Does land abundance explain African institutions?

Fenske, James

University of Oxford

December 2011

DOES LAND ABUNDANCE EXPLAIN AFRICAN INSTITUTIONS?

JAMES FENSKE[†]

ABSTRACT. I show that abundant land and scarce labor shaped African institutions before colonial rule. I test a model in which exogenous land quality and endogenous population determine the existence of land rights and slavery. I use cross-sectional data on a global sample of societies to demonstrate that, as in the model, land rights occurred where land quality was high and where population density was greatest. Slavery existed where land was good and population density was intermediate. The model can explain institutional differences across regions, but not within regions. I present suggestive evidence that this is due to institutional spillovers.

1. INTRODUCTION

The "land abundance" view of African history is an influential explanation of the economic institutions that existed on the continent before colonial rule (Austin, 2008a; Hopkins, 1973; Iliffe, 1995). This theory holds that, since uncleared land was freely available, land had no price, and rights to land were ill-defined. Because cultivators would not become free workers, coerced and household labor substituted for wage employment. Lagerlöf's (2010) model of "slavery and other property rights" mirrors these arguments. In this paper, I use cross-sectional data on a sample of global societies to test this view. I show that land rights and slavery existed in those regions predicted by the model, but that institutional spillovers prevent the model from predicting differences within broad geographic regions.

[†]DEPARTMENT OF ECONOMICS, UNIVERSITY OF OXFORD
E-mail address: `james.fenske@economics.ox.ac.uk`.
Date: December 8, 2011.
I would like to thank my advisors Timothy Guinnane, Benjamin Polak, and Christopher Udry for their guidance. I would also like to thank Tayo Adesina, Achyuta Adhvaryu, Gareth Austin, Reena Badiani, Benjamin Chabot, Adrian de la Garza, Rahul Deb, Shatakshee Dhongde, Oded Galor, Nils-Petter Lagerlöf, Naomi Lamoreaux, Giuseppe Moscarini, Nathan Nunn, Sheilagh Ogilvie, Florian Ploeckl, Mark Rosenzweig, Mir Salim, Veronica Santarosa, Ed Vytlacil, Warren Whatley, Ademola Yakubu, and the participants of the Economic History Association Annual Meeting, the Canadian Network for Economic History, the NEUDC, the LSE African Economic History Seminar, and seminars at the University of Warwick, Princeton University, Washington University in St. Louis, Northwestern University, the University of Rochester, the University of Toronto, Boston University and the University of Oxford for their comments and advice. Extra thanks are due to Christian Dippel and Nathan Nunn for their generous sharing of maps.

The pre-colonial institutions explained by the "land abundance" view constrained colonial powers (e.g. Austin (2008b)). As a result, pre-colonial institutions affect current performance in Africa (e.g. Gennaioli and Rainer (2007); Tertilt (2005)). Land rights and slavery, in particular, continue to shape outcomes in Africa and in the rest of the world. Land tenure shapes investment incentives (Goldstein and Udry, 2008), labor-supply (Field, 2007), and violence (Andre and Platteau, 1998). Nunn (2008a) shows that those African countries that exported the most slaves are comparatively poor today. Within the Americas, legacies of slavery explain differences in income across countries and U.S. counties (Engerman and Sokoloff, 1997; Nunn, 2008b), as well as long term racial gaps in education and income (Miller, 2008; Sacerdote, 2005).

The "land abundance" view of African history argues that the continent's geography has given it an abundance of land relative to labor. Geographic features, such as continental orientation, ruggedness, settler mortality, suitability for specific crops, and other biogeographic endowments predict contemporary institutional differences across countries (Easterly and Levine, 2003; Engerman and Sokoloff, 1997; Nunn and Puga, 2011). The model I test, from Lagerlöf (2010), similarly allows geography to shape institutions. There are two critical variables determining land rights and slavery. The first is exogenous land quality. This increases the returns to landownership, compensates for the inefficiencies of slavery, and sustains greater populations in the steady state. The second is population, which responds to the geographic and institutional environments. It shapes the relative values of land and labor and the relative costs of free and forced workers.

I test this thesis. I use data on a cross-section of global societies from Murdock's (1967) *Ethnographic Atlas* to support a model of land rights and slavery in which the land-labor ratio determines the institutions that exist. I find that the model correctly predicts that land rights and slavery were found in those societies that occupied the best land, and that greater population densities were correlated with rights over land. Slavery was present when population densities were intermediate, as in the model. While the model predicts differences across regions, it is not capable of predicting differences within regions.

In Section 2, I outline the literature in African history on how land abundance has shaped economic institutions. I present the basic features of the model and its testable implications. In Section 3, I describe the data used and lay out the econometric specifications. In Section 4, I report the results of these tests. In Section 5, I show that these results are robust to different measures of the institutional outcomes, alternative proxies for land quality and historical population density, and the possible endogeneity of land quality. I also argue that this theory of land rights and slavery better explains the data than some prominent alternatives, including ecological risk as a cause of common property, and suitability for specific crops as a predictor of slavery. In Section 6, I demonstrate that the model is unable to predict differences within regions, and that there is substantial spatial correlation in institutional outcomes. In Section 7 I conclude.

2. The land abundance view of African history

2.1. **The literature.** A first-order task in African history is explaining the continent's long-run differences from the rest of the world. The starting point of the land abundance view is the difference in settlement patterns between pre-colonial Africa and the rest of the world. Herbst (2000, p. 16) estimates the population density of Sub-Saharan Africa in 1900 at 4.4 persons per Sq. Km, contrasted with 38.2 for South Asia, 45.6 for China, and 62.9 for Europe.[1] Explanations of low African population densities stress geographic factors, the disease environment, and historical factors such as the slave trades (Mahadi and Inikori, 1987, p. 63-64). This sparse settlement, Hopkins (1973, p. 23-27) argues, shaped institutions, because Africans "measured wealth and power in men rather than in acres."[2] Here, I outline the implications for land rights and slavery.

Austin (2009, p. 33) argues that African land was often "easily and cheaply accessible in institutional terms"; pre-colonial authorities were eager to attract "more people with whom to subdue nature and, if necessary, their neighbors," so that strangers could generally acquire land indefinitely for token payments. These payments were made solely to acknowledge the sovereignty of the local authorities. Citizens were given land virtually freely. Austin (2008a, p. 591-594) notes that 'islands' of intensive agriculture have existed in Africa where insecurity has created artificial land scarcity and in specific locations of exceptional value. These had minerals, trees, market access, or suitability for particular crops. Against these views, Spear (1997, p. 154-157) argues that population density cannot explain individual cases. While on Mount Meru both the Arusha and the Meru intensified their agriculture as population rose, the less densely settled Meru did so more readily. Berry (1988), similarly, has noted that inheritance rules, tenancy contracts, and labor arrangements often prevent tree crops from leading to individualized land tenure in West Africa. Thornton (1992, p. 75-76) suggests that ownership of land results from legal claims, not population pressure.

[1] His estimate for North Africa is 9.4 persons per Sq. Km.

[2] Austin (2008a, p. 589) argues that Hopkins was the first to make this analysis systematic; earlier writers on Africa did account for the existence of slavery, for example, by noting Africa's land abundance – see Dowd (1917).

For Austin (2008a, p. 606-610), scarcity of labor explains African use of forced labor. He builds on Nieboer (1900) and Domar (1970), who argue that coercion is cheaper than paying a wage when labor is scarce and wages high. Slavery was prevalent in much of Africa even prior to the Atlantic slave trade (Fage, 1969). The use of underpopulation to explain African slavery is controversial. Writers such as Kopytoff and Miers (1977, p. 68-69), Lovejoy (1978, p. 349), or Miers and Klein (1998, p. 4-5) have stressed that slaves were employed in non-economic uses, distributed by non-market means, and that colonial rulers turned a blind eye to slavery for political reasons. Kopytoff (1987, p. 46) and Goody (1980, p. 26-31) add that dependents must be "seduced" rather than coerced, so slavery can only exist in complex societies and states with "well-developed systems of compulsion."

I clarify this literature and test its claims. With the Lagerlöf (2009) model as a reasonable formalization, the "land abundance" view can be reconciled with some of the objections it has faced. While high wages resulting from population density explain the preference for slavery over free labor under certain conditions in the model, there are also under conditions in which population is too sparse for slavery to be worthwhile, corresponding with the less complex societies in Africa that have poorly developed systems of compulsion. More importantly, I test the land abundance view in a global sample of societies. I show that the institutional effects of population and agricultural productivity follow regular patterns even if they cannot explain every case. I show that the presence of slavery is systematically related to the economic value of slaves and to population.

2.2. **Model.** I test the model of "slavery and other property rights" from Lagerlöf (2009). This is for two reasons. First, his model echoes the arguments made by historians, making explicit the testable implications of their views. Greater population lowers average product, which is shared equally in an egalitarian regime. This creates incentives to create rights over land. Similarly, the relative costs of land rights and slavery are determined by the competitive wage, which is itself a function of population size. If population pressure increases labor supply and depresses the wage, free labor becomes profitable relative to keeping slaves.

Second, his model extends the "land abundance" literature. If population is sufficiently low, slavery will not exist, since population pressure has not adequately depressed the returns to an egalitarian sharing of output while the opportunity costs of wasting labor on coercion remain high. This reconciles the land abundance view with the critiques of Kopytoff and Goody. In addition, the quality of land determines both the relative profitability of institutional regimes for a given population and the level of population that can be supported. This variable has been generally neglected by the Africanist literature. Lagerlöf (2009) makes the concept of "land abundance" more precise; it is the availability of cultivable land relative to both population and productivity that matters. Here, I briefly sketch the basic elements of the model and state its testable implications.

The model takes a society in period t with a population P_t of non-elite agents and a comparatively small elite that does not work. The elite chooses institutions. Output Y_t depends on land M, land-augmenting productivity $\tilde{A}_t$, and the labor used L_t:

$$Y_t = (M\tilde{A}_t)^{\alpha} L_t^{1-\alpha} \equiv A_t^{\alpha} L_t^{1-\alpha}. \tag{1}$$

At the beginning of each period, the elite chooses between three regimes based on which one yields them the greatest profits π_t^i, where i denotes one of three institutional regimes. The first is egalitarianism. Under this arrangement, there are no land rights or slavery. The elite and the non-elite each receive average product, and so:

$$\pi_t^E = \left(\frac{A_t}{P_t}\right)^{\alpha}. \tag{2}$$

The second possible outcome is slavery. Here, the elite enclose the entire land, creating rights over it. They enslave S_t slaves from the population, paying them only subsistence income $\bar{c}$. Each slave requires γ guards, who are also paid $\bar{c}$, and so the elite payoff is:

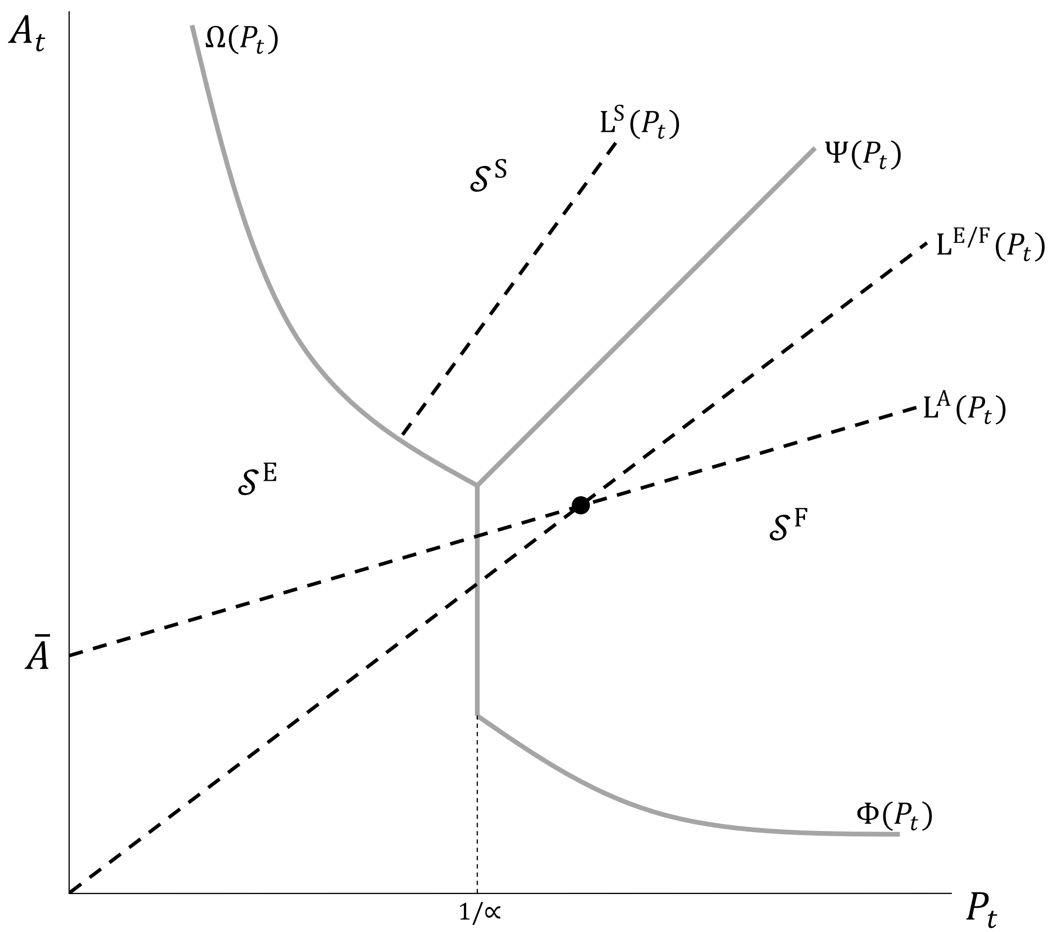
$$\pi_t^S = \max_{S_t \le P_t/(1+\gamma)} \{A_t^{\alpha} S_t^{1-\alpha} - (1+\gamma)\bar{c}S_t\}. \tag{3}$$

The third possible outcome is free labor. Again, the elite enclose the entire land. Now, however, they hire members of the population at a competitive wage w_t, which depends on P_t. The elite's payoff is:

$$\pi_t^F = \max_{L_t \ge 0} \{A_t^{\alpha} L_t^{1-\alpha} - w_t L_t\}. \tag{4}$$

Lagerlöf (2009) shows that the state space in A_t and P_t can be divided into three sets: $\mathcal{S}^E$, in which the elite prefer egalitarianism; $\mathcal{S}^S$, in which they prefer slavery, and; $\mathcal{S}^S$, in which they prefer free labor. The boundaries of these regions are defined by three functions of P_t: $\Psi(P_t)$, $\Omega(P_t)$, and $\Phi(P_t)$.[3] These are depicted in Figure 1.

