National Institutions and Self-Insurance*

Raphael Godefroy (Université de Montréal)
Joshua Lewis (Université de Montréal)

March 2020

Abstract

This paper studies the relationship between national institutions and farming practices in Africa. The analysis exploits detailed geospatial data to compare outcomes across nearby plots exposed to the same underlying agroclimatic risk but belonging to adjacent countries with different national institutions. We develop a novel approach to measure agroclimatic crop risk without information on local prices, and use this methodology to assess how formal institutions affect decisions to cultivate "safer" versus "riskier" crops. The results show significant cross-border differences in outcomes. In countries with worse national institutions, farmers grew lower risk crops, diversified land across more different crops, devoted more total land to agriculture, and were more likely to hold livestock. The findings cannot be attributed to cross-border differences in market access. Instead, the patterns are consistent with a setting in which differences in the expropriation of wealth and non-farming income affect how farmers respond to agroclimatic risk.

JEL Codes: O10, O43, O13, N17

*We thank Treb Allen, Francesco Amodio, Dwayne Benjamin, Tim Guinnane, Fabian Lange, Frank Lewis, Stelios Michalopoulos, Aloysius Siow, and seminar participants at McGill, Queen's, Université de Montréal, Yale, Columbia, University of Toronto, and CIRANO for valuable comments and suggestions. Financial support from SSHRC, Université de Montréal, and Yale University is gratefully acknowledged.
Contact: Raphael Godefroy: raphael.godefroy@umontreal.ca. Joshua Lewis: joshua.lewis@umontreal.ca.

1 Introduction

In his book *Structure and Change in Economic History*, Douglass North argues that a major function of good institutions is to mitigate the consequences of economic uncertainty. This paper examines North's claim in Africa, where there are wide differences in the quality of formal institutions, and where rural producers face extreme income risk due to climatic shocks. We study how national institutions influence farmers' responses to underlying agroclimatic risk.

Our analysis uses detailed geospatial data to compare self-insurance practices across nearby plots that fall within a common historical ethnic homeland, but belong to different countries with different formal institutions. We compare outcomes across plots that face the *same* underlying agroclimatic uncertainty but are subject to *different* levels of national property rights protection, as measured by the '*Rule of Law*' (Kauffman, Kraay, and Matruzzi, 2008). The protection of property rights is a central feature of North's argument, and the '*Rule of Law*' index is a commonly used measure of institutional quality. Our within-ethnic homeland estimation strategy was first adopted by Michalopoulos and Papaioannou (2013a, 2013b), and was also used by Anderson (2018).

We explore the effects of national institutions on a range of agricultural outcomes including total farmland, decisions over which crops to cultivate, and cattle ownership. Detailed data from the Global Agro-Ecological Zones (GAEZ) project provide information on total farmland (measured by satellite land cover data) and cultivation by crop at the 5 arc-minute plot-level (roughly 10km by 10km at the equator).[1] We also assemble a new district-level dataset on cultivated area by crop, based on a comprehensive subnational land use database compiled by Monfreda et al. (2008). This district-level dataset is not subject to measurement error from downscaling, so it provides a validation check for the plot-level analysis.[2] Finally, we obtain data on household cattle ownership from the Demographic and Health Surveys (DHS), along

[1]GAEZ calculates crop-specific cultivation by plot using downscaling techniques based on satellite land cover, local land and climatic characteristics, and agricultural statistics.

[2]Measurement error in the dependent variable (crop choice) will bias the plot-level estimates only if GAEZ data quality varies systematically across *both* countries *and* crops with different risk profiles (see Section 4.2).

with geographic information that allows us to identify households located near national borders.

To study the link between institutions and farmers' crop choices, we must overcome two empirical challenges. First, without information on local crop prices it is impossible to identify how climatic fluctuations affect revenue, and thereby distinguish crops that are more or less vulnerable to agroclimatic shocks (Allen and Atkin, 2018). Second, the decision to grow a particular crop may depend on the entire covariance structure of crop returns, making it impossible to rank crops based on their own individual risk profiles. To overcome the first issue, we use information on annual potential crop yields from 1986 to 2000 to calculate the joint crop failure distribution on every plot, a measure of crop-specific revenue risk that does not depend on local prices.[3] Crop failure is a major source of rural income risk in Africa (Fafchamps et al., 1998; Kazianga and Udry, 2006). Moreover, on the vast majority of plots, the empirical crop failure distribution is *statewise dominant*, allowing us to ignore the covariance structure (Figure 1).[4] Given statewise dominance across crops and under standard assumptions on prices, risk averse farmers will prefer to grow crops with a lower probability of failure (see Appendix B). We rank crops individually from "safest" to "riskiest" based on their probability of failure, and combine this ranking with data on cultivation by crop to study how formal institutions affect decisions to grow crops that are more or less vulnerable to agroclimatic shocks.

The results reveal systematic cross-border differences in farmers' crop choices. In countries with worse national institutions, farmers were substantially more likely to grow low-risk crops and they diversified land across more different crops. Meanwhile, we find no significant differences in the probability of cultivating high-risk crops. The inclusion of crop-specific fixed effects controls does not alter the results, suggesting that these crop choice patterns reflect farmers' self-insurance motivations, rather than some other crop-specific attribute that was differentially valued across national borders. Farmers who faced worse national institutions also devoted more

[3] GAEZ calculates annual potential yields by crop at the plot-level, regardless of whether the crop is actually cultivated. These potential yield estimates are derived based on a host of local agroclimatic conditions along with hundreds of crop-specific parameters that capture how each crop responds to environmental factors.

[4] For any two crops, j and $j+1$, j statewise dominates $j+1$ if j provides a positive yield in all states in which $j+1$ in nonzero. In this scenario, there is no scope to exploit the covariance structure to insure failure risk of one crop against another as in the standard portfolio optimization problem (Markowitz, 1952).

land to cassava, millet, pulses, groundnut, and sorghum; crops that are traditionally valued for their drought-resistant properties (FAO, 1997; McCann, 2005).

We also uncover systematic cross-border differences in total farmland. Comparing adjacent areas that belong to a common historical ethnic homeland and face similar underlying agroclimatic conditions, we estimate that better property rights enforcement is associated with a 17 to 20 percent decrease in the likelihood that a plot was cultivated.

Finally, we estimate significant cross-border differences in cattle ownership that mirror the crop choice patterns. In countries with worse national institutions, individuals were significantly more likely to own cattle. We find similar effects among the subset of rural residents who owned land, suggesting that these results cannot be explained by differential selection of workers out of agricultural. The estimated cross-border differences in cattle ownership are consistent with the use of bullock sales as a household risk-coping strategy in less developed countries (e.g., Rosenzweig and Wolpin, 1993; Dercon, 1998).

The empirical results are robust to a range of alternate specifications including regression discontinuity models that compare outcomes across plots in the immediate vicinity of the border. The findings are robust to different covariates, including controls for underlying land productivity and the replacement of border-ethic homeland fixed effects controls with border fixed effects. Moreover, the consistency of the broad patterns across several different data sources – GAEZ plot-level data, Monfreda district-level data, satellite land cover data, and survey data from DHS – lends further credibility to the findings.

Despite these systematic differences in farmer decision-making, we find no relationship between national institutions and crop yields. These results support previous research showing no systematic link between institutional quality and local development in Africa (Michalopoulos and Papaioannou, 2013a). They also demonstrate how standard measures of development may fail to capture important effects of national institutions on local economic activity.

We explore a variety of mechanisms that might account for the relationship between national institutions and agricultural outcomes. The evidence rules out mechanisms related to cross-border differences in market access, consistent with the porous nature of African borders (Aker et al., 2014). We find no significant differences in the cultivation of traditional export crops such as bananas, coffee, and sugar. In fact, farmers in countries with better institutions were less likely to grow revenue-maximizing crops according to national and international prices. Moreover, we estimate systematically *larger* cross-border differentials in diversification in areas where underlying land quality is *worse*. This result contrasts sharply with a trade mechanism, which would predict that the effects be concentrated on productive land, where the gains to specialization are greatest. We also find no significant differences in the quality of *cultivated* land across borders, so the observed land-use patterns cannot be attributed to farmers sorting onto worse quality land. The estimated effects are also similar across a range of specifications and distance intervals, including regression discontinuity models that compare outcomes in the immediate vicinity of the border, providing further confidence that the results do not reflect differences in land quality, knowledge of farming practices, or market access.

To interpret the results of the empirical analysis, we develop a simple model of agricultural decision-making in which better property rights enforcement reduces expropriation of wealth and income.[5] A key insight from the model is that the relationship between property rights and farming practices depends on *what* is expropriated: wealth, farming income, or non-farming income. The results of the empirical analysis – increased agricultural land, low-risk crop cultivation, and crop diversification in areas with weak institutions – are consistent with expropriation of non-farming income, which reduces the return to formal employment causing agents to allocate more time to agricultural production. Because a larger portion of household income is vulnerable to agroclimatic shocks, farmers will allocate more land to safer crops that offer lower expected returns. At the same time, there will be no effect on the decision to grow high-risk crops, which depends solely on the expected return. Expropriation of farming income, by contrast, raises the return to formal employment, which should *decrease* land in agricultural and

[5]We model expropriation as a tax on households (e.g., Tilly, 1985; Murphy, Shleifer, and Vishny, 1993; Grossman and Kim, 1995), although the impact can be interpreted more broadly to include any channel through which weak national institutions lower household earnings and assets.

reduce the share of land allocated to lower risk crops; predictions that contradict the empirical findings. Finally, expropriation of wealth reduces the buffer against agroclimatic shocks, causing households to allocate land to lower risk crops. The predicted impact on total agricultural land is negative (or zero), since farmers may shift time to formal employment to limit exposure to agroclimatic risk. As a result, wealth expropriation alone cannot account for the observed patterns in farmland across borders.

The results demonstrate how property rights in urban and semi-urban areas can shape rural decision-making, even if their direct influence does not extend into the hinterland (Herbst, 2000). The observed patterns may reflect the fact that World Bank measures of institutional quality are largely based on the conditions in urban rather than rural areas.[6] In many African countries, the expropriation rates of farming and non-farming income may have diverged, due to the widely different policies towards agricultural and non-agricultural sectors (Bates, 1981). Since formal institutions may be related to other national characteristics, the results need not capture the causal impact of Rule of Law, per se. Instead, they establish that cross-country differences in urban property rights enforcement are systematically related to decision-making in outlying rural areas, and that these effects cannot be attributed to either historical ethnocultural factors or land quality differences across countries. These findings provide a counterpoint to recent evidence demonstrating how *rural* economic conditions influence *urban* structural transformation (Bustos et al., 2016).

Our results support the viewpoint of Douglass North and a number of economists and political scientists that the political and legal organization of a society are fundamental determinants of how agents respond to economic uncertainty. The literature has largely focused on the distortions arising from investment decisions (Banerjee, Gertler, and Ghatak, 2002; Jacobi, Li, and Rozelle, 2002; Goldstein and Udry, 2008; Donovan, 2018). Our results highlight an alternative channel: agents' willingness to engage in risky ventures. To the extent that these findings

[6]The World Banks' 'Rule of Law' index is based on information provided by external expert reports on the country's commercial and business conditions. Prior to the addition of information from the IFAD Rural Sector Performance Assessments in 2004, there was no information on the rural policy environments.

apply to other sectors, our results offer a new perspective on the low rates of entrepreneurship in underdeveloped nations (McMillan and Woodruff, 2002; World Bank, 2016).

More broadly, our paper contributes to a large body of research that links formal institutional arrangements to economic development. Since John Locke and Adam Smith, scholars have recognized the potential for institutions to shape economic outcomes. La Porta et al. (1998; 1999) and Acemoglu, Johnson, and Robinson (2001; 2002) demonstrate the importance of institutional arrangements established under European colonization for long-term development. In contrast, Michalopoulos and Papaioannou (2013a and 2013b) find that, on average, differences in the quality of national institutions do not translate into differences in economic development across African borders. Rather than focusing on a single measure of local development, we build on this literature by studying the broader effects of institutions on a range of local economic outcomes. The results show that standard development measures – GDP per capita and light intensity – may fail to capture important changes in local economic activity.

