and the boll weevil shock. While culture transformed in response to changes in labor intensity, we also find that historical agricultural patterns had a lasting imprint that influences geographic variation in individualism today.

Martin Fiszbein
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215
and NBER
fiszbein@bu.edu

Yeonha Jung
Sungkyunkwan University
yhjung0408@skku.edu

Dietrich Vollrath
University of Houston
201C McElhinney Hall
Houston, TX 77204
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). An emerging body of research links individualistic cultures to low labor intensity in agriculture (see Talhelm, 2020; Ang, 2019). We provide new evidence on this link using rigorous empirical strategies that leverage rich subnational variation in the United States. We estimate the effects of historical labor intensity in agriculture on historical and present-day individualism exploiting climate-induced variation in crop mix. Moreover, we study the link between *changes* in agriculture and *changes* in culture, shedding new light on the process of cultural formation.

Agricultural labor intensity may have influenced individualism in various ways. High labor intensity may be associated with interdependence in the production process and thus higher returns to coordination and cooperation, favoring non-individualistic cultural formation. It may also be related to a higher population density and tighter social ties. Furthermore, labor intensity reflects the elasticity of output with respect to labor; a high elasticity implies that marginal products of labor are similar to average products and that non-marginal reductions in labor input lead to large social losses, thereby possibly fostering a sense of equality and community.

To measure labor intensity in agriculture at the county level, we use data on hours of work per hectare for various crops from historical sources that capture standard techniques in the 19th and early 20th century. We combine this information with data on the acreage devoted to each crop to compute an overall county-level measure 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). This is consistent with some 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.

As part of our identification strategy, we exploit 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.

We find that a one standard deviation increase in labor intensity in 1900 is associated with reductions of 0.29 standard deviations in individualism in 1910. These IV estimates are larger in magnitude than the ordinary least squares (OLS) estimates, although in all cases we obtain strongly significant associations. The results survive an array of robustness checks that include controls for the value of agricultural production and various demographic characteristics. They also hold *within* different regions of the United States, e.g. South and non-South, and for both white and black populations.

Next, we analyze the process of cultural formation. We show that historical *changes* in labor intensity led to *changes* in individualism in two exercises that leverage different sources of identifying variation in labor intensity—the mechanization of agriculture in the late 19th century and boll weevil infestation in the early 20th century. Beyond confirming the cross-sectional results, these exercises contribute to our understanding of cultural formation.

The mechanization of agricultural production differentially impacted the labor requirements of each crop, which induced county-level variation in labor intensity depending on 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 reduction of 0.36 standard deviations in individualism during the same period. These first-differenced results are not affected by unobservable time-invariant characteristics of U.S. counties, which might affect the cross-sectional results. We also confirm that the cultural shift due to changes in labor intensity is robust to the alternative characteristics of agricultural mechanization, such as the amount of agricultural capital acquired.

The boll weevil, an insect that feeds on cotton, infested the U.S. South between the 1890s and the 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. The shocks led to shifts from cotton, a very labor intensive crop, to alternative crops with different degrees of labor intensity, depending on the agricultural suitabilities of each county. Our results show that the arrival of the boll weevil induced a 0.11 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.

The historical relationship between cultural traits and agricultural labor intensity persists today. Based on Google Trends data, we find that counties with higher labor intensity in 1900 display a lower search interest in "unique," a typical example of an individualistic word (Twenge et al., 2012; Greenfield, 2013), relative to "common" as an antonym. This finding is corroborated by a greater search interest in team sports relative to individual sports. We also provide additional evidence based on social attitudes. Higher labor intensity in the past predicts higher voter turnout, a common measure of civic capital, as well as higher values of the social capital index. This pattern is accompanied by greater support for redistributive policies, consistent with the notion of "horizontal" collectivism, which emphasizes equality across members of the group (Singelis et al., 1995).

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., 2018), 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 (2017) 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, 2017), and the origin of religious communities (Ager and Ciccone, 2018). We provide rigorous new evidence on the roots of individualism in agriculture.

Our analysis fits into a strand of literature that studies both cultural persistence and cultural change, not viewing them as contradictory but rather as two aspects of cultural evolution (see Giuliano and Nunn, 2017). 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 systemically related to fundamental determinants of cultural background.

We 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., Beck-Knudsen, 2019; Olsson and Paik, 2016; Hoang-Anh et al., 2018; Bazzi et al., 2018; Roland, 2020) and are closely connected to those focusing on agricultural labor intensity (Talhelm et al., 2014; Talhelm, 2020; Talhelm and English, 2020; Ang, 2019).

The link between agricultural labor intensity and individualistic cultures was established in Talhelm et al. (2014) on cultural differences between the rice-growing and wheat-growing areas of China. Talhelm (2020) and Talhelm and English (2020) provide further evidence and establish a cross-country association between rice specialization and individualism. Ang (2019) anticipates our approach of going beyond rice and documents an association between agricultural labor intensity and individualism at the cross-country level and for 186 pre-industrial societies. Our study complements previous evidence by exploiting U.S. county-level data, the largest cross-section studied so far, and by introducing an identification strategy that exploits climate-induced variation in crop mix. Most importantly, while previous studies link historical to modern outcomes, we document not only such a relationship but also contemporaneous historical associations and further explore the connection between *changes* in labor intensity and in individualism. This novel type of evidence brings the causal nature of the association closer and sheds further light on the process of cultural formation.

We also make a contribution to the measurement of agricultural intensity. Conceptually, our definition of labor intensity links it to the marginal product of labor and the elasticity of output with respect to labor for different crops, similar to the analysis of Vollrath (2011), Eberhardt and Vollrath (2018), Johnson and Vollrath (2020), and Fouka and Schlaepfer (2015). One of our core data sources for measuring labor intensity in agriculture is the Census source used by Ang (2019), but we further draw on alternative sources to validate the data on crop-specific labor requirements and to expand the set of crops included in the overall measures of agricultural labor intensity across U.S. counties. Our contribution to measuring labor intensity at the county-level is important for our study and may also be useful for research that focuses on how this dimension of agricultural societies is related to aspects other than individualism, such as the link with preferences about leisure and work (see, e.g., Fouka and Schlaepfer, 2015) or the process of structural change (see, e.g., Vollrath, 2011).

The remainder of the paper is organized as follows. Section 2 explains how we measure agricultural labor intensity and individualism. In Section 3, we present a brief conceptual framework. Section 4 shows the cross-sectional results, and Section 5 provides evidence using changes in labor intensity due to mechanization and the boll weevil shock. Section 6 documents the long-term relationship between labor intensity and modern cultural outcomes. Section 7 concludes.

2 Measuring Labor Intensity and Individualism

2.1 Agricultural Labor Intensity

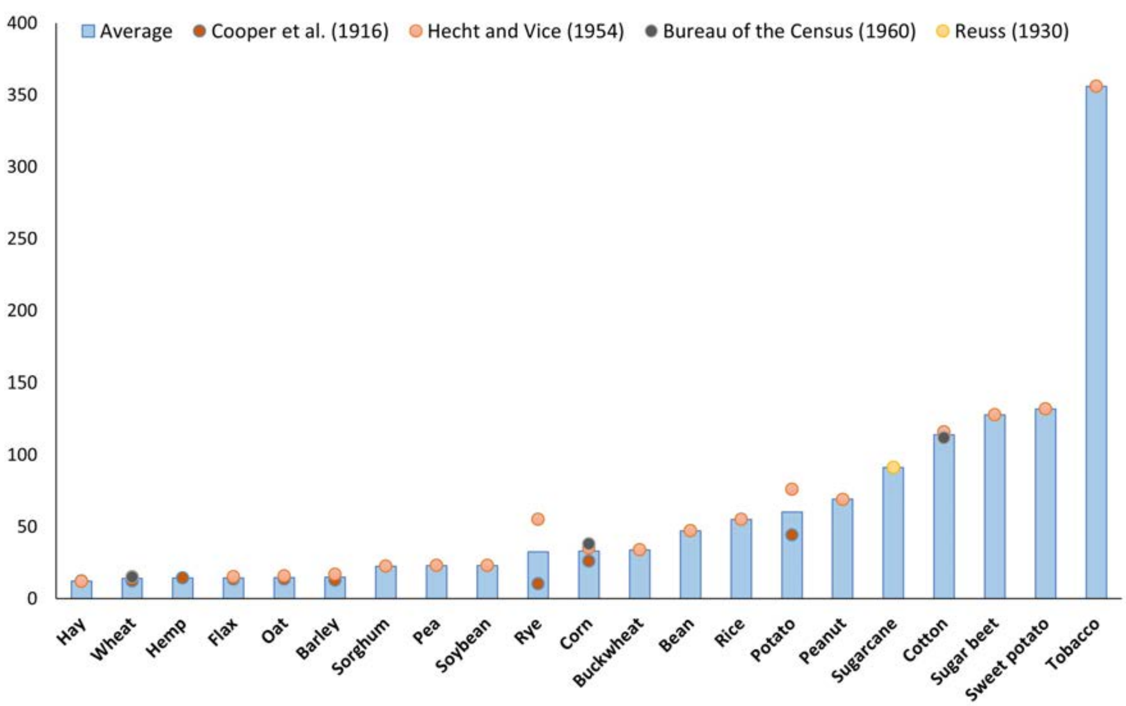
We measure labor intensity by the number of man-hours per acre used to produce a specific crop, based on data from historical sources described below. These crop-specific labor intensity measures are then combined with the actual acreage share of crops in a county (or the potential share based on geoclimatic 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).