FIGURE 1. Institutional regions and dynamics



The slavery region, $\mathcal{S}^S$, is where $A_t \geq \max\{\Psi(P_t), \Omega(P_t)\}$, and $P_t > (1+\gamma)^{1-\alpha}$. $A_t \geq \Psi(P_t)$ implies that population given A_t is still sufficiently low that w_t is high relative to the cost of keeping slaves. $A_t \geq \Omega(P_t)$ implies that population is sufficiently dense that the average product under egalitarianism has fallen, while high productivity also ensures the elite is willing to waste some labor on guarding slaves in order to take a greater share

3

$$\Psi(P_t) = \left(\frac{\bar{c}(1+\gamma)^{1-\alpha}}{1-\alpha(1+\gamma)^{1-\alpha}}\right)^{\frac{1}{\alpha}} P_t$$

$$\Omega(P_t) = \left(\frac{\bar{c}(1+\gamma)^{1-\alpha} P_t^{1+\alpha}}{P_t - (1+\gamma)^{1-\alpha}}\right)^{\frac{1}{\alpha}}$$

$$\Phi(P_t) = \left(\frac{1}{1-\alpha}\right)^{\frac{1}{1-\alpha}} \left(\frac{\bar{c}}{1-\alpha}\right)^{\frac{1}{\alpha}} P_t^{-\frac{\alpha}{1-\alpha}}$$

of output for themselves. The opportunity cost of these guards is particularly high when population is very low, which explains both the slope of $\Omega(P_t)$ and the condition that $P_t > (1+\gamma)^{1-\alpha}$.

$\mathcal{S}^F$ is the free labor region, in which $\Phi(P_t) \leq A_t \leq \Psi(P_t)$ and $P_t > 1/\alpha$. $P_t > 1/\alpha$ ensures that population is great enough that the average product has fallen, making enclosure worthwhile. $A_t \leq \Psi(P_t)$ occurs when population growth pushes down wages sufficiently relative to the costs of keeping slaves. The condition that $\Phi(P_t) \leq A_t$ is of less interest, driven by an assumption that the wage is bounded below by $\bar{c}$. $\mathcal{S}^E$ occurs in the remainder of the state space, where average product and the counterfactual wage are both relatively high.

The dynamics of the model are Malthusian and Boserupian. They are Malthusian in that fertility is increasing in income. Two upward-sloping zero population growth lines exist – one under slavery and one under both egalitarianism and free labor. To the left of these, income is high and population is growing. To the right, income is low and population is falling. These are shown in Figure 1 as $\mathbf{L}^{E/F}(P_t)$ and $\mathbf{L}^S(P_t)$.[4]

The dynamics are Boserupian in that agricultural technology in period $t+1$ has an intercept of $\bar{A}$ and depends positively on both A_t and P_t. Lagerlöf (2009) takes $\bar{A}$ as the "minimum level of agricultural technology," and I interpret it as exogenous land quality. The result is an upward-sloping zero-technological-growth line $\mathbf{L}^A(P_t)$.[5] Above this, productivity degrades, while below this it improves. This is also shown in Figure 1. A steady state exists where either $\mathbf{L}^{E/F}(P_t)$ or $\mathbf{L}^S(P_t)$ intersects $\mathbf{L}^A(P_t)$. Figure 1 depicts a steady state in the free labor region.

2.3. **Tests.** What are the testable implications of this model and, by extension, the land abundance view? First, land quality $\bar{A}$ should positively predict the existence of land rights and slavery. Land rights do not exist under egalitarianism, and if $\bar{A}$ is too low, it is impossible to support a steady state under either regime. Similarly, $\bar{A}$ must be high in order for a steady state to exist with slavery. However, since larger values of $\bar{A}$ can

4

$$\mathbf{L}^{E/F}(P_t) = \left(\frac{q}{\beta}\right)^{\frac{1}{\alpha}} P_t$$

$$\mathbf{L}^S(P_t) = (1+\gamma)^{\frac{1-\alpha}{\alpha}} \left(\frac{q}{\beta} + \bar{c}\right)^{\frac{1}{\alpha}} P_t$$

5

$$\mathbf{L}^A(P_t) = \bar{A} + D^{\frac{1}{\theta}} P_t.$$

support steady states in both the slavery and free labor regions, the relationship between $\bar{A}$ and slavery is expected to be weaker than for land rights. Second, population density, which I take as corresponding to P_t in the model, will predict land rights and slavery. While this is an endogenous variable, this is still a correlation implied by the model and one of its implications. For land rights to exist, P_t must be greater than the cutoffs implied by $\Omega(P_t)$, $1/\alpha$, and $\Phi(P_t)$. For slavery to exist, P_t must be great enough that enclosure of land is worthwhile and the opportunity costs of coercion are not too high, but also sparse enough that wages are not too low. It must, then, be between the cutoffs implied by $\Omega(P_t)$ and $\Psi(P_t)$. It is the implied relationships between land quality, population density, land rights, and slavery that I test in assessing the "land abundance" view.

2.4. **Other implications.** In addition to land quality, $\bar{A}$ and the state variable, P_t, which I use to test the land abundance view, there are seven exogenous parameters in the model. I do not use these to derive additional tests. Three parameters, β, q and α, do not yield clear predictions for the existence of land rights or slavery. The remaining parameters, γ, D, θ and $\bar{c}$, do give clear comparative statics (see the appendix). The difficulty is that these parameters are not directly observed in geographic data in the same sense that $\bar{A}$ and P_t are directly observed. The geographic controls I am able use are will affect several parameters at once. Ruggedness, for example, may increase the cost of guarding slaves (γ) through the mechanisms identified by Nunn and Puga (2011), but will be also impact technological parameters (θ, D, and α) directly, since feasible agricultural systems differ between flat and rugged areas.

3. Data and Specifications

In this section, I outline how I test the two predictions of the model described above. I use a cross section of data on 1,206 societies. In Section 3.1 I detail the specific econometric specifications. In Section 3.2, I describe the sources of data on institutions, the proxies for the variables $\bar{A}$ and P_t in the model, and the additional controls that I include.

3.1. **Specifications.** The first prediction of the model is that raising $\bar{A}$ will make it possible for steady states to exist with land rights or slavery. I test this by estimating:

$$y_i = \alpha + \beta_A A_i + x_i' \gamma + \epsilon_i, \tag{5}$$

where y_i is an outcome of interest for soceity i, A_i is a proxy for land quality (analogous to $\bar{A}$ in the model), x_i is a vector of geographical controls, and ϵ_i is random error. (5) is estimated as a probit with heteroskedasticity-robust standard errors. I expect that $\beta_A > 0$ when y_i is an indicator for land rights or slavery.

The second implication of the model is that land rights exist at higher levels of P_t, while slavery exists at intermediate levels of P_t. I test these by estimating:

$$y_i = \alpha + \beta_p \ln(1 + p_i) + x_i'\gamma + \epsilon_i, \tag{6}$$

and

$$y_i = \alpha + \beta_{p1} \ln(1 + p_i) + \beta_{p2}(\ln(1 + p_i))^2 + x_i'\gamma + \epsilon_i, \tag{7}$$

where (abusing notation) y_i, x_i, and ϵ_i are defined as in (5). p_i is population density, the proxy used for P_t. The functional form comes from visual inspection of the data – slavery peaks towards the left hand side of the distribution, while a strict logarithmic specification gives undue influence to very sparsely settled societies. These are also estimated as probit models. I expect that $\beta_p > 0$, $\beta_{p1} > 0$, and $\beta_{p2} < 0$. I estimate equations with log population density and land quality separately, since the correlation between the two ($\rho = 0.36$) inhibits joint tests. I do not report results of the land rights regression with the quadratic term, since an inverse-U relationship is not anticipated by the model. If I do estimate the equation with the quadratic term, that term not statistically significant.

3.2. **Data.**

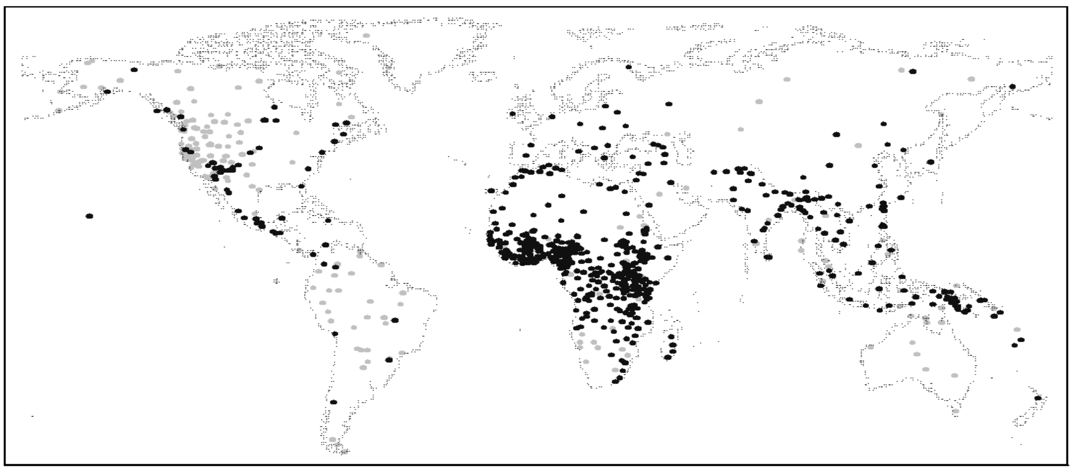
I use two types of data to test the ability of the model to explain institutional differences across societies. Details of the variables used and their sources are contained in the appendix. The first covers institutions, and is taken from Murdock's (1967) *Ethnographic Atlas*. Published in 29 installments of the journal *Ethnology* between 1962 and 1980, the Atlas is a database of 1267 societies from around the world. It contains categorical variables describing several institutional and cultural features of these societies, usually at the time of first contact with Europeans. From this sample, I remove 2 duplicate observations (the Chilcotin and Tokelau), 8 societies from before 1500 (Ancient Egypt, Aryans, Babylonia, Romans, Icelander, Uzbeg, Khmer, Hebrews), and 51 for which land quality information is missing (mostly small Pacific islands). This leaves a base sample of 1206 societies. 801 of these have data on land rights, 1041 on slavery.

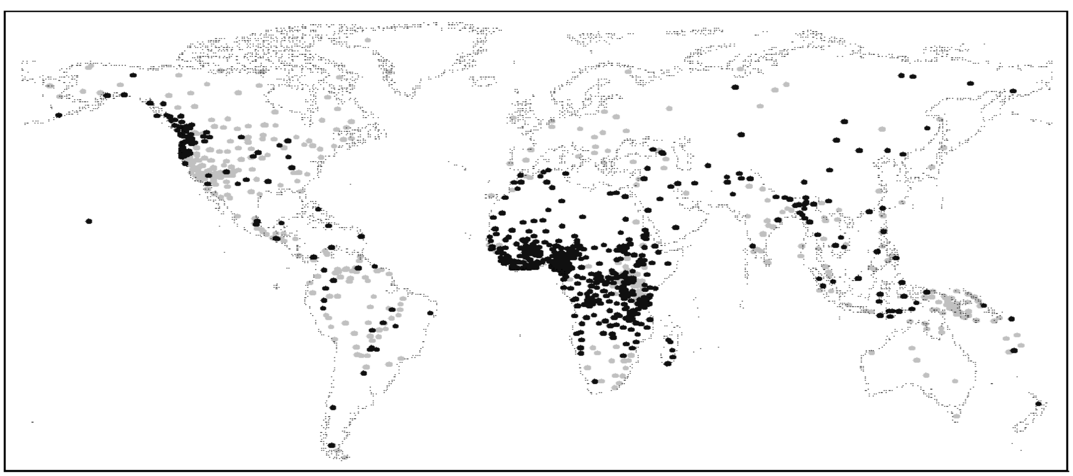
TABLE 1. Summary statistics

[Table 1 here]

I use variables from the *Ethnographic Atlas* to construct binary variables for whether land rights or slavery exist. Summary statistics for these are given in Table 1. For each society, I observe land rights and slavery together at a single point in time. I map these variables by the latitude and longitude coordinates of each society in Figure 2.

FIGURE 2. Land rights and slavery





Land rights are on top, slavery on bottom. Black circles indicate presence, grey circles absence.

Why use this data? The principal justification is availability. This is the only source of cross-cultural information on land rights and slavery of which I am aware that has global scope. In addition, the measures in the *Ethnographic Atlas* are internally consistent, having been compiled by the same author.

The second type of data used includes features of the natural environment. I join these to the data from the *Ethnographic Atlas* using one of five map sources. First, I join African societies to one or more ethnic groups mapped by Murdock (1959) in his "Tribal Map of Africa." Second, I merge First Nations groups in the United States and Canada with the maps that begin the volumes the *Handbook of North American Indians* (Heizer and Sturtevant, 1978), digitized for the United States by Dippel (2010) and for Canada by myself. Third, I join ethnic groups from the rest of the world to Global Mapping International's (GMI) detailed World Language Mapping System. Fourth, if no match can be found in the GMI map, I use the less detailed Geo-Referencing Ethnic Groups (GREG) map created by Weidmann et al. (2010). Finally, if no suitable match can be found in any of these, I match groups with modern administrative boundaries manually. For example, the Nunivak are matched to Nunivak Island. Not all societies are matched exactly. Of 1,267, 100 were matched to a different group indicated in the same location while 76 were matched to a larger group of which they form a smaller part (such as the Efik to the Ibibio). A full table of matches are given in the web appendix.

Once these matches are formed, geographic raster data is joined to them by taking the average of the raster points within an ethnic group's territory. Summary statistics for these variables are presented in Table 1. Two of these controls are of particular importance – land quality and population density.

3.2.1. *Land quality.* The variable used to capture land quality is based on Fischer et al.'s (2002) measure of combined climate, soil and terrain slope constrains. This is re-scaled as a standard normal variable between 0 and 1, with larger values indicating an absence of environmental constrains on rainfed agriculture. This is treated as a proxy for the variable $\bar{A}$ in the model.

The constraints measure was constructed as part of the Food and Agriculture Organization's Global Agro-Ecological Zones (FAO-GAEZ) project.[6] This measure is not particular to any particular crop or technology, and is a non-additive combination of three components:

(1) *Climate constraints:* The coldness constraint is "moderate" if there are fewer than 180 days with an average temperature below 5°C, and "severe" if there are fewer than 120. Aridity constraints are moderate if there are less than 120 days with an average temperature below 5°C during which moisture conditions are adequate to permit crop growth and severe if there are less than 60.

[6] See `http://www.iiasa.ac.at/Research/LUC/GAEZ/index.htm`

(2) *Soil constraints:* Five characteristics of soils are considered. These are depth, fertility, drainage, texture and chemical constraints. "Medium" and "shallow" depth are moderate and severe constraints, respectively. "Medium" and "low" fertility are treated similarly as moderate and severe constraints. "Poor" drainage is a severe constraint. Sandy and stony soils are severe constraints, and cracking clay is a moderate constraint. Salinity, sodicity, and gypsum are severe chemical constraints.
(3) *Terrain slope constraints:* Terrain slopes greater than 8% are "moderate" constraints, and slopes greater than 30% are "severe."

Climate constraints and soil texture are exogenous. It is possible that societies that developed slavery or rights over land were able to avoid degrading soil depth or fertility. I add the direct measures of soil depth and fertility constraints as additional controls as a robustness check. An additional advantage of this constraints-based measure is that it is not based on expected yields in contemporary agriculture, in which greater crop diversity would be available than for many of the societies at the time they are recorded in the *Ethnographic Atlas*.

3.2.2. *Population density.* All historical reconstructions of population are guesses. One book on estimates for pre-Columbian America is entitled "Numbers from Nowhere" (Henige, 1998). The principal measure that I use for historical population density is taken from the History Database of the Global Environment (HYDE) version 3.1, which has been previously used by Bluedorn et al. (2009). This is raster data on historical population, covering the years 1500, 1600, and every ten years since 1700. For each ethnic group, I measure historical population density as the average of the raster points within its territory for the year of observation recorded in the *Ethnographic Atlas*.[7]

These population estimates were assembled by the Netherlands Environmental Assessment Agency, and the details of their construction are reported by Bouwman et al. (2006), Klein Goldewijk et al. (2010) and Klein Goldewijk (2005). This data source takes as its base a map of 3441 administrative units from 222 countries. Historical data are then reconstructed on this base map using Lahmeyer (2004), Helders (2000), Tobler (1995), several local studies, interpolation, and back projection. The data are reported on a five minute grid.