The paper also contributes to the literature on uncertainty in agriculture. Households cope with income uncertainty through precautionary savings (Deaton, 1990, 1991; Fafchamps et al., 1998), remittances from urban family members (Rapoport and Docquier, 2006), delayed technological adoption (Dercon and Christiaensen, 2011, Donovan, 2018), crop diversification (Kurosaki, and Fafchamps, 2002; Di Falco, and Chavas, 2009, Nicola, 2015); and cattle holdings (Rosenzweig and Wolpin, 1993; Dercon, 1998). There is ongoing debate of the effectiveness of these strategies in comparison to formal insurance (e.g., Binswanger-Mkhize, 2012; Mobarak and Rosenzweig, 2013). The lack of consensus may, in part, stem from the fact that the populations studied differ widely in both the underlying risk and institutional quality. Our cross-border research design allows us to compare economic outcomes across agents exposed to common agroclimatic risk. Moreover, detailed outcome data allow us to explore these relationships across a large number of countries in Africa. Our findings complement macroeconomic evidence that misallocation in agricultural is an important driver of cross-country income differences (Restuccia, Yang, and Zhu, 2008; Adamopoulos and Restuccia, 2018).

2 Data

2.1 Potential Yields, Land Use, and Crop Yields

We obtain information on potential yields, land use, and crop yields from the Global Agro-Ecological Zones (GAEZ) project, which is organized under the Food and Agriculture Organization's (FAO) and the Institute for Applied Systems Analysis.

The GAEZ project provides estimates of annual potential yields from 1960 to 2000 for different crops in every 5 arc-minute grid cell worldwide, regardless of whether the crop is actually grown. Potential crop yields depend on invariant geographic conditions such as soil type, elevation, and land gradient, and time-varying climatic conditions including rainfall, temperature, humidity, wind speed, and sun exposure.[7] These geographic and climatic factors are combined with a vector of hundreds of crop-specific parameters, derived from the agronomic literature, to capture how each crop responds to local environmental conditions. GAEZ calculates potential crop yields under different input choices. We focus on "high input" potential yields from "rain-fed" water supply systems, which account for more than 85 percent of agricultural land.[8]

We use annual potential crop yields to calculate the historical frequency of potential crop failures and rank crops according to their sensitivity to local growing conditions. We calculate the frequency of potential crop failures for the period 1986 to 2000, given the disproportionate influence of recent history in the formation of subjective probability (Viscusi and Zeckhauser, 2006; Gallagher, 2014). Farmers faced substantial risk of crop failure, and the median crop failed every five years.[9] Given the *statewise dominance* structure of crop failure risk (Figure

[7]The GAEZ project is particularly careful in its treatment of weather conditions. Annual potential crop yields are derived based on an aggregation of daily weather conditions, and the model captures how potential yields of each crop are affected by weather conditions throughout the growing cycle.

[8]"High input" technologies are mechanized, market oriented farm management systems that adopt fallow and conservation measures, and use high yielding seed varieties, and chemical pest controls. We explore the sensitivity of the findings to production under "low input" technologies.

[9]In non-failure years, however, there was much less variation in potential yields: the median coefficient of variation among positive potential crop yields was just 9 percent. The revenue impact of this yield volatility in non-failure years may have been further offset by local price adjustments (Allen and Atkin, 2018).

1), we rank crop from "riskiest" to "safest" based on their individual failure risk. This ranking is consistent with the fact that under standard assumptions on price, risk averse farmers will prefer to grow crops with a lower probability of failure (Appendix B).

GAEZ provides information on total farmland in 2000 at the 5 arc-minute grid cell level, based on satellite land cover imaging. GAEZ also provides grid cell data on cultivated area by crop in 2000, based on census agricultural statistics compiled by Monfreda et al. (2008) (see Section 2.2). These crop-specific estimates are obtained through downscaling techniques that use satellite land cover data and potential crop yields to assign agricultural statistics to a finer geographic resolution.[10] GAEZ uses a similar approach to construct grid cell measures of crop yields in 2000. We use these data to create a number of agricultural outcomes (see Section 2.5).

2.2 Land Use by Crop

We construct a new dataset on cultivated area by crop across 2,009 subnational districts in 40 African countries. This dataset addresses concerns with GAEZ's estimates of cultivated area by crop, whose quality may vary across countries and across crops.[11] Our district-level dataset addresses these limitations, since information on cultivated area by crop is reported at the same geographic scale at which the agricultural statistics were originally collected. We are also able to explore the sensitivity of the results to excluding both large reporting districts and crops for which subnational data is unavailable.

Our dataset is derived from Monfreda et al. (2008), a comprehensive database on harvested area for 175 crops compiled from agricultural census statistics across 150 countries at subnational units, and 19,751 units two levels below the country. The primary sources for these data were national census statistics, supplemented with additional information from agricultural

[10]GAEZ's downscaling algorithm is: 1) calculate total cultivated by grid cell based on satellite land cover datasets, 2) estimate suitability of individual crops on the cultivated land within each grid cell, 3) combine these estimates with other spatial information such as distance to market, population density, and farm system, 4) use the previous constraints to allocate national and subnational data on land use by crop to the grid cell level.

[11]The accuracy of downscaling depends on the size of the reporting districts in the underlying agricultural data, which varies across countries. Similarly, the quality of the GAEZ estimates may vary across crops, since subnational data on land use is often available only for a subset of crops.

surveys. The published Monfreda et al. (2008) database reports downscaled estimates of total cultivated area by crop at the 5 arc-minute grid cell level, along with administrative unit from which the agricultural data was collected (county/district, state/province, or national). We re-aggregate the grid cell estimates back to the original administrative districts. The resulting district-level dataset provides information on the total area cultivated by crop.

Figure 2 displays a map of the administration districts for the agricultural database. The level of geographic detail is greater in areas of agricultural production, whereas the larger political units typically cover regions in the Sahara. The median district area is 6,950 km^2, approximately four times the size of the median U.S. county. Subnational data is reported for 40 of 42 countries, and available for a large number of crops (Table A.1).

2.3 Cattle Ownership

Data on cattle ownership are from the Demographic and Health Surveys (DHS), a series of nationally representative household surveys. In supplementary modules, households were questioned about cattle ownership. The DHS also collects geographic information, allowing us to link respondents to national borders. We use the GPS coordinates of participating households to identify survey clusters located within 100 kilometers of a national border. For each country, we use the largest available survey wave between 2006 to 2016. The resulting sample consists of 63,289 households in 21 countries that span 35 borders.

2.4 National Institutions and Ethnic Partitions

Data on national institutions are from the World Bank's Governance Matters Database (Kauffman, Kraay, and Matruzzi, 2008). These data are based on an aggregate of various institutional quality measures that the World Bank categorizes based on principal components methods. Our main analysis is based on measures of *Rule of Law*, which ranges from -2.5 to +2.5. This variable reflects institutional factors, such as the quality of the judiciary and the level of property rights enforcement, that have been found to be particularly relevant for land development (Alston, Libecap, and Schneider, 1996; De Soto, 2000), and farmers' investment decisions (Goldstein and Udry, 2008; Banerjee, Gertler, and Ghatak, 2002; Jacoby, Li, and Rozelle, 2002). To limit concerns of reverse causality, we rely on this measure in 1996, the first period in which it was recorded.

Information on historical ethnic homelands is from Murdock's (1959) Tribal Map of Africa. Drawing on numerous anthropological sources, Murdock (1959) identifies the historical spatial distribution of more than 500 ethnic homelands in Africa. Following Michalopoulos and Papaioannou (2013a, 2013b), we combine these data with contemporary African country borders to identify ethnic homelands that were partitioned into two different countries.

2.5 Sample Selection and Outcome Variables

Our main sample consists in observations (grid cell plots, subnational districts, or DHS households) located within 100 kilometers of a national border on land with non-zero agricultural potential.[12] We construct several measures that capture farmers' willingness to grow crops that are more or less vulnerable to agroclimatic shocks: whether a 'low' or 'high' risk crop is grown (at the plot-level) and the fraction of cultivated land devoted to 'low' or 'high' risk crops (at the district-level). To measure agricultural land use, we construct an indicator for whether a plot is cultivated. We measure crop diversification as the number of different crops grown per hectare of cultivated land, and construct a series of dummy variables for whether individual crops are cultivated. We also construct indicators for cattle ownership among DHS households. Finally, we calculate actual yields per hectare by crop, at the plot-level.

[12]We drop plots with zero potential yields across all crops that are primarily located in the Sahara.

3 Empirical Strategy

3.1 Border/Ethnic Fixed Effects Estimation Strategy

Our first estimation strategy compares outcomes across plots within 100 km of the border that belong to a common historical ethnic homeland but fall within neighboring countries. We estimate the following regression model:

$$Y_{iebc} = \alpha_0 + \beta Rule\ of\ Law_{bc} + X'_{iebc}\gamma + \lambda_{eb} + \epsilon_{iebc}, \tag{1}$$

where Y_{iebc} denotes outcome on plot i that falls in the historical territory of ethnicity e, near border segment, b, in country c. The term X_{iebc} represents a vector plot-level covariates for potential yields by crop, to control for any differences in land and climatic conditions across plots on either side of the border. The term λ_{eb} denotes a vector of historical ethnic homeland-border fixed effects that control for persistent cultural factors which might continue to influence economic activity today. The variable of interest, $Rule\ of\ Law_{bc}$, is a dummy equal to one if institutional quality – the Rule of Law – in country c is higher than its neighbor at border b. The coefficient of interest, β, captures the average difference in outcomes between neighboring countries with better and worse institutions, based on comparisons across nearby plots within a common ethnic homeland that face similar agroclimatic conditions.

We also report estimates of the (local) impact of national institutions at the border using a regression discontinuity (RD) approach (see Imbens and Lemieux, 2008; Lee and Lemieux, 2010). We adopt the following parametric specification:

$$Y_{iebc} = \alpha_0 + \delta Rule\ of\ Law_{bc} + F(Dist_{iebc}) + X'_{iebc}\gamma + \lambda_{eb} + \epsilon_{iebc}. \tag{2}$$

In addition to the baseline controls, equation (2) includes an RD-polynomial, $F(Dist_{iebc})$, that controls for a third-order polynomial in distance to the border, whose coefficients are allowed to differ on either side of the national border. The coefficient of interest, δ, captures the impact of institutions on economic outcomes in the immediate vicinity of the national border.

To further explore treatment heterogeneity, we estimate a generalized version of equation (1) based on the following specification:

$$Y_{iebc} = \alpha_0 + \sum_{j=1}^{4} \beta^j Rule\ of\ Law_{bc} \times I(Dist_{iebc} = j) + \sum_{j=1}^{4} \rho^j I(Dist_{iebc} = j) + X'_{iebc}\gamma + \lambda_{eb} + \epsilon_{iebc}, \tag{3}$$

where the term $j \in 1, 2, 3, 4$ identifies four distance bins (<25 km, 25-50 km, 50-75 km, and 75-100 km) from the national border. The variable $I(Dist_{iebc} = j)$ is a dummy equal to one if the plot falls within a particular distance bin. Equation (3) allows the impact of national institutions to vary by distance to the border. For example, β^3, captures the cross-border difference in outcomes on plots 50-75 km from the national border.

Two final estimation details are worth noting. First, regressions for extensive margin outcomes (plot and crop cultivation) are unweighted to estimate the average effect for a plot. Regressions for intensive margin outcomes (crop diversification and crop yields per hectare) are weighted by total farmland to estimate the average effect per hectare of farmland. Second, standard errors are two-way clustered across both country and historical ethnic boundaries to account for arbitrary spatial correlation along both dimensions, following the approach of Cameron, Gelbach, and Miller (2011).

