DISCUSSION PAPER SERIES

DP17889

HOUSEHOLD FINANCE AT THE ORIGIN: HOME OWNERSHIP AS A CULTURAL HERITAGE FROM AGRICULTURE

Guillaume Vuillemey

ASSET PRICING

CEPR

ISSN 0265-8003

HOUSEHOLD FINANCE AT THE ORIGIN: HOME OWNERSHIP AS A CULTURAL HERITAGE FROM AGRICULTURE

Guillaume Vuillemey

Discussion Paper DP17889
Published 08 February 2023
Submitted 08 February 2023

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

HOUSEHOLD FINANCE AT THE ORIGIN: HOME OWNERSHIP AS A CULTURAL HERITAGE FROM AGRICULTURE

Abstract

I show that home ownership decisions across countries and individuals are shaped by a cultural heritage from agriculture. For centuries, dominant assets in pre-industrial economies were either land or cattle. Consequently, the type of farming prevailing locally shaped preferences and believes about the relative value of immovable and movable assets. This cultural heritage had long-lasting consequences. Today, individuals originating from societies with a history of crop agriculture -- where the dominant asset was land -- are more likely to be homeowners. For identification, I rely both on home ownership decisions of second-generation immigrants in the US and on instrumental variables.

JEL Classification: G51

Keywords: Homeownership, Culture, Persistence, agriculture

Guillaume Vuillemey - vuillemey@hec.fr

HEC Paris School of Management and CEPR

Household Finance at the Origin: Home Ownership as a Cultural Heritage from Agriculture

Guillaume Vuillemey*

February 8, 2023

Abstract

I show that home ownership decisions across countries and individuals are shaped by a cultural heritage from agriculture. For centuries, dominant assets in pre-industrial economies were either land or cattle. Consequently, the type of farming prevailing locally shaped preferences and believes about the relative value of immovable and movable assets. This cultural heritage had long-lasting consequences. Today, individuals originating from societies with a history of crop agriculture – where the dominant asset was land – are more likely to be homeowners. For identification, I rely both on home ownership decisions of second-generation immigrants in the US and on instrumental variables.

*HEC Paris and CEPR. Email: vuillemey@hec.fr.

1 Introduction

For centuries, pre-industrial societies were dominated by two main types of assets: land or cattle. Both were thought of respectively as leading types of immovable and movable assets. Both produced cultures, myths, representations, and systems of political power based on them. There were "land-based" societies (such as feudal Europe) and "cattle-based" societies (such as many East African societies), depending on the type of farming that was locally dominant.

In this paper, I provide causal evidence of a *cultural heritage from agriculture*. My main hypothesis is that individuals from societies traditionally dominated by crop agriculture – in which the main asset was land and immovable assets were thus relatively better considered – are still inclined to desire more immovable assets, and are thus more likely to be home owners. I confirm that this cultural heritage can explain an economically meaningful part of the persisting variation in home ownership rates across both countries and individuals.

To justify this hypothesis, I start by providing ethnographic evidence that societies in which the dominant form of farming was based either on land or on cattle often produced a whole culture around these assets. Theoretically, this cultural background can be seen as providing individuals with preferences and believes about the relative value or "safety" of immovable or movable assets. Theory also provides rationales why cultural backgrounds can display high persistence from one generation to the next, even in the face of a changing environment (Bisin and Verdier, 2000, 2001).

Empirically, I start by showing that cross-country home ownership patterns in the OECD – the largest dataset in which measurement is similar across countries – are consistent with my hypothesis. I measure the historical prevalence of crop agriculture for each country with data on land use: I construct a continuous variable, called *CropShare*, which measures the relative importance of cropland and pasture areas on average over the period from 1780 to 2010. In the OECD sample, I find an economically large effect: a one standard-deviation increase in *CropShare* is associated with an increase in home ownership rate of about 6 percentage points, which amounts to one half of the cross-country standard deviation (equal to 0.120).

While consistent with my hypothesis, these results cannot be given any causal interpretation. The main concern is that countries are heterogeneous along many dimensions, notably institutions in a broad sense: beyond a cultural heritage, there could be a number of institutions that could have led simultaneously to a history of crop agriculture in a country and to more home ownership today. In sum, the first main identification challenge is to isolate "culture" from "institutions".

To address it, I use the so-called "epidemiological approach" pioneered by Fernandez and Fogli (2009) (see Fernandez, 2011, for a survey). The idea is to "fix" institutions – that is, to focus on within-country variation – and to study decisions by individuals with heterogeneous cultural backgrounds. Specifically, I study the home ownership decisions of second-generation immigrants in the US, that is, individuals born in the US but whose both parents are born abroad. As opposed to first-generation immigrants, who have been directly exposed to the institutions of their country of origin, second-generation immigrants have only been indirectly exposed to this country via cultural transmission. I use data from the March Supplement of the Current Population Survey, which is currently the only dataset in the US in which respondents are asked about the country of birth of their parents.

Across a variety of specifications, I confirm that individuals whose parents migrated from a country historically dominated by crop agriculture – in which culture built around immovable assets – are significantly more likely to be home owners. This holds after including several fixed effects and controls for standard determinants of home ownership.

While these results causally identify a persistent effect of the country of origin's culture on home ownership decisions, it remains partially unclear whether this cultural effect is indeed an heritage from agriculture, or reflects some other cultural traits that could be correlated with the prevalence of crop agriculture across countries.

I address this concern using instrumental variables (IV). Specifically, an satisfactory instrument in this context would shift the likelihood of adopting farming practices based either on land or cattle, but would be unrelated to other cultural traits. As instruments satisfying such conditions, I use two biochemical properties of soils which are conducive to fertility. According to the Food and Agriculture Organization (FAO), the two properties that are most relevant for soil fertility are the percentage of subsoil organic carbon and subsoil pH (a measure of acidity). I thus use detailed raster data on global soils from the Harmonized World Soil Database to construct country-level measures of organic carbon and pH. I use these variables to instrument *CropShare*, that is, the prevalence of crop agriculture relative to cattle grazing. Regressions are estimated using two-stage least squares.

The IV regression results confirm my hypothesis: an economically meaningful part of the cross-individual variation in home ownership can be causally interpreted as a cultural heritage from agriculture. In plain terms, households attach greater value to housing if they grew up in a society that traditionally attached greater cultural, political, or even religious value to land and immovable assets, as opposed to cattle and movable assets. This finding contributes to explain the considerable cross-country variation in household portfolios, which has often puzzled economists so far (Badarinza et al., 2016).

Related literature

This paper relates to two main strands of the literature. First, a large number of papers have demonstrated the impact of culture on economic outcomes (Guiso et al., 2006; Algan and Cahuc, 2010). A common theme is that culture shapes individuals' believes and preferences and can be extremely persistent, including over centuries (Voigtländer and Voth, 2012) and when institutions change. For example, Alesina et al. (2013) show that culturally transmitted gender norms persist after individuals with heterogeneous backgrounds migrate. Giuliano (2007) shows that culture determines living arrangements, notably the fraction of young adults living with their parents, while Fernandez and Fogli (2009) shows that it affects women's work and fertility behavior.

Second, there is a large literature in household finance that seeks to understand cross-country or cross-individual variation in portfolio allocations, including home ownership decisions. This literature is largely surveyed by Campbell (2006) and Badarinza et al. (2016). The main cultural factor that has been related to household finance is trust or "social capital". Guiso et al. (2004) show that Italian households living in high-social-capital areas are more likely to use checks, invest less in cash and more in stocks. Relatedly, Guiso et al. (2008) show that less trusting individuals are less likely to buy stocks, while El-Attar and Poschke (2011) show that they are more likely to invest in housing. Other papers pointing to a role of culture on household finance include Huber and Schmidt (2022) and Haliassos et al. (2016). Relative to these papers, I provide evidence in favor of a novel channel through which culture affects households financial decisions. Beyond culture, households' financial and home ownership choices have been shown to depend notably on exposure to certain institutions (Osili and Paulson, 2008), on experiences (Malmendier and Nagel, 2011) or on social interactions (Hong et al., 2004).

2 Historical background and main hypothesis

I start by providing a brief historical background on the cultural legacy of agriculture. Understanding cultural views about the respective valuation and safety of movable and immovable assets across societies helps justifying my main hypothesis on home ownership.