Figure 1: Crop-specific labor intensities from different data sources



Note: 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.

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 A2 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, and 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 may suggest that recorded man hours per acre measure 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 hectare 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.

The argument above is formalized in Appendix C, where we also connect our results to others that estimate historic crop-specific production functions (Fouka and Schlaepfer, 2015). 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 our empirical analysis to be meaningful are 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 4.2 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:

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

$\text{Crop share}_{j,c}$ 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. In short, Intensity_c is the weighted average of crop-specific labor intensity using the acreage share of each crop as the weight.

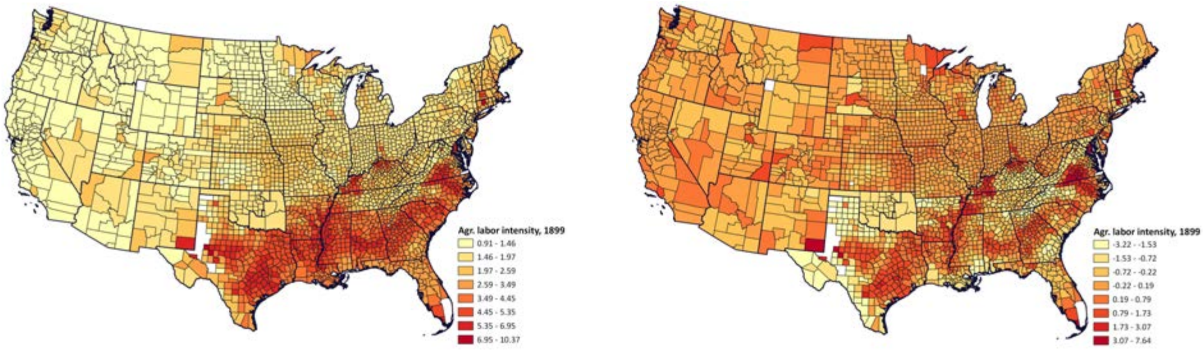
Figure 2 shows the spatial distribution of Intensity_c (Panel a) and the deviations from the state-level means (Panel b). While there was stark regional contrast, with the South characterized by much higher levels of labor intensity, there was also significant within-state variation everywhere, which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

Appendix Table A2 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 Sections 3.2 and 5.1.[1]

[1] Appendix Figure A1 plots the crop-specific labor requirements from Wright (1899) for both production methods against our 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.

Figure 2: Agricultural labor intensity, 1900

(a) Raw figures **(b) Deviations from state-level means**



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

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 suggest that the differences in naming patterns reflect the individualistic culture of frontier settlement. Bazzi et al. (2018) also rely on name-based measures to investigate the link between frontier experience and individualistic culture in the US. Ogihara et al. (2015) show that popular pronunciations of baby names in Japan decreased between 2004 and 2013, and that this pattern is closely associated with a decrease in collectivism indices.

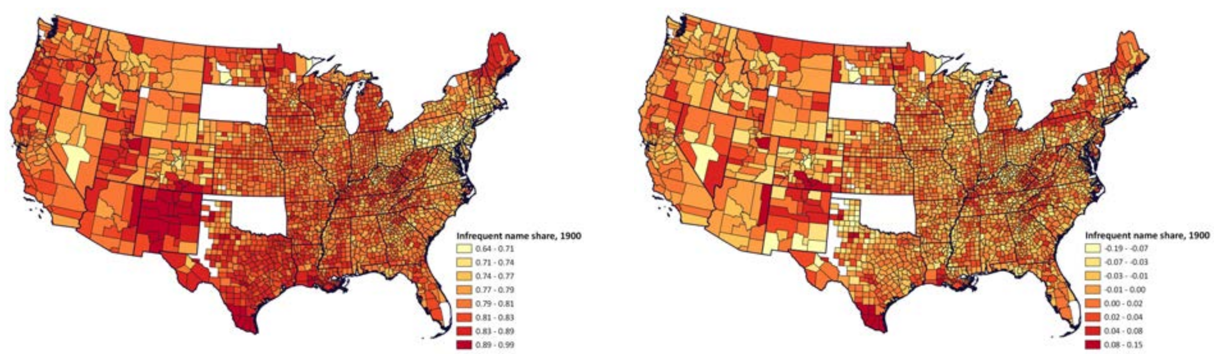
We measure county-level individualism in the historical period using the complete count census data. We compute the share of children aged 0-9 whose names are out of the top 10 most frequent names. This threshold is common in the social psychology literature, but different thresholds lead to similar results.

To avoid bias due to different naming patterns by race, we identify 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 shows the top 10 most frequent names of whites and blacks at the national 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%.

We restrict the sample to native white and black children with native parents to eliminate the divergent naming patterns of recent immigrants. The frequency of names is calculated at the U.S. Census region level for each decade's birth cohort, but the results are robust to other geographical levels for computing the top-10 names. While the baseline computation of the share of children with infrequent names includes both blacks and whites, Section 4.3.2 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.

Figure 3: Share of infrequent names in 1900

(a) Raw figures **(b) Deviations from state-level means**



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

3 Conceptual Framework

3.1 How Labor Intensity May Shape Culture

There are various mechanisms that could drive the relationship between labor intensity and individualism. Prior to the empirical investigations, this section discusses multiple channels in cultural formation that may have been relevant: the stability of labor use throughout the year, the scale of the labor force involved in the production process or at given steps, and the degree of complexity in production. These multiple dimensions associated with labor intensity may have favored the diffusion of non-individualistic norms through utilitarian intergenerational cultural transmission (e.g., Doepke and Zilibotti, 2014) and/or through cultural group selection (e.g., Henrich, 2004). While examining the relevance of the various forces discussed here is beyond the scope of our empirical analysis, discussing concrete mechanisms is still important for motivating why hours per acre could be associated with individualism.

To see how higher agricultural labor intensity is associated with factors promoting non-individualistic culture, it is convenient to note that our measure of labor intensity, $\frac{hours}{acre}$, can be expressed as $\frac{hours}{worker} \times \frac{workers}{acre}$. We use this expression to organize ideas throughout this section.

A key driver of differences in $\frac{hours}{worker}$ across crops was variation in labor needs over the course of the growing season. Insofar as workers put in broadly similar work hours per day, then yearly $\frac{hours}{worker}$ proxies for (the inverse of) seasonality. The grains and hay that were prominent on northern farms required concentrated labor at the time of planting and harvest, but little labor between those steps. Compared to those grains, crops such as maize and cotton involved more attention throughout the growing season (e.g. multiple rounds of cultivation), but the greater flexibility in planting and harvesting did not require a concentrated labor effort at either stage (Earle, 1978; Reid, 1979). Earle made the distinction between "few-day" northern grain crops and "multiple-day" southern staples based on these characteristics.

For multiple-day crops, it made sense to have a dedicated group of workers (family and/or hired-in) throughout the year. This would in turn imply more stable interactions, plausibly fostering cooperation by making workers interact through a repeated game. Few-day crops did not engender the same need for dedicated workers, and instead usually relied on occasional hired-in workers. Few-day crops thus created limited possibilities for building a collective identity among workers, while multiple-day crops created many such opportunities. A similar connection is made by Talhelm when discussing differences in individualism between "few-day" wheat and "multiple-day" paddy rice production (Talhelm et al., 2014; Talhelm, 2020).

The other factor in the decomposition above, $\frac{workers}{acre}$, is a crude proxy for production scale. High values of $\frac{workers}{acre}$ could lead to higher population density for any given level of $\frac{hours}{worker}$ (high values in this factor due to low seasonality in labor requirements would have implications in the same direction). In turn, high population density may be associated with stronger social connections (Chay and Munshi, 2012). Empirically, though, we do not find that population density can account for the link between agricultural labor intensity and culture (see Table 2).

Beyond an effect through population density, $\frac{workers}{acre}$ could more directly influence culture through its implications for the organization of agricultural production, which may have involved two different forms of interdependence—a basic form created by the size of production teams at given tasks and a stronger form created by complexity and coordination requirements in the production process. Both larger teams at specific steps and a greater number of distinct steps in the production process would imply a larger scale as proxied by $\frac{workers}{acre}$.

Consider first why the size of the production team at a given step of production might raise the returns to communitarian cultural norms. Some crops, such as wheat and barley, required large crews of workers during planting and harvesting. Coordination needs among individual workers, which may favor non-individualistic norms, could be increasing in team size. Even with limited needs for coordination, larger teams may foster communitarian norms if they were associated with a higher elasticity of output with respect to labor. A higher elasticity would imply larger proportional reductions in output from non-marginal reductions in labor input due to, for example, shirking, absenteeism, low effort, or poor health. Communitarian norms would limit such output losses.

While the number of workers at each step could imply some degree of interdependence, the number of steps—a proxy for the complexity of the production process—could involve an additional, perhaps stronger form of interdependence. With strong complementarities across steps, like in the "O-ring" perspective of Kremer (1993), more steps imply more possibilities for failure created by lack of skill or effort in any given part of the process.[2] A high number of steps thus entails greater dependence of workers at a given step on the work of those at other steps. The organization of production for complex crops could then favor the adoption of non-individualistic norms.

The idea that the complexity of agricultural production influences individualism is consistent with the theories of Talhelm et al. (2014), Talhelm and Oishi (2018), Talhelm (2020), and Talhelm and English (2020) regarding paddy rice versus wheat agriculture in China. Large-scale irrigation networks were not as common in the U.S., although there were some projects consistent with the cultural theories discussed here. Arrington and May (1975) wrote 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 our evidence does not rely on such extreme examples, it does suggest that similar forces were at work in the U.S. during this time period.