3.2 Identification and Potential Yields Across National Borders

Each of the three regression models compare relative outcomes across plots in neighboring countries with different formal institutions. The identifying assumption is that agricultural activity would have been similar among these nearby plots absent any national differences associated with the quality of institutions. In particular, we require that national borders were drawn independently of local conditions relevant for current agricultural outcomes. This assumption is supported by a large historical narrative documenting the various factors that shaped the formation of African borders in the late nineteenth century (see Asiwaju 1985; Wesseling 1996; Herbst 2000; Englebert 2009). At the time of colonization the interior of the continent was largely unexplored, and European colonizers drew African borders with little knowledge of local geographic conditions. Although the European powers partitioned land for the creation of colonies and protectorates, as opposed to future independent countries, the original borders have remained almost entirely unchanged since independence.

The arbitrary partitioning of African countries is supported by growing body of empirical research. Michalopoulos and Papaioannou (2013a, 2016) document the plausibly exogenous formation in African borders, and show that there are no systematic differences across a range of factors related to economic potential through country borders. Similarly, Alesina, Easterly, and Matuszeski (2011) demonstrate that eighty percent of African country borders follow longitude and latitude lines, more than any other continent. The arbitrary assignment of African borders has also been used by a number of other researchers as a source of quasi-experimental variation (e.g., Miguel, 2004; Cogneau and Moradi, 2014).

To further assess the validity of our approach, we estimate the relationship between national institutional quality and potential crop yields. These outcomes incorporate all location-specific factors relevant for agricultural yields, so should capture any cross-border differences to farmers' decision-making. Table 1 reports the results for the ten most widely cultivated crops. The coefficient estimates are all small and statistically insignificant, suggesting no systematic cross-border differences in land quality. These broad patterns are robust to a range of alternative specifications. RD estimates, based on equation (2), show no significant differences in potential yields at the border (Table A.2). The estimates are similar when ethnic homeland-border fixed effects are replaced by border fixed effects (Table A.3), suggesting that potential yield varied little across national borders regardless of historical ethnic homelands. Together, these results provide confidence that African borders were drawn independently from local characteristics, supporting our identifying assumption.[13]

[13] These patterns contrast with stark differences in land quality across modern European borders, consistent with Europe's long history of disputed territories (results available upon request).

4 Results

4.1 Land Use

Table 2 reports the estimated differentials in land use across national borders in Africa. We estimate the effects of Rule of Law on an indicator for whether the plot is cultivated for the each of the three empirical specifications (equations 1-3). Across the various specifications, the results show a clear link between national institution quality and agricultural land (cols. 1-3). The estimates are all negative and statistically significant, implying more cultivation in countries with weaker property rights. These cross-country differences hold at various distances intervals, including among plots in the immediate vicinity of the national border.

The effects on cultivation are sizeable. Combining the results from column (1) with the average difference in measured institutional quality across neighboring countries (0.45), we calculate that a one point increase in the Rule of Law index – roughly the gap between countries at the 25th percentile (e.g., Chad or Nigeria) and countries at the 75th percentile (e.g., Malawi or Mali) – decreases the probability of cultivation by 28 percent.

We exploit heterogeneity in average land quality across ethnic homeland-border pairs to study how the effects varied according to underlying land quality. We re-estimate the three empirical specification separately for ethnic homeland-border pairs with above- versus below-median land quality. The results reveal systematic differences in effects across the two subsamples (Table 2, cols. 4-9). In areas with more productive land, we estimate large and significant effects of national institutions on land use, whereas we find no significant cross-border differences in areas with lower quality land. Because the differences in underlying land productivity may be correlated other factors relevant for agricultural activity, these heterogeneous patterns cannot be given a causal interpretation.[14] Nevertheless, the absence of extensive margin land use effects in areas with low quality land, provides a useful subsample on which to study the

[14]Notably, the patterns are robust to controls for size of the cross-country gap in the Rule of Law index.

effects on farmers' intensive margin decisions.

The broad patterns reported in Table 2 are robust to a range of alternative specifications. For instance, Table A.4 reports results from models that replace border-ethnic homeland fixed effects with border fixed effects. These estimates capture the average differences in outcomes across borders regardless of historical ethnic homeland boundaries, so will not be biased by measurement error in the original Murdock maps (see Herbst, 2000; Cogneau and Dupraz, 2015). The effects are similar in sign, significance, and magnitude.

4.2 Crop Choice

We examine the impact of national institutions on farmers' crop choices. The analysis proceeds in two steps. First, we use annual information on potential crop yield failures from 1986 to 2000 to rank crop according to their historical probability of failure. We use this information to construct indicators for low-risk and high-risk crops at every ethnic-border homeland pair, based on how individual crops respond to local agroclimatic conditions. Second, we estimate cross-border differences in farmers' willingness to cultivate the two different types of crops. We estimate these models at both the plot-level – comparing the probability that different crop types were cultivated, and at the district-level – comparing the amount of land devoted to the different crop types.

4.2.1 Plot-level Results

Figure 3 graphs the average likelihood that a plot is cultivated with high- and low-risk crops, by distance to the national border. We report these averages by 5km distance bin width along with the predicted probability based on a third-order RD polynomial. The relatively high institutional quality countries are depicted on right-hand side. There are no discernible border-distance patterns in countries with worse institutions, whereas the probability of cultivation declines with distance to the border in countries with better institutions. Figure 3(a) shows a clear discontinuity at the national border, with less cultivation of low-risk crops in countries with better quality institutions. The discontinuity is large in magnitude. The polynomial fit predicts a 14 percentage point difference in the likelihood of crop cultivation at the national border. Figure 3(b) shows some evidence of a corresponding cross-border difference in the probability of high-risk crop cultivation, although the discontinuity is less pronounced.

Table 3 reports the regression results from equations 1-3 for the two crop risk categories. The estimates mirror the patterns displayed in Figures 3. In countries with worse national institutions, farmers were significantly more likely to grow low-risk crops. Meanwhile, we find no statistically significant relationship between the Rule of Law and the probability of growing high-risk crops. The overall patterns and effect sizes are similar in magnitude across the three different estimating equations.[15] Consistent with these patterns, we find that farmers in countries with weaker property rights were more likely to grow drought-resistant crops such as cassava, millet, pulses, groundnut, potato, and sorghum (Table A.6).

The effects on low-risk crop cultivation could reflect either cross-country differences in self-insurance practices or other crop-specific attributes, such as a calorie content, that were valued differently across countries. To shed light on this question, we re-estimate versions of equation 1 that include crop-specific fixed effects controls. Identification is based on differences in the agroclimatic crop risk ranking across borders-ethnic homeland pairs arising from local agroclimatic conditions. These models capture the relative frequency of low- and high-risk crop cultivation, holding constant other specific crop attributes that are common across borders. The estimated effects are similar in sign, significance, and magnitude to the baseline estimates (Table 3, cols. 4, 8). These results provide further evidence that responses to agroclimatic risk differed systematically across national borders, and that farmers exposed to worse institutions devoted more land to safer crops.

[15]We find the same patterns in models that replace border-ethnic homeland fixed effects with border fixed effect (Table A.5).

4.2.2 District-level Results

We estimate the impact of national institutions on the fraction of land devoted to high- and low-risk crops across nearby districts in neighboring countries. We focus on districts within 100 km of the national border, and estimate modified versions of equation 1 that replace the border-ethnic homeland fixed effect with a border fixed effect.[16] The dependent variable is the log ratio of total land devoted to high-risk crops relative to total land devoted to both high- and low-risk crops.[17] While conceptually similar to the plot-level analysis, the district-level estimates are not subject to measurement error concerns due to downscaling (see Section 2.2).

Table 4 reports the results for *Rule of Law* from equation (1). Even numbered columns report the results for the subset of crops for which land use was reported at the subnational level. Columns 1-2 report the results for the baseline set of districts. To further ensure similarity in the underlying land and agroclimatic conditions, columns 3-6 report the results for a restricted sample of border-district pairs that had the same high- and low-risk crops. In columns 5-6, districts are weighted by total land area to capture the average effect at the plot-level.

Across the various specifications, we estimate a large and statistically significant relationship between the Rule of Law and crop shares. In countries with worse institutions, farmers devoted relatively more land to low-risk crops. Restricting to crops observed at the district-level reduces the sample by one third, nevertheless, the estimated coefficients remain large and statistically significant. When we focus on border-district pairs with the same set of high- and low-risk crops, the coefficient estimates increase in magnitude and significance. Together, these findings support the main conclusions from the plot-level analysis that cross-country differences in farmers' crop choices were systematically related to the quality of national institutions.

[16]Given the imperfect geographic overlap, we cannot uniquely assign districts ethnic homelands.

[17]Less than 3 percent of districts reported zero cultivation of high- or low-risk crops.

4.3 Crop Diversification

Table 5 reports the effects of Rule of Law on crop diversification, measured as the log number of crops per hectare of farmland at the plot-level. Across the different specifications, we find large and statistically significant cross-border differences in diversification (cols. 1-3). Farmers in countries with worse institutions grew more crops per hectare. These patterns hold even among producers operating in the immediate vicinity of the border.

The crop diversification analysis is based on cultivated plots. As a result, differential selection onto agricultural land could bias the estimates. For example, farmers exposed to better institutions might discontinue operations on marginal quality land, which could increase specialization independently of agroclimatic risk. To explore this issue, Table 5 (cols. 4-9) report the effects on diversification separately in areas with below- and above-median land quality. The effects on diversification are concentrated in areas with lower underlying land quality, where there were no extensive margin changes in farmland (Table 2, cols. 4-9). To further explore this issue, we use GAEZ data on potential crop yields *on cultivated land* to test whether national institutions were systematically related to the quality of land on which farmers' operated. The results show no significant cross-country differences in the quality of cultivated land, providing further evidence against a selection mechanism (Table A.7).[18]

4.4 Cattle Ownership

Table 6 reports the effects of Rule of Law on cattle ownership, based on household-level data from the DHS.[19] We report estimates from equation 1 for households within 100 km of the border (cols. 1-3) and within 50 km of the border (cols. 4-6). The results reveal systematic cross-border differences: households in countries with better national institutions were 13 to 15 percent less likely to own cattle. The estimates are similar among households located within

[18]The absence of cross-border differences in potential yields can largely be attributed to homogeneity in plot land quality within the 100 km bandwidth (results available upon request).

[19]Because observations span the period 2006 to 2016, we use the values of the Rule of Law index in 2006 to measure relative property rights across countries. The results are not sensitive to this specification.

50 km of the border and unaffected by individual covariates. We also find similar results when we focus on households that owned farmland (cols. 2-3, 5-6), suggesting that the cross-border differentials cannot be attributed to selection out of agriculture.

Taken together, the results in sections 4.1 to 4.4 reveal important differences in farmers' decision-making across countries with different quality institutions. The relationships between national institutions crop choices, land use, and cattle ownership are consistent with the high level of agroclimatic insecurity that agricultural producers face (Kazianga and Udry, 1996). In countries with poor institutions, rural producers often lack access to credit and insurance markets and there is limited government-provided security (Binswanger and Rosenzweig, 1986). As a result, rural producers may optimally engage in a range of self-insurance practices in an effort to mitigate the consequences of agroclimatic risk.

4.5 Crop Yields

Table 7 reports the effects of Rule of Law on crop yield. For each crop, we measure total yields (in tons) per hectare of cultivated. Despite the stark cross-border differences in land use, crop choice, and cattle ownership across country borders, we find no systematic link between institutional quality and crop yields. The estimates are small in magnitude and statistically insignificant.[20] The results are consistent with Michalopoulos and Papaioannou (2013a), who find no link between national institutions and local development, as measured by luminosity.