2.1 The cultural legacy of agriculture

Representations according to which there fundamentally exists two types of assets – immovables and movables – are at least several millennia old. For example, Benveniste (2016) shows that, instead of a simple term that would designate total "wealth", Greeks from the Homeric period (1200 to 800 BC) were using distinct terms for movable and immovable wealth.

There is strong reason to think that the first concepts of immovable and movable assets were given substantive meaning based on prevailing farming practices. In particular, the association between immovable assets and cattle can be seen in many examples. Benveniste (2016, Book I, Chap. 3) shows that the Greek terms designating "sheep" and "movable wealth" are derived from the same root. Similarly, a number of terms related to movable wealth across European languages are derived from the Latin "pecus" – such as "pecuniary" in English –, which means "cattle". In common law countries, movable assets are legally called "chattel", a term which derives from the same root as "cattle". Similarly, the term "capital", which was historically used to designate financial wealth, as opposed to real estate, derives from the Latin "capitālis", which meant "head of cattle".

While the linguistics confirms a close association between early concepts of assets and farming practices, the relation goes much beyond. It is often an entire culture that built around either land or cattle. Both of them have been associated with divinities, myths, legal representations and systems of power which often favored one type of wealth at the expense of others. The ethnographic and historical literature on them is enormous and cannot be surveyed here in its entirety. I will simply provide two examples.

European countries, from Ancient Greece, through the Roman era, and until the Middle Ages, have tended to be land-based societies: political power is associated with land holding in Greek cities, in Rome, as well as throughout the feudal society until the French Revolution, while movable wealth was often despised (Ellul, 2013). Writing about the pre-modern society, Dumont (1977, p. 5) notes that "In the traditional type of society, immovable wealth (estates) is sharply distinguished from movable wealth (money, chattels) by the fact that rights in land are enmeshed in the social organization in such a manner that superior rights accompany power over men. Such rights or 'wealth', appearing essentially as a matter of relations between men, are intrinsically superior to movable wealth, which is disparaged, as is natural in such a system for a mere relation between men and things." For example, a common saying in French medieval law is "res mobilis, res vilis" ("movable asset, vulgar asset"). Instead, throughout the feudal period, the same Latin term ("dominium") designates power over land and the power over people, and is positively connoted.

The opposite hierarchy is found in other societies, in which cattle is the most valued asset. This is notably the case in East African societies. In his classical study on the Nuer, Evans-Pritchard (1940) called them a "cattle people". In another famous study, Herskovits (1926) writes about the "cattle complex" in East Africa, that is, the mix of myths, representations and political structures that are based on cattle. He cites evidence that "among the Nuer, wealth is judged entirely by the number of cattle and sheep a man possesses" (p. 257). In contrasts, other tribes that practice agriculture are regarded with contempt. Galaty and Bonte (2020) further note that, in these societies, political power is derived from cattle, not from land: "in full-fledged pastoral aristocracies, cattle are distinctly associated with kingship." They also note that the pastoral specialization of East African societies is very old, as it is already attested in the 3rd millenia BC.

These brief elements confirm that key assets in pre-industrial societies were much more than assets. They were cultural objects surrounded with myths, representations and power structures. The question I seek to answer is whether this cultural background still affects the way societies perceive immovable or movable assets, even after most institutions surrounding land and cattle have stopped to exist.

2.2 Hypothesis

My main hypothesis is based on the idea that societies in which land was the dominant asset – in which political power and representations were also based on land – have endowed individuals with a cultural background leading them to attach a relatively greater value to real estate.

Hypothesis 1. *Individuals from cultures with a history of crop agriculture are more likely to be home owners.*

This hypothesis requires that culture has a persistent component, even in the face of changing infrastructures and economic environment. Evidence for this persistence has been demonstrated in a variety of contexts (Guiso et al., 2004; Voigtländer and Voth, 2012). If persistence is large enough, preferences and believes about the relative value of immovable and movable assets, inherited from an agricultural past, could still shape financial decisions long after agriculture stopped being the main economic sector in many countries.

2.3 Measurement

The formulation of Hypothesis 1 has implications for measurement. First, as understood here, culture is slow-moving. Therefore, it is expected to explain primarily average differences across groups of individuals with heterogeneous backgrounds, as opposed to deviations from the average. For example, Hypothesis 1 is a candidate to explain why home ownership is on average higher in country A than in country B, but not to explain why home ownership rises in a specific year in country A or B. To smooth away the effect of such short-term deviations, I work whenever possible with long-term averages for both the home ownership rate and for the variable measuring the prevalence of crop agriculture.

Second, Hypothesis 1 guides the construction of the main independent variable. To measure the relative predominance of farming based either on land or cattle – that is, on immovable or movable assets – I obtain data on global land use broken down, at the country level, between cropland and pasture. I then define the main independent variable, *CropShare*, as

$$CropShare_c = \frac{\text{Cropland area}_c - \text{Pasture area}_c}{\text{Cropland area}_c + \text{Pasture area}_c}, \tag{1}$$

that is, a number between -1 (if a country c has only pasture) and 1 (if a country c has only cropland). One advantage of this measure is that, by focusing only on the relative shares of cropland and pasture areas, it is not distorted by the fact that some countries have vast pieces of land that are used neither for cropland nor for pasture. To compute the variable *CropShare*, I rely on data from Taylor and Rising (2021), further discussed in Appendix A.2. These data measure the relative share of cropland and pasture over the period from 1780 and 2010, and is thus well-suited to assess the relative importance of land and cattle in farming. While the *CropShare* variable is extremely persistent over time – as illustrated in Figure 2, which plots *CropShare* in 2010 against *CropShare* in 1800 for all sample countries[1] – I nonetheless with a long-term average to ensure that I capture a structural element of each country's farming model.

The cross-country distribution of the variable *CropShare* is plotted in Figure 1 and

[1]The pairwise correlation between *CropShare* in 2010 and *CropShare* in 1800 equals 84.3%, and the one between *CropShare* in 2010 and *CropShare* in 1900 equals 93.0%.

mapped in Figure 4. As can be seen, there is considerable cross-country heterogeneity, as *CropShare* spans the full range of values from close to -1 (countries such as Saudi Arabia or Mongolia) to close to 1 (countries such as India or Myanmar). I now explore the potential of this variable to explain home ownership decisions across countries and individuals.

3 Cross-country evidence

A first attempt to assess the validity of Hypothesis 1 is to study the relation between crop agriculture and home ownership in a cross-section of countries. To do so, a challenge is that of measurement: there is generally no unique method to compute home ownership rates across countries. To my knowledge, the best data so far are those by the OECD, covering 41 countries, which have the benefit of homogeneous measurement across countries.[2] For each country, I compute home ownership as the sum of outright ownership and ownership with a mortgage, and focus on the average rate over the 2010-2020 period. As illustrated in Figure 3, there is considerable heterogeneity in the cross-section: home ownership rates range from close to 40% (42.0% in Colombia or 43.4% in Switzerland) to above 90% (91.0% in Lithuania or 96.3% in Romania). As reported in Panel A of Table 1, there is also large heterogeneity in cropland shares across the subset of OECD countries, since the 10th and the 90th percentiles respectively equal -0.533 and 0.665 (for a variable ranging theoretically from -1 to 1).

To test whether this cross-country heterogeneity is consistent with Hypothesis 1, I

[2]For additional details on the datasets and variables, see Appendix A.

start by estimating

$$OwnRate_c = \alpha + \beta \cdot CropShare_c + FE_{reg} + FE_{inc} + \epsilon_c, \tag{2}$$

where $OwnRate_c$ is the average home ownership rate in country c and $CropShare_c$ is the measure of cropland share defined in Equation (1). FE_{reg} and FE_{inc} are respectively fixed effects at the regional level (as defined by the United Nations, that is, Americas, Asia, Europe and Oceania) and at the income group level (i.e., high income or upper middle income).

The estimation results, together with robust standard errors, are reported in Panel B of Table 1. Columns (1) to (3) are estimated using our baseline measure of *CropShare*, that is, country-level data from Taylor and Rising (2021). The coefficients show a positive association between cropland share and the home ownership rate. Across specifications without and with fixed effects, the estimate is statistically significant at the 5% or 1% levels. In terms of magnitude, it is also economically large: a one standard-deviation increase in *CropShare* is associated with an increase in home ownership rate of about 6 percentage points, which amounts to one half of the cross-country standard deviation (equal to 0.120). Therefore, countries in which crop agriculture has been the dominant form of farming over the past 200 years tend to have significantly higher home ownership rates today – even so agriculture is now a minor economic sector in most OECD countries.[3]

In columns (4) to (6), I reproduce the same regressions, with an alternative measure of *CropShare*, computed from OECD data on land use (see Appendix A.1 for details).