3.2 Labor Intensity, Seasonality, and Complexity at the Crop Level

We can illustrate the associations of labor intensity with seasonality, scale, and complexity at the crop level for 13 relevant crops using data from the *Thirteenth Annual Report of the Commissioner of Labor* (Wright, 1899). This exceptionally detailed report provides descriptions of the production process for different crops during the late 19th century, including information on the various production steps, numbers of workers, their roles, and hours.[3]

Labor intensity, $\frac{hours}{acre}$, is strongly correlated with both $\frac{hours}{workers}$ and $\frac{workers}{acres}$ at the crop level. Figure 5a plots $\left(\frac{hours}{acre}\right)$ from Table A2, in logs, against (a) $log\left(\frac{hours}{worker}\right)$ and (b) $log\left(\frac{workers}{acre}\right)$ for all crops available from the Commissioner's Report. These two variables in the horizontal axes of are the two terms in the decomposition of labor intensity, which proxy for (a) (the inverse of) seasonality and (b) production scale.

Either or both seasonality and scale are important when considering the relative labor intensity of specific crops. For instance, cotton and sweet potato have much higher levels of labor intensity than barley and rye mainly due to differences in $\frac{hours}{worker}$. The distinction between "multiple-day" staple crops and "few-day" grains is the key driver of the differences in these cases. Multiple-day

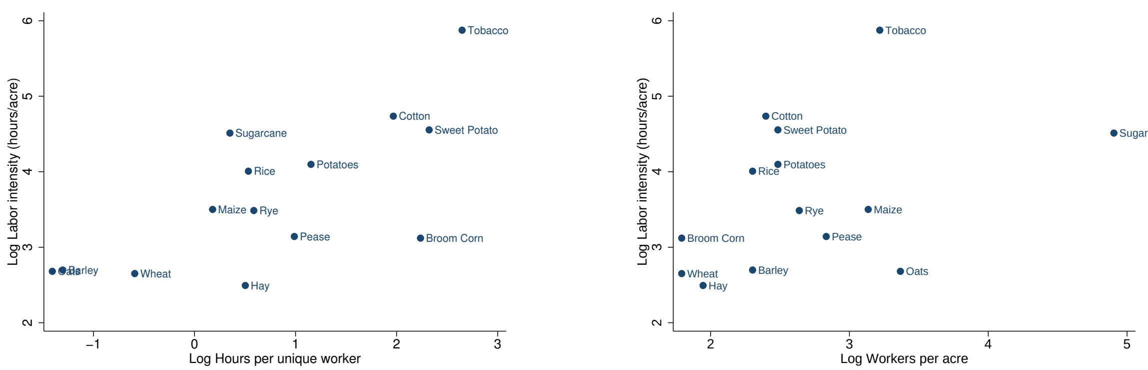
[2] To illustrate "O-ring" complementarities, consider a 90% chance of success at each production step. A crop with 5 steps (e.g. wheat) then had a 60% chance of success, while a crop with 15 steps (e.g., maize) would have a 21% chance of success.

[3] Here we use the data for the "machine method" (Wright, 1899, p.25), which corresponds to a combination of hand labor and animal-powered implements, prior to the widespread adoption of tractors and powered implements in the 20th century (Olmstead and Rhode, 2001). In Section 5, we provide further details.

crops entailed extensive work of workers throughout the year, implying they were more tightly tied to the farm and one another. In other cases, the key dimension was scale: for instance, sugarcane's high labor intensity was not due to its level of $\frac{hours}{workers}$, which were comparable to those of maize, rye, and hay, but to its extremely high level of $\frac{workers}{acre}$. Both dimensions are important in other comparisons. For example, potatoes' high labor intensity relative to wheat was driven both by more hours per worker and more workers per acre.

Figure 4: Labor intensity, seasonality and scale

(a) Labor intensity and hours per worker **(b) Labor intensity and workers per acre**

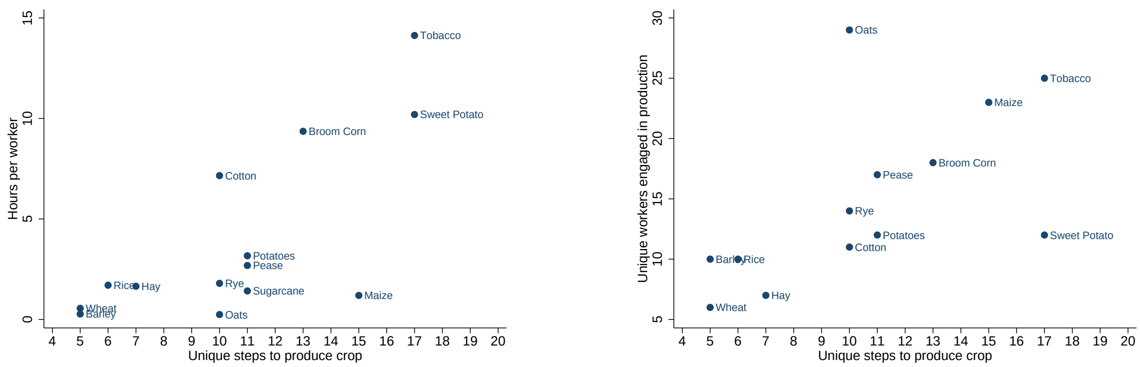


Notes: Measures of hours per worker and workers per acre are based on the Thirteenth Annual Report of the Commissioner of Labor (Wright, 1899). Log hours per acre (labor intensity) are from Appendix Table A2.

Both dimensions of labor intensity emphasized above—scale and seasonality—are in turn plausibly associated with the number of steps in production—another aspect of agricultural production that can shape cultural norms. Scale, as proxied by $\frac{workers}{acres}$, depends positively on the average team size across steps and the number of steps required in production. Crops with many required steps also tend to be low seasonality crops, since steps are commonly laid out in a sequential (rather than simultaneous) manner over the year. Figure 5b shows that both $\frac{hours}{workers}$ and $\frac{workers}{acres}$ (displayed in logs in panels a and b, respectively) are both positively associated with the number of required steps across crops. The patterns observed with data from the Commissioner of Labor's 1900 Report are consistent with the textual evidence about the production process from Robinson (1863).

Figure 5: Seasonality, scale, and complexity

(a) Hours per worker and steps **(b) Workers per acre and steps**



Notes: Measures of hours per worker and workers per acre are based on the Thirteenth Annual Report of the Commissioner of Labor (Wright, 1899). Log hours per acre (labor intensity) are from Appendix Table A2.

While our proxies for seasonality, scale, and complexity cannot be aggregated from the crop level to the county level in a straightforward manner, the crop level associations documented here provide concrete illustrations of their links with agricultural labor intensity.[4]

3.3 Cross-Regional and Within-Region Comparisons

Two key dimensions of agricultural labor intensity that we discussed above—seasonality and scale—have been emphasized in two salient economic theories of slavery. Earle (1978) and Reid (1979) argue that slavery was not profitable under high seasonality, while Engerman and Sokoloff (1997, 2002) hold that slavery was favored by the presence of scale economies. According to these ideas, the South's suitability for crops that required large production teams throughout the year was conducive to the region's history of slavery.[5]

These theories are consistent with our emphasis on the influence of labor intensity on culture. Culture and institutions may both be influenced by labor intensity. Slavery and later forms of labor coercion were plausibly a (morally abominable) substitute for cultural mechanisms favoring cooperation. This speculative comment on cross-regional differences in institutions is beyond the scope of our empirical analysis, which focuses within-region variation. However, these differences are key to context. They may have influenced the responsiveness of culture to agricultural labor intensity and/or shaped culture through different but related mechanisms.

[4]For our measure of labor intensity, $\frac{hours}{acre}$, we simply average the crop-level figures using county-level acreage shares as weights. However, for $\frac{hours}{worker}$ and $\frac{workers}{acre}$, a meaningful aggregation to the county level would require knowing if workers were involved in the production of multiple crops. For steps, we would need to know if the common steps for multiple crops are performed jointly.

[5]The link between specialization in plantation crops and reliance on enslaved workers had exceptions, for example, the mixed farms in Virginia that grew wheat with enslaved workers tied to the farm year-round, as described by Wright (2003).

Our subnational analysis, where we leverage variation across counties controlling for state fixed effects, compares levels of labor intensity and levels of individualism for comparison units that are located within a similar macro-institutional context. In other words, we identify how labor intensity shapes culture based on variations *within* regions (in fact, more narrowly, within states). Of course, within-region effects could be heterogeneous depending on the institutional context, but our results show that the relationship between labor intensity and individualism holds within both the North and South when estimated separately. This is the salient takeaways for us. While sharply heterogeneous institutions existed in different regions of the U.S., agricultural labor intensity shaped cultures within all regions, suggesting that the forces we accentuate may be of general relevance.

In terms of identifying variations in agricultural labor intensity that we leverage, our within-region approach implies that we are mostly not directly comparing crop specialization patterns at the extremes of the range but rather intermediate cases with either one of the extremes. For instance, while Earle (1978) and Reid (1979) compare "few-day" grain crops in the north and "multiple-day" staples in the south, our estimates are identified from comparing these extremes not directly but rather with intermediate cases. Maize provides a useful intermediate comparison, as it was grown throughout the U.S. during this period. Maize used 1.19 man hours per acre, about 4.4 times higher than barley did, but only one-sixth the requirement for cotton. Earle (1978) notes that maize production "demanded exceptional attention" from plowing through the third cultivation, well above what wheat or barley required. Moreover, because maize could be left on the stalk, it did not have the concentrated demand for labor at harvest that wheat or barley did. Maize harvesting and husking were labor intensive but could be spread over many months (Bidwell and Falconer, 1925). In sum, maize would have had higher returns to collectivism than wheat or barley but lower than cotton. This in turn means that there is relevant variation in labor intensity within all geographic regions of the United States.