My key estimates of land quality and population density are plotted together in Figure 3. In addition to summary statistics in Table 1, I present the percentiles of the HYDE

[7]For computational reasons, I use data from each 50 year interval, imputing intermediate years exponentially.

population data and the two principal alternatives to this data, described below. These range from nearly zero persons per square mile for several groups in the Mato Grosso and interior Amazon, to over 3,000 persons per square mile for the Okinawans of Japan.[8]

3.2.3. *Alternative measures of population density.* Because historical population reconstruction unavoidably inexact, it is important to show that the results can be obtained using alternatives to the HYDE estimates. Because the HYDE estimates are the only raster data of which I am aware on historical population, I adopt a simple method to estimate alternative spatially disaggregated historic population densities for the societies in my data. I begin with raster data on population density in 1995 for each of these ethnic groups and combine it with other estimates of historical population densities for the broader regions within which these groups are located. Specifically, my alternative estimates of historical population density take the form:

$$\text{Historical population density} = \text{Population density in 1995} \times \frac{\text{Regional density at the date of observation}}{\text{Regional density in 1995}}. \tag{8}$$

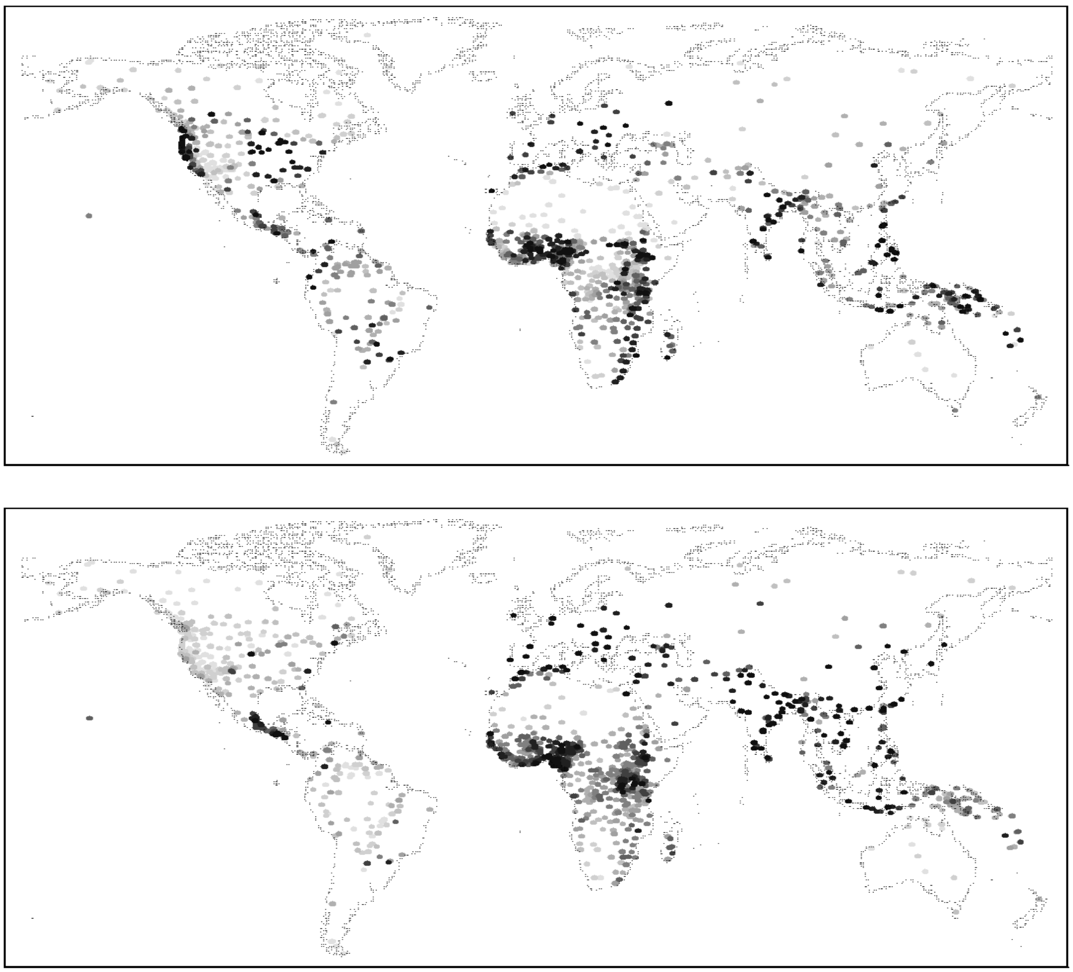
Critically, this assumes that the relative distribution of population has not changed within regions over time. If the Tamil were 1.37 times as dense as the entirety of the broad region "India" in 1995, this ratio is pushed back to 1880, the date at which they are observed. While only a first-order approximation, this approach is preferable to using the unweighted regional densities directly. McEvedy and Jones (1978), for example, assign a single population density to all of Canada. To treat the population densities of the Inuit and Ojibwe groups in the data as equal would be implausible, and would introduce substantial measurement error on the right hand side.[9] In addition to these two main alternatives, I use the 1995 densities directly, the historical regional densities directly, and for roughly 175 societies I have access to independent estimates of their population densities from the *Standard Cross-Cultural Sample* (SCCS) of Murdock and White (1969). A derivative of the *Ethnographic Atlas*, the SCCS contains a larger number of variables for a smaller sample of societies.

[8]This appears to be a mis-measurement due to over-representation of Naha in the original raster data; administrate records give a modern density of just above 1,500 persons per square mile. Results are robust to excluding the Okinawans.

[9]Ruff (2006) suggests that the Northeast had a population density at contact roughly seven times that of the Arctic. The method used here assigns the Ojibwe a historic population density of 2.20 per square mile and the Copper Eskimo a population density of 0.31 per square mile – a roughly seven-fold difference.

GIS data on population in 1995 is readily available from the FAO-GAEZ. I use two sources of regional estimates. The first is McEvedy and Jones (1978), who create estimates at regular intervals for 163 regions of the world. There are, however, well-known problems with these data (Austin, 2008b; Hopkins, 2009). Thus, I also use the ARVE Group's estimates, constructed by Krumhardt (2010).[10] She divides the world into 209 regions, and gives population estimates for hundred year periods between 6050 BC and 1850 AD.

FIGURE 3. Land quality and historic population density



Land quality is on top, population density on bottom. Darker colors indicate higher values; the ranges of both are given in Table 1.

3.2.4. *Other controls.* In addition, I control for several other factors that may determine the existence of land rights and slavery. These are re-scaled as standard normal variables

[10]See `http://ecospriv4.epfl.ch/index.php?dir=pub/&file=pop_landuse_data.tar.gz`

in the regressions. Definitions and sources are outlined in the appendix. These are the presence of a major river, distance to the coast, elevation, the percentage of the society's territory in which malaria is endemic, precipitation, ruggedness, temperature, date of observation, absolute latitude, share desert, and an indicator for whether the society derives most of its income from fishing. Summary statistics are given in Table 1. These geographic controls are intended to capture variation in the parameters of the model that I am unable to measure directly, i.e. $\bar{c}$, β, α, γ, q, D, and θ.

Access to a major river and distance from the coast proxy for water-borne diseases that affect the cost of fertility (q). These distances also capture the presence of trade, which affects both the production function (α) and the relative utility of fertility (β) through the menu of goods that are traded and through technological transfer. They will affect the cost of slavery (γ) through what uses exist for slaves and whether they can be punished by sale for export. Proximity to markets also affects β through the usefulness of children as substitutes for insurance and savings. Elevation is related to the disease environment, and hence q. It also affects the range of available crops and technologies, and hence α.

Malaria affects the cost of fertility (q) through child mortality and the cost of slavery (γ) via slave mortality. It may also alter the physical cost of effort in adults. Malaria and absolute latitude are both also for the unobservable features of the tropics that make their institutions systematically different from those in other parts of the world.[11] Precipitation determines what crops can be grown, shaping the production function (α). African growing seasons and diseases are constrained by the seasonal availability of moisture. Areas with low rainfall are also those most susceptible to drought. Ruggedness, as discussed above, shapes γ, α and D. Temperature affects the physical cost of effort, and hence γ. In hostile environments such as deserts, it is more difficult for slaves to flee; γ is lower. Temperature affects q through nutrition and disease.

The analysis generally treats the ethnic groups in the sample as if they are in their steady states. The date of observation, then, is a proxy for the degree of European influence and other institutional contamination that pushes societies towards a Westernized recognition of land rights and abolition of slavery. If these societies were to be viewed at another point in time, there is no guarantee that they would possess the same institutions or population densities.

[11]I report results excluding absolute latitude in Table A3 in the appendix. Results are similar without this control, though the quadratic term on population density becomes marginally insignificant in the slavery equation. The population density at which slavery peaks, roughly 116 persons per square mile, is between the 90^{th} and 95^{th} percentiles of the distribution.

TABLE 2. Main results

[Table 2 here]

The control for fishing is included with the Pacific Northwest in mind. Here, groups such as the Haida used slaves in fishing and hunting, and were able to capture them using canoe raids (Donald, 1997). This region is well known for having a relatively high surplus and developed material culture despite the lack of importance of agriculture. For groups with easy access to fish and an economy not centered on agriculture, land quality may be a poor measure of $\bar{A}$ in the model. This is why Nieboer (1900) formulated his view in terms of the abundance of resources in general.

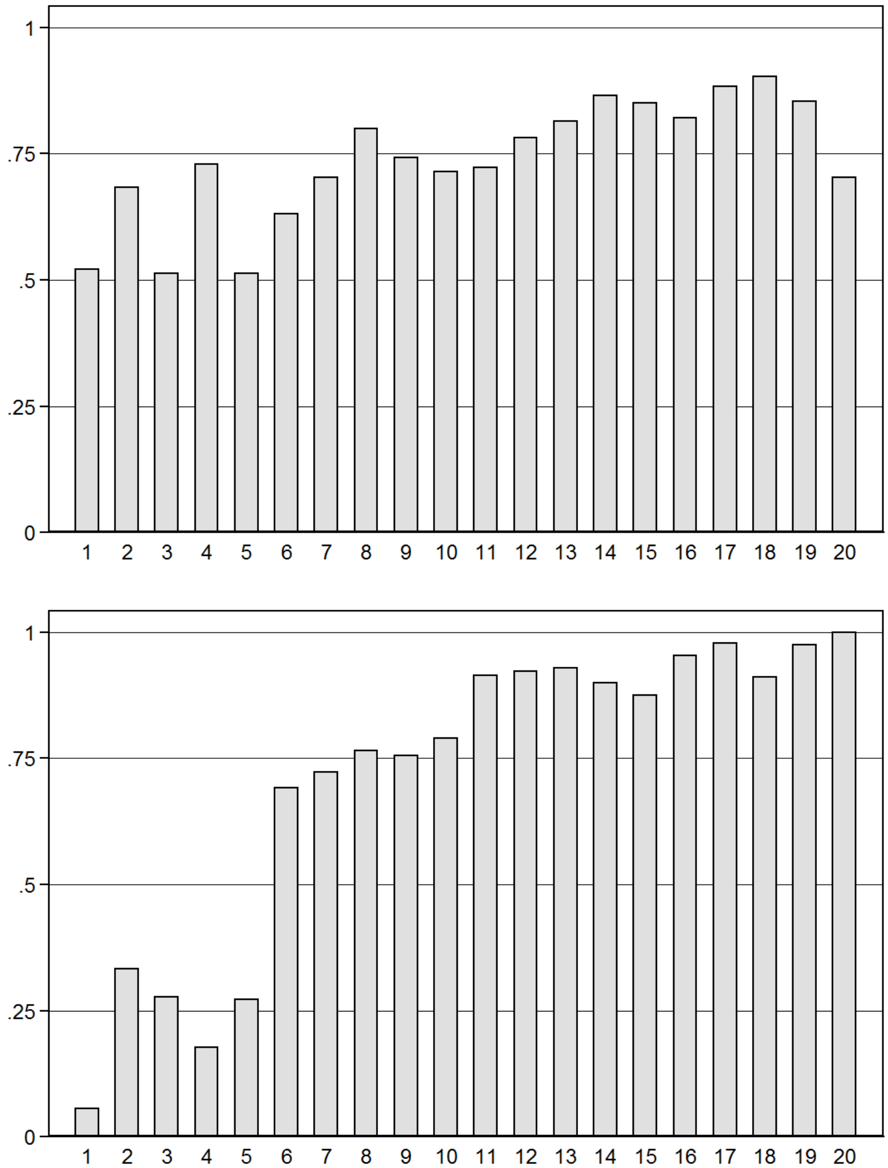
4. RESULTS

4.1. **In pictures.** It is useful, first, to know whether the correlations predicted by the model are apparent in the raw data. I divide the sample into percentiles of land quality and historic population density. In Figure 4, I report the fraction of societies in each percentile that have land rights. In Figure 5, I do the same for slavery. The raw correlations are as predicted. Land rights are positively related to land quality and population density. Slavery is positively correlated with land quality, though this is weaker than the relationship for land rights. In the model, the existence of multiple steady states helps explain this. Further, slavery is most prevalent in societies with intermediate population densities.

4.2. **Regressions.** In Table 2, I report the results of estimating (5), (6) and (7). Specifically, I report marginal effects. For land quality, these can be interpreted as the effects of a one standard deviation improvement. When additional controls are added, the results suggest that a one standard deviation improvement in land quality raises the probability that land rights exist by roughly 4.6%. Interpreting the coefficient on population density as an elasticity, a 1% increase in population density is associated with a 0.124% increase in the chance that land rights exist.[12] A one standard deviation increase in land quality predicts a 4.7% increase in the chance of slavery. While the coefficients on the quadratic term for population density are less easy to interpret, the inverted-U probability profile visible in Figure 5 is visible here. The level of historic population density at which slavery peaks, given by $e^{\beta_{p1}/(2\beta_{p2})}$, is reported in the table. With controls, this is 71 persons per square mile, which is between the 85^{th} and the 90^{th} percentiles of the data.

[12]This is a reasonable approximation, though not strictly correct, because the normalization is log(1+pop. den.), not log(pop. den.).