[3] As reported in Panel A of Table 1, for the median OECD country, only 3.4% of workers are in the agricultural sector.

These data have the additional benefit of covering a few extra European countries not included in the global dataset from Taylor and Rising (2021).[4] Once reestimated, the results remain statistically significant, albeit at lower levels (10% or 5%) and with a slightly lower economic magnitude: a one standard-deviation increase in *CropShare* is associated with an increase in home ownership rate of about 4.5 percentage points, which amounts to one third of the cross-country standard deviation.

While consistent with Hypothesis 1, these results cannot be interpreted as causal evidence that a history of crop agriculture led to a culture valuing real estate relatively more than other assets. The main potential confounding factor is coming from institutions (e.g., land regulation, ownership regimes, mortgage design) in a broad sense. This would specifically be a concern if societies that are more likely to adopt certain types of institutions are also more likely to favor crop agriculture over cattle farming.

4 Identification using second-generation immigrants

I turn to the first main element of my identification strategy. Studying the home ownership decisions of second-generation immigrants with heterogeneous cultural backgrounds allows me to isolate the role of culture from that of other confounding factors, notably institutions.

4.1 Identification strategy

To isolate the role of culture, as opposed to institutions, my strategy is to fix institutions and to study variation in financial decisions across individuals with hetero-

[4] As discussed in Appendix A.2, the correlation between the two series is very high (88.8%).

geneous cultural backgrounds. As in Fernandez and Fogli (2009), I rely on the fact that a cultural background is "portable" while institutions remain attached to specific countries. Consequently, the financial decisions of immigrants can be particularly informative about the causal role of culture.

Focusing on the home ownership decisions of first-generation immigrants (that is, individuals born abroad who are now living in a new country) would not alleviate all endogeneity concerns. Indeed, individuals born abroad have been exposed to other institutions, and this exposure itself may have lifetime consequences even after moving (Osili and Paulson, 2008). To overcome this concern, as in Fernandez and Fogli (2009) or Alesina et al. (2013), I focus on home ownership decisions of second-generation immigrants, that is, individuals born in a country, but whose parents were born abroad and moved to this country. Second-generation immigrants have never been directly exposed to the institutions of their parents' country of origin; the only exposure they retain to this country is through the cultural background that may have been transmitted via parents.

Datasets in which one can observe both some household finance finance variables – including home ownership –, the country of birth, and the country of birth of both parents, are rare. In the US, the only dataset with such information is the March Supplement of the US Current Population Survey (also called Annual Social and Economic Supplements), publicly available from the US Census. For my tests, I obtain the latest vintage of the dataset, corresponding to the year 2022. It contains a broad range of economic and social data from 152,732 persons belonging to 89,197 households.

To get the tightest identification possible, I focus on second-generation immigrants whose both father and mother are born outside the US. This leaves me with a sample of

2,720 individuals, whose parents are born in 113 countries. For each of them, I obtain a dummy variable equal to one if they are homeowners, as well as other demographic and economic variables that may affect home ownership. In Table 2, I provide descriptive statistics on these variables (Panel A) and list the 15 most represented countries of origin (Panel B). I am then able to match the variable *CropShare* for 98 countries of origin, representing 2,675 persons.[5]

4.2 Baseline estimation

On the sample of second-generation immigrants, I estimate

$$\begin{aligned} Owner_i =& \alpha + \beta \cdot CropShareParents_i + \gamma \cdot Controls_i \\ &+ FE_{state} + FE_{metropolitan} + FE_{marital} + FE_{educ} + \epsilon_i, \end{aligned} \tag{3}$$

where $Owner_i$ equals one if individual i is a home owner. The main dependent variable, $CropShareParents_i$, is defined as the average value of $CropShare$ for i's mother's and the father's countries of origin. This approach avoids taking a stance on whether culture transmits primarily through the father or the mother, which may differ across cultures.[6]

The approach taken in Equation (3) alleviates two additional concerns. First, I include a number of fixed effects for states, metropolitan areas, marital statuses, and levels of education. These fixed effects mitigate the concern that immigrants from certain origins may be over- or under-represented in certain geographical areas

[5] All these persons are defined as the "reference person" in the CPS data, that is, I exclude other persons in the household, beyond the respondent. Some variables available only at the household level are matched to data at the person level. See Appendix A.3 for details.

[6] In our sample, the country of birth of the father and of the mother is the same for 87.7% of individuals.

or have certain characteristics that are conducive to home ownership. Second, since Equation (3) is estimated at the individual level, I also include a number of personal characteristics that are known to be associated with home ownership. The vector of controls includes the following variables: age, age squared, sex, the number of persons in the household, and the logarithm of the household's income.

Estimates are reported in Table 3, together with robust standard errors. From column (1) to (4), I sequentially make the specification more complete. In column (1), I report an estimate for all individuals whose parents are born outside the US, regardless of whether they are themselves born in the US. In column (2), I further restrict to persons born in the US, that is, second-generation immigrants. Both regressions yield a positive estimate of β, statistically significant at the 1% level. Its magnitude is fairly comparable to the one found using cross-country data in the OECD. In column (3), I add all fixed effects, and in column (4), I add all control variables. This cuts the magnitude of the effect by about one half, but the estimate remains statistically significant at the 10% level. The economic magnitude remains sizable: a one-unit increase in *CropShare* generates a 3.3 percentage points increase in the probability of being homeowner (compared to an average home ownership rate of 55.2% in the sample). These results point to a causal effect of culture on home ownership, driven by historical exposure to agriculture.

4.3 Robustness

I next explore the robustness of the most complete specification – the one in column 4 of Table 3 – to alternatives, and present the results in Table 4. Column (1) addresses the concern that the results could be entirely driven by second-generation immigrants from Mexico, which are over-represented in the sample of countries of origin (see Panel B of Table 2). It re-estimates Equation (3) after excluding individuals whose parents were born in Mexico. If anything, the results are larger in magnitude and more significant (at the 5% level).

In column (2), I address the issue that the results could be coming from individuals working directly in the agricultural sector in the US. Indeed, a concern could exist that individuals coming from societies with an agricultural background are more likely to continue working in agriculture once in the US, and that agriculture requires ownership of real estate. Re-estimating the baseline model after excluding individuals working in agriculture does not affect the results.

In column (3), I use an even tighter fixed effect strategy, by including $State \cdot Metropolitan$ fixed effects, rather than $State$ and $Metropolitan$ fixed effects separately. This allows me to study variation within metropolitan or non-metropolitan areas of any given state. The baseline estimate of β is economically and statistically unchanged.

Finally, in column (4), I cluster standard errors by income groups. Theoretically, clustering should be done along dimensions in which the sampling may not reflect random draws from the entire population (Abadie et al., 2023). In the case of the Current Population Survey, a common concern is that certain income groups (particularly the richest) are over-sampled. Once again, clustering by income groups does not significantly affect the results.

5 Identification using instrumental variables

Next, I turn to the second main element of my identification strategy. By instrumenting the prevalence of crop agriculture at the country level with biochemical properties of the soil, I can shed more precise light on the mechanism behind the results.

5.1 Instrumental variables

At this stage, I have causally identified an effect of culture on home ownership decisions. Nonetheless, one could still question whether this cultural effect is indeed a cultural heritage from agriculture, i.e., views about the relative desirability of movable or immovable assets originating from the fact that, for long periods of time, dominant assets were either land or cattle. An alternative could be that the cultural effect is due to other social or political characteristics that are systematically correlated with crop agriculture across countries – but not originating from crop agriculture itself.

The instrumental variable approach that I now introduce is designed to pin down more precisely the mechanism at play. Specifically, one wants instruments that shift the extent of crop agriculture, but that are unrelated to any other social, cultural or political characteristic in a society.

To this end, I instrument the relative prevalence of crop agriculture with biochemical properties that are conducive to soil fertility. The idea is that crop agriculture requires a certain degree of soil fertility to be viable. In regions where these properties do not exist, the main way for people to obtain food is to rely more on cattle and products derived from cattle (indeed, herbivores eat many plants that humans cannot digest and produce digestible food out of them, including meat, butter or milk).