4 Labor Intensity and Individualism Across U.S. Counties

This section examines the contemporaneous relationship between agricultural labor intensity and individualism. First, we find a strong negative correlation between actual labor intensity and the share of infrequent names. Second, to address bias due to endogenous crop choice, we employ a measure of potential labor intensity as an IV. Last, we document that our findings survive a wide range of robustness checks.

4.1 Estimating Equation

Our estimating equation is

$$y_{c,1910} = \alpha + \beta \text{Intensity}_{c,1900} + \gamma' \boldsymbol{X}_c + \mu_s + \epsilon_{c,t} \tag{2}$$

where $y_{c,1910}$ denotes the proportion of infrequent names measured in year 1910 for county c, $\text{Intensity}_{c,1900}$ is the index of agricultural labor intensity measured in 1900, and $\boldsymbol{X}_c$ is the vector of predetermined conditions that could have impacted both labor intensity and cultural characteristics simultaneously. The vector 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).

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 4.3.2 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.

4.2 Instrumental Variable Strategy

Crop choice is endogenous. A natural concern about the OLS results is that individualism could influence the composition of production and thus the Intensity_c index. 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 (2017), which adopts a conditional logit framework of McFadden (1974). 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 $\mathbf{A}_c$.[6] 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

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

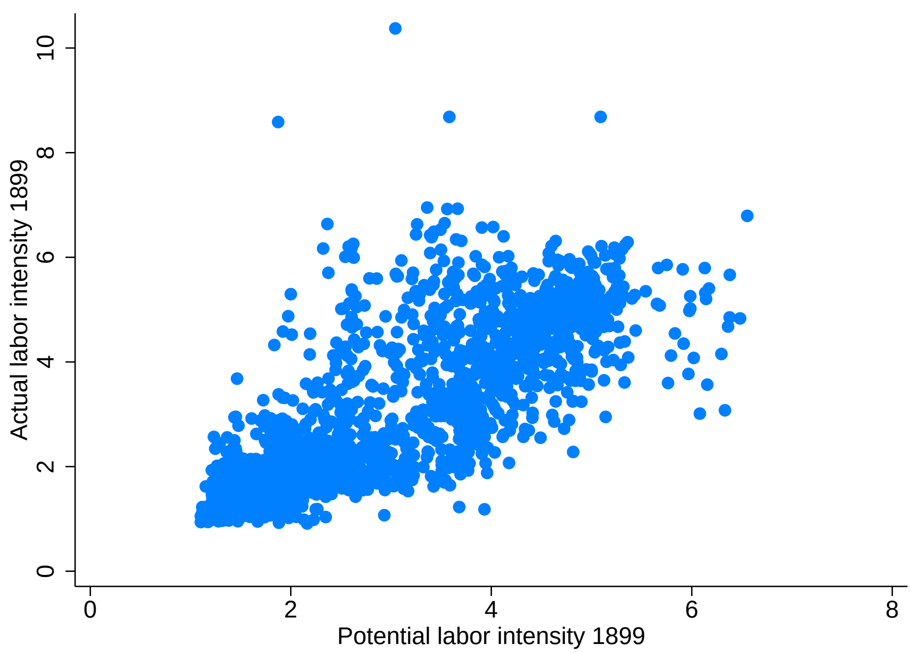
which is the fitted value for a county from the FML estimation. Thus, it strips out the non-geographic variation present in the residuals.

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 A2.

Figure 6: Correlation between the actual and potential labor intensity



Note: 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.

[6]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.

Figure 6 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, but there is variation in actual labor intensity around that potential, which may reflect endogenous crop choice based on factors other than inherent productivity. In the empirical work that follows, we use the measure of potential intensity as an instrument for the actual labor intensity, and see that the relationship between labor intensity and individualism holds.

4.3 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. This relationship survives a battery of robustness checks. We find that the negative effects of agricultural labor intensity on individualism are stable against dividing the sample by race or region, and including additional controls. This suggests that our findings do not stem from a particular context of U.S. history.

4.3.1 Baseline Estimation

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 0.20 standard deviation decrease in the share of infrequent names in 1910. Illustratively, a one standard deviation difference in labor intensity corresponds to the difference between Gadsden County and Suwannee County in Florida, which are approximately 80 miles apart. The most prevalent crop in the two counties was corn, with acreage shares of 65.4% and 80.0% in Gadsden and Suwannee, respectively. The cotton share in Gadsden was 10%, whereas Suwannee reported no cotton production. Consequently, Gadsden had a labor intensity 0.98 standard deviations higher than Suwannee, and also reported a share of infrequent names that was 0.78 standard deviations lower than Suwannee.

The IV estimation supports the strong negative effects of labor intensity on individualism. Panel 2 shows that the IV estimates are significant and negative with larger magnitudes than the OLS coefficients. The size difference might result from attenuation bias or potential downward bias in the OLS estimates. For example, higher labor intensity could have been fostered by regional laws and regulations that support low-wage labor markets. If such institutional environments were correlated with non-individualistic tendencies, then the OLS estimates could be biased downward.

The larger size of the IV estimates could reflect a differentially higher effect of labor intensity on culture for counties with higher levels of cultural malleability (for a discussion of hardwired and fluid cultures, see Acemoglu and Robinson, 2021). Higher cultural malleability would imply higher responsiveness of agricultural production patterns (in particular, actual labor intensity) to the geo-climatic conditions captured in the IV, and would also likely imply stronger effects of labor intensity on culture.

Table 1: Agricultural labor intensity and individualism

Dep. var: Infrequent name share 1910				
	(1)	(2)	(3)	(4)
	Panel 1: OLS estimates			
Labor intensity 1900	-0.222***	-0.216***	-0.199***	-0.195***
	(0.038)	(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			
Labor intensity 1900	-0.291***	-0.291***	-0.284***	-0.291***
	(0.072)	(0.073)	(0.067)	(0.066)
F-stat	135.40	129.75	85.65	82.87
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.

4.3.2 Robustness to Additional Controls

As the potential labor intensity is constructed from exogenous crop suitability, the results of the IV regressions mitigate the concerns of omitted variables. Nevertheless, some initial socioeconomic conditions might confound the interpretation of the results. Though the additional controls could be in part results of labor intensity and thus “bad controls”, robustness to their inclusion would support a causal interpretation in Section 4.3.1.

Table 2 documents the OLS and IV estimation results with various additional controls. Column (2) controls for overall population density, which could foster the production of labor-intensive crops and may foster more collectivist behavior through channels unrelated to agricultural labor intesnity. Moreover, the path dependence of population density can lead to bias in the long-run (Bleakley and Lin, 2012). In all specifications, however, the inclusion of population density does not change the estimates.

Table 2: Robustness to additional controls

Dep. var: Infrequent name share 1910	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
					Panel 1: OLS Estimation						
Labor intensity 1900	-0.188***	-0.193***	-0.198***	-0.185***	-0.187***	-0.163***	-0.186***	-0.242***	-0.187***	-0.187***	-0.215***
	(0.034)	(0.034)	(0.033)	(0.034)	(0.034)	(0.033)	(0.034)	(0.036)	(0.034)	(0.034)	(0.035)
R-squared	0.60	0.61	0.67	0.60	0.60	0.62	0.60	0.67	0.60	0.60	0.69
Observations	2732	2717	2732	2732	2732	2732	2731	2726	2726	2726	2711
					Panel 2: IV Estimation						
Labor intensity 1900	-0.277***	-0.290***	-0.279***	-0.265***	-0.275***	-0.323***	-0.272***	-0.337***	-0.277***	-0.277***	-0.336***
	(0.064)	(0.064)	(0.062)	(0.063)	(0.064)	(0.064)	(0.065)	(0.064)	(0.064)	(0.064)	(0.066)
F-stat	81.63	80.82	81.87	82.12	80.84	87.35	81.05	82.91	82.16	82.57	79.34
Observations	2732	2717	2732	2732	2732	2732	2731	2726	2726	2726	2711
Population density		√									√
Urban population ratio			√								√
Share of farmland				√							√
Average farm size					√						√
Average farm yields						√					√
Crop concentration							√				√
Agricultural employment share								√			√
Share of elderly									√		√
Sex ratio										√	√
Predetermined controls	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

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. 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.

Column (3) controls for the urbanization rate to confirm that economic development status is not a valid omitted variable. Columns (4)-(7) consider agricultural properties other than labor intensity: the share of farmland among the total county area, average farm size, yield per farm acre, and crop concentration. We control for the share of farmland in consideration that the prevalence of agriculture alone could be a source of collectivism.[7] By including the average farm size and average farm yields, we also test whether the effects of labor intensity are confounded by scale economies or productivity of crop production. We also control for the extent of crop concentration, which is measured by the Herfindahl–Hirschman index of crop acreage shares, to mitigate potential bias from homogeneous conditions of agricultural production. In all cases, the estimates are not sensitive to the additional controls. Columns (8)-(10) show that the baseline estimates are robust to other demographic characteristics including the proportion of females, elderly, and agricultural employment. Column (11) includes all additional controls, and the results

[7] 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.

do not change significantly.

4.3.3 Heterogeneity by Region and Race

Heterogeneity across regions and races might challenge the generalization of the cultural effects of agricultural labor intensity. To address this concern, 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.