FIGURE 4. Land rights by percentiles of land quality and hist. pop. density

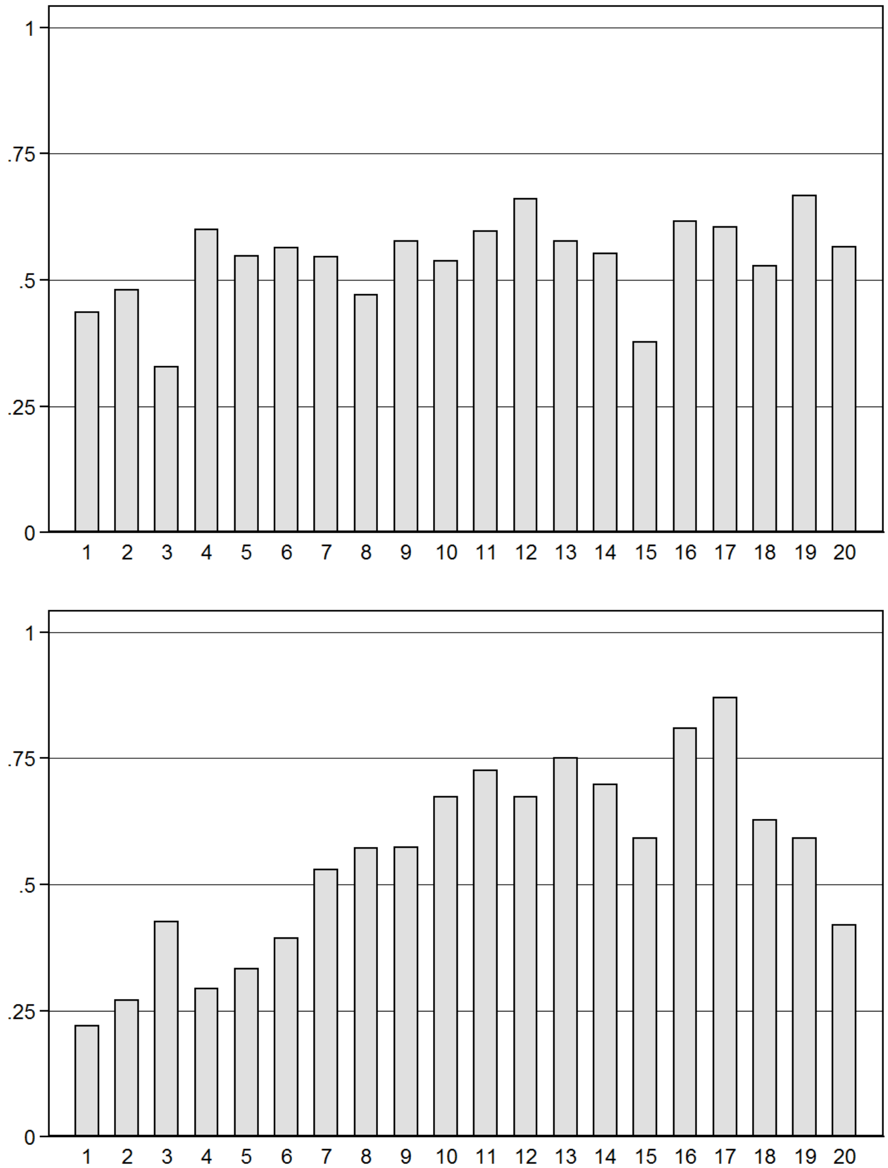


The y axis is the percentage of societies with land rights. The top picture divides this by twenty percentiles of land quality, each representing 5% of the sample. The bottom picture is divided by similar percentiles of population density.

5. ROBUSTNESS: WHAT THE MODEL CAN EXPLAIN

In this section, I show that the results in Section 4 are robust to several possible objections. I show that they can be replicated using alternative measures of land rights and slavery, and that the measures used for the dependent variables are correlated with other measures of these in other samples not large enough to be used for replicating the results. Second, I show that similar results can be obtained using different estimates

FIGURE 5. Slavery by percentiles of land quality and hist. pop. density



The y axis is the percentage of societies with slavery. The top picture divides this into twenty percentiles of land quality, each representing 5% of the sample. The bottom picture is divided by similar percentiles of population density.

of both population density and land quality. Third, I show that the results generally survive additional checks for the importance of influential observations, for the possible endogeneity of land quality, and for alternative clustering of the standard errors. Fourth, while I acknowledge that the data provide only limited scope for testing the model against alternative explanations of land rights and slavery, I argue that the model performs well against notable competing theories.

TABLE 3. Alternative measures of the dependent variables

[Table 3 here]

5.1. **Alternative measures of the dependent variables.** Land rights and slavery are sharp indicators of the existence of these institutions. Land rights in particular exist for some 74% of societies in the data, but do not necessarily capture differences in how well defined these rights are. I begin by demonstrating that my measure of land rights is positively correlated with $v1726$ in the *SCCS*, an indicator for whether land is mostly private. Because $v1726$ is only available for 80 societies, I am not able to replicate the econometric analysis with it. The results of regressing the existence of land rights on $v1726$ are positive and significant, as reported in column (1) of Table 3.

Next, I use an indicator for whether the inheritance of land is patrilineal as an alternative measure of land rights.[13] Following Goody (1969), this captures the degree to which the control of real property is directed towards the nuclear family. Roughly, this is one step along the transition from weakly defined to strongly defined rights in land. Similarly, I use an indicator for whether land is inherited by sons.[14] I show in columns (2) through (5) of Table 3 that both of these are positively related to land quality and population density, conditional on the other controls.

For slavery, I make similar tests. First, I show in column (6) of Table 3 that the main measure of slavery is correlated with an indicator constructed from $v919$ of the *SCCS* for the existence of large-scale slaveholding. In columns (7) through (10), I show that the results can be mostly replicated by constructing alternative measures of slavery from the *Ethnographic Atlas*. Slavery is recorded as either "absent" (1), "incipient or nonhereditary" (2) "reported but type not identified" (3), or "hereditary and socially significant" (4). I create a "slavery above incipient" dummy for whether $V70 > 2$, and a "hereditary slavery" dummy for $V70 = 4$. The positive conditional correlation between land quality and non-incipient slavery is still apparent, as are the hump-shaped relationships with population density, though the link between land quality and hereditary slavery is small and statistically insignificant.

5.2. **Alternative measures of land quality and population density.** I validate the use of the land quality measure by showing that it is strongly correlated with three alternative measures of land quality contained in the *SCCS* – $v921$, $v924$ and $v928$. This is reported in columns (1) through (3) of Table 4.

[13]Like the indicator for land rights, this is constructed using V74: Inheritance Rule for Real Property (Land). This is equal to 1 if V74=4, V74=5, V74=6, or V74=7.

[14]This is equal to 1 if V74=7.

TABLE 4. Alternative measures of land quality and population density

[Table 4 here]

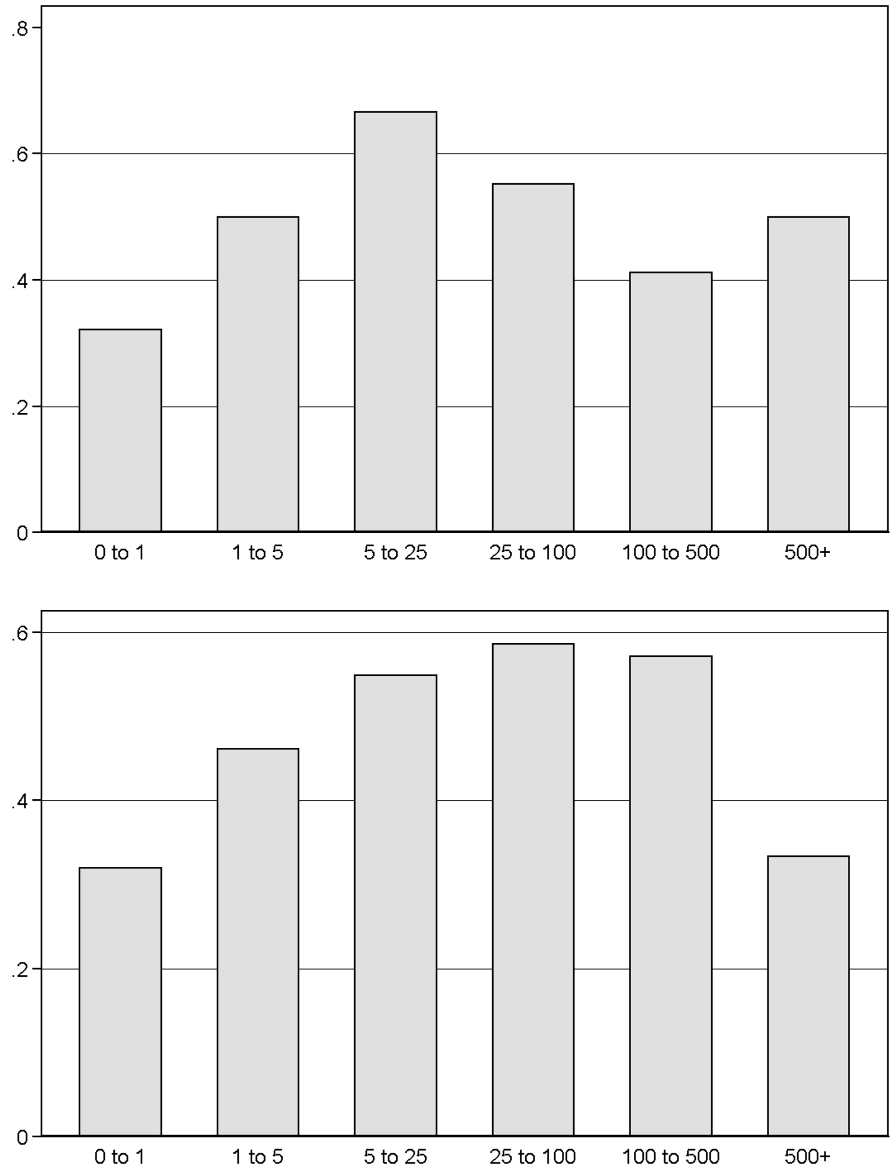
Because of the uncertainties involved in reconstructing historical population density, I replicate the results with alternative proxies. In column (4) of Table 4, I show that the main measure of population density is correlated with an indicator of land shortage ($v1720$) from the *SCCS*. In Columns (5) through (10), I show that the main results for land rights and slavery can be replicated with three alternative measures of population density – density in 1995, and densities computed using McEvedy and Jones (1978) or ARVE as in (8). If the principal measure of historical population density is replaced for Canada and the United States with the estimates reported in Ruff (2006), the results (not reported) are very similar to those given in Table 2. If I do not weight these regional densities, I find an inverse-U but insignificant relationship with slavery using the McEvedy and Jones (1978) estimates, and a significant inverse-U using the ARVE data (not reported).

In addition, there are two variables in the *SCCS* ($v64$ and $v1130$) that create independent estimates of the population densities of several societies in the data. These are not continuous measures, but instead categorize the societies into bins. While there are not enough observations and the data are too coarse to replicate the econometric analysis, I show in Figure 6 that these alternative measures have inverse-U relationships with slavery.

Though any historical population density estimate is untrustworthy on its own. The various measures here agree on two conclusions. Land rights have existed where population was densest, and slavery was most likely at intermediate values of population density, as in the model and consistent with the literature on African history.

5.3. **Other robustness checks.** In Table 5, I test whether the results are sensitive to the inclusion of influential observations and sub-samples. I begin by re-estimating the results by Ordinary Least Squares (OLS), and computing both leverage and dfbeta statistics for the variables of interest. In the slavery quadratic, this calculated for the linear term. In columns (1) through (4) and (11) through (14), it is clear that the results do not depend on including these observations. In columns (5), (6), (15), and (16), I replicate the results excluding both North and South America. The results are unchanged excepting that the relationship between slavery and land quality becomes small and insignificant. This is surprising, as slavery within the Americas was most prominent in areas of the Pacific Northwest where agriculture was unimportant.

FIGURE 6. Slavery by bins of population density in the *SCCS*



The y axis is the percentage of societies with slavery. The top picture divides this by population density bins according to $v64$, while the bottom picture does so following $v1130$.

Dropping Europeans (columns 7-8 and 17-18) and their offshoots does not change the results by much. Excluding non-agricultural societies (columns 9-10 and 19-20) completely eliminates the relationship between land quality and land rights, suggesting that this is driven by better land quality permitting the existence of settled agriculture. This highlights a mechanism by which societies move from $\mathcal{S}^E$ to $\mathcal{S}^S$, rather than providing evidence against the model.

TABLE 5. Influential observations

[Table 5 here]

TABLE 6. Possible endogeneity of land quality

[Table 6 here]

No sub-set of the data are determining all of the results. In Table A1 in the appendix, I list the most ten most influential societies by dfbeta for each of the major coefficients of interest. The societies that drive the relationships between land rights and the two variables of interest are not concentrated in any one region. For the correlation of land quality with slavery, four of the ten most influential societies are slave-holding American Indian groups concentrated in the fertile parts of the prairies. For the positive coefficient on the log of historical population density, three of the most influential societies are moderately populated slave-owning societies of Central Asia. Three of the ten most influential societies for the negative coefficient on the quadratic population term are in densely-settled Northern India and do not possess slaves.

I am not concerned here with possible reverse causation of population density. The model expects that population growth will respond to institutions, and I am only testing a correlation between two endogenous variables. I am, however, potentially concerned about the endogeneity of land quality. The FAO measure is an index of several constraints, of which soil depth and soil fertility may be potentially human-caused. In Table 6, I address this concern by controlling directly for these components. If the entire relationship between the variables of interest and land quality can be explained by correlation with these potentially endogenous components, it is evidence that the causal inference may be spurious. The results show, however, that the result survives separating land quality into its separate parts.

Finally, I have reported results with heteroskedasticity-robust standard errors. How sensitive is the statistical inference to correlations in the errors within possible clusters of observations? I address this question in Table 7, clustering the standard errors by ethnographic region (of which there are 60), by the principal country of the ethnic group, or by that country's global region as classified by the UN.[15] The "robust" errors are the baseline results. The results are generally stable, though clustering by country pushes the p. values of the population quadratic to 0.15. The major exception is that

[15]Ethnic groups were classified according to the location of their centroid, and then obvious errors were corrected manually. For example, the centroids for the Japanese and Annamese fall outside of Japan and Vietnam, respectively.

TABLE 7. Alternative clusters (p. values)

[Table 7 here]

slavery is not significantly related to land quality if the results are made robust to arbitrary correlation by country or by UN region. This foreshadows the results of Section 6, suggesting that there are strong correlations in institutions within broad regions.

5.4. **Other theories of land rights.** In this section, I contrast the model with other explanations of land rights. The two most influential are those of Boserup (1965) and Demsetz (1967). Boserup (1965) argues that exogenous population increase is the principal driver of agricultural intensification and more permanent tenure. This is the intuition captured by the Lagerlöf (2009) model; an increase in P_t pushes down the average product of land under egalitarianism, creating incentives for the elite to enclose it. Other formalizations of this argument have captured these changes as the selection of a new, more intensive production technology in response to changes in the relative scarcity of land and labor (e.g. Hayami (1997); Quisumbing and Otsuka (2001)).

The strong correlation of population density and land rights identified in Section 4 supports this view. Further, this result supports the model against informal critiques of Boserup (1965) that have argued that population pressure can lead to multiple outcomes, including open access (e.g. Baland and Platteau (1998)). The tendency in the data is for population pressure to be associated with more defined rights. These data do not, of course, allow the model to be tested against all possible mechanisms by which population density is positively correlated with the existence of rights over land. Specific alternatives can be evaluated. It may be supposed, for example, that population density affects land rights only through the existence of states. Re-estimating (6) separately for stateless societies and those with states, however, gives a positive and significant coefficient on historic population density in both sub-samples (not shown).[16]

Demsetz (1967), by contrast, focuses on trade. He argues that land rights internalize externalities when the gains outweigh the costs. This drives enclosure of the commons in the formal treatments of Hotte et al. (2000) or Copeland and Taylor (2009), and explains the empirical results of Bogart and Richardson (2010). It is similar to the greater effort expended in defending rights over more valuable resources predicted by models of the economics of conflict (e.g. Baker (2003); Grossman and Kim (1995)). These motivations are not inconsistent with the land abundance view, and are incorporated into Austin's (2008a) account of it. The data do not include direct information on trade. The

[16]"Stateless" societies and those with states are defined according to $V33$ of the *Ethnographic Atlas*, which captures jurisdictional hierarchy above the local community

TABLE 8. Alternative theories

[Table 8 here]

two controls that best capture trade in the data – distance from the coast and access to a major river – do not significantly predict the existence of land rights in Table 2.