What explains the similarity in observed crop yields across borders? One possibility is that farmers in countries with worse institutions were able to achieve similar yields through increased in labor inputs. If these farmers faced a binding subsistence food requirement, they might simultaneously divert land to less productive crops to mitigate agroclimatic risk *and* increase labor inputs to compensate for the forgone output. Alternatively, the results could reflect worker selection: in countries with better institutions, relatively high urban wages may

[20]These patterns are consistent across a range of alternate specifications, including models based on log crop yields and regressions that replace border-ethnic homeland fixed effects with border fixed effects (results available upon request).

draw the most productive workers out of agriculture, lowering average yields (Lagakos and Waugh, 2013).

The absence of cross-country differences in crop yield contrasts with previous work showing a strong link between property rights enforcement and farmers' investment decisions (Goldstein and Udry, 2008; Banerjee, Gertler, and Ghatak, 2002; Jacoby, Li, and Rozelle, 2002). In accounting for these differences, it should also be noted that measure Rule of Law index is based on urban expropriation. To the extent that these urban property rights did not extend to the hinterlands (Herbst, 2000), farmers might have been unwilling to invest in land development.

5 Interpretation

5.1 Expropriation and Farming Practices

The empirical results show a strong link between formal institutions, land use, and crop choices. To interpret these patterns, we formalize a model, in which risk-averse agents choose how to allocate time between farming and non-farming labor, and how to allocate land across different crops (see Appendix - Section B). We apply this framework to study how better quality national institutions affect farming practices. We model expropriation as a tax on household wealth, non-farming income, or farming income (e.g. Tilly, 1985; Murphy, Shleifer, and Vishny, 1993; Grossman and Kim, 1995), and assume that better property rights enforcement reduces the rates of expropriation.[21] The analysis compares individuals who face a common underlying risk of crop failure and have access to the same output markets, but are subject to different levels of property rights enforcement.

The model delivers the following predictions regarding the impact of wealth and non-farming income expropriation on agricultural outcomes:

[21] The effects of national institutions on wealth and income need not arise directly through expropriation, but could reflect any channel through which better quality institutions contribute to higher income or wealth levels. Our analysis abstracts from formal crop insurance, which was virtually non-existent in the African setting. Allowing for this insurance channel would reinforce the main predicted relationships.

(1) Greater *expropriation of non-farming income* leads to (a) an increase in the number of different crops grown, (b) an increase in the total agricultural land, (c) an ambiguous effect on crop diversification (per unit of farmland).

(2) Greater *expropriation of wealth* leads to (a) an increase in the number of different crops grown, (b) a decrease (or no change) in the total agriculture land, (c) an increase in crop diversification.

(3) Newly cultivated crops under 1(a) and 2(a) will have a lower risk of failure.

Greater expropriation of *non-farming income* lowers the return to formal employment, causing individuals to allocate more time to farm labor, which increases total agricultural land. Because a larger share of earnings is vulnerable to agroclimatic shocks, individuals will devote land to less risky crops. The net effect on diversification (number of crops per unit of farmland) is ambiguous and depends on the relative size of these two offsetting effects. Greater expropriation of *wealth* should cause a similar reallocation to 'safer' crops, since household assets act as a buffer against agroclimatic shocks. Meanwhile, the predicted impact on total agricultural land is negative (or zero), since a decline in household assets may cause individuals to increase formal employment to reduce exposure to agroclimatic risk. Finally, greater expropriation of *farming income* raises the return to formal employment, which leads to a decrease in farm labor and total agricultural land. Since a smaller share of earnings is subject to agroclimatic shocks, individuals will have less need to cultivate low-risk crops.

These model's predictions underscore the role of wealth and non-farming income in shielding farmers from the consequences of agroclimatic risk. In addition, the model predicts that expropriation will influence the decision to grow crops with low failure risk, but have no impact on the decision to grow high-risk crops. The latter decision is governed solely by the expected return: if a crop provides a large enough return it will be grown everywhere, regardless of household wealth or non-farming income.

A key insight from the model is that the effects depend crucial on *which* assets are expropriated. This distinction has empirical relevance, since the Rule of Law index is derived from urban property rights indicators. Interestingly, the empirical patterns are consistent with effects arising through both expropriation of both non-farming income and household wealth. The relative decline in total agricultural land is consistent with non-farming income expropriation, while the cross-border differences in low-risk crop cultivation could be driven by either wealth or non-farming income expropriation. We also find no systematic cross-border differences in high-risk crop cultivation, consistent with the theoretical predictions.

5.2 Markets Access and Agricultural Outcomes

The systematic cross-border differences in outcomes – crop choice, agricultural land, and livestock ownership – are consistent with better national institutions reducing the prevalence of self-insurance practices in agriculture. National institutions may also influence agricultural decision-making through market access. For example, countries with high quality institutions may have better rural infrastructure, providing greater access to national and international markets. In this section, we explore whether the results can be attributed to cross-border differences in market access. We explore three tests for a market access mechanism: (1) effects by border distance, (2) effects by land productivity, and (3) effects on export-oriented crops.

We explore heterogeneous effects by border distance (equation 3). The estimates are broadly similar at different distance intervals, and we find large cross-border outcomes differences even among plots within a narrow window of the national border (Tables 2, 3, 5, col. 3). In contrast, under a market access mechanism, the coefficient estimates should increase with border distance, given the limited reach of transportation infrastructure to the country's periphery (Storeygard, 2016), and the porous nature of African borders (Aker et al., 2014).

The heterogeneous effects by land quality also contrast with a market access mechanism, which predicts larger effects of national institutions in areas with more productive land, where the gains to specialization are be greater. Instead, the effects on diversification are concentrated in areas with *lower* quality land (Table 5, cols. 4-9).

Finally, we study the relationship between national institutions and decisions to grow cash crops, which should be particularly responsive to changes in market conditions. Table A.8 (cols. 1-4) report the results. We find no positive effects on cocoa, coffee, tea, groundnut, bananas, palm oil, sugar, and coconut; crops that are traditionally grown for sale on domestic or foreign markets (FAO, 1997; Achterbosch et al., 2014; Chauvin et al., 2012). Farmers in countries with better institutions were also significantly *less* likely to grow revenue-maximizing crops according to international prices (cols. 6-7).

6 Conclusion

This paper studies the impact of national institutions on local economic activity in Africa. Drawing on detailed geospatial data on a range of agricultural outcomes, we compare outcomes across producers who faced similar agroclimatic conditions but were exposed to different formal institutions. We find that the quality of national institutions had substantial effects on economic decision-making in rural areas: worse institutions are associated with more land in agriculture, increased cultivation of low-risk crops, and more cattle ownership.

Our findings highlight the interaction between formal institutions, underlying risk, and economic decision-making. The results demonstrate how better quality institutions can shield individuals from the consequences of economic risk, consistent with the views of Douglass North. While the efficiency gains from property rights and contract enforcement may be especially large in agriculture, where producers are exposed to high levels of agroclimatic risk, the findings may apply more broadly to other sectors. Understanding the relationship between property rights enforcement, entrepreneurship, and growth may be an interesting area for future research.

Finally, our findings highlight how standard measures of development may fail to capture differences in economic activity at the local level. These limitations may soon be addressed, given newly available sources of geospatial data on a variety of local outcomes. Combining these measures with machine-learning techniques may allow researchers to develop better proxies for economic development at the local level.

References

Acemoglu, D., S. Johnson, and J. A. Robinson. (2001). "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*, 91(5): 1369-1401.

Acemoglu, D., S. Johnson, and J. A. Robinson. (2002). "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics*, 107(4): 1231-1294.

Achterbosch, T.J., S. van Berkum, and G. W. Meijerink. (2014). *Cash Crops and Food Security: Contributions to Income, Livelihood Risk and Agricultural Innovation.* LEI Report, LEI Wageningen UR (University & Research centre).

Aker, J. C., M. W. Klein, S. A. O'Connell, and M. Yang. (2014). "Borders, Ethnicity, and Trade." *Journal of Development Economics*, 107: 1-16.

Alesina, A., W. Easterly, and J. Matuszeski. (2011). "Artificial States." *Journal of the European Economic Association*, 9(2): 246-277.

Allen, T. and D. Atkin. (2018). "Volatility, Insurance, and the Gains from Trade." *Mimeo.*

Alston, L. J., G. D. Libecap, and R. Schneider. (1996). "The Determinants and Impact of Property Rights: Land Titles on the Brazilian Frontier." *Journal of Law, Economics, and Organization*, 12: 25-61.

Anderson, S. (2018). "Legal Origins and Female HIV." *American Economic Review*, 108(6): 1407-1439.

Asiwaju. A. (1985). "The Conceptual Framework." in *Partitions Africans*, pp. 1-18. St. Martin Press, New York, NY.

Banerjee, A. V., P. J. Gertler, and M. Ghatak. (2002). "Empowerment and Efficiency: Tenancy Reform in West Bengal." *Journal of Political Economy*, 110: 239-280.

Bates, R. H. (1981). *Markets and States in Tropical Africa: The Political Basis of Agricultural Policies.* University of California Press, Berkeley and Los Angeles, CA.

Binswanger, H.P., and M.R. Rosenzweig. (1986). "Behavioral and Material Determinants of Production Relations in Agriculture." *Journal of Development Studies*, 22(3), 503-539.

Binswanger-Mkhize, H. P. (2012). "Is There Too Much Hype About Index-Based Agricultural Insurance?" *Journal of Development Studies*, 48(2): 187-200.

Bustos, P., B. Caprettini, and J. Ponticelli. (2016). "Agricultural Productivity and Structural Transformation: Evidence from Brazil." *American Economic Review*, 106(6): 1320-1365.

Cameron, C. A., J. Gelbach, and D. Miller. (2011). "Robust Inference with Multi-Way Clustering." *Journal of Business and Economic Statistics*, 28(2): 238-249.

Chauvin, N. P., F. Mulangu, and G. Porto. (2012). "Food Production and Consumption Trends in Sub-Saharan Africa: Prospects for the Transportation of the Agricultural Sector." UNDP, Regional Bureau for Africa.

Cogneau, D. and Y. Dupraz. (2015). "Historical Institutions and Economic Development in Africa. A Selective and Critical Survey of Recent Literature." Histoire & Mesure. 30(1): 103-134.

Cogneau, D. and A. Moradi. (2014). "Borders that Divide: Education and Religion in Ghana and Togo Since Colonial Times." *Journal of Economic History*, 74(3): 694-729.

De Soto, H. (2000). *The Mystery of Capital*, Basic Books, New York, NY.

Deaton, A. (1990). On Risk, Insurance and Intra-Village Consumption Smoothing. *Research Program in Development Studies*, Princeton University, N.J.

Deaton, A. (1991). Savings and Liquidity Constraints. *Econometrica*, 59(5), 1221-1248.

Dercon, S. and L. Christiaensen. (2011). "Consumption Risk, Technology Adoption, and Poverty Traps: Evidence from Ethiopia." *Journal of Development Economics*, 96(2): 159-173.

Di Falco, S., and J.P. Chavas. (2009). "On crop biodiversity, risk exposure and food security in the highlands of Ethiopia." *American Journal of Agricultural Economics*, 91, 599-611.

Donovan, K. (2018). Agricultural risk, intermediate inputs, and cross-country productivity differences. *Mimeo.*

Englebert, P. (2009). *Africa, Unity, Sovereignty and Sorrow*. Lynne Rienner Publishers Inc., Boulder, CO.

Fafchamps, M., C. Udry, and K. Czukas. (1998). Drought and Saving in West Africa: Are Livestock a Buffer Stock? *Journal of Development Economics*, 55(2), 273-305.

FAO. (1997). "Agriculture, Food and Nutrition for Africa: A Resource Book for Teachers of Agriculture. Chapter 2: Agriculture, Food Security and Nutrition." FAO Corporate Document Repository, Rome.