The sharp historical contrasts in institutions and social stratification between regions are also likely sources of heterogeneity in how agricultural patterns influence culture. 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) show the robustness of the estimates to using only the southern sample; dropping the southern states also does not significantly affect the results, as shown in Columns (3) and (4). These suggest that the relationship between agricultural labor intensity and individualism is not confounded by differences in demographic structures or historical backgrounds between the southern and non-southern regions. In Panel 2, we reproduce the estimates for blacks and whites separately and find similar results for both groups.

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
Labor intensity 1900	-0.108***	-0.228***	-0.545***	-0.908**
	(0.038)	(0.088)	(0.129)	(0.423)
F-stat		69.96		36.11
R-squared	0.56		0.64	
Observations	1215	1215	1517	1517
	Panel 2: Heterogeneity by Race			
	Whites		Blacks	
	OLS	IV	OLS	IV
Labor intensity 1900	-0.186***	-0.247***	-0.102**	-0.261***
	(0.034)	(0.064)	(0.041)	(0.100)
F-stat		80.70		29.69
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.

5 Effects of Shocks to Labor Intensity

The cross-sectional results above provide evidence that agricultural labor intensity influenced culture, but this relationship might be driven by unobserved county characteristics. We are able to address this concern by analyzing the link between *changes* in labor intensity and *changes* in individualism. In particular, 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 sweeping process of mechanization in agricultural production entailed heterogeneous changes in 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 and the spatial and temporal patterns of the spread of infestations 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.

5.1 Mechanization

In the late 19th century, American agriculture experienced sweeping technological change and rapid productivity growth (Kendrick et al., 1961; Evenson, 1978; Olmstead et al., 2008). Part of this process involved mechanization and declines in labor inputs per acre, with differences across crops.

Table 4: Required man-hours per acre for producing wheat by operations

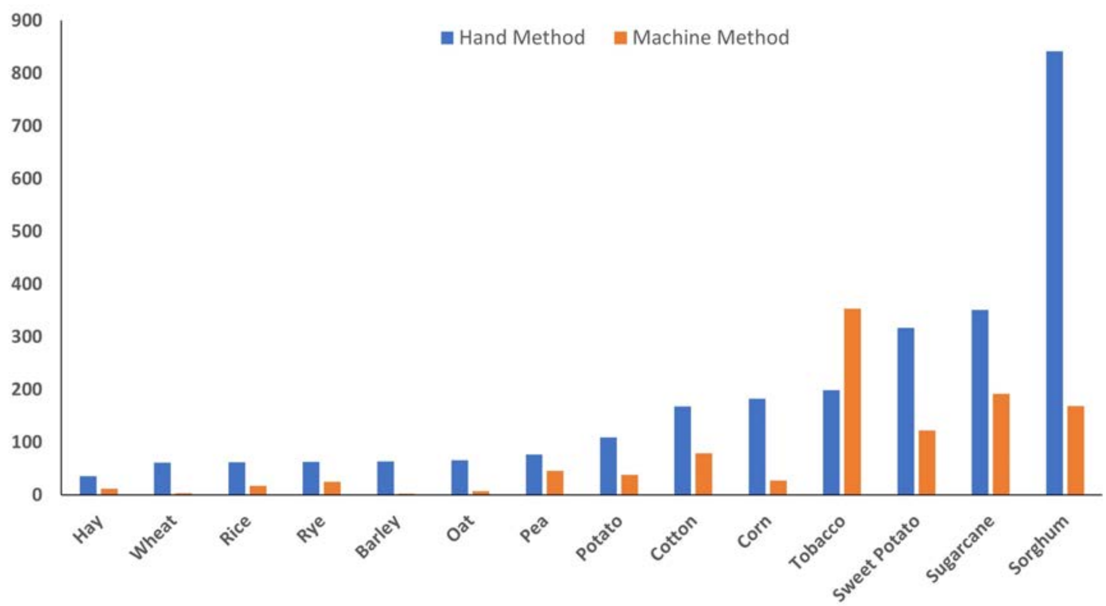
Hand-Method		Machine-Method	
Operations	Man-Hours	Operations	Man-Hours
Breaking ground	6h 40m	Breaking ground	1h
Sowing seed	1h 15m	Sowing seed	15m
Pulverizing	2h 30m	Pulverizing	12m
Reaping, binding, and shocking	20h	Reaping, thrashing, and sacking	1h
Hauling sheaves to barn	4h	Hauling to granary	52m
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.

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 and after mechanization, based on data from representative farms. To illustrate, Table 4 shows the labor requirements for wheat under the "hand method" and "machine method"—the terms used by the report to refer to pre- and post-mechanization practices.[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 7 displays the labor requirements for the 12 crops included in the report under the pre- and post-mechanization methods.[9]

Figure 7: Labor requirements before and after mechanization



Note: 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).

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.

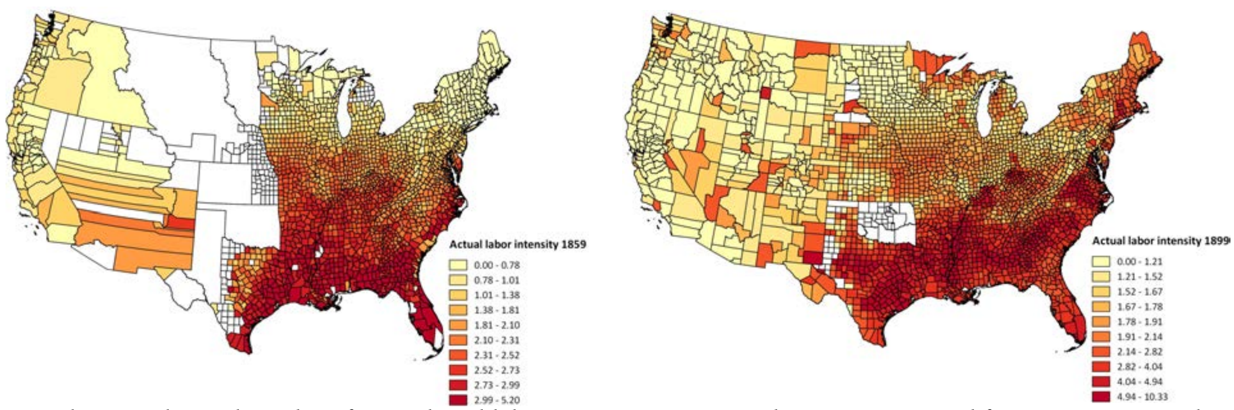
[8]While the report used the terms throughout (even in its title), the use of hands and machines was not exclusive to each method. The hand method was defined as "*the primitive method of production which was in vogue before the general use of automatic or power machines*," while the machine method corresponds 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). It is also worth recalling that the main results hold if we use a measure of labor intensity based only on data from the report.

Figure 8 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.

Figure 8: Agricultural labor intensity in 1860 and 1900

Labor intensity 1860, Hand-Method **Labor intensity 1900, Machine-Method**



Note: 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.

We use the two measures of labor intensity for each county c at different points in time: $\text{Intensity}_{c,1860}$ is computed with the hand labor requirements and $\text{Intensity}_{c,1900}$ with the machine labor requirements.[10] [11] To Our estimating equation is

$$y_{c,t+10} = \alpha + \beta \text{Intensity}_{c,t} + \gamma X_{c,t} + \theta_c + \theta_{st} + \epsilon_{c,t}, \tag{5}$$

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.[12] $X_{c,t}$ includes additional time-varying controls that could be associated with agricultural mechanization. Most relevant here is that we are able to control for county fixed effects (θ_c) that absorb unobserved county-level time-invariant characteristics. θ_{st} denotes state-year fixed effects. The sample consists of counties in which all variables are available in both periods.

The results are displayed in Table 5. Changes in labor intensity are strongly (negatively) associated with changes in the prevalence of individualism. According to the OLS estimate in

[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 E illustrates more details about the harmonization process with a different example.

[11]$\text{Intensity}_{c,1860}$ is harmonized to the 1900 county boundaries.

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

Column (1), a one standard deviation reduction in labor intensity through mechanization was associated with a 0.28 standard deviation increase in the share of infrequent names. Column (2) indicates that the estimate is not confounded by alternative county characteristics that are potentially correlated with agricultural mechanization. In particular, the robustness to the value of farm machinery suggests that the negative correlation is not a direct outcome of mechanization itself but captures a consequence through changes in labor intensity.

Table 5: Mechanization, labor intensity and changes in individualism

Dep. var: Infrequent name share				
	(1)	(2)	(3)	(4)
	OLS		IV	
Labor intensity	-0.228***	-0.213***	-0.357***	-0.344***
	(0.038)	(0.037)	(0.105)	(0.104)
F-stat			80.85	75.86
R-squared	0.05	0.07		
Observations	3586	3532	3586	3532
Additional time-varying controls				
Farm productivity		√		√
Average farm size		√		√
Value of farm machinery		√		√
County fixed effects	Y	Y	Y	Y
State-year 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 share of non top-10 names is computed for counties whose number of relevant observations (individuals) is no less than 30.

Using potential labor intensity as an IV, Columns (3) and (4) 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 production patterns 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 additional time-varying controls, and the effect sizes are slightly larger than the OLS estimates.

5.2 The Boll Weevil Shocks

We also assess 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. Therefore, we 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 in this section 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 551 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 4. We include state-year fixed effects, $\theta_{s,t}$, which absorb any state-level responses to the boll weevil, and county fixed effects, θ_c, which control for time-invariant county characteristics. Finally, $\boldsymbol{X}_{c,t}$ is a vector of time-varying controls that are correlated with the extent of cotton farming and can also influence local individualism.