Beyond these two influential theories, there is a literature on the enclosure of common property (e.g. Baland and Francois (2005); Baland and Platteau (2003); Grantham (1980); Lueck (1994); Netting (1976); Ostrom (1991); Runge (1986)). These works identify several benefits of common property that help explain why it survives. These include scale economies, risk pooling, exclusion and effort costs, and equity concerns. None of these explanations exclude land abundance as an explanation of weakly defined land rights. Most of these variables are not available in the data, making it impossible to test the land abundance view against them. The exception to this is risk. In Table 8, I add a measure of ecological risk, including the coefficient of variation of annual rainfall over the period 1950-1999.[17] In column (1), there is a significant negative coefficient; without additional controls, it does appear that added risk helps explain a lack of rights to land, though this does not diminish the direct effect of land quality. Adding log population density or the full set of controls, however, leads the effect of risk to become insignificant.

5.5. **Other theories of slavery.** Several theoretical analyses of slavery and coercion exist (e.g. Barzel (1977); Bergstrom (1971); Canarella and Tomaske (1975); Findlay (1975); Genicot (2002)). Some, e.g. Conning (2004), formally capture the same intuition as Nieboer (1900) or Domar (1970), which is the basis of the Lagerlöf (2009) model. In this section, I contrast the results outlined in Section 4 with two other broad explanations of slavery.

First, several theories emphasize coerced workers' outside options. These include what Acemoglu and Wolitzky (2011) refer to as the "neo-Malthusian" explanations of the decline of serfdom. North and Thomas (1971), for example, hold that serfs voluntarily exchanged their labor for protection from lords. These payments were in inputs rather than money because of the limited nature of output markets. Critically, North and Thomas (1971) suggest that lower population densities after the black death led lords

[17]Data here come from the University of Delaware Center for Climatic Research.

to compete for labor, weakening serfdom.[18] Several models find that worse outside options for workers increase the degree of coercion in labor contracts (Beber and Blattman, 2011; Chwe, 1990; Naidu and Yuchtman, 2011). Similarly, Acemoglu and Wolitzky (2011) find that labor scarcity has two effects, raising coercion through a Domar-type increase in the price of output, but also reducing coercion by improving workers' outside options.

If workers' outside options help explain the existence of slavery, this does not rule out labor scarcity as an explanation. Workers' outside options are not directly observed in the data, making it impossible to test the model against this explanation directly. Following the narrative of Nunn and Puga (2011), ruggedness is expected to improve the outside option of slaves by making it easier for them to flee. Contrary to this intuition, however, the marginal effect ruggedness on slavery in Table 2 is positive.

There are four reasons North and Thomas (1971) cannot explain Africa. First, Fenoaltea (1975), who demonstrates that they err in treating serfdom as voluntary, underestimate the transactions costs in labor contracts, misidentify the historical trends that acted on the manorial system, and overemphasize the rigidity of "custom." Second, I show land quality and population density at low levels are positively associated with slavery. North and Thomas (1971) predict these would promote the development of trade and markets, lessening the need for contracts to be written in labor dues. Third, their prediction that trade will discourage the use of serfs runs counter to African history, in which external trade spurred greater use of slaves in production (e.g. Lovejoy (2000) or Law (1995)). Finally, there is no evidence that African slaves received payments that approximated their marginal products. In many cases, slaveowners had to be compelled to receive manumission payments.

The second set of theories I address argue that, in certain contexts, slavery is more productive than free labor, which explains its use. For Fenoaltea (1984), this occurs where "pain incentives" are effective and detailed care by the worker is unnecessary. Fogel and Engerman (1974) link the exceptional productivity of slaves in the American south to economies of scale that could only be achieved through gang labor, an activity so grueling that free men could not be induced to take part at any price. Engerman and Sokoloff (1997), similarly, argue that the cultivation of crops with economies of scale is more conducive to slavery. Hanes (1996) explains the concentration of slaves in rural and domestic production by invoking the high turnover costs in these industries.

[18]Writers such as Inikori (1999) have suggested that African "slaves" held a position closer to that of the European serf. In the model, slaves differ from free laborers in that they are coerced workers whose price does not depend on the local supply of labor. The severity of slavery is not important to this conceptual distinction.

These arguments again cannot alone explain slavery in Africa, even if they can explain it in other contexts. First, there is no evidence that slaves were used in production in sectors systematically different than those dominated by free peasants. Studies of slavery in individual African societies frequently make reference to slave labor and free labor working in the same tasks. Austin (2005) notes gold and kola production in Asante were both carried out by free people, pawns, *corvée* labor, slaves, and descendants of slaves. Uchendu (1979) shows for Igbo society that "[i]n domestic activities ... no operation was strictly reserved for slaves." Describing the Kerebe of Tanzania, Hartwig (1979) writes that masters often worked alongside their slaves, who performed the same tasks as their owners and their owners' wives. Second, the literature on the "legitimate commerce" period suggests that slaves were used in the activities where labor of all kinds was most productive; in the model this is consistent with a rise in $\bar{A}$, and does not require a different production function under slavery. The nineteenth century export markets for oils, ivory, ostrich feathers and other goods created higher returns to slave labor, and slavery within Africa intensified (Law, 1995; Lovejoy, 2000; Lynn, 1997) Third, African agriculture both past and present has been overwhelmingly characterized by diminishing or constant returns to scale (Hopkins, 1973). Without evidence of scale economies, an appeal to "pain incentives" is not necessary to explain slavery over and above a comparison of the costs of slavery to those of free labor.

The data only allow limited tests of the land abundance view against these arguments, since economies of scale or the detailed care needed in production are not observed. The FAO does, however, report the suitability of land for eight classes of rainfed crop – wheat, maize, cereals, roots/tubers, pulses, oil crops, sugar, and cotton. I test whether including these measures in the slavery regressions has any effect on the results. This tests the Lagerlöf (2009) model against the alternative that slavery is explained by its productivity in the production of specific crops. I report the results in Table 8. With one exception, none of these specific crops has a major impact on the coefficient on land quality. The exception is maize suitability, which weakens the effect of land quality, making it marginally insignificant. Since the main effect of maize is itself insignificant and the two variables have a high raw correlation of 0.44, this is likely the result of multicollinearity.

6. Heterogeneity: What the model cannot explain

In Table 9, I show a simple method to do away with most of the results presented so far: add fixed effects for the major ethnographic regions in the data. These are North America, South America, Africa, the Circum-Mediterranean, the Insular Pacific, and

TABLE 9. Results with region fixed effects

[Table 9 here]

TABLE 10. Galton's problem

[Table 10 here]

East Eurasia. There is still a relationship between population density and land rights, and the marginal effect of land quality on slavery has not fallen by much, but the other results have now disappeared completely. The model can predict differences across broad regions, but not within them.

Why? Anthropologists have a name for the diffusion of institutions across societies: "Galton's problem." Economists would refer to this as serial correlation or spatial dependence. I propose that the lack of robustness of the main results stems from spillovers. If a nearby society has slavery, it is almost impossible to avoid developing the institution or becoming slaves of your neighbor, regardless of prevailing land quality and population density. The existence of rights over land is an idea that can spread across societies, and can be used to defend claims against a rival group. Within the Lagerlöf (2009) model, these make sense as parameter shifts dependent on a neighboring group's institutions. The cost of slavery (γ), for example, should be lower if a slave who flees can only do so to another slave-holding society.

In Table 10, I provide suggestive evidence that these neighbor effects exist by estimating spatial lag and spatial error models. The spatial lag adds a term $\rho W y$ to the estimating equation. W is an $N \times N$ spatial weight matrix, in which each entry W_{ij} is the inverse of the distance between observation i and observation j, normalized so that its rows sum to 1 or 0. ρ captures whether the institutional outcome of one group will affect its neighbor's institutions. The reason this evidence is only suggestive is that ρ is not separately identified from localized unobservables. This is estimated as a linear probability model using maximum likelihood.[19] The spatial error model is similar. Now, the error term is given by $u = \lambda W u + \epsilon$, so that a society's random error may depend on the error terms for societies that are close to it.

In Table 10 it is clear that there is very strong spatial correlation in land rights. The Wald tests for ρ and λ are very large, even conditional on the observed controls. Once controls are added, none of the results concerning land quality survive. The results with population density fare better, but for slavery these are only marginally significant in

[19]In particular, I use the spatreg command in Stata.

the spatial lag model. While the model can explain differences across regions, it cannot explain differences within them, and the strong spatial correlation in institutional outcomes suggests this is due to neighbor effects.

I confirm the ability of the model to explain differences across regions in Figures 7 and 8. I show that the relationships between the averages of land quality and population density within an ethnographic region are correlated with the fraction of societies possessing land rights or slavery as the model predicts. The positive relationships of land rights with both land quality and population density are still apparent, and the inverse-U correlation between slavery and population density is still apparent. Only the correlation between slavery and land quality cannot be seen across regions in the data. Once again, the existence of multiple steady states can explain this. In Table A2, in the appendix, I report the regional means for these variables for the six major regions of the *Ethnographic Atlas*.

7. CONCLUSION AND DISCUSSION

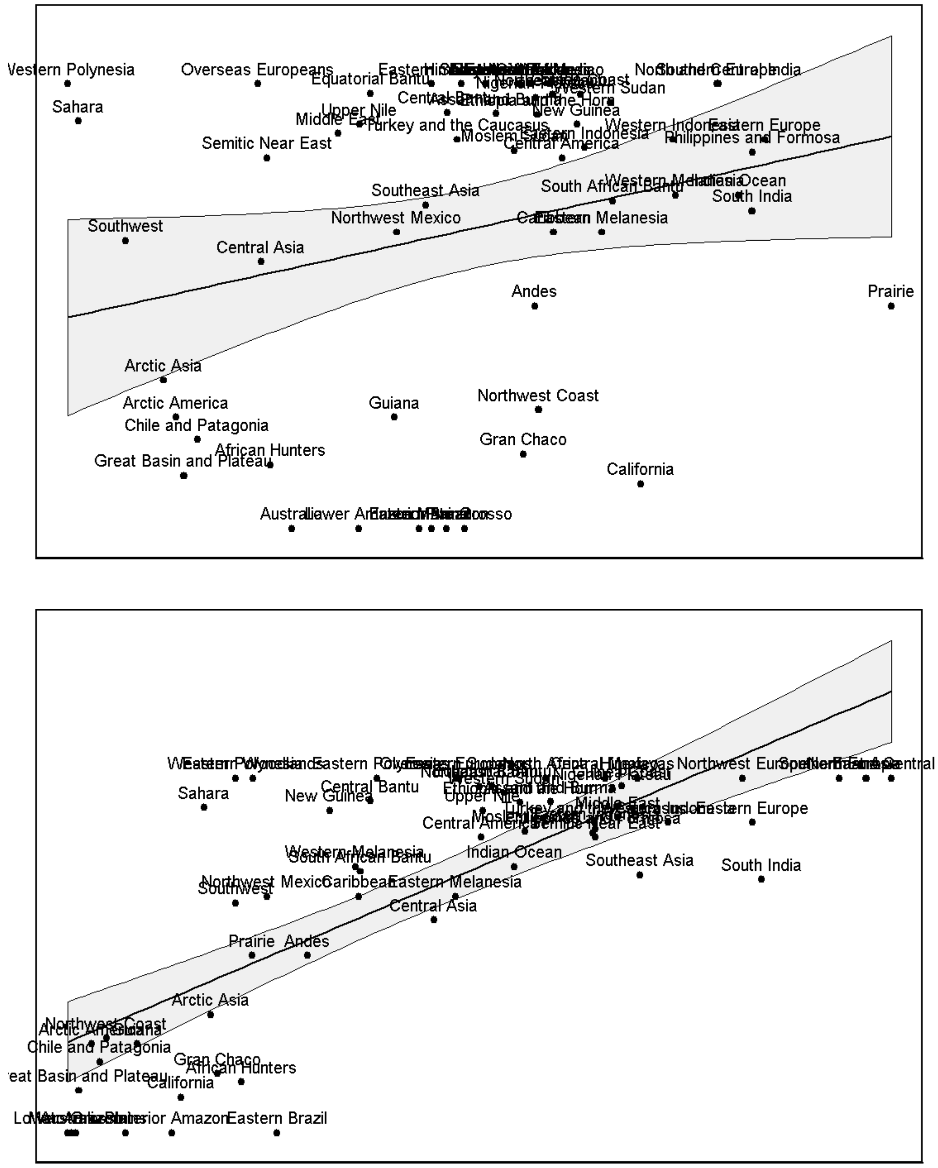
It appears then, that, the land abundance view performs reasonably well in predicting broad differences in the prevalence of land rights and slavery between Africa and the rest of the world, though not as well at predicting outcomes within regions. What of other institutions discussed by historians of Africa? The relative lack of state centralization and high rates of polygyny have also been tied to sparse population. Rulers were unable to tie subjects to the land and tax them, sought subjects and cattle, rather than territory, and had to contend with the ability of subjects to exit easily (Austin, 2004a,b). Goody (1976) argues that polygyny exists where allocating land to additional wives is less costly but their labor is valuable.

In Table 11, I replicate (5), (6) and (7) with states and polygamy as outcomes. The prevalence of states in the global sample mimics that of rights over land, rising monotonically with land quality and population density.[20] Polygyny, by contrast, mimics the pattern seen for slavery – its presence increases weakly with land quality, but is strongest at intermediate levels of population.[21] The relationships between polygyny and the controls of interest are not robust to the inclusion of additional controls – malaria ecology is sufficient to make either one insignificant. This suggests that the land abundance view may have some power to explain the relative prevalence of states, though its application to polygamy may be more limited.

[20] I measure state centralization as a dummy variable, equal to one if variable 33 in the *Ethnographic Atlas*, the levels of jurisdiction above the local, is greater than one.

[21] I measure polygyny as a dummy variable, equal to one if variable 9 in the *Ethnographic Atlas*, marital composition, is 3, 4, 5, or 6. This codes outcome 2, "Independent nuclear, occasional polygyny", as zero.