Biochemical soil properties have attractive properties in our perspective, since they can be treated as exogenous to a large extent: even though human activities can affect the properties of soils, they are extremely slow-moving. For example, it takes about 300 years for just 1 centimeter of soil to form. To further alleviate concerns that soil properties may be partially endogenous, I rely only on measures related to subsoil as opposed to topsoil.

The main issue that soil fertility raises is one of measurement. While biologists have long attempted to come up with a unique metric for soil fertility, this has proved impossible to obtain. Fertility exists only as a combination of several biochemical properties that need to be jointly satisfied. That said, not all properties are as important as others. According to the Food and Agriculture Organization (FAO), "organic carbon is together with pH [a measure of acidity], the best simple indicator of the health status of the soil. Moderate to high amounts of organic carbon are associated with fertile soils with a good structure " (FAO, 2009, p.14). This gives me a strong rationale to use measures of either subsoil organic carbon or subsoil pH suitability as instruments for the prevalence of crop agriculture, i.e., the variable *CropShare*.

To construct these instruments, I proceed in several steps. First, I obtain the Harmonized World Soil Database (HWSD) from the FAO. This dataset is a raster providing soil properties – including subsoil organic carbon and subsoil pH – for every 30 arc-second (approximately 1 km by 1 km) for the entire earth. I then assign a longitude and a latitude to each cell in the raster and map these coordinates to a country. I then compute average soil properties for each country in the world. Finally, I follow the guidelines by the FAO to define variables more precisely. Regarding subsoil organic carbon, I transform the raw percentage into a categorical variable that takes five values, using the thresholds suggested by the FAO (2009, p.14).[7] Regarding subsoil pH, its relation to soil fertility is not linear (as excess of acidity or basicity is detrimental to fertility). The FAO notes that values between 5.5 and 7.2 "are the best pH conditions for nutrient availability and suitable for most crops" (FAO, 2009, p.14). I thus compute a measure of average pH suitability for a country c as the absolute deviation from the center of this interval, that is

$$Suitability = -\left|pH_c - \frac{7.2 + 5.5}{2}\right|, \tag{4}$$

where pH_c is the average subsoil pH in country c. The negative sign ensures that a pH closer to the center of the $[5.5, 7.2]$ interval corresponds to greater suitability.

Figure 5 maps both subsoil organic carbon (Panel A) and subsoil pH (Panel B) at the global level. Beyond large cross-country variation in these two dimensions, visual comparison of these maps with Figure 4 gives preliminary indication that biochemical properties that make soils fertile are also conducive to a relatively greater reliance on crop agriculture over cattle grazing in farming.

5.2 IV estimation

The instrumental variables (IV) approach consists in instrumenting for $CropShareParents_i$ in Equation (3) with the soil properties of individual i's parents' countries of origin.

[7] See Appendix A.4 for details.

That is, I estimate the following first-stage equation

$$\begin{aligned} CropShareParents_i =& \phi + \mu \cdot SoilPropertiesParents_i + \eta \cdot Controls_i \\ &+ FE_{state} + FE_{metropolitan} + FE_{marital} + FE_{educ} + \varepsilon_i, \end{aligned} \tag{5}$$

where $SoilPropertiesParents_i$ is either the average subsoil organic carbon or the average subsoil pH suitability in i's parents' countries of origin. I then use a two-stage least squares estimation to obtain the coefficients of interest.

To assess the relevance of the instruments, I first report estimates from the first-stage regressions, in Panels A of Tables 5 (for subsoil organic carbon) and 6 (for subsoil pH). For both instruments, I report the same set of regressions as for the uninstrumented results. Across specifications and instruments, I find statistically significant estimates for the coefficient μ, at the 1% level, and with the expected signs: better biochemical properties of the subsoil are indeed conducive to a higher share of land devoted to crop agriculture as opposed to cattle grazing. I additionally report F-statistics, which are high in all cases. In the most complete regressions, with all fixed effects and controls, the F-statistics are respectively equal to 10.86 (when the instrument is subsoil organic carbon) and 27.09 (when the instrument is subsoil pH suitability). This is me confidence about the relevance of the instruments.

The results from the second-stage are then reported in Panels B of Tables 5 (for subsoil organic carbon) and 6 (for subsoil pH). While I reproduce all baseline regressions for completeness, I focus the interpretation on the most complete specification, in column (4). With both instruments, I find statistically significant coefficients, at the 1% level, and of relatively similar magnitudes (0.141 vs. 0.120) – which is reassuring. The magnitudes are strikingly similar to those obtained in the cross-country regressions of Table 1 (Panel B, columns 1 to 3), and are larger than those obtained in similar regressions estimated with OLS (Table 2, column 4). This magnitude in large: a one-unit increase in *CropShare* yields a change in the probability to be home owner by 12 to 14 percentage points.

The IV results help us clarify the mechanism and the interpretation of the previous findings. The practice of crop agriculture produced a culture in which land is the dominant asset (as opposed to cattle-based societies, in which the dominant asset is movable). There is ample historical evidence that the value attached to land (resp. cattle) permeated into culture, as briefly summarized in Section 2.1. My estimates show that this cultural background still matters today for household finance, even though many economies are no longer dominated by farming: countries which have been used to attach most value to immovable assets – due to their role in agriculture — are still significantly more likely to have high ownership rates. The same holds true at the individual level, even after households migrate to countries with different institutions.

6 Conclusion

For most households, choosing to own a home is the single most important financial decision they take in their lifetime. When they do so, real estate becomes their dominant asset. In this paper, I use several identification strategies to show that this decision is significantly influenced by a cultural heritage from agriculture. For centuries, dominant assets in pre-industrial economies were either land or cattle, so that preferences and views about the relative value of immovable and movable assets were largely shaped by the prevailing type of farming. In particular, societies dominated by crop agriculture tended to view land as the preferred and "safest" asset. Today, individuals originating from societies with a history of crop agriculture are significantly more likely to own a home, even after they migrate.

These results open the question whether the cultural heritage of agriculture also explains other financial decisions by households, beyond home ownership – such as a differential valuation for movable real assets (e.g., precious metals) or financial assets. Testing these additional hypotheses would require more granular data on portfolio allocation, and is left for future research. Another interesting avenue would be to directly relate the cultural heritage of agriculture to believes about assets. For the US, Adelino et al. (2018) report believes that are puzzling from the vantage point of financial theory: 71% of US households believe that housing is safe, while this percentage is only 55% for bonds. Unfortunately, standard cross-country datasets on believes and values, such as the World Values Survey, currently do not include questions on assets – so that it remains hard to identify the role of culture in these broader believes.

References

Abadie, A., S. Athey, G. W. Imbens, and J. M. Woolridge (2023). When should you adjust standard errors for clustering? *Quarterly Journal of Economics 138*, 1–35.

Adelino, M., A. Schoar, and F. Severino (2018). Perception of house price risk and homeownership. *Working paper*.

Alesina, A., P. Giuliano, and N. Nunn (2013). On the origins of gender roles: Women and the plough. *Quarterly Journal of Economics 128*, 469–530.

Algan, Y. and P. Cahuc (2010). Inherited trust and growth. *American Economic Review 100*, 2060–2092.

Badarinza, C., J. Y. Campbell, and T. Ramadorai (2016). International comparative household finance. *Annual Review of Economics 8*, 111–144.

Benveniste, E. (2016). *Dictionary of Indo-European Concepts and Society.* HAU Books.

Bisin, A. and T. Verdier (2000). “Beyond the melting pot”: Cultural transmission, marriage, and the evolution of ethnic and religious traits. *Quarterly Journal of Economics 115*, 955–988.

Bisin, A. and T. Verdier (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory 97*, 298–319.

Campbell, J. Y. (2006). Household finance. *Journal of Finance 61*, 1553–1604.

Dumont, L. (1977). *From Mandeville to Marx: The Genesis and Triumph of Economic Ideology.* University of Chicago Press.

El-Attar, M. and M. Poschke (2011). Trust and the choice between housing and financial assets: Evidence from Spanish households. *Journal of Finance 15*, 727–756.

Ellul, J. (2013). *Histoire des institutions.* PUF. 3 vol.

Evans-Pritchard, E. E. (1940). *The Nuer: A Description of the Modes of Livelihood and Political Institutions of a Nilotic People.* Clarendon Press.

FAO, . (2009). *Harmonized World Soil Database – Version 1.1.* FAO.