The vector of time-varying controls includes two demographic characteristics. First, we control for the share of blacks in the population in each period. Blacks represented large shares of cotton tenant farmers (Aiken, 1998), and insofar as the boll weevil induced differential out-migration by race (see Clay et al., 2020), it could have affected cultural attitudes by changing the racial composition. In addition, we control for average literacy rate. Given that cotton tenant farmers had lower levels of education than average farm laborers (Jung, 2020), the boll weevil shock may have led to changes in the average levels of education, which may in turn be associated with changes in individualism.

Table 6: The boll-weevil shock and changes in individualism

Dep. var: Infrequent name share				
	(1)	(2)	(3)	(4)
Boll weevil	0.118**	0.097**	0.114**	0.101**
	(0.048)	(0.047)	(0.043)	(0.044)
Share of blacks		0.252***		0.170*
		(0.083)		(0.096)
Average literacy			-0.149***	-0.129**
			(0.046)	(0.051)
R-squared	0.48	0.49	0.49	0.49
Observations	1653	1653	1653	1653
County fixed effects	Y	Y	Y	Y
State-year fixed effects	Y	Y	Y	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 6 shows that the arrival of the boll weevil increased the share of infrequent names by 0.12 standard deviations. The results are consistent with the prediction that shifts away from cotton, presumably toward less labor-intensive crops, would have increased the prevalence of individualism.

The logic of the previous results suggests that the effects of the boll weevil shocks would be larger in locations where farmers shifted more to less labor intensive crops. We conduct a second exercise that examines this prediction about the heterogeneous effects of the negative shocks on cotton depending on the labor-intensity of the other local crops. We estimate the following equation:

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

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

This exercise also mitigates the possible concern that the shifts away from cotton induced by boll weevil shocks may have affected individualism through channels other than reduced labor intensity. In particular, in the postbellum period, large cotton plantations were divided into small tenancies operated by poor farm families (Alston and Higgs, 1982; Ransom and Sutch, 2001). If the characteristics of tenant farming were correlated with local culture, the estimates of Equation 6 may not capture the role of labor intensity, but the interaction term in Equation would.

The results in Table 7 indicate significant positive effects of boll weevil shocks 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 0.06 standard deviation increase in individualism. The coefficients are robust to controlling for the share of blacks and average literacy.

Table 7: Heterogeneous effects of boll-weevil shocks depending on crop mix

Dep. var: Infrequent name share				
	(1)	(2)	(3)	(4)
Boll weevil	0.107**	0.086*	0.103**	0.089*
	(0.049)	(0.049)	(0.045)	(0.046)
Boll weevil x Less intensive crops	0.057**	0.057**	0.059**	0.059**
	(0.026)	(0.026)	(0.025)	(0.026)
Less intensive crops	-0.020	-0.019	-0.026	-0.024
	(0.028)	(0.028)	(0.029)	(0.029)
Share of blacks		0.254***		0.170*
		(0.085)		(0.097)
Average literacy			-0.151***	-0.130**
			(0.045)	(0.051)
R-squared	0.48	0.49	0.49	0.49
Observations	1653	1653	1653	1653
County fixed effects	Y	Y	Y	Y
State-year fixed effects	Y	Y	Y	Y

Notes: Except the dummy variable of the 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.

6 Long-Term Effects of Historical Labor Intensity

Even if short-run cultural changes result from heterogeneous temporal shocks, the extent of individualism rooted in fundamental factors may remain persistent. Having established that changes in agricultural labor intensity had significant effects on individualism over relatively short periods of time, we now turn to whether the patterns established around 1900 persisted. This section shows that agricultural labor intensity in 1900 had long-run effects on culture that remain significant in the 21th century. Taken together, our findings in this section and the previous one suggest that cultural change and persistence are both important aspects of the same process.

Measuring contemporary individualism requires alternative proxies since Census data on names are not available beyond 1940. We rely on two approaches. First, we use Google Trends data on online search interest across metropolitan areas and consider the relative search frequency of (i) individualistic vs communitarian terms and (ii) individual sports vs team sports. Second, we construct county-level measures that capture individualism less directly, but offer a useful complement with finer geographic coverage and broader implications about social preferences.

Our proxies for individualism based on Google Trends data are inspired by studies in social psychology. First, we consider the relative search volume of the term "unique" compared to "common" from 2004 (when the data become available) to the present. We choose the term "unique" as a typical example of an individualistic word (see, e.g. Twenge et al., 2012; Greenfield, 2013), and "common" as an antonym. Measuring relative interest in these terms provides a direct proxy for individualistic orientation, although naturally there could be confounders, for example, other possible connotations or reasons for searching a term.

A second proxy based on Google Trends data is the relative search interest in team sports. Several studies find a positive relationship between team sports and collectivist culture in the context of team performance (Hartenian, 2003; Gundlach et al., 2006) or athletes' satisfaction (McCutcheon and Ashe, 1999), and House et al. (2004) further argue that a preference for team sports can be a good proxy for collectivist values. In this view, we consider the relative search volume for the three popular team sports (football, baseball, and basketball) compared to the six most popular individual sports (tennis, golf, boxing, skating, swimming, and wrestling).[13]

Google search interest from 2004 to the present is available at the level of Designated Metro Areas (DMA). This gives a cross section of units larger than counties, and requires harmonizing the county-level data with the DMA boundaries. Figure A2 shows the map of the harmonized actual labor intensity in 1900. Relative search volumes take values between 0 and 100, indicating proportions relative to the highest search queries among the entered terms for the designated period and regions.

Columns (1) and (2) of Table 8 reveal that people in locations with higher historical labor intensity have lower search interests in "unique" relative to "common." A one standard deviation increase in labor intensity is associated with a 0.29-0.46 standard deviation decrease in interest in a salient individualistic term relative to a collectivistic one. Columns (3) and (4) show a similar pattern in which metro regions with higher labor intensity search team sports more heavily. Depending on the specifications, a one standard deviation increase in labor intensity is associated

[13] Six individual sports are chosen because of the dominant search volume for football, basket ball, and baseball, but using three individual sports does not alter the results. The estimates are also robust to using the principal components of team sports and individual sports as outcome variables, or each sport separately.

with a 0.21-0.48 standard deviation increase in the relative search volume of team sports.[14]

Table 8: Labor intensity and online search interest

Dep. var: Relative search interest from Google Trends				
	(1)	(2)	(3)	(4)
	"Unique" vs. "Common"		Team sports vs. Individual sports	
	OLS	IV	OLS	IV
Labor intensity 1900	-0.289**	-0.462***	0.219**	0.476***
	(0.145)	(0.170)	(0.086)	(0.129)
F-stat		158.06		154.39
R-squared	0.47		0.71	
Observations	200	200	199	199
Agricultural land suitability	Y	Y	Y	Y
Geo-climatic controls	Y	Y	Y	Y
Region fixed effects	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. 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.

The persistent relationship is also observed at the county-level. We consider three measures of social preferences related to individualism. The first is voter turnout, a common measure of civic culture. Since voting is an engagement in public affairs, higher turnout can be interpreted as partly driven by greater emphasis on collective values. Along these lines, Putnam et al. (2000) show that turnout is strongly correlated with group membership, civic engagement, and informal sociability. From a similar but broader perspective, we also consider social capital, as proxied by an index from Rupasingha et al. (2006). The index comprises measures of election turnout as well as associational density, Census response rates, and the number of non-profit organizations. Thus, while social capital is conceptually not a direct indicator of collectivism, the index value may be interpreted as an informative proxy for the extent of social engagement. Finally, we consider a measure of preferences in support for welfare spending, using data from the Congressional Cooperative

[14] While sports events have the advantage of being comparable within a consistent category, socioeconomic characteristics related to a sports event could still be a potential concern. In particular, the individual sports considered in Table 8 include the so called country club sports such as tennis and golf. Given that country club sports are in general enjoyed by high-income groups, greater search interest in individual sports might be biased due to regional economic conditions. In this regard, we test the robustness of the results when dropping tennis and golf from the group of individual sports.

Election Study (CCES).[15] This proxy is closely associated with "horizontal collectivism," a notion proposed by Singelis et al. (1995). In addition to identifying the self as a part of the collective, horizontal collectivism sees all members as the same and stresses equality within the group, in contrast to vertical collectivism, which is more tolerant of inequalities.

Table 9: Agricultural labor intensity, social attitudes, and individualism in the long-run

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var:	Avg. turnout 2008-2020		Social capital index		Preferences for welfare spending	
	OLS	IV	OLS	IV	OLS	IV
Labor intensity 1900	0.080**	0.275***	0.080**	0.234***	0.110**	0.374***
	(0.037)	(0.103)	(0.034)	(0.085)	(0.049)	(0.116)
F-stat		83.14		83.08		57.08
R-squared	0.43		0.51		0.06	
Observations	2766	2766	2765	2765	2055	2055
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 shows standardized coefficients with robust standard errors clustered at 60mi×60mi grid cells in parentheses. The survey-based outcome in Columns (5) and (6) excludes counties whose number of relevant observations is less than 5. 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 9 shows that higher labor intensity in 1900 predicts stronger civic culture and preferences for redistribution in the long-run. First we examine the relationship between voting behavior and labor intensity. Columns (1) and (2) demonstrate a positive association between agricultural labor intensity in 1900 and average turnout in the last four presidential election. According to the IV specification, counties with a one standard deviation higher labor intensity display a 0.28 standard deviation higher turnout.[16] Similarly, Columns (3) and (4) reveal the positive effects of labor intensity on social capital. Based upon the constitution of the index, the results are interpreted as indicating that historical labor intensity has promoted social participation and activities in the long-run.