FIGURE 7. Land rights, land quality and population density across regions



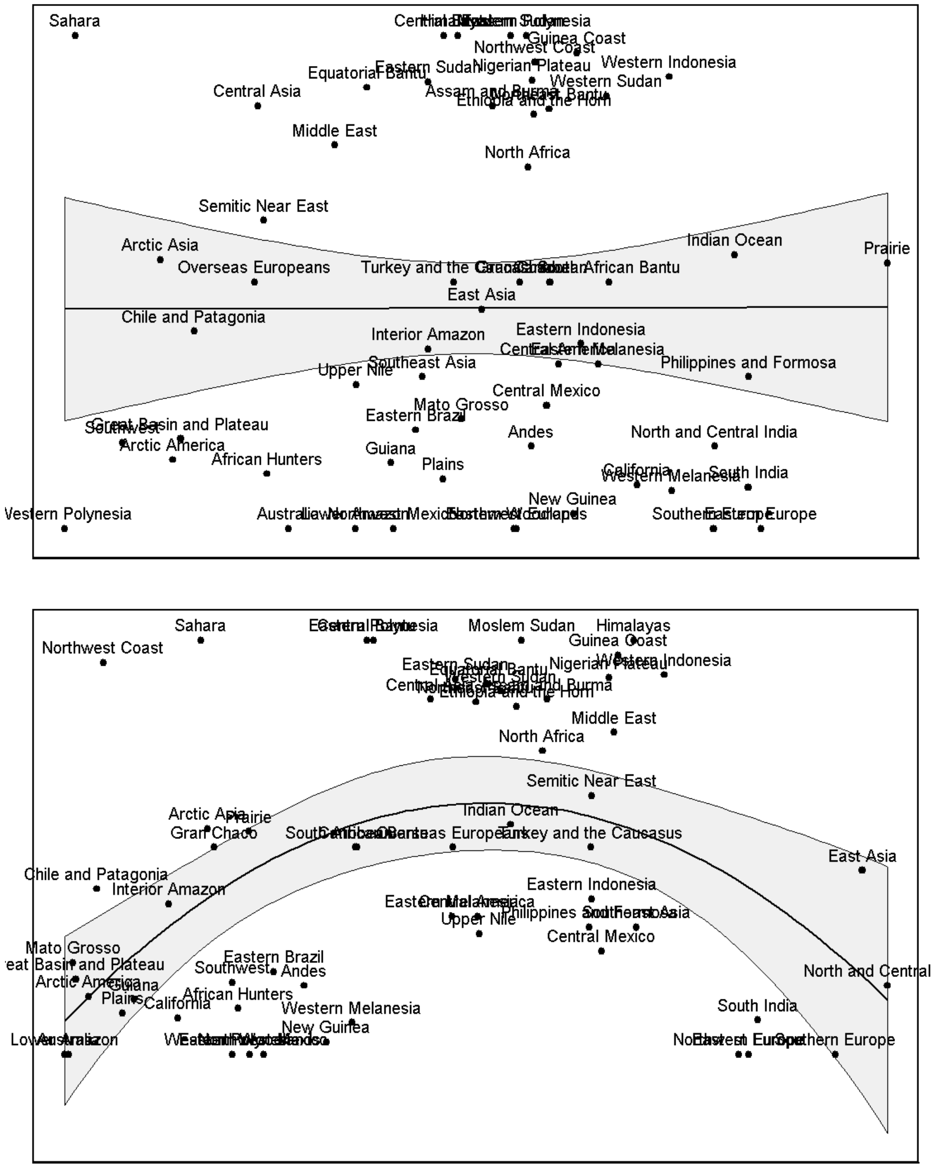
The y axis is the percentage of societies within a region with land rights. The x axis in the top picture is average land quality for the region. In the bottom picture it is average log population density.

TABLE 11. Other outcomes

[Table 11 here]

Bad institutions are one of the fundamental causes of African poverty, and the institutions that exist on the continent currently have been shaped by those that existed prior to colonial rule. I have addressed a theme in the economics literature – how geography affects institutions – by looking in depth at one hypothesis from the literature on African

FIGURE 8. Slavery, land quality and population density across regions



The y axis is the percentage of societies within a region with slavery. The x axis in the top picture is average land quality for the region. In the bottom picture it is average log population density.

history. I find that African land tenure and slavery have been decisively shaped by the continent's abundance of land and scarcity of labor. I find that this perspective explains much about institutions across a global cross-section of societies, but that neighbor effects weaken its ability to predict differences within them.

These tests have made several points that must be taken into account in understanding the impacts of under-population on African institutions. First, when both productivity and population are low, the opportunity cost of coercion is high, and the benefit to creating estates is low. This explains why slavery is less common among the most sparsely populated societies. Africa appears not as the least populous region in the sample, but as one that of medium density. While it is comparatively more prone to slavery than Europe or South Asia, there the is more slavery on the continent than in many parts of the Americas. Second, greater land quality (as well as access to trade), will encourage increased reliance on slavery conditional on population. This explains why some of the most agriculturally prosperous though densely populated regions in Africa, such as Sokoto, also used slaves most intensively (cf. Hill (1985)). Finally, there are substantial institutional spatial correlations across African societies relating to land rights and slavery. These revisions to the current thinking allow the "land abundance" perspective to better explain institutions and are borne out in comparative data.

References

Acemoglu, D. and Wolitzky, A. (2011). The economics of labor coercion. *Econometrica*, 79(2):555–600.

Andre, C. and Platteau, J.-P. (1998). Land relations under unbearable stress: Rwanda caught in the Malthusian trap. *Journal of Economic Behavior and Organization*, 34(1):1–47.

Austin, G. (2004a). Markets with, without, and in spite of states: West Africa in the pre-colonial nineteenth century. *Working Paper*.

Austin, G. (2004b). *Sub-Saharan Africa: land rights and ethno-national consciousness in historically land abundant economies*, pages 276–293. Land Rights, Ethno-nationality and Sovereignty in History. Routledge, London.

Austin, G. (2005). *Labour, land, and capital in Ghana: from slavery to free labour in Asante, 1807-1956*. University of Rochester Press, Rochester, NY.

Austin, G. (2008a). Resources, techniques, and strategies south of the Sahara: revising the factor endowments perspective on African economic development, 1500-2001. *The Economic History Review*, 61(3):587–624.

Austin, G. (2008b). The "reversal of fortune" thesis and the compression of history: Perspectives from African and comparative economic history. *Journal of International Development*, 20(8):996–1027.

Austin, G. (2009). Factor markets in Nieboer conditions: Pre-colonial West Africa, c. 1500-c. 1900. *Continuity and Change*, 24(01):23–53.

Baker, M. J. (2003). An equilibrium conflict model of land tenure in hunter-gatherer societies. *Journal of Political Economy*, 111(1):124–173.

Baland, J. M. and Francois, P. (2005). Commons as insurance and the welfare impact of privatization. *Journal of Public Economics*, 89(2-3):211–231.

Baland, J.-M. and Platteau, J.-P. (1998). Division of the commons: A partial assessment of the new institutional economics of land rights. *American Journal of Agricultural Economics*, 80(3):644–650.

Baland, J. M. and Platteau, J. P. (2003). *Economics of common property management regimes*, volume 1 of *Handbook of Environmental Economics*, pages 127–190. Elsevier.

Barzel, Y. (1977). An economic analysis of slavery. *Journal of Law and Economics*, 20(1):87–110.

Beber, B. and Blattman, C. (2011). The logic of child soldiering and coercion. *Unpublished working paper, Yale University*.

Bergstrom, T. (1971). On the existence and optimality of competitive equilibrium for a slave economy. *The Review of Economic Studies*, 38(1):23–36.

Berry, S. (1988). Property rights and rural resource management: The case of tree crops in West Africa. *Cahier des sciences humaines*, 24(1):3–16.

Bluedorn, J., Valentinyi, A., and Vlassopoulos, M. (2009). The long-lived effects of historic climate on the wealth of nations. *Centre for Economic Policy Research Working Paper*.

Bogart, D. and Richardson, G. (2010). Property rights and parliament in industrializing Britain. *NBER Working Papers*.

Boserup, E. (1965). *The conditions of agricultural growth: The economics of agrarian change under population pressure*. Aldine Pub. Co., New York.

Bouwman, A., Kram, T., and Klein Goldewijk, K. (2006). *Integrated modelling of global environmental change: an overview of Image 2.4*. Netherlands Environmental Assessment Agency Bilthoven,, The Netherlands.

Canarella, G. and Tomaske, J. (1975). The optimal utilization of slaves. *The Journal of Economic History*, 35(3):621–629.

Chwe, M. (1990). Why were workers whipped? Pain in a principal-agent model. *The Economic Journal*, 100(403):1109–1121.

Conning, J. (2004). The causes of slavery or serfdom and the roads to agrarian capitalism: Domar's hypothesis revisited. *Hunter College: Department of Economics*.

Copeland, B. and Taylor, M. (2009). Trade, tragedy, and the commons. *The American Economic Review*, 99(3):725–749.

Demsetz, H. (1967). Toward a theory of property rights. *The American Economic Review*, 57(2):347–359.

Dippel, C. (2010). Forced Coexistence and Economic Development: Evidence from Native American Reservations. *Working Paper*.

Domar, E. D. (1970). The causes of slavery or serfdom: A hypothesis. *The Journal of Economic History*, 30(1):18–32.

Donald, L. (1997). *Aboriginal slavery on the northwest coast of North America*. Univ of California Press.

Dowd, J. (1917). Slavery and the slave trade in Africa. *Journal of Negro History*, 2(1):1–20.

Easterly, W. and Levine, R. (2003). Tropics, germs, and crops: How endowments influence economic development. *Journal of Monetary Economics*, 50(1):3–39.

Engerman, S. L. and Sokoloff, K. L. (1997). *Factor endowments, institutions, and differential paths of growth among new world economies*, pages 260–304. How Latin America Fell Behind. Essays on the Economic Histories of Brazil and Mexico, 1800-1914. Stanford University Press.

Fage, J. D. (1969). Slavery and the slave trade in the context of West African history. *Journal of African History*, 10(3):393–404.

Fenoaltea, S. (1975). The rise and fall of a theoretical model: The manorial system. *Journal of Economic History*, 35(2):386–409.

Fenoaltea, S. (1984). Slavery and supervision in comparative perspective: A model. *The Journal of Economic History*, 44(3):635–668.

Field, E. (2007). Entitled to work: Urban property rights and labor supply in Peru. *Quarterly Journal of Economics*, 122(4):1561–1602.

Findlay, R. (1975). Slavery, incentives, and manumission: a theoretical model. *The Journal of Political Economy*, 85(5):923–933.

Fischer, G., Shah, M., Velthuizen, H., and Nachtergaele, F. O. (2002). *Global Agro-Ecological Assessment for Agriculture in the 21st Century: Methodology and results*. International Institute for Applied Systems Analysis/Food and Agriculture Organization of the United Nations, Laxenburg and Rome.

Fogel, R. W. and Engerman, S. L. (1974). *Time on the cross: The economics of American Negro slavery*. Little, Brown and Company, Boston.

Genicot, G. (2002). Bonded labor and serfdom: a paradox of voluntary choice. *Journal of Development Economics*, 67(1):101–127.

Gennaioli, N. and Rainer, I. (2007). The modern impact of precolonial centralization in Africa. *Journal of Economic Growth*, 12(3):185–234.

Goldstein, M. and Udry, C. (2008). The profits of power: Land rights and agricultural investment in Ghana. *Journal of Political Economy*, 116(6):981–1022.

Goody, J. (1969). Inheritance, property, and marriage in Africa and Eurasia. *Sociology*, 3(1):55–76.

Goody, J. (1976). *Production and reproduction: A comparative study of the domestic domain*. Cambridge University Press, Cambridge.

Goody, J. (1980). *Slavery in time and space*, pages 16–42. Asian and African Systems of Slavery. University of California Press, Berkeley and Los Angeles.

Grantham, G. W. (1980). The persistence of open-field farming in nineteenth-century France. *Journal of Economic History*, 40(3):515–531.

Grossman, H. I. and Kim, M. (1995). Swords or plowshares? A theory of the security of claims to property. *The Journal of Political Economy*, 103(6):1275–1288.

Hanes, C. (1996). Turnover cost and the distribution of slave labor in Anglo-America. *The Journal of Economic History*, 56(2):307–329.

Hartwig, G. W. (1979). *Changing forms of servitude among the Kerebe of Tanzania*, pages 261–287. Slavery in Africa: historical and anthropological perspectives. University of Wisconsin Press.

Hayami, Y. (1997). *Development economics: From the poverty to the wealth of nations*. Oxford University Press, USA.

Heizer, R. and Sturtevant, W. (1978). Handbook of North American Indians.

Helders, S. (2000). *The world gazetteer*.

Henige, D. (1998). *Numbers from nowhere: the American Indian contact population debate*. University of Oklahoma Press.

Herbst, J. (2000). *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton University Press.

Hill, P. (1985). *Comparative West African Farm-slavery Systems*, volume 2 of *Slaves and Slavery in Muslim Africa*, pages 33–50. Frank Cass, London.

Hopkins, A. G. (1973). *An economic history of West Africa*. Columbia University Press, New York.

Hopkins, A. G. (2009). The new economic history of Africa. *The Journal of African History*, 50(02):155–177.

Hotte, L., Long, N., and Tian, H. (2000). International trade with endogenous enforcement of property rights. *Journal of Development Economics*, 62(1):25–54.

Iliffe, J. (1995). *Africans: the history of a continent*. Cambridge University Press, Cambridge.

Inikori, J. E. (1999). *Slaves or Serfs? A Comparative Study of Slavery and Serfdom in Europe and Africa*, pages 49–75. The African Diaspora: African Origins and New World Identities. Indiana Univ Press.

Klein Goldewijk, K. (2005). Three centuries of global population growth: A spatial referenced population (density) database for 1700–2000. *Population & Environment*, 26(4):343–367.

Klein Goldewijk, K., Beusen, A., and Janssen, P. (2010). Long-term dynamic modeling of global population and built-up area in a spatially explicit way: Hyde 3.1. *The Holocene*, 20(4):565.

Kopytoff, I. (1987). *The internal African frontier: The making of African political culture*, pages 3–84. The African frontier: The reproduction of traditional African societies. Indiana University Press.

Kopytoff, I. and Miers, S. (1977). African 'slavery' as an institution of marginality. *Slavery in Africa: historical and anthropological perspectives*, pages 3–81.

Krumhardt, K. (2010). Methodology for worldwide population estimates: 1000 BC to 1850.

Lagerlöf, N. (2010). Pacifying monogamy. *Journal of Economic Growth*, pages 1–28.

Lagerlöf, N. P. (2009). Slavery and other property rights. *Review of Economic Studies*, 76(1):319–342.

Lahmeyer, J. (2004). Populstat database. growth of the population per country in a historical perspective, including their administrative divisions and principal towns.

Law, R. (1995). *From Slave Trade to "Legitimate" Commerce*. Cambridge University Press, Cambridge.

Lovejoy, P. E. (1978). Plantations in the economy of the Sokoto Caliphate. *Journal of African History*, 19:341–368.

Lovejoy, P. E. (2000). *Transformations in Slavery: A History of Slavery in Africa*. Cambridge University Press.

Lueck, D. (1994). Common property as an egalitarian share contract. *Journal of Economic Behavior & Organization*, 25(1):93–108.

Lynn, M. (1997). *Commerce and economic change in West Africa: The palm oil trade in the nineteenth century*. Cambridge University Press, Cambridge.

Mahadi, A. and Inikori, J. E. (1987). *Population and capitalist development in precolonial West Africa: Kasar Kano in the nineteenth century*, pages 62–73. African Population and Capitalism: Historical Perspectives. The University of Wisconsin Press, Madison, Wisconsin.

McEvedy, C. and Jones, R. (1978). *Atlas of world population history*. Penguin books Harmondsworth, UK.

Miers, S. and Klein, M. A. (1998). Introduction. *Slavery & Abolition*, 19(2):1–15.

Miller, M. (2008). Essays on race and the persistence of economic inequality.

Murdock, G. and White, D. (1969). Standard cross-cultural sample. *Ethnology*, 8(4):329–369.

Murdock, G. P. (1959). *Africa: Its Peoples and Their Culture History*. McGraw-Hill.

Murdock, G. P. (1967). *Ethnographic Atlas*. University of Pittsburgh Press.

Naidu, S. and Yuchtman, N. (2011). Coercive contract enforcement: Law and the labor market in 19th century industrial Britain. *Working paper*.