Fernandez, R. (2011). Does culture matter? In J. Benhabib, M. O. Jackson, and A. Bisin (Eds.), *Handbook of Social Economics*, Volume 1A. North Holland.

Fernandez, R. and A. Fogli (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics 1*, 146–177.

Galaty, J. G. and P. Bonte (2020). *Herders, Warriors, And Traders: Pastoralism In Africa.* Routledge.

Giuliano, P. (2007). Living arrangements in Western Europe: Does cultural origin matter? *Journal of the European Economic Association 5*, 927–952.

Guiso, L., P. Sapienza, and L. Zingales (2004). The role of social capital in financial development. *American Economic Review 94*, 526–556.

Guiso, L., P. Sapienza, and L. Zingales (2006). Does culture affect economic outcomes? *Journal of Economic Perspectives 20*, 23–48.

Guiso, L., P. Sapienza, and L. Zingales (2008). Trusting the stock market. *Journal of Finance 63*, 2557–2600.

Haliassos, M., T. Jansson, and Y. Karabulut (2016). Incompatible European partners? Cultural predispositions and household financial behavior. *Management Science 63*, 3780–3808.

Herskovits, M. J. (1926). The cattle complex in East Africa. *American Anthropologist 28*, 230–272.

Hong, H., J. D. Kubik, and J. C. Stein (2004). Social interaction and stock market participation. *Journal of Finance 59*, 137–163.

Huber, S. J. and T. Schmidt (2022). Nevertheless, they persist: Cross-country differences in homeownership behavior. *Journal of Housing Economics 55*, 101804.

Malmendier, U. and S. Nagel (2011). Depression babies: Do macroeconomic experiences affect risk taking? *Quarterly Journal of Economics 126*, 373–416.

Osili, U. O. and A. L. Paulson (2008). Institutions and financial development: Evidence from international migrants in the United States. *Review of Economics and Statistics 90*, 498–517.

Taylor, C. A. and J. Rising (2021). Tipping point dynamics in global land use. *Environmental Research Letters 16*, 125012.

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

Table 1: Cropland share and home ownership – OECD countries

This table studies the relation between cropland share and home ownership rates in a sample of OECD countries. Panel A provides descriptive statistics on the home ownership rate, on two measures of the variable *CropShare*, and on the employment share in agriculture. Panel B estimates several specifications of Equation (2), without and with region or income group fixed effects. Robust standard errors are in parentheses. *, ** and *** denote respectively statistical significance at the 10%, 5% and 1% levels. Additional details on the data sources and on the construction of the variables are provided in Appendix A.

Panel A : Descriptive statistics

	10th	25th	Mean	Median	75th	90th	St. dev.	Obs.
Own. rate (in %)	0.578	0.662	0.730	0.727	0.811	0.893	0.120	40
Crop. share – Global data	-0.533	-0.245	0.106	0.167	0.543	0.665	0.497	35
Crop. share – OECD data	-0.706	-0.192	0.205	0.342	0.543	0.898	0.536	40
Employment in agriculture (in %)	0.016	0.021	0.045	0.034	0.049	0.106	0.037	35

Panel B : Cross-country regressions

	Own. rate	Own. rate	Own. rate	Own. rate	Own. rate	Own. rate
Crop. share – Global data	0.120**	0.164***	0.126**			
	(0.045)	(0.053)	(0.048)			
Crop. share – OECD data				0.084**	0.091*	0.097**
				(0.036)	(0.049)	(0.044)
Robust std. error	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	No	Yes	No	No	Yes	No
Inc. group FE	No	No	Yes	No	No	Yes
Adj. R2	0.201	0.250	0.222	0.119	0.108	0.128
Obs.	35	34	35	40	38	39

Table 2: Descriptive statistics – CPS dataset

This table provides descriptive statistics on the individual level data, from the Current Population Survey, that I use in the main regressions. Panel A reports the distribution of the variables used in the analysis. Panel B list the 15 largest countries of origin of second-generation immigrants' parents, as well as the number of fathers and mothers from these origins. Additional details on the data sources and on the construction of the variables are provided in Appendix A.

Panel A : Moments of the main variables

	10th	25th	Mean	Median	75th	90th	St. dev.	Obs.
Owner	0.000	0.000	0.552	1.000	1.000	1.000	0.497	2675
Avg. Crop share – Parents	-0.481	-0.481	-0.142	-0.451	0.157	0.706	0.489	2645
Subsoil org. carbon – Parents	2.000	2.000	2.391	2.000	2.500	3.000	0.844	2555
Subsoil pH suitability – Parents	-1.025	-0.972	-0.745	-0.821	-0.482	-0.255	0.324	2555
Age	24.000	29.000	43.153	39.000	54.000	72.000	17.843	2675
N. persons in household	1.000	2.000	2.837	2.000	4.000	5.000	1.639	2675
Log(Income)	9.908	10.648	11.148	11.282	11.850	12.344	1.259	2641

Panel B : Countries of origin – Top-15

Country	Share Obs.	N. Fathers	N. Mothers
Mexico	0.371	978	985
Puerto Rico	0.063	167	167
China	0.043	112	116
Philippines	0.042	109	111
India	0.028	74	72
Italy	0.026	69	67
Salvador	0.026	63	72
Viet Nam	0.025	66	68
Cuba	0.023	64	58
Canada	0.018	43	51
Dominican Republic	0.018	45	49
Germany	0.017	44	47
Poland	0.016	42	41
Republic of Korea	0.015	41	41
Guatemala	0.014	39	37

Table 3: Cropland share and home ownership by individuals – Baseline

This table provides estimates of Equation (3) at the individual-level. The dependent variable is a dummy variable equal to one for home owners. The main independent variable is the average value of *CropShare* – measuring the relative importance of cropland and pasture, as defined in Equation (1) – for the individual's parents' countries of origin. The regressions are estimated in a sample of second-generation immigrants from the Current Population Survey, that is, individuals born in the US but whose both parents are born outside the US. Robust standard errors are in parentheses. *, ** and *** denote respectively statistical significance at the 10%, 5% and 1% levels. Additional details on the data sources and on the construction of the variables are provided in Appendix A.

	Owner	Owner	Owner	Owner
Avg. Crop share – Parents	0.077***	0.079***	0.036*	0.033*
	(0.008)	(0.020)	(0.020)	(0.020)
Age				0.013***
				(0.003)
Age squared				-0.000
				(0.000)
Sex				-0.014
				(0.018)
N. persons in household				0.038***
				(0.007)
Log(Income)				0.047***
				(0.009)
Robust std. error	Yes	Yes	Yes	Yes
State FE	No	No	Yes	Yes
Metropolitan FE	No	No	Yes	Yes
Marital status FE	No	No	Yes	Yes
Education level FE	No	No	Yes	Yes
Adj. R2	0.007	0.006	0.147	0.207
Obs.	12234	2645	2627	2594

Table 4: Cropland share and home ownership by individuals – Robustness

This table provides evidence for the robustness of the estimates of Equation (3) in Table 3. The dependent variable is a dummy variable equal to one for home owners. The main independent variable is the average value of *CropShare* – measuring the relative importance of cropland and pasture, as defined in Equation (1) – for the individual's parents' countries of origin. The regressions are estimated in a sample of second-generation immigrants from the Current Population Survey, that is, individuals born in the US but whose both parents are born outside the US. Column (1) excludes individuals who have at least one parent born in Mexico. Column (2) excludes individuals working in the agricultural sector. Column (3) includes $State \cdot Metropolitan$ fixed effects. In column (4), standard errors are clustered at the income group level, while they are robust in other columns. Standard errors are in parentheses. *, ** and *** denote respectively statistical significance at the 10%, 5% and 1% levels. Additional details on the data sources and on the construction of the variables are provided in Appendix A.