In addition, higher collectivism in high labor intensity counties appears tied to the horizontal dimension emphasizing equality within the community. Columns (5) and (6) show that a higher

[15] CCES is a series of nationally representative surveys with different cross-sectional samples for each year. We pool the five waves in 2010, 2012, 2014, 2016, and 2018 to construct the county-level measure. After translating the relevant CCES question into a dummy variable, we take the county-level average for all waves.

[16] The number of total votes is from MIT Election Data and Science Lab, and the county-level turnout is calculated using the voting age population from Ruggles et al. (2021).

labor intensity induced greater support for increasing welfare spending. A one standard deviation increase in labor intensity is associated with a 0.11-0.37 standard deviation increase in the proportion of the population agreeing to expand welfare spending. Considering that support for redistributive policies can be an indicator of the horizontal orientation of culture (Singelis and Triandis, 1998; Triandis, 2018), the results constitute suggestive evidence for the conjecture that high labor intensity leads to greater horizontal collectivism.

7 Conclusion

The distinct labor needs of crops create geographical variations in agricultural labor intensity. Prior work has argued that labor intensity is crucial in the formation of individualistic versus collectivistic cultures. We establish this relationship in the context of U.S. counties around 1900. In particular, we find that lower labor intensity in agricultural production is associated with a more individualistic culture, as measured by the share of unique names among children. These cross-sectional results are robust to including a wide range of controls. Moreover, we confirm this association not only in the cross-section, but also in terms of changes.

Our analysis of cultural changes in response to shocks is a key aspect of this contribution. The cross-sectional results provide supportive evidence for the association between labor intensity and culture, but the patterns could be driven by unobserved county characteristics. This is a common concern for studies on deeply-rooted factors in comparative development that rely on cross-sectional variation across geographical units of observation, as regressors of interest may be correlated with time-invariant features displaying spatial autocorrelation. Exploiting plausibly exogenous time variation allows us to directly address this concern.

We use two sources of time variation in agricultural labor intensity. First, we use Census sources to construct measures of changes in labor intensity due to mechanization. Second, we study the shocks of the boll weevil in the U.S. South, which induced changes in the types of crops grown and hence to the labor intensity of agriculture. In both situations, we find that decreases in labor intensity are associated with increases in individualism, which is consistent with the cross-sectional results and show that they are not driven by unobserved fixed factors at the county level.

These results further confirm that early agricultural characteristics were key in forming cultural attitudes, and given the importance of cultural attitudes for overall development, were key in long-run development. While this finding suggests that historical differences in agricultural conditions could explain some persistent differences in living standards, our time-varying results show that these cultural attitudes are malleable. Agricultural characteristics may be powerful in forming culture, but they are not an unchanging destiny. Nevertheless, we are able to show that contemporaneous outcomes such as social attitudes toward redistribution and online search patterns remain strongly associated with agricultural labor intensity a century earlier. These long-run results are consistent with recent findings that cultural attitudes shaped by agriculture have significant long-run effects.

The mechanism linking agricultural labor intensity and individualism is not something tha can be precisely identified. Given the context of the United States, our results are useful in establishing that the mechanism is not driven by the comparison of paddy rice production to other dry cereal crops but that the relationship is more general. Labor intensity may tell us about the need for labor coordination among workers, or it may be informative about the relative contribution of each individual worker to overall output and thus their view of one another as peers, but at the moment, such observations are speculative. Given the strong relationship we find between labor intensity and individualism, further exploration of its mechanism is warranted.

References

Acemoglu, D. and Robinson, J. A. (2021). Culture, institutions and social equilibria: A framework. Technical report, National Bureau of Economic Research.

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.

Arrington, L. J. and May, D. (1975). ”a different mode of life”: Irrigation and society in ninettenth-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. (2018). Frontier Culture: The Roots and Persistence of Rugged Individualism in the United States. Boston University - Department of Economics - Working Papers Series dp-302, Boston University - Department of Economics.

Beck-Knudsen, A. S. (2019). Those who stayed: Individualism, self-selection and cultural change during the age of mass migration. *Unpublished Manuscript*.

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.

Bidwell, P. W. and Falconer, J. I. (1925). *History of Agriculture in the Northern United States, 1620-1860*. Carnegie Institute, Washington, D.C.

Bleakley, H. and Lin, J. (2012). Portage and path dependence *. *The Quarterly Journal of Economics*, 127(2):587–644.

Buggle, J. and Durante, R. (2017). Climate Risk, Cooperation, and the Co-Evolution of Culture and Institutions. CEPR Discussion Papers 12380, C.E.P.R. Discussion Papers.

Buggle, J. C. (2017). Irrigation, Collectivism and Long-Run Technological Divergence. Cahiers de Recherches Economiques du Département d'Econométrie et d'Economie politique (DEEP) 17.06, Université de Lausanne, Faculté des HEC, DEEP.

Chay, K. and Munshi, K. (2012). Black networks after emancipation: evidence from reconstruction and the great migration. *Unpublished working paper*.

Clay, K., Schmick, E. J., and Troesken, W. (2020). The boll weevil's impact on racial income gaps in the early twentieth century. Technical report, National Bureau of Economic Research.

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.

Doepke, M. and Zilibotti, F. (2014). Culture, entrepreneurship, and growth. In *Handbook of economic growth*, volume 2, pages 1–48. Elsevier.

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.

Fiszbein, M. (2017). Agricultural diversity, structural change, and long-run development: Evidence from the u.s. Technical Report 23183, National Bureau of Economic Research.

Fouka, V. and Schlaepfer, A. (2015). Agricultural labor intensity and the origins of work ehtics. Working paper.

Galor, O. and Özak, Ö. (2016). The Agricultural Origins of Time Preference. *American Economic Review*, 106(10):3064–3103.

Giuliano, P. and Nunn, N. (2017). Understanding cultural persistence and change. Technical report, National Bureau of Economic Research.

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.

Gundlach, M., Zivnuska, S., and Stoner, J. (2006). Understanding the relationship between individualism–collectivism and team performance through an integration of social identity theory and the social relations model. *Human relations*, 59(12):1603–1632.

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).

Hartenian, L. S. (2003). Team member acquisition of team knowledge, skills, and abilities. *Team Performance Management: An International Journal*.

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. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior & Organization*, 53(1):3–35.

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. (2018). The origins of cultural divergence: Evidence from a developing country. *Unpublished Manuscript*.

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). *IIASA, Laxenburg, Austria and FAO, Rome, Italy*.

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

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. et al. (1961). Productivity trends in the united states. *Productivity trends in the United States.*

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.

McCutcheon, L. E. and Ashe, D. (1999). Can individualists find satisfaction participating in interactive team sports? *Journal of Sport Behavior*, 22(4).

McFadden, D. (1974). *Conditional logit analysis of qualitative choice behavior*, 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.

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. (2001). Reshaping The Landscape: The Impact And Diffusion Of The Tractor In American Agriculture, 1910-1960. *The Journal of Economic History*, 61(3):663–698.

Olmstead, A. L., Rhode, P. W., et al. (2008). Creating abundance. *Cambridge Books*.

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

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.

Putnam, R. D. et al. (2000). *Bowling alone: The collapse and revival of American community*. Simon and schuster.

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

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.

Robinson, S. (1863). *Facts for Farmers, also for the Family Circle*. Johnson and Ward, New York.

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

Ruggles, S. et al. (2020). Ipums usa: Version 10.0, doi: 10.18128/d010. *V10. 0*.

Ruggles, S., Kugler, T., Schroeder, J., and Manson, S. (2021). National historical geographic information system: Version 16.0. *Minneapolis: University of Minesota*.

Rupasingha, A., Goetz, S. J., and Freshwater, D. (2006). The production of social capital in us counties. *The journal of socio-economics*, 35(1):83–101.

Singelis, T. M. and Triandis, H. C. (1998). Training to recognize individual differences in collectivism and individualism within culture. *International Journal of Intercultural Relations*, 22(1):35–47.

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. Socio-Ecological Psychology.

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*.

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.

Triandis, H. C. (2018). *Individualism and collectivism*. Routledge.

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 (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.

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

Wright, G. (2003). Slavery and american agricultural history. *Agricultural History*, 77(4):527–552.

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

Appendices

A Data Appendix

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 A1 shows the the top-10 most frequent names of white and black children at the national level in 1910 and 1940.

Table A1: 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	Marry
2	John	Ruth	James	Annie	John	Betty	John	Dorothy
3	James	Helen	Willie	Mattie	William	Dorothy	Robert	Annie
4	George	Midlred	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). The main specification also controls for quadratic terms to reflect non-linear effects.

Terrain elevation: County-level average of the median elevation at 0.5 arc-min resolution from IIASA and FAO (2012). The main specification also controls for quadratic term to reflect non-linear effects.

Terrain slope: County-level average of the terrain slope index at 0.5 arc-min resolution from IIASA and FAO (2012). The main specification also controls for quadratic term to reflect non-linear effects.

Latitude/Longitude: Latitudinal/Longitudinal distance from the equator, calculated from the centroid of each county using shapefiles from Ruggles et al. (2020). The main specification also controls for quadratic terms to reflect non-linear effects.

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: As shown in Section 6, we choose the pair of "Unique" and "Common," which 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. By construction, there is a perfect negative correlation between the RSVs of "Unique" and "Common".