Netting, R. M. (1976). What alpine peasants have in common: Observations on communal tenure in a Swiss village. *Human Ecology*, 4(2):135–146.

Nieboer, H. J. (1900). *Slavery as an industrial system: Ethnological researches*. M. Nijhoff, The Hague.

North, D. C. and Thomas, R. P. (1971). The rise and fall of the manorial system: A theoretical model. *The Journal of Economic History*, 31(4):777–803.

Nunn, N. (2008a). The long-term effects of Africa's slave trades. *The Quarterly Journal of Economics*, 123(1):139–176.

Nunn, N. (2008b). Slavery, inequality, and economic development in the Americas: an examination of the Engerman-Sokoloff Hypothesis. *Institutions and economic performance*, pages 148–181.

Nunn, N. and Puga, D. (2011). Ruggedness: The blessing of bad geography in Africa. *Forthcoming in The Review of Economics and Statistics*.

Ostrom, E. (1991). *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press.

Quisumbing, M. A. R. and Otsuka, K. (2001). *Land, Trees, and Women: Evolution of Land Tenure Institutions in Western Ghana and Sumatra*. IFPRI.

Ruff, C. (2006). *Population size, contact to Nadir*. Handbook of North American Indians: Volume 3, Environment, Origins, and Population. Smithsonian Institution, London.

Runge, C. F. (1986). Common property and collective action in economic development. *World Development*, 14(5):623–635.

Sacerdote, B. (2005). Slavery and the intergenerational transmission of human capital. *Review of Economics and Statistics*, 87(2):217–234.

Spear, T. T. (1997). *Mountain farmers: Moral economies of land and agricultural development in Arusha and Meru*. James Currey; University of California Press, Oxford; Berkeley, Calif.

Tertilt, M. (2005). Polygyny, fertility, and savings. *Journal of Political Economy*, 113(6):1341–1371.

Thornton, J. (1992). *Africa and Africans in the Making of the Atlantic World, 1400-1680*. Cambridge University Press.

Tobler, W. (1995). *The global demography project: National Center for Geographic Information & Analysis (US)*. National Center for Geographic Information and Analysis Santa Barbara.

Uchendu, V. C. (1979). *Slaves and slavery in Iboland, Nigeria*, pages 121–132. Slavery in Africa: historical and anthropological perspectives. University of Wisconsin Press.

Weidmann, N., Rød, J., and Cederman, L. (2010). Representing ethnic groups in space: A new dataset. *Journal of Peace Research*, 47(4):491.

Table 1: Summary statistics and percentiles of population density

	(1) Mean	(2) s.d.	(3) Min	(4) Max	(5) N	Pct.	(6) HYDE Estimate	(7) MJ Base	(8) ARVE Base
Any slavery	0.54	0.50	0	1	1,041	5	0.03	0.29	0.05
Any land rights	0.74	0.44	0	1	801	10	0.09	0.60	0.18
Land quality	1.33	0.90	0	3.98	1,206	15	0.28	1.21	0.28
Date observed	1,905	53.0	1,500	1,965	1,206	20	0.57	1.93	0.50
Historic pop density	42.8	141	0	3,627	1,206	25	1.22	2.58	0.94
Precipitation	1,263	858	12.6	6,164	1,206	30	2.42	3.87	2.00
Temperature	7,203	2,774	35.5	10,830	1,206	35	3.71	5.08	3.87
Absolute latitude	20.7	17.0	0.017	78.1	1,206	40	5.82	6.64	7.13
Pct. malarial	0.17	0.20	0	0.69	1,206	45	7.75	8.29	10.21
Dist. to coast	4.26	3.87	0	16.5	1,206	50	10.14	10.14	14.81
Elevation	167	9.61	141	230	1,206	55	12.63	13.16	19.39
Major river	0.28	0.45	0	1	1,206	60	15.85	17.38	24.35
Ruggedness	121,122	132,811	137	977,941	1,206	65	20.17	22.60	32.30
Share desert	0.11	0.26	0	1	1,206	70	26.07	29.88	40.24
Mostly fishing	0.069	0.25	0	1	1,206	75	35.40	39.17	54.98
						80	47.25	53.66	76.03
						85	63.27	71.90	105.73
						90	96.09	116.78	151.86
						95	164.72	198.82	246.07

Notes: Variable definitions in text.

Table 2. Main results

	(1)		(2)		(3)		(4)	
	Any land rights							
Land quality	0.091***	(0.017)	0.046***	(0.018)				
ln(1+pop. den.)					0.161***	(0.010)	0.124***	(0.011)
Precipitation			-0.047**	(0.021)			-0.022	(0.021)
Temperature			-0.028	(0.030)			-0.057*	(0.032)
Date observed			0.050***	(0.019)			0.004	(0.016)
Share desert			0.010	(0.018)			0.039**	(0.016)
Dist. to coast			-0.023	(0.018)			0.009	(0.018)
Elevation			-0.007	(0.019)			0.002	(0.018)
Pct. malarial			0.174***	(0.026)			0.127***	(0.025)
Ruggedness			0.064***	(0.017)			0.030*	(0.016)
Absolute latitude			-0.107***	(0.033)			-0.088**	(0.037)
Major river			-0.031	(0.034)			-0.079**	(0.034)
Mostly fishing			-0.125*	(0.074)			0.029	(0.071)
Observations		801		801		801		801
	(5)		(6)		(7)		(8)	
	Any slavery							
Land quality	0.040***	(0.015)	0.047**	(0.021)				
ln(1+pop. den.)					0.286***	(0.029)	0.109***	(0.038)
ln(1+pop. den.) sqrd.					-0.042***	(0.005)	-0.013*	(0.007)
Precipitation			-0.063**	(0.025)			-0.047*	(0.026)
Temperature			0.218***	(0.038)			0.217***	(0.038)
Date observed			-0.049***	(0.019)			-0.065***	(0.021)
Share desert			0.033	(0.022)			0.030	(0.020)
Dist. to coast			0.048**	(0.023)			0.054**	(0.024)
Elevation			0.013	(0.022)			0.014	(0.023)
Pct. malarial			0.386***	(0.030)			0.360***	(0.030)
Ruggedness			0.135***	(0.021)			0.121***	(0.022)
Absolute latitude			0.111***	(0.042)			0.136***	(0.043)
Major river			0.089**	(0.042)			0.083*	(0.042)
Mostly fishing			0.388***	(0.078)			0.416***	(0.081)
Observations		1,041		1,041		1,041		1,041
Pop. den. at peak slavery						30.76		71.14

Notes: ***Significant at 1%, **Significant at 5%, *Significant at 10%. Robust standard errors in parentheses. All regressions are probit, with marginal effects reported.

Table 3: Alternative measures of the dependent variables

	(1)	(2)	(3)	(4)	(5)
	Any land rights	*Land is patrilineal*		*Land inherited by sons*	
Land mostly private (v1726 in SCCS)	0.301**				
	(0.144)				
Land quality		0.050**		0.051**	
		(0.023)		(0.022)	
ln(1+pop. den.)			0.123***		0.183***
			(0.015)		(0.016)
Observations	80	801	801	801	801
Other cont.	N	Y	Y	Y	Y

	(6)	(7)	(8)	(9)	(10)
	Any slavery	*Slavery above incipient*		*Hereditary slavery*	
Large scale slaveholding (v919 in SCCS)	0.538***				
	(0.166)				
Land quality		0.037*		0.014	
		(0.021)		(0.015)	
ln(1+pop. den.)			0.218***		0.100***
			(0.042)		(0.029)
ln(1+pop. den.) sqrd.			-0.027***		-0.011**
			(0.007)		(0.005)
Observations	166	1,041	1,041	1,041	1,041
Other cont.	N	Y	Y	Y	Y

Notes: ***Significant at 1%, **Significant at 5%, *Significant at 10%. Robust standard errors in parentheses. All regressions are probit, with marginal effects reported. Other controls are as in the table of main results.

Table 4: Alternative measures of land quality and population density

	(1)	(2)	(3)	(4)		
	Land quality (v921)	*Land quality (v924)*	*Land quality (v928)*	*Land scarcity (v1720)*		
Land quality	1.677***	0.703***	0.871***			
	(0.254)	(0.112)	(0.107)			
ln(1+pop. den.)				0.080***		
				(0.030)		
Observations	172	172	172	79		
R-squared	0.223	0.196	0.274			
Other cont.	N	N	N	N		
	(5)	(6)	(7)	(8)	(9)	(10)
	Any land rights			*Any slavery*		
Ln pop density 1995	0.098***			0.123***		
	(0.010)			(0.046)		
squared.				-0.013**		
				(0.006)		
ln(1 + pop. den.) - MJ Base		0.124***			0.122***	
		(0.011)			(0.040)	
squared.					-0.012*	
					(0.007)	
ln(1 + pop. den.) - ARVE Base			0.124***			0.109***
			(0.011)			(0.038)
squared.						-0.013*
						(0.007)
Observations	801	801	801	1,041	1,041	1,041
Other cont.	Y	Y	Y	Y	Y	Y

Notes: ***Significant at 1%, **Significant at 5%, *Significant at 10%. Robust standard errors in parentheses. All regressions are probit, excepting columns 1-3, with marginal effects reported. Columns 1-3 are OLS. Other controls are as in the table of main results.

Table 5: Influential observations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Any land rights									
Land quality	0.059***	0.086***			0.032***		0.036*		0.004	
	(0.018)	(0.017)			(0.012)		(0.018)		(0.010)	
ln(1+pop. den.)			0.117***	0.123***		0.045***		0.121***		0.029***
			(0.013)	(0.012)		(0.008)		(0.012)		(0.006)
Observations	744	744	749	758	597	597	781	781	610	610
Dropped	Lev LQ	DfB LQ	Lev PD	DfB PD	Americas	Americas	Europeans	Europeans	NonAg	NonAg
Other cont.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
	Any slavery									
Land quality	0.039*	0.046**			0.010		0.055***		0.044*	
	(0.022)	(0.023)			(0.023)		(0.021)		(0.024)	
ln(1+pop. den.)			0.124***	0.178***		0.106**		0.121***		0.113**
			(0.042)	(0.042)		(0.049)		(0.039)		(0.044)
ln(1+pop. den.) sqrd.			-0.012	-0.022***		-0.013*		-0.013*		-0.012*
			(0.008)	(0.007)		(0.007)		(0.007)		(0.007)
Observations	983	973	980	974	687	687	1,016	1,016	759	759
Dropped	Lev LQ	DfB LQ	Lev PD	DfB PD	Amer.	Amer.	Euro.	Euro.	NonAg	NonAg
Other cont.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Notes: ***Significant at 1%, **Significant at 5%, *Significant at 10%. Robust standard errors in parentheses. All regressions are probit, with marginal effects reported. Other controls are as in the table of main results.

Table 6: Possible endogeneity of land quality

	(1)	(2)	(3)	(4)
	Any land rights		*Any slavery*	
Land quality	0.069***	0.056***	0.061***	0.084***
	(0.018)	(0.019)	(0.022)	(0.027)
Soil depth constraints	0.095***		0.053**	
	(0.019)		(0.021)	
Soil fertility constraints		0.020		0.074***
		(0.018)		(0.026)
Observations	801	801	1,041	1,041
Other cont.	Y	Y	Y	Y

Notes: ***Significant at 1%, **Significant at 5%, *Significant at 10%. Robust standard errors in parentheses. All regressions are probit, with marginal effects reported. Other controls are as in the table of main results.

Table 7: Alternative clusters (p. values)

		(1)	(2)	(3)	(4)
		Any land rights		*Any slavery*	
	Clustering				
Land quality	Robust	0.00	0.01	0.01	0.03
	Ethno. Region	0.02	0.05	0.01	0.02
	Country	0.00	0.03	0.37	0.08
	UN Region	0.00	0.08	0.34	0.10
ln(1+pop. den.)	Robust	0.00	0.00	0.00	0.00
	Ethno. Region	0.00	0.00	0.00	0.00
	Country	0.00	0.00	0.00	0.04
	UN Region	0.00	0.00	0.00	0.01
ln(1+pop. den.) sqrd.	Robust			0.00	0.05
	Ethno. Region			0.00	0.05
	Country			0.00	0.15
	UN Region			0.00	0.09
Observations		801	801	1,041	1,041
Other cont.		N	Y	N	Y

Notes: All regressions are probit, with marginal effects reported. Other controls are as in the table of main results.

Table 8: Alternative theories

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Any land rights							
Land quality	0.079***		0.045**					
	(0.017)		(0.018)					
Rainfall CV	-0.311**	0.042	-0.120	0.009				
	(0.143)	(0.113)	(0.138)	(0.138)				
ln(1+pop. den.)		0.162***		0.124***				
		(0.010)		(0.012)				
Observations	801	801	801	801				
Other cont.	N	N	Y	Y				
	Any slavery							
Other crop	*Wheat*	*Maize*	*Cereals*	*Roots/ tubers*	*Pulses*	*Oil crops*	*Sugar*	*Cotton*
Land quality	0.041*	0.036	0.048**	0.057***	0.039*	0.071***	0.049**	0.041*
	(0.023)	(0.023)	(0.023)	(0.022)	(0.022)	(0.023)	(0.021)	(0.022)
Other crop suitability	0.012	0.036	-0.004	-0.045*	0.032	-0.074***	-0.033	0.042*
	(0.021)	(0.025)	(0.028)	(0.025)	(0.024)	(0.028)	(0.025)	(0.025)
Observations	1,041	1,041	1,041	1,041	1,041	1,041	1,041	1,041
Other cont.	Y	Y	Y	Y	Y	Y	Y	Y
ln(1+pop. den.)	0.106***	0.103***	0.109***	0.111***	0.104***	0.122***	0.110***	0.109***
	(0.039)	(0.039)	(0.038)	(0.038)	(0.039)	(0.039)	(0.038)	(0.038)
ln(1+pop. den.) sqrd.	-0.012*	-0.012*	-0.013*	-0.012*	-0.013*	-0.014**	-0.013*	-0.013**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Other crop suitability	0.008	0.036	0.004	-0.042	0.032	-0.059**	-0.037	0.048*
	(0.021)	(0.024)	(0.027)	(0.026)	(0.024)	(0.026)	(0.026)	(0.025)
Observations	1,041	1,041	1,041	1,041	1,041	1,041	1,041	1,041
Other cont.	Y	Y	Y	Y	Y	Y	Y	Y
Corr(Land qual., suit.)	0.36	0.44	0.49	0.37	0.42	0.50	0.21	0.30

Notes: ***Significant at 1%, **Significant at 5%, *Significant at 10%. Robust standard errors in parentheses. All regressions are probit, with marginal effects reported. Other controls are as in the table of main results.

Table 9: Results with region fixed effects

	(1)	(2)	(3)	(4)
	Any land rights		*Any slavery*	
Land quality	0.014		0.037	
	(0.018)		(0.023)	
ln(1+pop. den.)		0.103***		0.017
		(0.013)		(0.042)
ln(1+pop. den.) sqrd.				-0.005
				(0.007)
Observations	801	801	1,041	1,041
Other cont.	Y	Y	Y	Y

Notes: ***Significant at 1%, **Significant at 5%, *Significant at 10%. Robust standard errors in parentheses. All regressions are probit, with marginal effects reported. Other controls are as in the table of main results.