Gallagher, J. (2014). "Learning about an Infrequent Event: Evidence from Flood Insurance Take-Up in the United States." *AEJ: Applied Economics*, 6(3): 206-233.

Grossman, H., and M. Kim. (1995). "Swords or Plowshares? A Theory of the Security to Property." *Journal of Political Economy*, 103(6): 1275-1288.

Goldstein, M. and C. Udry. (2008). "The Profits of Power: Land Rights and Agricultural Investment in Ghana." *Journal of Political Economy*, 116: 981-1022.

Herbst, J. (2000). *States and Power in Africa*. Princeton University Press, Princeton, NJ.

Imbens, G. W., T. Lemieux. (2008). "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics*, 142(2): 615-635.

Jacoby, H. G., G. Li, and S. Rozelle. (2002). "Hazards of Expropriation: Tenure Insecurity and Investment in Rural China." *American Economic Review*, 92: 1420-1447.

Kazianga, H. and C. Udry. (2006). Consumption Smoothing? Livestock, Insurance and Drought in Rural Burkina Faso. *Journal of Development Economics*, 79, 413-446.

Kaufmann, D., A. Kraay, and M. Mastruzzi. (2008). "Governance Matters VIII: Aggregate and Individual Governance Indicators, 1996-2007," *Policy Research Working Paper Series 4654. The World Bank*.

Kurosaki, T. and M. Fafchamps. (2002). "Insurance Market Efficiency and Crop Choice in Pakistan." *Journal of Development Economics*, 67(2): 419-453.

La Porta, R., F. Lopez-de-Silanes, and A. Shleifer. (1997). "The Economic Consequences of Legal Origins." *Journal of Economic Literature*, 46(2): 285-332.

La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R. Vishny. (1998). "Law and Finance." *Journal of Political Economy*, 106(6): 1113-1155.

Lagakos, D. and M. E. Waugh. (2013). "Selection, Agriculture, and Cross-Country Productivity Differences." *American Economic Review*, 103(2): 948-980.

Lee, D. S., and T. Lemieux. (2010). "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48(2): 281-355.

Markowitz, H. (1952). "Portfolio Selection." *Journal of Finance*, 40, 561-572.

McCann, J. C. (1999). *Green Land, Brown Land, Black Land: An Environmental History of Africa: 1800-1990.* James Currey, Cambridge, MA.

McCann, J. C. (2005). *Maize and Grace: Africa's Encounter with a New World Crop, 1500-2000.* Harvard University Press, Cambridge, MA.

McMillan, J. and C. Woodruff. (2002). "The Central Role of Entrepreneurship in Transition Economies." *Journal of Economic Perspectives*, 16(3): 153-170.

Michaloupoulos, S. (2012). "The Origins of Ethnolinguistic Diversity." *American Economic Review*, 102(4): 1508-1539.

Michaloupoulos, S., and E. Papaioannou. (2013a). National Institutions and Subnational Development in Africa. *Quarterly Journal of Economics*, 129(1): 151-213.

Michaloupoulos, S., and E. Papaioannou. (2013b). Pre-Colonial Ethnic Institutions and Contemporary African Development. *Econometrica*, 81(1): 113-152.

Michaloupoulos, S., and E. Papaioannou. (2016). "The Long-Run Effects of the Scramble for Africa." *American Economic Review*, 106(7): 1802-1848.

Miguel, E. (2004). “Tribe or Nation? Nation-Building and Public Goods in Kenya Versus Tanzania.” *World Politics*, 56(3): 327-362.

Mobarak, M. A and M. R. Rosenzweig. (2013). “Informal Risk Sharing, Index Insurance, and Risk Taking in Developing Countries.” *American Economic Review*, 103(3): 375-380.

Monfreda, C., N. Ramankutty, and J. A. Foley. (2008). “Farming the Planet: 2. Geographic Distribution of Crop Areas, Yields, Physiological Types, and Net Primary Production in the Year 2000.” *Global Biogeochemical Cycles*, 22, GB1022.

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

Murphy, K., A. Shleifer, and R. Vishny. (1993). “Why is Rent-seeking so Costly?” *American Economic Review*, 83(2): 409-414.

Nicola, F. de. (2015). “The Impact of Weather Insurance on Consumption, Investment, and Welfare.” *Quantitative Economics*, 6: 637-661.

Rapoport, H. and F. Docquier. (2006). “The Economics of Migrants’ Remittances.” in *Handbook of the Economics of Giving, Altruism and Reciprocity*, vol. 1: 1135-1198.

Rosenzweig, M.R., and K.I. Wolpin. (1993). Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income countries: Investments in Bullocks in India. *Journal of Political Economy*, 101(2), 223-244.

Storeygard, A. (2016). “Farther on Down the Road: Transportation Costs, Trade and Urban Growth.” *Review of Economic Studies*, 83(3): 1263-1295.

Tilly, C. (1985). “War Making and State Making as Organized Crime.” In: Evans, P. (Ed.) *Bringing the State Back In*. Cambridge University Press.

Viscusi, W. K. and R. J. Zeckhauser. (2006). “National Survey Evidence on Disaster and Relief: Risk Beliefs, Self-Insurance, and Compassion.” *Journal of Risk and Uncertainty*, 33(1): 13-36.

Wesseling, H. L. (1996). *Divide and Rule: The Partitions of Africa, 1880-1914*. Praeger, Amsterdam, Netherlands.

7 Figures and Tables

Figure 1: **An Example of a Plot with *Statewise Dominant* Failure Risk**

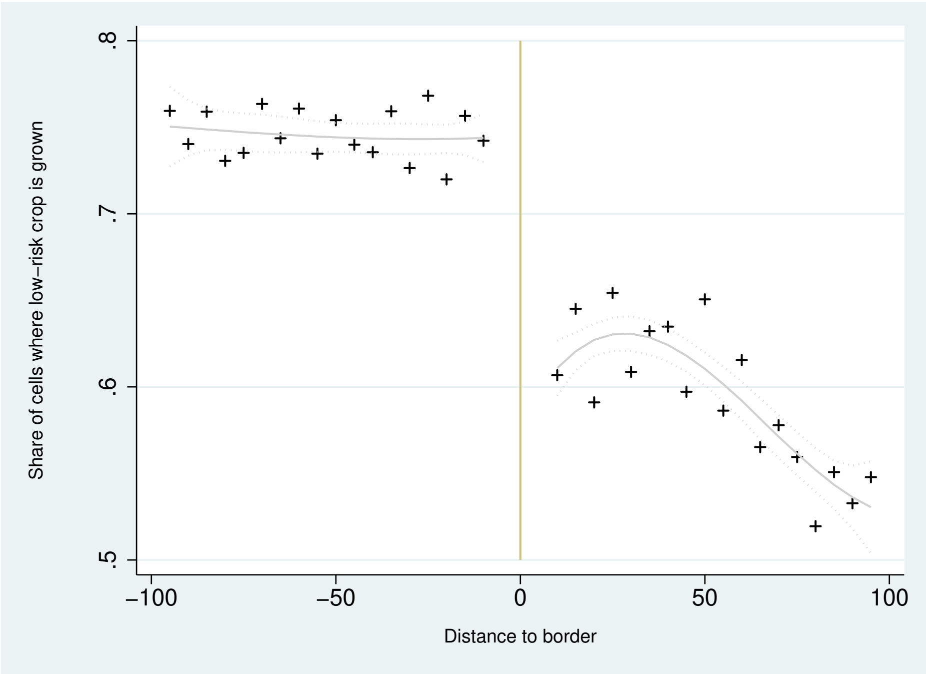
Year	Cassava	Rapeseed	Banana	Potato	Sorghum	Phaseolus
1989	10.547	3.384	5.421	7.606	2.485	2.694
1998	10.987	3.336	2.775	7.19	2.394	0
1994	10.823	3.169	4.429	6.845	2.247	0
1986	10.411	3.264	2.202	7.525	2.355	0
1992	10.607	3.208	2.543	7.283	2.329	0
1990	10.827	3.156	2.914	6.07	1.945	0
1997	10.485	3.181	2.38	5	0	0
1993	11.3	3.346	3.245	6.48	0	0
1996	11.12	3.306	3.645	7.329	0	0
1991	11.139	3.28	4.635	7.224	0	0
1987	11.078	3.199	3.539	7.066	0	0
1995	10.825	3.117	5.338	0	0	0
1988	10.947	3.181	3.099	0	0	0
1999	9.662	3.053	0	0	0	0
2000	8.248	0	0	0	0	0

Notes: This figure reports an example of annual potential crop yields from 1986 to 2000 for the plot at latitude-longitude (-1.875, 34.958). Potential crop yields are reported for GAEZ's high input / rain fed technology. This empirical *statewise dominant* crop failure risk holds across **83 percent** of sample plots, although the specific crops and ranking differ across plots.

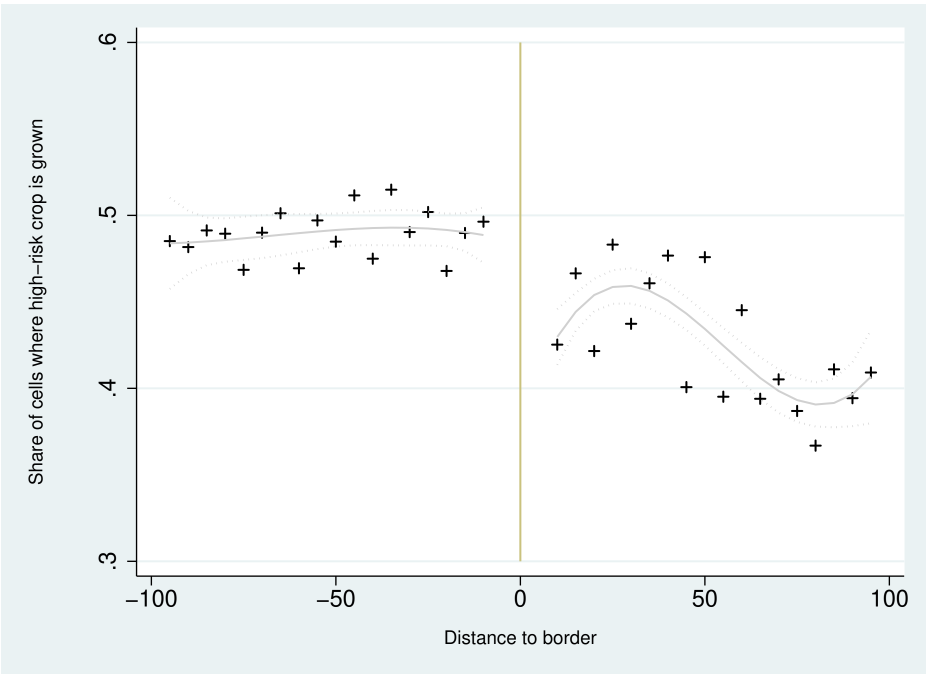
Figure 2: **District-level data on cultivated land by crop**

Notes: The district-level data were compiled from Monfreda et al. (2008). See text for discussion.

Figure 3: **Probability of Crop Cultivation, by Distance to the National Border**



(a) Crops with **low risk** of potential yield failure



(b) Crops with **high risk** of potential yield failure

Notes: This figure graphs the probability that plots are cultivated with low-risk and high-risk crops by distance to the national border. Low- and high-risk crops are identified based on the historical frequency of potential yield failure from 1986 to 2000 at the border-ethnic homeland pair. Relatively high institutional quality countries are depicted on the right-hand side of the figure. Average probabilities are grouped by 5km distance bins. The solid line depicts the predicted probability based on a third-order RD polynomial (dashed lines denote the 95% confidence intervals).