	Owner	Owner	Owner	Owner
Avg. Crop share – Parents	0.047**	0.034*	0.033*	0.033*
	(0.023)	(0.020)	(0.020)	(0.019)
Age	0.020***	0.013***	0.013***	0.013***
	(0.004)	(0.003)	(0.003)	(0.004)
Age squared	-0.000***	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Sex	-0.010	-0.016	-0.012	-0.014
	(0.023)	(0.018)	(0.018)	(0.017)
N. persons in household	0.031***	0.038***	0.038***	0.038***
	(0.009)	(0.007)	(0.007)	(0.006)
Log(Income)	0.032***	0.047***	0.046***	0.047*
	(0.010)	(0.009)	(0.009)	(0.023)
Excluding Mexico	Yes	No	No	No
Excluding agriculture	No	Yes	No	No
Robust std. error	Yes	Yes	Yes	No
Inc. group cluster	No	No	No	Yes
State FE	Yes	Yes	No	Yes
Metropolitan FE	Yes	Yes	No	Yes
State*Metropolitan FE	No	No	Yes	No
Marital status FE	Yes	Yes	Yes	Yes
Education level FE	Yes	Yes	Yes	Yes
Adj. R2	0.231	0.208	0.206	0.207
Obs.	1641	2578	2584	2594

Table 5: Cropland share and home ownership by individuals – Subsoil organic carbon as instrument

This table reports estimates from our instrumental variables regression. Panel A reports the first-stage regressions (Equation 5), in which the average value of *CropShare* – measuring the relative importance of cropland and pasture, as defined in Equation (1) – for the individual's parents' countries of origin is instrumented using average subsoil organic carbon in these countries. Panel B reports the second-stage regressions, in which the dependent variable is a dummy variable equal to one for home owners. The regressions are estimated in a sample of second-generation immigrants from the Current Population Survey, that is, individuals born in the US but whose both parents are born outside the US. Robust standard errors are in parentheses. *, ** and *** denote respectively statistical significance at the 10%, 5% and 1% levels. Additional details on the data sources and on the construction of the variables are provided in Appendix A.

Panel A: First-stage regressions

	Crop. share	Crop. share	Crop. share	Crop. share
Subsoil org. carbon – Parents	0.083***	0.099***	0.065***	0.060***
	(0.006)	(0.011)	(0.013)	(0.013)
Age				-0.013***
				(0.003)
Age squared				0.000***
				(0.000)
Sex				-0.032*
				(0.018)
N. persons in household				-0.013**
				(0.006)
Log(Income)				0.015
				(0.011)
Robust std. error	Yes	Yes	Yes	Yes
State FE	No	No	Yes	Yes
Metropolitan FE	No	No	Yes	Yes
Marital status FE	No	No	Yes	Yes
Education level FE	No	No	Yes	Yes
Adj. R2	0.015	0.028	0.170	0.182
F stat.	202.59	74.26	26.60	10.86
Obs.	11852	2555	2537	2506

Panel B: Second-stage regressions

	Owner	Owner	Owner	Owner
Instr. crop. share – Parents	0.412***	0.429***	0.312***	0.141***
	(0.070)	(0.122)	(0.046)	(0.049)
Age				0.019***
				(0.003)
Age squared				-0.000***
				(0.000)
Sex				-0.023
				(0.019)
N. persons in household				0.052***
				(0.006)
Log(Income)				0.066***
				(0.010)
Robust std. error	Yes	Yes	Yes	Yes
State FE	No	No	Yes	Yes
Metropolitan FE	No	No	Yes	Yes
Marital status FE	No	No	Yes	Yes
Education level FE	No	No	Yes	Yes
Obs.	11852	2555	2537	2506

Table 6: Cropland share and home ownership by individuals – Subsoil pH as instrument

This table reports estimates from our instrumental variables regression. Panel A reports the first-stage regressions (Equation 5), in which the average value of *CropShare* – measuring the relative importance of cropland and pasture, as defined in Equation (1) – for the individual's parents' countries of origin is instrumented using average subsoil pH in these countries. Panel B reports the second-stage regressions, in which the dependent variable is a dummy variable equal to one for home owners. The regressions are estimated in a sample of second-generation immigrants from the Current Population Survey, that is, individuals born in the US but whose both parents are born outside the US. Robust standard errors are in parentheses. *, ** and *** denote respectively statistical significance at the 10%, 5% and 1% levels. Additional details on the data sources and on the construction of the variables are provided in Appendix A.

Panel A: First-stage regressions

	Crop. share	Crop. share	Crop. share	Crop. share
Subsoil pH suitability – Parents	0.413***	0.411***	0.321***	0.308***
	(0.014)	(0.029)	(0.030)	(0.031)
Age				-0.014***
				(0.003)
Age squared				0.000***
				(0.000)
Sex				-0.026
				(0.018)
N. persons in household				-0.010
				(0.006)
Log(Income)				0.015
				(0.010)
Robust std. error	Yes	Yes	Yes	Yes
State FE	No	No	Yes	Yes
Metropolitan FE	No	No	Yes	Yes
Marital status FE	No	No	Yes	Yes
Education level FE	No	No	Yes	Yes
Adj. R2	0.068	0.073	0.197	0.207
F stat.	881.33	205.17	113.33	27.09
Obs.	11852	2555	2537	2506

Panel B: Second-stage regressions

	Owner	Owner	Owner	Owner
Instr. crop. share – Parents	0.074**	0.119	0.273***	0.120***
	(0.032)	(0.074)	(0.043)	(0.045)
Age				0.018***
				(0.003)
Age squared				-0.000***
				(0.000)
Sex				-0.024
				(0.018)
N. persons in household				0.051***
				(0.006)
Log(Income)				0.067***
				(0.010)
Robust std. error	Yes	Yes	Yes	Yes
State FE	No	No	Yes	Yes
Metropolitan FE	No	No	Yes	Yes
Marital status FE	No	No	Yes	Yes
Education level FE	No	No	Yes	Yes
Obs.	11852	2555	2537	2506

Figure 1: Cropland share – Global data

This figure plots an histogram of the main independent variable *CropShare*, defined in Equation (1), computed in the cross-section of countries at the global level. Negative values correspond to countries dominated by pasture and positive values to countries dominated by cropland. Additional details on the data sources and on the construction of the variables are provided in Appendix A.

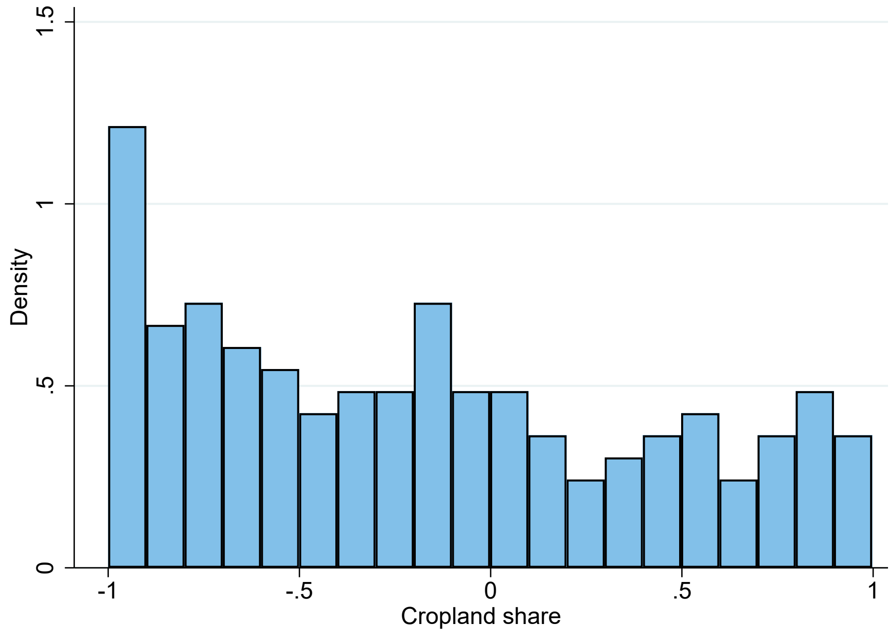


Figure 2: Persistence of the cropland share – Global data

This figure plots the variable *CropShare* computed in 2010 against the same variable computed in 1800. This variable, defined in Equation (1), measures the relative importance of cropland and pasture in land use. Each observation corresponds to a distinct country. Additional details on the data sources and on the construction of the variables are provided in Appendix A.