Table 10: Galton's problem

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Spatial error model	*Any land rights*				*Any slavery*			
Land quality	0.031**		0.017		0.006		0.015	
	(0.014)		(0.015)		(0.014)		(0.015)	
ln(1+pop. den.)		0.083***		0.081***		0.087***		0.047
		(0.010)		(0.011)		(0.029)		(0.030)
ln(1+pop. den.) sqrd.						-0.018***		-0.009*
						(0.005)		(0.005)
Wald test (λ=0)	726.5	387.7	580.6	233.9	1556	1567	668.0	689.6
Observations	801	801	801	801	1,041	1,041	1,041	1,041
Other cont.	N	N	Y	Y	N	N	Y	Y

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Spatial lag model	*Any land rights*				*Any slavery*			
Land quality	0.029**		0.017		0.011		0.009	
	(0.013)		(0.013)		(0.012)		(0.014)	
ln(1+pop. den.)		0.069***		0.074***		0.067***		0.033
		(0.008)		(0.009)		(0.023)		(0.028)
ln(1+pop. den.) sqrd.						-0.014***		-0.007
						(0.004)		(0.005)
Wald test (ρ=0)	618.2	186.3	207.2	69.87	1517	1003	243.2	246.6
Observations	801	801	801	801	1,041	1,041	1,041	1,041
Other cont.	N	N	Y	Y	N	N	Y	Y

Notes: ***Significant at 1%, **Significant at 5%, *Significant at 10%. Robust standard errors in parentheses. All regressions are probit, with marginal effects reported. Other controls are as in the table of main results.

Table 11: Other outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Any state centralization				*Polygyny is usual*			
Land quality	0.057***		0.052***		0.044***		0.013	
	(0.016)		(0.019)		(0.015)		(0.019)	
ln(1+pop. den.)		0.157***		0.180***		0.266***		-0.042
		(0.011)		(0.014)		(0.029)		(0.038)
ln(1+pop. den.) sqrd.						-0.046***		-0.002
						(0.006)		(0.007)
Observations	1,075	1,075	1,075	1,075	1,172	1,172	1,172	1,172
Other cont.	N	N	Y	Y	N	N	Y	Y

Notes: ***Significant at 1%, **Significant at 5%, *Significant at 10%. Robust standard errors in parentheses. All regressions are probit, with marginal effects reported. Other controls are as in the table of main results.

APPENDICES NOT FOR PUBLICATION

1. Data appendix

1.1. **Definitions and sources of ethnographic data.** Six variables are computed from the *Ethnographic Atlas*:

(1) *Any slavery* is an indicator for whether $V70$, "Type of Slavery" is greater than 1, which indicates "absence or near absence." The other categories grouped together are "incipient or nonhereditary," "reported but type not identified," and "hereditary and socially significant."

(2) *Any land rights* is an indicator for whether $V74$, "Inheritance Rule for Real Property (land)" is greater than 1, which indicates "absence of individual property rights." The other categories grouped together are, first, "matrilineal (sister's sons)," second, "other matrilineal heirs," third, "children, with daughters receiving less," fourth, "children, equally for both sexes," fifth, "other patrilineal heirs," and, sixth, "patrilineal (sons)."

(3) *State centralization* is an indicator for whether $V33$, "Jurisdictional Hierarchy Beyond Local Community" is greater than zero. This variable ranges from zero levels to four levels.

(4) *Usual polygyny* is an indicator for whether $V9$ "Marital Composition: Monogamy and Polygamy,' is equal to 3 (Preferentially sororal, same dwelling), 4 (Preferentially sororal, separate dwelling), 5 (Non-sororal, separate dwellings), or 6 (Non-sororal, same dwelling). This excludes categories 1 (Independent nuclear, monogamous), 2 (Independent nuclear, occasional polygyny) and 7 (Independent polyandrous families).

(5) *Date observed* is the rough date at which the information on the society was recorded. Where this is missing, I impute it using the average value for other ethnic groups within the society's ethnographic region, of which there are 60 in the final sample. This is variable $V102$ in the atlas.

(6) *Mostly fishing* is an indicator for whether the society's percentage dependence on fishing is greater than 50%, computed from $V3$ in the *Ethnographic Atlas*.

A revised version of the Atlas has been made available for download in SPSS format by J. Patrick Gray at `http://eclectic.ss.uci.edu/~drwhite/worldcul/`. This is the version used for the present study.

Five variables are used from the *Standard Cross-Cultural Sample*. These are:

(1) *v64*: This is "Population Density." This divides societies into categories of < 1, $1-5$, $5-25$, $26-100$, $101-500$, and > 500 persons per square mile.
(2) *v921*: This is "Agricultural Potential 1: Sum of Land Slope, Soils, Climate Scales." It rates land quality on a scale between 4 and 23.
(3) *v924*: This is "Suitability of Soils for Agriculture". It ranges soils on a scale from 0 to 8, from "very poor" to "very fair."
(4) *v928*: This is "Agricultural Potential 2: Lowest of Land Slope, Soils, Climate Scales." It ranges soils on a scale from 0 to 8, from "very poor" to "very fair."
(5) *v1130*: This is "Population Density." This divides societies into categories of < 1, $1-5$, $5-25$, $26-100$, $101-500$, and > 500 persons per square mile.
(6) *v1720*: This is "land shortage," computed by converting the variable "Causes of Land Shortage" to 0 if there is "no land shortage" and 1 if there is "population pressure", "territorial invasions," or both.

The SCCS is available online at `http://eclectic.ss.uci.edu/~drwhite/sccs/`.

1.2. **Definitions and sources of population data.** The principal data for historic population density, from HYDE, are online at `ftp://ftp.mnp.nl/hyde/hyde31_final/`.

The raster data for population density in 1995, used to compute alternative historic population density, are downloaded from `http://www.iiasa.ac.at/Research/LUC/GAEZ/index.htm`. Regional estimates from **?** are **?** are not available online. The regional estimates from **?** are available at `http://ecospriv4.epfl.ch/index.php?dir=pub/&file=pop_landuse_data.tar.gz`. The UNEP estimates for Africa in 1960 can be downloaded from `http://na.unep.net/metadata/unep/GRID/AFPOP60.html`.

In constructing regional estimates from **?** and **?**, I impute values for intermediate years or years between the end of the data and 1995 exponentially. Where **?** report country-level estimates in recent years and broad regions (e.g. "The Sahel States") in earlier years, I divide the population among countries according to their ratio in the earliest year that they are separately reported. Population density estimates for 1995 were obtained from the World Bank for country-level regions, and populstat.info for sub-national regions.

1.3. **Definitions and sources of GIS data.** The main measure of *land quality* has been explained in detail in the text. The **?** alternative measure of land quality is available at `http://www.sage.wisc.edu/iamdata/`.

Major river: This is a dummy that equals one if a river with a rank of at least 6 according to the North American Cartographic Information Society (NACIS) intersects the ethnic group's territory. The data are taken from `http://www.naturalearthdata.com/`.

Dist. to coast: This is average distance from each point in the ethnic group's territory to the nearest point on the coast, in decimal degrees, calculated in ArcMap.

Elevation: This is average elevation for the ethnic group. Raster data are provided by the NACIS, downloaded from `http://www.naturalearthdata.com/`.

Pct. malarial: This is the fraction of the society's territory in which malaria is endemic, according to the Malaria Atlas Project, downloaded from `http://www.map.ox.ac.uk/data/`.

Precipitation: This is average annual precipitation (mm). Because some societies are too small for a raster point to fall within their territory, I impute missing data using the nearest raster point. The data are downloaded from `http://www.iiasa.ac.at/Research/LUC/GAEZ/index.htm`.

Ruggedness: This is a measure of terrain ruggedness used by **?**. This measures the elevation distance between a raster cell and its neighbors. The data are downloaded from `http://diegopuga.org/data/`.

Temperature: This is the accumulated temperature on days with mean daily temperature above 0°C, computed using monthly data from 1961 to 2000 collected by the Climate Research Unit (CRU) of the University of East Anglia. I treat 55537 is as an error code and drop these points. I impute missing values using the nearest raster point. The data are downloaded from `http://www.iiasa.ac.at/Research/LUC/GAEZ/index.htm`.

Absolute latitude: This is the absolute value of the latitude of the society's centroid.

Rainfall C.V.: This is the coefficient of variation of annual total rainfall over the period 1950-1999 for the point in the rainfall data closest to the ethnic group's centroid. These data are taken from the University of Delaware's Center for Climatic Research, and are downloaded from `http://climate.geog.udel.edu/~climate/html_pages/archive.html`.

Other crop suitability. This is average suitability for the chosen rain-fed crop, according to plates 29 through 36 of the FAO's GAEZ project. For comparison, the general land quality measure used in this paper is computed from plate 28. For each of these crops, suitability is a scale between 0 and 8, where 0 indicates very high suitability and 8 indicates non-suitability. This is re-scaled so that larger values indicate greater suitability, and is (like the other controls) converted to a standard normal variable for the regressions. The data are downloaded from `http://www.iiasa.ac.at/Research/LUC/GAEZ/index.htm`.

2. Other implications of the model

The additional parameters are $\bar{c}$, subsistence consumption, β, the exponent on fertility in the utility function, α, the exponent on productivity-augmented land in the production function, γ, the marginal cost of guarding slaves, q, the marginal cost of fertility, D, the coefficient on population and existing technology in the equation for technological progress, and θ, the exponent on population in the equation for technological progress.

Three parameters of the model give ambiguous predictions. An increase in β will make $\mathbf{L}^S(P_t)$ and $\mathbf{L}^{E/F}(P_t)$ flatter. This will not have any effect for a given A_t and P_t. In the steady state, this can make slavery appear if $\mathbf{L}^S(P_t)$ rotates downwards to intersect $\mathbf{L}^A(P_t)$, but can also make slavery disappear if $\mathbf{L}^S(P_t)$ rotates downwards so that it no longer intersects $\mathcal{S}^S$. For the same reason, the steady state effect on land rights will be ambiguous. The effects of an increase in q will be the opposite of those for an increase in β, and will be indeterminate for the same reasons. An increase in α can lead $\Psi(P_t)$ and $\Phi(P_t)$ to rotate upwards or downwards depending on the values of all the parameters.

The remaining parameters give clearer predictions. An increase in γ will cause both $\Psi(P_t)$ and $\Omega(P_t)$ to become steeper, shrinking the slavery region. Both in the steady state and for given values of A_t and P_t, this makes both slavery and land rights less likely. Raising θ will make $\mathbf{L}^A(P_t)$ rotate downwards. This will have no impact for a particular A_t and P_t. It will make a steady state under slavery less likely, since $\mathbf{L}^A(P_t)$ may tilt downwards so that it no longer intersects $\mathbf{L}^S(P_t)$. It will also push the intersection of $\mathbf{L}^{E/F}(P_t)$ and $\mathbf{L}^A(P_t)$ (if there is one) to the left, which may move the steady state from $\mathcal{S}^F$ to $\mathcal{S}^E$. This will make slavery and land rights less likely in the steady state. Raising D will have opposite effects. An increase in $\bar{c}$ will lead $\Psi(P_t)$, $\Omega(P_t)$ and $\Phi(P_t)$ to become steeper, shrinking both the free labor and slavery regions. It will also cause $\mathbf{L}^S(P_t)$ to become steeper. For a given A_t and P_t and in the steady state, this will make slavery and land rights less likely.

Table A1: Influential observations

Dep. variable	*Land rights*		*Slavery*			
RHS variable	*Land quality*		*Land quality*			
	Omaha	Prairie	Pawnee	Prairie		
	Cheremis	Eastern Europe	Fox	Prairie		
	Guanche	North Africa	Miami	Prairie		
	Yalunka	Western Sudan	Omaha	Prairie		
	Mbuti	African Hunters	Matakam	Nigerian Plateau		
	Ramcocame	Eastern Brazil	Makah	Northwest Coast		
	Djuka	Guiana	Bachama	Moslem Sudan		
	Walbiri	Australia	Macassare	Western Indonesia		
	Fang	Equatorial Bantu	Gude	Nigerian Plateau		
Dep. variable	*Land rights*		*Slavery*		*Slavery*	
RHS variable	*ln(1+pop. den.)*		*ln(1+pop. den.)*		*ln(1+pop. den.) sqrd.*	
	Yucatecma	Central America	Obostyak	Arctic Asia	Obostyak	Arctic Asia
	Bribri	Central America	Chahar	Central Asia	Bengali	North and Central India
	Guanche	North Africa	Menomini	Prairie	Uttarprad	North and Central India
	Keraki	New Guinea	Goldi	Arctic Asia	Goldi	Arctic Asia
	Banaro	New Guinea	Turkmen	Central Asia	Balinese	Eastern Indonesia
	Cheremis	Eastern Europe	Bengali	North and Central India	Santal	North and Central India
	Gidjingal	Australia	Hamyan	North Africa	Chahar	Central Asia
	Kaoka	Western Melanesia	Buryat	Central Asia	Kerala	South India
	Yanomamo	Lower Amazon	Mzab	Sahara	Menomini	Prairie

Notes: For the regressions with the full set of controls in Table 2, these are the most influential observations in terms of dfbeta.

Table A2: Outcomes by Major Region

Region	*% of societies with:* Any land rights	Any slavery	*Mean of:* Land quality (N (0,1))	Historic population density	N
Africa	0.93	0.83	0.09	2.97	414
Circum-Mediterranean	0.92	0.70	0.00	3.21	157
East Eurasia	0.83	0.54	0.07	3.87	123
Insular Pacific	0.73	0.24	0.46	2.43	119
North America	0.29	0.27	-0.33	0.77	284
South America	0.27	0.27	-0.05	1.23	109

Notes: Each entry reports the fraction of societies in the region that possess land rights or slavery, or the average land quality (normalized to be N(0,1)) or historic population density in the region. "N" refers to the sample for which historic population density is not missing.

Table A3. Main results without absolute latitude

	(1)		(2)		(3)		(4)	
	Any land rights				*Any slavery*			
Land quality	0.044**	(0.017)			0.049**	(0.021)		
ln(1+pop. den.)			0.124***	(0.012)			0.087**	(0.037)
ln(1+pop. den.) sqrd.							-0.009	(0.006)
Precipitation	-0.025	(0.020)	-0.007	(0.021)	-0.083***	(0.025)	-0.072***	(0.025)
Temperature	0.045**	(0.022)	0.005	(0.020)	0.139***	(0.027)	0.125***	(0.028)
Date observed	0.051***	(0.019)	0.004	(0.016)	-0.051***	(0.019)	-0.067***	(0.022)
Share desert	0.008	(0.017)	0.036**	(0.016)	0.036	(0.022)	0.031	(0.020)
Dist. to coast	-0.011	(0.018)	0.016	(0.018)	0.038*	(0.023)	0.041*	(0.024)
Elevation	-0.005	(0.019)	0.005	(0.018)	0.010	(0.022)	0.011	(0.023)
Pct. malarial	0.196***	(0.025)	0.142***	(0.024)	0.362***	(0.027)	0.334***	(0.027)
Ruggedness	0.070***	(0.016)	0.036**	(0.016)	0.129***	(0.021)	0.115***	(0.022)
Major river	-0.043	(0.034)	-0.088**	(0.034)	0.101**	(0.042)	0.097**	(0.043)
Mostly fishing	-0.176**	(0.071)	-0.014	(0.067)	0.437***	(0.077)	0.469***	(0.081)
Observations	801		801			1,041		1,041
Pop. den. at peak slavery								116.1

Notes: ***Significant at 1%, **Significant at 5%, *Significant at 10%. Robust standard errors in parentheses. All regressions are probit, with marginal effects reported.