Relative search interests in sports: Table 8 uses the RSVs of 9 sports events. Since Football shows the highest search queries, the RSVs of the others are normalized to that of Football by construction.

Other Long-Run Outcomes

Turnout in the presidential elections: County-level total votes are from MIT Data Science Data Lab, and turnout rates are computed using the voting age population from Ruggles et al. (2020). For each election year, the 5-year American Community Survey estimates in the corresponding period are used.

Social Capital Index: The county-level social capital index is available from Rupasingha et al. (2006) for 2009 and 2014. We use the latter for our estimations.

Support for welfare spending: In 2014, 2016, and 2018, the CCES asked whether respondents prefer their legislature to increase or decrease welfare spending. Based on the questions, we create a dummy variable that equals 1 if the answer is "Greatly increase." The estimation results hardly change even if we also treat "Slightly increase" as 1.

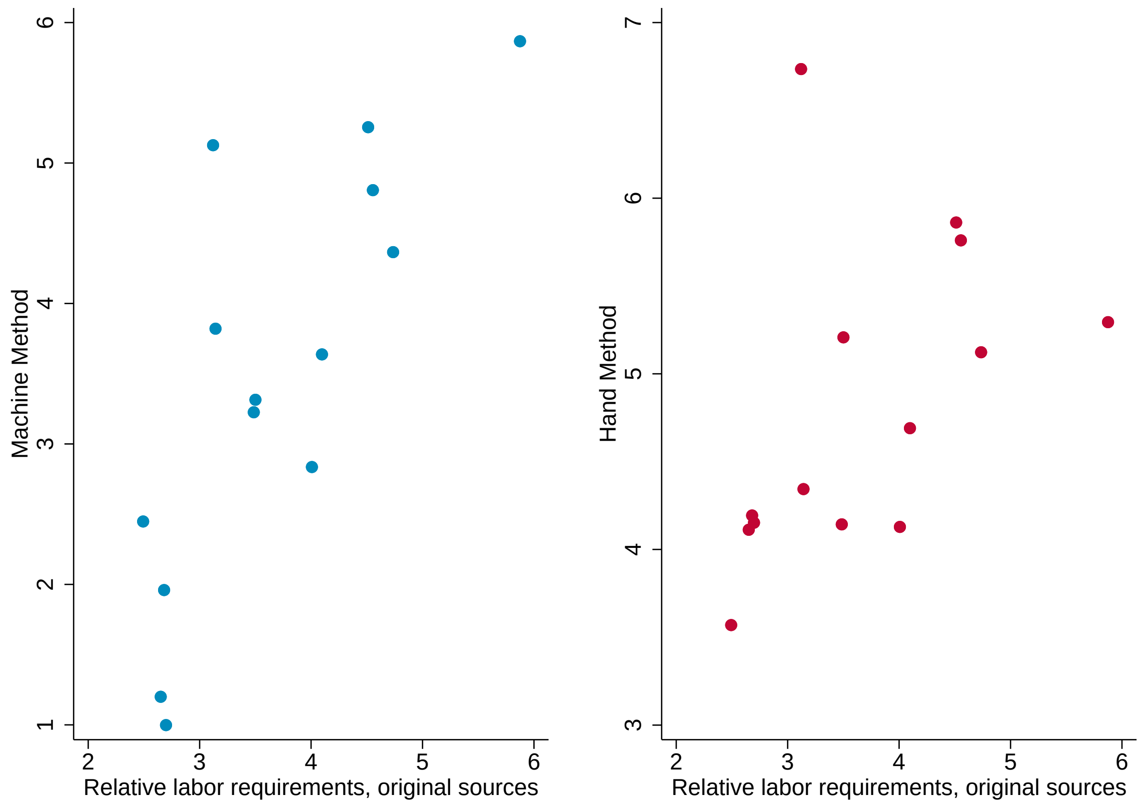
B Agricultural Labor Requirements

Table A2: 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) 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	841.5	168.5
Soybean	-	23.2	-	-	23.2	-	-
Sugar beet	-	128.0	-	-	128.0	-	-
Sugarcane	-	-	-	91.1	91.1	-	-
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	14.6	61.1	3.32

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

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



Note: 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.

C 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

$$Y_i = A_i X_i^{\alpha_i} L_i^{1-\alpha_i} \tag{8}$$

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

$$\alpha_a \frac{p_a Y_a}{X_a} = \alpha_b \frac{p_b Y_b}{X_b} \tag{9}$$

and

$$(1-\alpha_a)\frac{p_a Y_a}{L_a} = (1-\alpha_b)\frac{p_b Y_b}{L_b}. \tag{10}$$

Write the labor market condition as

$$(1-\alpha_a)\frac{p_a Y_a}{X_a}\frac{X_a}{L_a} = (1-\alpha_b)\frac{p_b Y_b}{X_b}\frac{X_b}{L_b}, \tag{11}$$

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

$$\frac{L_a/X_a}{L_b/X_b} = \frac{(1-\alpha_a)/\alpha_a}{(1-\alpha_b)/\alpha_b}. \tag{12}$$

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 Schlaepfer (2015), 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 Schlaepfer (2015) require in their approach.

D Robustness to Different Data Sources

Section 5.1 utilizes Wright (1899) to measure agricultural labor intensity under hand- and machine-based production methods. While the estimation results show the dynamic relationship between agricultural labor intensity and individualism effectively, it should be verified that the labor intensity computed from Wright (1899) also shows consistent patterns cross-sectionally.

Table A3: Agricultural labor intensity and individualism: Robustness to different measures

Dep. var: Infrequent name share 1940						
	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline Index		From Wright (1899)		Baseline & Wright (1899)	
	OLS	IV	OLS	IV	OLS	IV
Actual labor intensity 1899	-0.188***	-0.277***	-0.138***	-0.296***	-0.180***	-0.264***
	(0.034)	(0.064)	(0.032)	(0.070)	(0.033)	(0.058)
F-stat		81.63		71.54		86.37
R-squared	0.60		0.60		0.60	
Observations	2732	2732	2732	2732	2732	2732
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: 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.

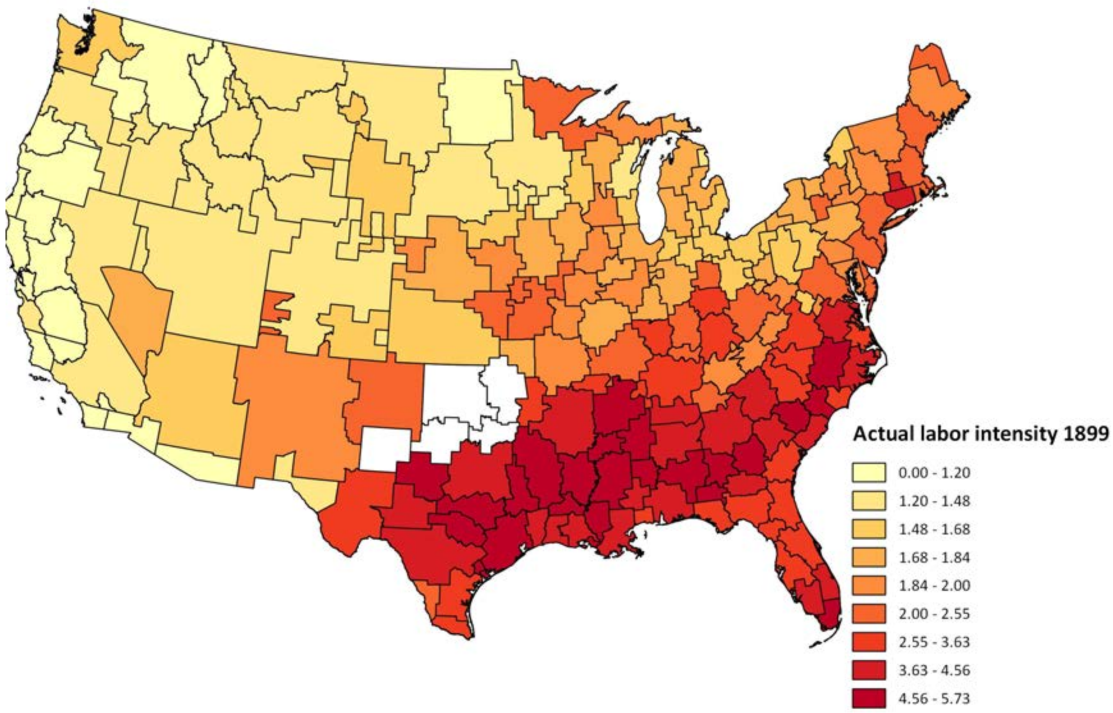
Table A3 documents that the negative effects of labor intensity on individualism are not sensitive to different data sources. Columns (1) and (2) report the relationship between individualism and our baseline index of labor intensity under the OLS and IV specifications. Alternatively, Columns (3) and (4) use the labor intensity computed from Wright (1899) as an explanatory variable. To be consistent with the base period of the original index, we employ the machine-based production method to compute the labor intensity. In all specifications, the cross-sectional relationship between labor intensity and individualism hardly changes. This suggests that despite the scale differences across the data sources, they provide consistent information on the relative labor requirements of crops. In Columns (5) and(6), we compute agricultural labor intensity utilizing all the available data sources, and the results show almost identical patterns.[17]

[17] Using the crops available from Wright (1899), we average the relative labor requirements of the all data sources to

E Agricultural Labor Intensity Harmonized with DMA Boundaries

For the analyses based on Google Trends data in Section 6, 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 A2: Agricultural labor intensity harmonized with the DMA boundaries



Note: 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.

compute the index of labor intensity