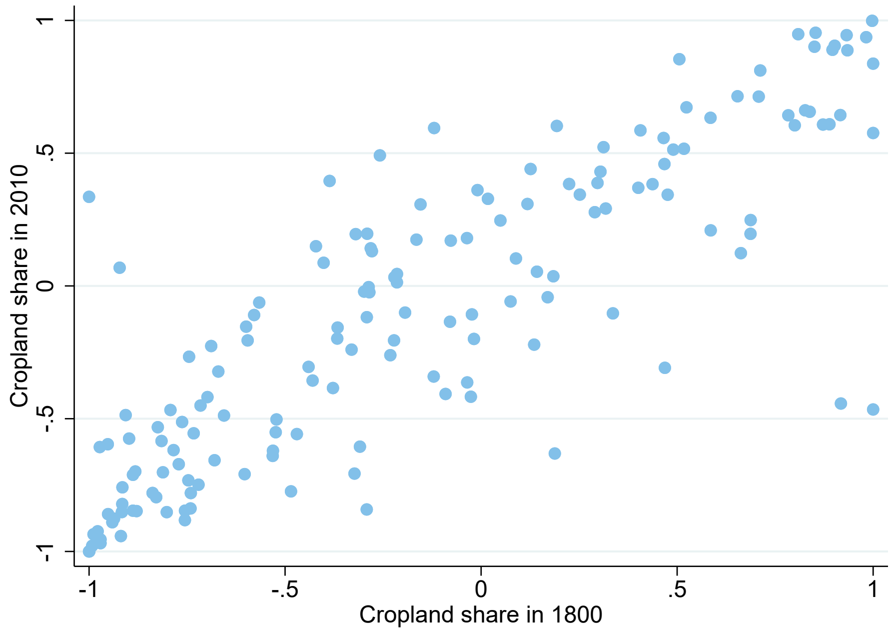


Figure 3: Home ownership rate – OECD countries

This figure plots the average home ownership rate for OECD countries over the period from 2010 to 2020. The data combines outright ownership and ownership with a mortgage. Additional details on the data sources and on the construction of the variables are provided in Appendix A.

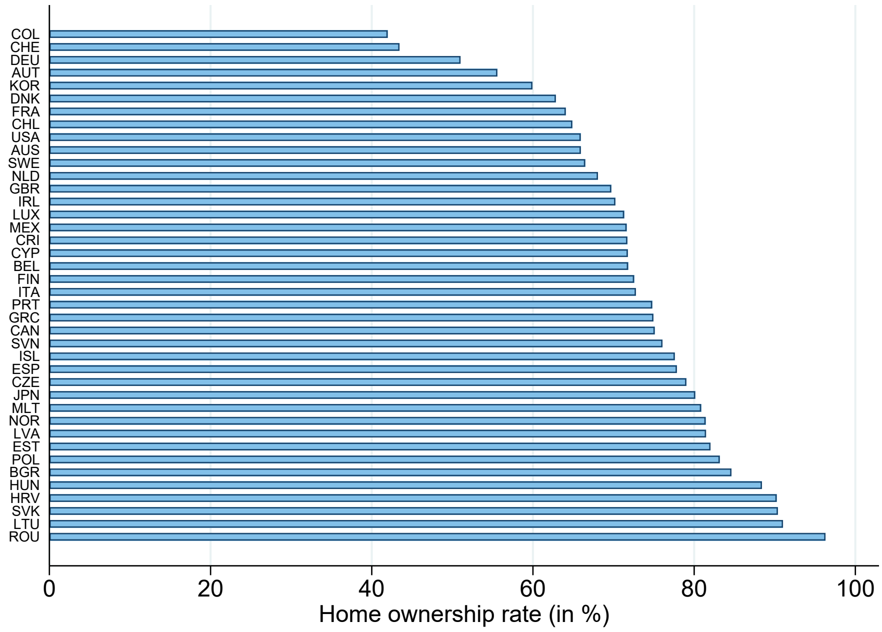


Figure 4: Cropland share – Global data

This figure maps data on cropland share, measured by the variable *CropShare*, at the country level. This variable is defined in Equation (1). Values closer to -1 (respectively 1) correspond to countries in which pasture (respectively cropland) dominates in relative terms. Additional details on the data sources and on the construction of the variables are provided in Appendix A.

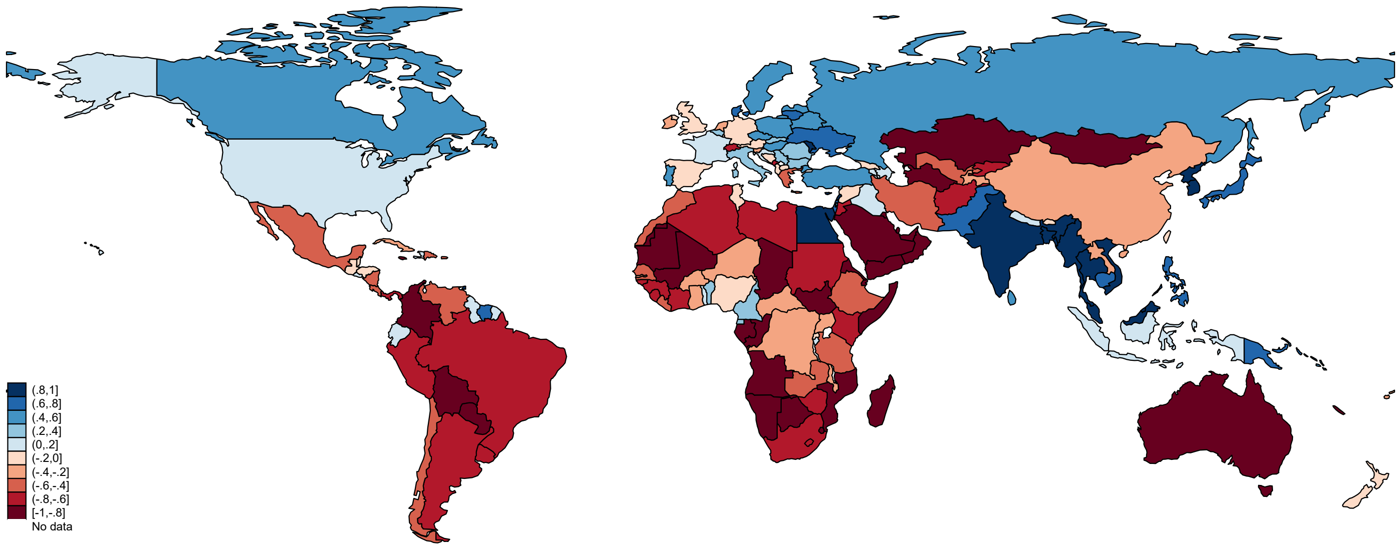
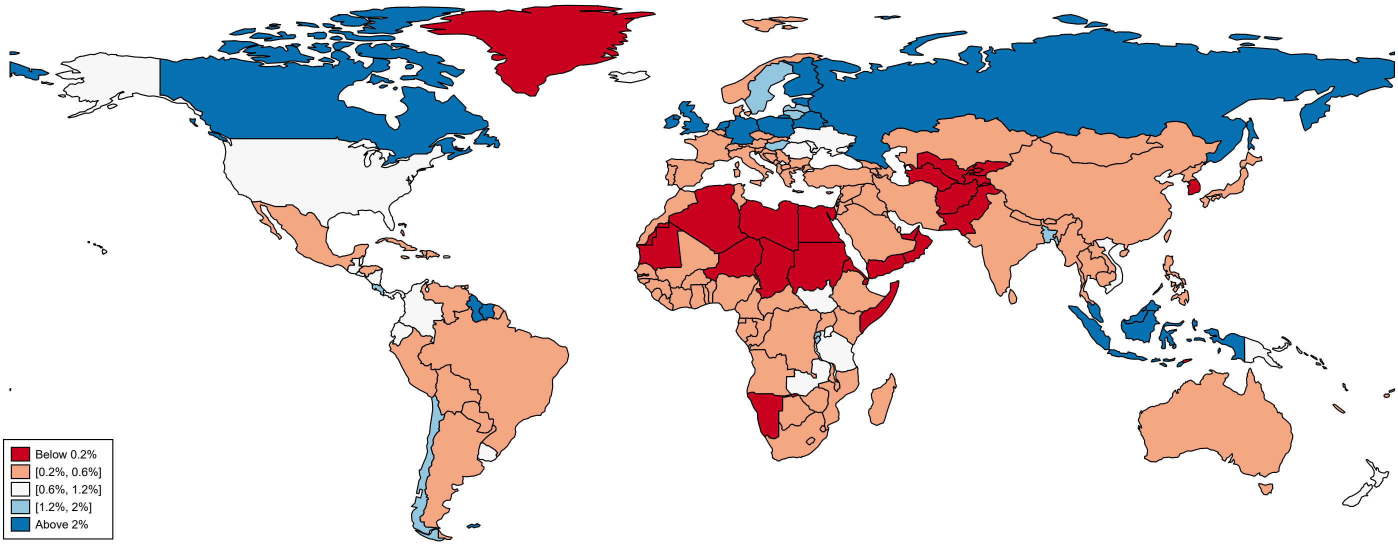