Table 1: **Potential yields at the border**

	Dependent variable: Potential yield									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Rule of Law	0.138	-0.079	-0.049	-0.024	-0.149	-0.082	-0.286	-0.072	0.068	-0.219
	(0.458)	(0.066)	(0.053)	(0.074)	(0.200)	(0.118)	(0.526)	(0.057)	(0.121)	(0.174)
Border x Ethnic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.807	0.877	0.692	0.635	0.680	0.678	0.701	0.669	0.731	0.697
# Observations	51280	51280	51280	51280	51280	51280	51280	51280	51280	51280
Mean Dep. Var.	9.43	0.77	1.33	1.87	5.17	2.03	13.8	1.60	1.44	3.98
Crop	Cassava	Cocoa	Cotton	Groundnut	Maize	Millet	Sweet Potato	Phaseolus	Rice	Sorghum

NOTES. This table reports the results of equation (1) separately for each of the ten most widely cultivated crops. The dependent variable is potential yield measured in tons per hectare. All regression models are unweighted. Each cell corresponds to a different crop. All models include controls for the interaction of border-country pair and ethnic homeland fixed effects. The sample comprises all plots located between 10 and 100 km from the country border. Standard errors are two-way clustered for both country and ethnic homeland, following the approach of Cameron, Gelbach, and Miller (2011). ***, **, * denote significance at the 1%, 5%, and 10% level.

Table 2: **Land use**

	Dependent variable: Plot is cultivated								
	All plots			*Plots with below-median land quality*			*Plots with above-median land quality*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Rule of Law	-0.126**	-0.140**		-0.084	-0.084		-0.168**	-0.201**	
	(0.057)	(0.063)		(0.064)	(0.081)		(0.073)	(0.078)	
Rule of Law, 25 km			-0.110**			-0.085			-0.136*
			(0.051)			(0.058)			(0.070)
Rule of Law, 50 km			-0.116**			-0.082			-0.150**
			(0.056)			(0.066)			(0.072)
Rule of Law, 75 km			-0.137**			-0.082			-0.196**
			(0.063)			(0.076)			(0.075)
Rule of Law, 100 km			-0.167**			-0.088			-0.254**
			(0.071)			(0.074)			(0.098)
Border x Ethnic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	No	Yes	No	No	Yes	No	No	Yes	No
Border Distance FE	No	No	Yes	No	No	Yes	No	No	Yes
R2	0.431	0.431	0.431	0.507	0.507	0.507	0.330	0.333	0.333
# Observations	51280	51280	51280	25683	25683	25683	25597	25597	25597
Mean Dep. Var.	0.68	0.68	0.68	0.63	0.63	0.63	0.73	0.73	0.73

NOTES. This table reports the effects of *Rule of Law* on the probability of plot cultivation from equations (1)-(3). The dependent variables are an indicator for whether the plot is cultivated. Cols. 1-3 report the results for the full sample of plots, cols. 4-6 report the results for plots with below-median land quality, cols. 7-9 report the results for plots with above-median land quality. Cols. 1,4,7 report the results of the estimation of Equation (1), cols. 2,5,8 report the results of the estimation of Equation (2), and cols. 3,6,9 report the results of the estimation of Equation (3). All regression models are unweighted. All models include controls for the interaction of border-country pair and ethnic homeland fixed effects. The sample comprises of plots located between 10 and 100 km from the country border. Standard errors are two-way clustered for both country and ethnic homeland, following the approach of Cameron, Gelbach, and Miller (2011). ***, **, * denote significance at the 1%, 5%, and 10% level.

Table 3: **Probability of growing low- and high-risk crops**

	Dep var: Low-risk crop is grown				*Dep var: High-risk crop is grown*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rule of Law	-0.129**	-0.144**		-0.125**	-0.063	-0.054		-0.059
	(0.056)	(0.064)		(0.055)	(0.051)	(0.063)		(0.046)
Rule of Law, 25 km			-0.113**				-0.046	
			(0.051)				(0.046)	
Rule of Law, 50 km			-0.119**				-0.061	
			(0.056)				(0.051)	
Rule of Law, 75 km			-0.142**				-0.073	
			(0.063)				(0.062)	
Rule of Law, 100 km			-0.175**				-0.092	
			(0.070)				(0.067)	
Border x Ethnic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	No	Yes	No	No	No	Yes	No	No
Border distance FE	No	No	Yes	No	No	No	Yes	No
Crop FE	No	No	No	Yes	No	No	No	Yes
R2	0.423	0.424	0.424	0.448	0.361	0.362	0.362	0.385
# Observations	51280	51280	51280	51280	51280	51280	51280	51280
Mean dep. var.	0.67	0.67	0.67	0.67	0.46	0.46	0.46	0.46

NOTES. This table reports the effects of *Rule of Law* on the probability of cultivating low- and high-risk crops from equations (1)-(3). The dependent variable is an indicator variable equal to 1 if the crops with the lowest (resp. highest) probability of zero yield are cultivated, and zero otherwise (see text for details). All regression models are unweighted. All models include controls for the interaction of border-country pair and ethnic homeland fixed effects. The sample comprises all plots located between 10 and 100 km from the country border. Standard errors are two-way clustered for both country and ethnic homeland, following the approach of Cameron, Gelbach, and Miller (2011). ***, **, * denote significance at the 1%, 5%, and 10% level.

Table 4: **Share of land cultivated with high- vs low-risk crops**

	Dependent variable: *Log high-risk cropland / high-risk cropland + low-risk cropland*					
	(1)	(2)	(3)	(4)	(5)	(6)
Rule of Law	0.656**	0.685**	1.028***	1.023**	0.916*	1.160***
	(0.320)	(0.312)	(0.345)	(0.401)	(0.483)	(0.414)
Border FE	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.606	0.851	0.629	0.871	0.674	0.858
# Observations	479	314	338	223	338	223
# Country pairs	55	41	44	35	44	35
Mean dep. var.	-3.3	-3.4	-3.2	-3.7	-3.2	-3.7
Same high- & low-risk crops across borders	No	No	Yes	Yes	Yes	Yes
Subset of crops observed at the district level	No	Yes	No	Yes	No	Yes
Weights	No	No	No	No	Yes	Yes

NOTES. This table reports the effects of *Rule of Law* on the relative area devoted to high- and low-risk crops across districts in Africa, based on equation (1). All models include country pairs fixed effects. The sample comprises districts within 100 km of the national border. The dependent variable is logarithm of the fraction of area devoted to the highest risk crops relative to the sum of area devoted to highest and lowest risk crops. Cols. 1-4 are unweighted, cols. 5-6 are weighted by district area. Cols. 1-2 report the results for all border districts, while cols. 3-6 report the results for the subsample of border-pair districts that have the same low- and high-risk crops. Cols. 1, 3, and 5 report results across all crops, cols. 2, 4, and 6 restrict the sample to the subset of crops for which information on area cultivated was reported at the subnational district level. Standard errors are clustered at the country-pair level. ***, **, * denote significance at the 1%, 5%, and 10% level.

Table 5: **Crop Diversification**

	Dep var: Log number of crops per 1000 hectares of cultivated land								
	All plots			*Plots with below-median land quality*			*Plots with above-median land quality*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Rule of Law	-0.325**	-0.439*		-0.377*	-0.489*		0.518**	0.188	
	(0.150)	(0.221)		(0.188)	(0.283)		(0.220)	(0.253)	
Rule of Law, 25 km			-0.338*			-0.375*			0.353
			(0.171)			(0.218)			(0.219)
Rule of Law, 50 km			-0.311*			-0.386*			0.502**
			(0.163)			(0.206)			(0.237)
Rule of Law, 75 km			-0.318**			-0.416*			0.730***
			(0.139)			(0.222)			(0.215)
Rule of Law, 100 km			-0.289**			-0.302			0.642**
			(0.136)			(0.221)			(0.254)
Border x Ethnic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	No	Yes	No	No	Yes	No	No	Yes	No
Border Distance FE	No	No	Yes	No	No	Yes	No	No	Yes
R2	0.612	0.612	0.612	0.480	0.481	0.480	0.503	0.505	0.505
# Observations	34994	34994	34994	16273	16273	16273	18721	18721	18721
Mean Dep. Var.	3.31	3.31	3.31	2.89	2.89	2.89	3.68	3.68	3.68

NOTES. This table reports the effects of *Rule of Law* on crop diversification from equations (1)-(3). The dependent variable is the log of the number of crops per 1,000 hectares of cultivated land. Cols. 1-3 report the results for the full sample of plots, cols. 4-6 report the results for plots with below-median land quality, cols. 7-9 report the results for plots with above-median land quality. The sample comprises cultivated plots located between 10 and 100 km from the country border. In cols. 1-3, observations are weighted by the total area cultivated on the plot, in cols. 4-9, observations are unweighted. All models include controls for the interaction of border-country pair and ethnic homeland fixed effects. Standard errors are two-way clustered for both country and ethnic homeland, following the approach of Cameron, Gelbach, and Miller (2011). ***, **, * denote significance at the 1%, 5%, and 10% level.

Table 6: **Cattle ownership**

	Dep var: Household owns cattle					
	(1)	(2)	(3)	(4)	(5)	(6)
Rule of Law	-0.079**	-0.076**	-0.079**	-0.067**	-0.087***	-0.088***
	(0.034)	(0.030)	(0.030)	(0.032)	(0.026)	(0.026)
Border × Ethnic FE	Yes	Yes	Yes	Yes	Yes	Yes
Demographic covariates	No	No	Yes	No	No	Yes
R2	0.094	0.100	0.144	0.102	0.098	0.146
# Observations	63289	48520	48520	44018	34178	34178
# Country pairs	35	35	35	34	34	34
Mean Dep. Var.	0.58	0.58	0.58	0.57	0.57	0.57
Distance to border	$<$ 100 km	$<$ 100 km	$<$ 100 km	$<$ 50 km	$<$ 50 km	$<$ 50 km
Household owns farmland		Yes	Yes		Yes	Yes

NOTES. This table reports the effects of *Rule of Law* on cattle ownership from equations (1). The dependent variable is an indicator for whether the household reports owning cattle. Cols. 1-3 report results for DHS clusters located within 100 km of the national border, cols. 4-6 report results for clusters within 50 km of the national border. All models include controls for the interaction of border-country pair and ethnic homeland fixed effects. Individual demographic controls include age, education, household size, and gender of the household head. Standard errors are two-way clustered for both country and ethnic homeland, following the approach of Cameron, Gelbach, and Miller (2011). ***, **, * denote significance at the 1%, 5%, and 10% level.

Table 7: **Crop yields per hectare of cultivated land**

	Dep var: Crop yield per hectare									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Rule of Law	1.274	-0.124	0.041	-0.100	-0.088	-0.032	0.752	0.003	-0.085	-0.165
	(1.169)	(0.102)	(0.097)	(0.076)	(0.103)	(0.077)	(1.233)	(0.032)	(0.126)	(0.151)
Border x Ethnic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.512	0.781	0.730	0.629	0.711	0.647	0.747	0.673	0.715	0.683
# Observations	27907	10853	17598	24408	26414	19893	29118	32555	13483	22291
Mean Dep. Var.	8.29	0.50	0.86	0.86	1.26	0.67	7.15	0.23	1.36	0.87
Crop	Cassava	Cocoa	Cotton	Groundnut	Maize	Millet	Sweet potato	Phaseolus	Rice	Sorghum

NOTES. This table reports the effects of *Rule of Law* on crop yields for each of the 10 most widely cultivated crops from equation (1). The dependent variable is crop yield measured in tons per hectare of cultivated land. All models include controls for the interaction of border-country fixed effects. Regression are weighted by the area cultivated. The sample comprises plots located between 10 and 100 km from the country border, on which the crop is cultivated. Standard errors are two-way clustered for both country and ethnic homeland, following the approach of Cameron, Gelbach, and Miller (2011). ***, **, * denote significance at the 1%, 5%, and 10% level.