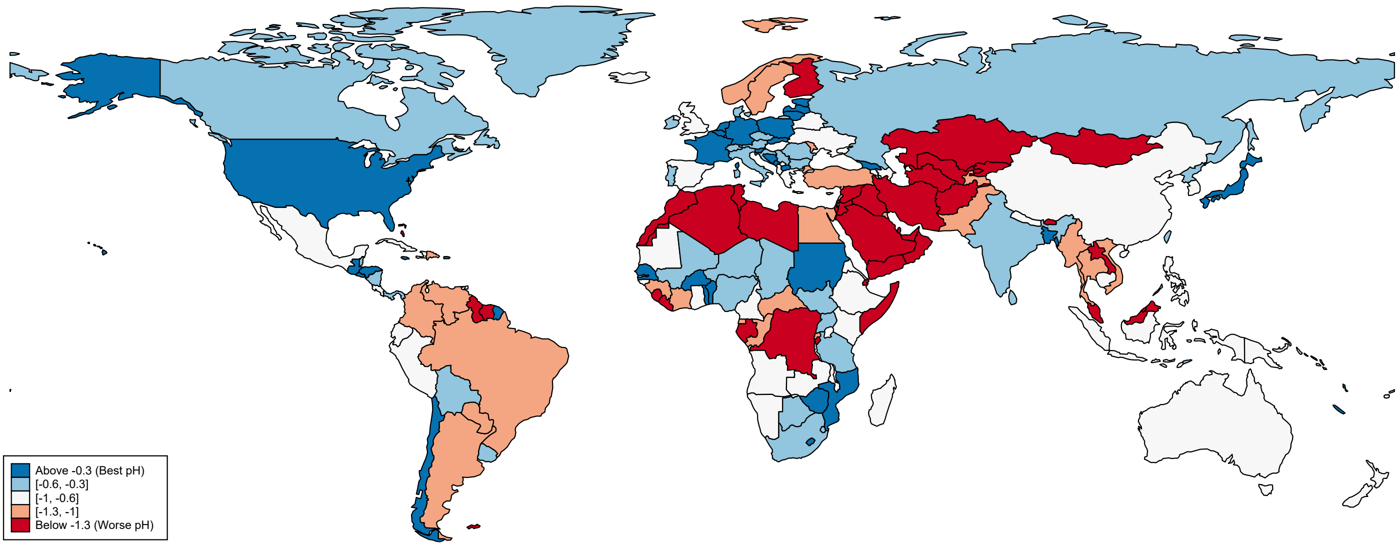
Figure 5: Soil properties – Global data

This figure maps data on average soil properties at the country level. Panel A maps the percentage of subsoil organic carbon and Panel B subsoil pH suitability, i.e., an absolute deviation from the ideal pH, as defined in Equation (4). Country-level data is obtained by collapsing highly detailed data from the Harmonized World Soil Database. Additional details on the data sources and on the construction of the variables are provided in Appendix A.

Panel A: Subsoil organic carbon



Panel B: Subsoil pH suitability



A Data sources and definition of the variables

This appendix provides additional information on the datasets used in the analysis and on the construction of the variables.

A.1 Data from the OECD

From the OECD, I retrieve three series of data. The main one contains housing tenure data from the Affordable Housing Database. It is available for the following 41 countries: Australia, Austria, Belgium, Bulgaria, Canada, Chile, Colombia, Costa Rica, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Malta, Mexico, Netherlands, Norway, Poland, Portugal, Romania, Slovak Republik, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States. Out of these countries, Turkey has missing data for home ownership with mortgages, and is thus excluded. The OECD provides data on housing tenures as a share of either households or total population. I use figures expressed as a share of population. For all countries, I compute home ownership rate as the sum of the two variables "Own outright" and "Owner with mortgage". I then take the average of this rate for all years in which it is available over the 2010-2020 period. When matching these data to data on land use, a few additional observations may be lost, which explains why the number of observations is between 34 and 40 in Table 1.

The second dataset obtained from the OECD contains data on agricultural land use by country. I retrieve data on the total area of cropland and pasture by country from 1985 to 2020. I compute the variable *CropShare* as the average of the ratio in Equation (1) for all years in which the data is available. These data are used only to confirm the results established using my main data on land use, described below in Appendix Section A.2.

The third dataset obtained from the OECD containts data on employment by activity for the year 2020. I compute the share of workers in agriculture as the number of workers in agriculture over total workers (agriculture, industry including construction, manufacturing, services).

A.2 Data on global land use

Data on global land use is obtained from Taylor and Rising (2021), as recompiled by Our World in Data under the header "Cropland and pasture per person, World". It contains information on pasture and cropland in hectares per person for the period between 1780 and 2010. I compute *CropShare* as the average of yearly observations for the ratio defined in Equation (1).

To test the reliability of these data, I compare it with data provided by the OECD on land use for 48 countries (as described in Appendix Section A.1). In this subsample of countries, the correlation between the two measures of *CropShare* is equal to 88.8%. This high correlation is reassuring, since it is obtained from two independent sources, and computed on two distinct time periods (1780-2010 for the global data, 1985-2020 for the OECD data).

A.3 Data from the Current Population Survey

From the US Census, I download datasets from the March Supplement (or Annual Social and Economic Supplements) of the 2022 Current Population Survey. For each household, I keep only the reference person (variable *perrp* equal to 40 or 41). The March Supplement contains datasets at both the household and the person levels. I match these datasets (using the variable *h_seq*) after retrieving the following variables. First, at the household level:

- **Owner**: This variable is a dummy that takes value 1 if the variable *h_tenure* (tenure) is equal to "Owned or being bought", and 0 otherwise.
- **State**: Variable *gestfips*.
- **Metropolitan**: Variable *gtmetsta*. The few observations for which the metropolitan status is missing are excluded.
- **Number of persons in the household**: Variable *h_numper*.
- **Log(Income)**: This is the logarithm of variable *htotval* (total household income).
- **Income group**: Variable *hhinc*.

Second, at the person level:

- **Country of birth of the father**: Variable *pefntvty*. It is then converted to the country's ISO code for matching with other datasets.
- **Country of birth of the mother**: Variable *pemntvty*. It is then converted to the country's ISO code for matching with other datasets.
- **Country of birth**: Variable *penatvty*.

- **Age**: Variable *a_age*. This variable is also used to compute the variable *Age squared*.
- **Sex**: Variable *a_sex*. This variables takes value 1 for males and 2 for females.
- **Marital status**: Variable *a_maritl*. This variable takes distinct values for (i) married - civilian spouse present, (ii) married - Armed Forces spouse present, (iii) married - spouse absent (exc. separated), (iv) widowed, (v) divorced, (vi) separated, (vii) never married.
- **Education level**: The variable *a_hga* is recoded to take four distinct values for educational attainment (i) below high school, (ii) above or equal to high school but below bachelor, (iii) bachelor, (iv) master or above, including professional school degrees.
- **Agriculture**: The variable *a_mjind* (major industry code) is recoded to take value 1 for persons working in the sector called "Agriculture, forestry, fishing, and hunting".

For my main sample, I keep only individuals born in the US whose father and mother are born outside the US.

A.4 Data from the Harmonized World Soil Database

From the Food and Agriculture Organization, I download the Harmonized World Soil Database (HWSD). These data consist of a raster providing a number of soil properties for every 30 arc-second (approximately 1 km by 1 km) for the entire earth. Using the QGIS software, I assign a point to each cell in the raster and obtain the longitude and latitude of each point, which I then map to country shapefiles in Stata. I then average data at the country level for each soil property. I repeat this sequence to obtain data on the following properties:

- **Subsoil organic carbon**: Variable *S_OC*. I use the guidelines provided by the HWSD *Technical Report and Instructions* (Version 1.1, dated March 2009, p. 14) to assign values to five distinct groups based on the percentage of subsoil organic carbon: (i) below 0.2%, (ii) between 0.2% and 0.6%, (iii) between 0.6% and 1.2%, (iv) between 1.2% and 2%, (v) above 2%.
- **Subsoil pH**: Variable *S_PH_H2O*. The pH values are then transformed to obtain a measure of pH suitability for agriculture, as described in Section 5.1.