A Additional Figures and Tables

Table A.1: **District-level data on cultivated land by crop**

Country	# of admin. districts (1)	Ave. district size (km^2) (2)	# crops reported at subnational admin. level (3)	Country	# of admin. districts (4)	Ave. district size (km^2) (5)	# crops reported at subnational admin. level (6)
Algeria	48	49,620	9	Lesotho	10	3,036	4
Angola	163	7,648	1	Liberia	15	7,425	7
Benin	76	1,482	13	Malawi	28	4,232	5
Botswana	27	21,545	5	Mali	50	24,804	12
Burkina Faso	45	6,089	9	Mauritania	13	79,285	4
Burundi	1	27,830	0	Morocco	56	12,694	6
Cameroon	10	47,544	11	Mozambique	130	6,166	6
Central African Rep.	17	36,646	9	Namibia	13	63,494	2
Chad	29	44,276	8	Niger	36	35,194	9
Congo	48	7,125	5	Nigeria	40	23,094	18
Congo DRC	38	61,707	10	Rwanda	14	1,914	9
Cote D'Ivoire	19	16,972	4	Senegal	14	14,052	14
Equatorial Guinea	1	28,051	0	Sierra Leone	14	5,124	10
Eritrea	6	19,600	1	South Africa	355	3,440	11
Ethiopia	72	15,338	4	Swaziland	4	4,341	6
Gabon	37	7,234	0	Tanzania	136	6,950	5
Gambia	6	1,730	1	Togo	21	2,704	13
Ghana	10	23,853	10	Tunisia	24	6,817	5
Guinea	34	7,231	9	Uganda	162	1,457	7
Guinea-Bissau	9	4,014	7	Zambia	72	10,453	10
Kenya	48	12,091	1	Zimbabwe	60	6,513	4
				Continent Average	48	11,827	7
				Continent Median	29	6,950	7

Notes: The district-level data were compiled from Monfreda et al. (2008). See text for discussion. Continent average and median values for district size and number of reported crops are weighted by the number of national administrative districts.

Table A.2: **Potential yields at the border (RD specification)**

	Dependent variable: Potential yield									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Rule of Law	-0.225	-0.056	-0.007	-0.030	0.109	-0.023	0.070	0.025	-0.101	0.004
	(0.431)	(0.046)	(0.061)	(0.077)	(0.237)	(0.097)	(0.253)	(0.070)	(0.112)	(0.174)
Border x Ethnic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.808	0.878	0.693	0.635	0.680	0.678	0.676	0.669	0.731	0.697
# Observations	51280	51280	51280	51280	51280	51280	51280	51280	51280	51280
Mean Dep. Var.	9.43	0.77	1.33	1.87	5.17	2.03	2.90	1.60	1.44	3.98
Crop	Cassava	Cocoa	Cotton	Groundnut	Maize	Millet	Potato	Phaseolus	Rice	Sorghum

NOTES. This table reports the results of equation (2) estimated separately for each crop in the set of the 10 most cultivated crops in the African sample. The dependent variable is the potential yield of the crop measured in tons per hectare. All models include controls for the interaction of border-country pair and ethnic homeland fixed effects. All regression models are unweighted. The sample comprises all plots located between 10 and 100 km from the country border. Standard errors are two-way clustered for both country and ethnic homeland, following the approach of Cameron, Gelbach, and Miller (2011). ***, **, * denote significance at the 1%, 5%, and 10% level.

Table A.3: **Potential yields at the border (border FE specification)**

	Dependent variable: Potential yield									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Rule of Law	0.070	-0.089	-0.039	-0.027	-0.197	-0.059	-0.275	-0.080	0.094	-0.210
	(0.467)	(0.084)	(0.060)	(0.085)	(0.254)	(0.133)	(0.397)	(0.070)	(0.130)	(0.205)
Border FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.715	0.777	0.581	0.510	0.536	0.569	0.576	0.535	0.559	0.573
# Observations	51280	51280	51280	51280	51280	51280	51280	51280	51280	51280
Mean Dep. Var.	9.43	0.77	1.33	1.87	5.17	2.03	2.90	1.60	1.44	3.98
Crop	Cassava	Cocoa	Cotton	Groundnut	Maize	Millet	Potato	Phaseolus	Rice	Sorghum

NOTES. This table reports the results of equation (1) estimated separately for each crop in the set of the 10 most cultivated crops in the African sample. The dependent variable is potential yield measured in tons per hectare. All regression models are unweighted. All models include controls for country pair fixed effects. The sample comprises all plots located between 10 and 100 km from the country border. Standard errors are two-way clustered for both country and border, following the approach of Cameron, Gelbach, and Miller (2011). ***, **, * denote significance at the 1%, 5%, and 10% level.

Table A.4: **Land use (border FE specification)**

	Dependent variable: Plot is cultivated								
	All plots			*Plots with below-median land quality*			*Plots with above-median land quality*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Rule of Law	-0.146***	-0.177***		-0.110	-0.132*		-0.177**	-0.221***	
	(0.053)	(0.060)		(0.066)	(0.073)		(0.069)	(0.077)	
Rule of Law, 25 km			-0.128**			-0.114*			-0.141*
			(0.050)			(0.059)			(0.070)
Rule of Law, 50 km			-0.126**			-0.100			-0.149**
			(0.053)			(0.068)			(0.068)
Rule of Law, 75 km			-0.158***			-0.107			-0.205***
			(0.058)			(0.078)			(0.065)
Rule of Law, 100 km			-0.203***			-0.129*			-0.273***
			(0.065)			(0.077)			(0.084)
Border FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	No	Yes	No	No	Yes	No	No	Yes	No
Border Distance FE	No	No	Yes	No	No	Yes	No	No	Yes
R2	0.304	0.305	0.305	0.355	0.355	0.355	0.248	0.254	0.254
# Observations	51280	51280	51280	25683	25683	25683	25597	25597	25597
Mean Dep. Var.	0.68	0.68	0.68	0.63	0.63	0.63	0.73	0.73	0.73

NOTES. This table reports the effects of *Rule of Law* on the probability of plot cultivation from equations (1)-(3). The dependent variables are an indicator for whether the plot is cultivated. Cols. 1-3 report the results for the full sample of plots, cols. 4-6 report the results for plots with below-median land quality, cols. 7-9 report the results for plots with above-median land quality. All regression models are unweighted. All models include controls for border-country pair fixed effects. The sample comprises of plots located between 10 and 100 km from the country border. Standard errors are two-way clustered for both country and ethnic homeland, following the approach of Cameron, Gelbach, and Miller (2011). ***, **, * denote significance at the 1%, 5%, and 10% level.

Table A.5: **Probability of growing low and high-risk crops (border FE specification)**

	Dep var: Low-risk crop grown			*Dep var: High-risk crop grown*		
	(1)	(2)	(3)	(4)	(5)	(6)
Rule of Law	-0.149***	-0.182***		-0.062	-0.062	
	(0.053)	(0.060)		(0.046)	(0.062)	
Rule of Law, 25 km			-0.130**			-0.051
			(0.050)			(0.046)
Rule of Law, 50 km			-0.127**			-0.065
			(0.053)			(0.048)
Rule of Law, 75 km			-0.163***			-0.070
			(0.057)			(0.051)
Rule of Law, 100 km			-0.211***			-0.065
			(0.064)			(0.059)
Border FE	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	No	Yes	No	No	Yes	No
Border distance FE	No	No	Yes	No	No	Yes
R2	0.295	0.296	0.296	0.276	0.276	0.276
# Observations	51280	51280	51280	51280	51280	51280
Mean dep. var.	0.67	0.67	0.67	0.42	0.42	0.42

NOTES. This table reports the effects of *Rule of Law* on the probability of cultivating low- and high-risk crops from equations (1)-(3). The dependent variable is an indicator variable equal to 1 if the crops with the lowest (resp. highest) probability of zero yield are cultivated, and zero otherwise (see text for details). All regression models are unweighted. All models include controls for border-country pair fixed effects. The sample comprises all plots located between 10 and 100 km from the country border. Standard errors are two-way clustered for both country and ethnic homeland, following the approach of Cameron, Gelbach, and Miller (2011). ***, **, * denote significance at the 1%, 5%, and 10% level.

Table A.6: **Crop choice**

	Dependent variable: Crop is grown					
	Cassava	Sorghum	Groundnut	Millet	Pulses	Potato
	(1)	(2)	(3)	(4)	(5)	(6)
Rule of Law	-0.105*	-0.106*	-0.123**	-0.089**	-0.153**	-0.195***
	(0.056)	(0.063)	(0.049)	(0.038)	(0.058)	(0.064)
Mean Dep. Var.	0.54	0.43	0.47	0.38	0.63	0.56
	Maize	Rice	Wheat	Barley	Other cereal	
	(7)	(8)	(9)	(10)	(11)	
Rule of Law	-0.104*	-0.017	-0.009	-0.047	-0.044	
	(0.061)	(0.033)	(0.011)	(0.050)	(0.037)	
Mean Dep. Var.	0.51	0.245	0.05	0.24	0.271	
	Banana	Coffee	Sugarcane	Oilpalm & other oil crops	Cocoa	
	(12)	(13)	(14)	(15)	(16)	
Rule of Law	0.028	-0.03	0.008	-0.088	-0.03	
	(0.024)	(0.034)	(0.035)	(0.058)	(0.034)	
Mean Dep. Var.	0.17	0.21	0.16	0.28	0.21	

NOTES. This table reports the effects of *Rule of Law* on the probability of crop cultivation from equations (1). The dependent variable is an indicator variable equal to 1 if the crop is grown. All regression models are unweighted. All models include controls for the interaction of border-country pair and ethnic homeland fixed effects. The sample comprises all plots located between 10 and 100 km from the country border. Standard errors are two-way clustered for both country and ethnic homeland, following the approach of Cameron, Gelbach, and Miller (2011). ***, **, * denote significance at the 1%, 5%, and 10% level.

Table A.7: **Potential yields on cultivated land**

	Dependent variable: Potential yield									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Rule of Law	0.432	0.038	-0.028	0.005	-0.374	-0.106	-0.886	-0.084	0.132	-0.230
	(0.357)	(0.051)	(0.070)	(0.100)	(0.340)	(0.120)	(0.542)	(0.087)	(0.083)	(0.267)
Border x Ethnic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.800	0.879	0.657	0.613	0.654	0.632	0.695	0.628	0.773	0.643
# Observations	34994	34994	34994	34994	34994	34994	34994	34994	34994	34994
Mean Dep. Var.	9.90	0.80	1.43	1.99	5.66	2.17	3.28	1.74	1.48	4.38
Crop	Cassava	Cocoa	Cotton	Groundnut	Maize	Millet	Potato	Phaseolus	Rice	Sorghum

NOTES. This table reports the results of equation (1) separately for each of the ten most widely cultivated crops. The dependent variable is potential yield measured in tons per hectare on *cultivated land*. Each observation is a plot weighted by the size of cultivated area. Each cell corresponds to a different crop. All models include controls for the interaction of border-country pair and ethnic homeland fixed effects. The sample comprises all cultivated plots located between 10 and 100 km from the country border. Standard errors are two-way clustered for both country and ethnic homeland, following the approach of Cameron, Gelbach, and Miller (2011). ***, **, * denote significance at the 1%, 5%, and 10% level.

Table A.8: **Crop choice (market-oriented crops)**

	Dep var: Crop is grown					
	(1)	(2)	(3)	(4)	(5)	(6)
Rule of Law	-0.030	-0.123**	-0.003	-0.088	-0.150**	-0.161**
	(0.034)	(0.049)	(0.035)	(0.058)	(0.062)	(0.062)
Border x Ethnic FE	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.500	0.414	0.408	0.426	0.464	0.425
# Observations	51280	51280	51280	51280	51280	51280
Mean Dep. Var.	0.21	0.47	0.23	0.28	0.60	0.63
Crop	Cocoa	Groundnut	Banana, coconut, sugar crops	Oilpalm and other oil crops	Best crop under US prices	Best crop under French prices

NOTES. This table reports the effects of *Rule of Law* on the probability of crop cultivation from equations (1). The dependent variable is an indicator variable equal to 1 if the crop is grown. All regression models are unweighted. All models include controls for the interaction of border-country pair and ethnic homeland fixed effects. The sample comprises all plots located between 10 and 100 km from the country border. Standard errors are two-way clustered for both country and ethnic homeland, following the approach of Cameron, Gelbach, and Miller (2011). ***, **, * denote significance at the 1%, 5%, and 10% level.

B Model

B.1 Assumptions and notations

We consider an agent's portfolio allocation problem with m assets. The first two assumptions reflect the fact that crop yields respond differently to agroclimatic conditions. These assumptions are consistent with the actual distribution of yields across crops and across years (see Figure 1).

Assumption 1 For any k, l, with $k < l$, the probability of positive yield of asset k *statewise stochastically dominates* the probability of positive yield of asset l, which means that, for any state of the world, asset l has a positive yield if and only if asset k has a positive yield as well. For simplicity, we assume that there are exactly $m+1$ states of the world, numbered from 0 to m, and that asset k has a positive yield if and only if the state of the world is in the subset $\{k, k+1, ..., m\}$. We denote p_k, with $p_k > 0$ if $k > 0$, the probability that state k occurs, with $\sum_{k=0}^{m} p_k = 1$.

Assumption 2 The yield of asset k is $q_k > 0$ in states k to m. To avoid discussions of knife-edge cases, we also assume that all assets have different expected revenues.

Assumption 3 Let $w > 0$ denote the agent's wealth. The agent's utility is the log of the sum of her wealth and the revenue from her portfolio.

Agent's problem The agent's problem is:

$$\underset{(\alpha_k)_{k=1}^{m}}{\text{Max}} \sum_{j=1}^{n} p_j \log\left(w + \sum_{i=1}^{j} \alpha_i q_i\right)$$

$$\textit{such that } 0 \leq \alpha_i \leq 1, \sum_{i=1}^{m} \alpha_i = 1$$

Notations We denote $V(w, q_1, ..., q_m, p_1, ..., p_m) \equiv \sum_{j=1}^{m} p_j \log\left(w + \sum_{i=1}^{j} \alpha_i^* q_i\right)$ the value of the objective function at the optimal solution of this problem.

We define $r_k \equiv q_k \sum_{i=k}^{m} p_i$, the expected revenue of asset k, and $w_{j,k} = \frac{\sum_{i=j} k{-}1 p_j q_j q_k}{r_k - r_j}$ if $1 \leq j < k \leq m$.

To solve the agent's problem, we analyze the interior solution first, and then the general solution.

B.2 Interior solution

Lemma 1. *For any n, w, and $(q_k, p_k)_{k=1}^{n}$ such that there exists an interior solution to the agent's problem, V can be written as the sum of two terms $V(w, q_1, ..., q_n, p_1, ..., p_n) = \sum_{j=1}^{n} p_j \log(w + q_1) + G(q_1, ..., q_n, p_1, ..., p_n)$, where G is a function of $(q_1, ..., q_n, p_1, ..., p_n)$ and not of w.*

Proof. The lemma is trivially true if $n = 1$, in which case $V(w, q_1, p_1) = p_1 \log(w + q_1)$. Suppose $n = 2$. We have, assuming an interior solution exists:

$$\alpha_1^* = \frac{(p_1+p_2)wq_1 + p_1q_1q_2 - p_2q_2w}{(p_1+p_2)q_1(q_2-q_1)},$$

and hence,

$$V(w, q_1, q_2, p_1, p_2) = (p_1+p_2)\log(w+q_1) + G(q_1, q_2, p_1, p_2),$$

where G is a function of (q_1, q_2, p_1, p_2), not of w.
Suppose by induction that, for any interior solution with $n-1, w$, and $(q_k, p_k)_{k=1}^{n-1}$ for which there exists an interior solution to the agent's problem,

$$V(w, q_1, ..., q_{n-1}, p_1, ..., p_{n-1}) = \sum_{j=1}^{n-1} p_j \log(w+q_1) + G(q_1, ..., q_{n-1}, p_1, ..., p_{n-1}),$$

where G is a function of $(q_1, ..., q_{n-1}, p_1, ..., p_{n-1})$ and not of w.
The agent's problem for n assets,

$$\begin{aligned} \underset{(\alpha_k)_{k=1}^n}{\text{Max}} \quad & \sum_{j=1}^{n} p_j \log\left(w + \sum_{i=1}^{j} \alpha_i q_i\right) \\ \textit{such that } & 0 \le \alpha_i \le 1, \sum_{i=1}^{n} \alpha_i = 1 \end{aligned}$$

is equivalent to the following problem:

$$\begin{aligned} \underset{\alpha,(\beta_k)_{k=2}^n}{\text{Max}} \quad & p_1\log(w+\alpha q_1) + \sum_{j=2}^{n} p_j \log\left(w + \alpha q_1 + (1-\alpha)\sum_{i=2}^{j} \beta_i q_i\right) \\ \textit{such that } & 0 \le \alpha \le 1, \sum_{i=2}^{n} \beta_i = 1 \end{aligned}$$

We are interested in the optimal α of the last problem, which is the solution of the following problem:

$$\begin{aligned} \underset{\alpha}{\text{Max}} \quad & p_1\log(w+\alpha q_1) + \sum_{j=2}^{n} p_j \log(1-\alpha) + V\left(\frac{w+\alpha q_1}{1-\alpha}, q_2, ..., q_n, p_2, ..., p_n\right) \\ \textit{such that } & 0 \le \alpha \le 1 \end{aligned}$$

If this problem has an interior solution, then, necessarily, $0 < \alpha^* < 1$ and the agent's problem with $n-1$ assets and parameters $\frac{w+\alpha^* q_1}{1-\alpha}$ and $(q_k, p_k)_{k=2}^n$ has an interior solution. In addition, the induction hypothesis implies that there exists a function G such that, where defined,

$$V(\frac{w+\alpha q_1}{1-\alpha}, q_2, ..., q_n, p_2, ..., p_n) = \sum_{j=2}^{n} p_j \log(\frac{w+\alpha q_1}{1-\alpha} + q_2) + G(q_1, ..., q_n, p_1, ..., p_n),$$

so the last problem is equivalent to the following:

$$\underset{\alpha}{\text{Max}} \quad p_1 \log(w+\alpha q_1) + \sum_{j=2}^{n} p_j \log\left(w+\alpha q_1 + (1-\alpha) q_2\right) + G(q_1, ..., q_n, p_1, ..., p_n)$$

such that $0 \leq \alpha \leq 1$

From the solution above, we thus have that

$$V(w, q_1, q_2, ..., q_n, p_1, p_2, ..., p_n) = \sum_{j=1}^{n} p_j \log(w+q_1) + G(q_1, ..., q_n, p_1, ..., p_n),$$

where G is a function of $(q_1, ..., q_n, p_1, ..., p_n)$, not of w, which proves the induction hypothesis for n, and the lemma. □

Lemma 2. *The agent's problem can be rewritten as follows:*

$$\underset{(\beta_k)_{k=1}^n}{\text{Max}} \quad \sum_{j=1}^{n} p_j \log\left(w + \sum_{i=1}^{j} \beta_i q_i \Pi_{t=1}^{j-1}(1-\beta_t)\right)$$

such that $0 \leq \beta_i \leq 1, \beta_n = 1$

For any interior solution, we have with the notations defined above, for any $k = 1...n-1$, $\beta_k^* = \frac{(r_{k+1}-r_k)(w_{k,k+1}-w_{k-1,k})}{r_k(q_{k+1}-q_k)}$, *where* $w_{0,1} \equiv w$.

Proof. Necessarily, the optimal β_1 in the agent's problem must solve:

$$\underset{\beta_1}{\text{Max}} \quad p_1 \log(w_{0,1} + \beta_1 q_1) + \sum_{j=2}^{n} p_j \log(1-\beta_1) + V(\frac{w_{0,1}+\beta_1 q_1}{1-\beta_1}, q_2, ..., q_n, p_2, ..., p_n)$$

such that $0 \leq \beta_1 \leq 1$

where, by definition,

$$V(w,q_2,...,q_n,p_2,...,p_n) \equiv \underset{(\beta_k)_{k=2}^n}{\text{Max}} \quad \sum_{j=2}^{n} p_j \log\left(w + \sum_{i=2}^{j} \beta_i q_i \Pi_{t=2}^{j-1}(1-\beta_t)\right)$$
$$\textit{such that} \quad 0 \leq \beta_i \leq 1, \beta_n = 1$$

Suppose, by induction, that necessarily, for some $k \leq n-2$, the optimal β_k in the agent's problem must solve:

$$\begin{aligned}
\underset{\beta_k}{\text{Max}} \quad & p_1 \log(w_{0,1} + \beta_1^* q_1) + \sum_{j=2}^{n} p_j \log(1-\beta_1^*) + ... \\
& + p_{k-1} \log(w_{k-2,k-1} + \beta_{k-1}^* q_{k-1}) + \sum_{j=k}^{n} p_j \log(1-\beta_{k-1}^*) \\
& + p_k \log(w_{k-1,k} + \beta_k q_k) + \sum_{j=k+1}^{n} p_j \log(1-\beta_k) \\
& + V(\frac{w_{k-1,k} + \beta_k q_k}{1-\beta_k}, q_{k+1}, ..., q_n, p_{k+1}, ..., p_n) \\
\textit{such that } & 0 \leq \beta_k \leq 1
\end{aligned}$$

Then, using Lemma 1 to rewrite the last term of the objective function, this problem becomes:

$$\begin{aligned}
\underset{\beta_k}{\text{Max}} \quad & p_1 \log(w_{0,1} + \beta_1^* q_1) + \sum_{j=2}^{n} p_j \log(1-\beta_1^*) + ... \\
& + p_{k-1} \log(w_{k-2,k-1} + \beta_{k-1}^* q_{k-1}) + \sum_{j=k}^{n} p_j \log(1-\beta_{k-1}^*) \\
& + p_k \log(w_{k-1,k} + \beta_k q_k) + \sum_{j=k+1}^{n} p_j \log(1-\beta_k) \\
& + \sum_{j=k}^{n} p_j \log\left(\frac{w_{k-1,k} + \beta_k q_k}{1-\beta_k} + q_{k+1}\right) + G(q_{k+1}, ..., q_n, p_{k+1}, ..., p_n) \\
\textit{such that } & 0 \leq \beta_k \leq 1
\end{aligned}$$

where G does not depend on w_k or β_k.
The first order condition implies that:

$$\beta_k^* = \frac{(r_{k+1} - r_k)(w_{k,k+1} - w_{k-1,k})}{r_k(q_{k+1} - q_k)}$$

In addition, $\frac{w_{k-1,k}+\beta_k^* q_k}{1-\beta_k^*} = w_{k,k+1}$, so that, necessarily, the optimal β_{k+1} in the agent's problem must solve:

$$\begin{aligned}
\underset{\beta_{k+1}}{\text{Max}} \quad & p_1 \log(w_{0,1} + \beta_1^* q_1) + \sum_{j=2}^{n} p_j \log(1-\beta_1^*) + ... \\
& + p_k \log(w_{k-1,k} + \beta_k^* q_k) + \sum_{j=k+1}^{n} p_j \log(1-\beta_k^*) \\
& + p_{k+1} \log(w_{k,k+1} + \beta_{k+1} q_{k+1}) + \sum_{j=k+2}^{n} p_j \log(1-\beta_{k+1}) \\
& + V(\frac{w_{k,k+1} + \beta_{k+1} q_{k+1}}{1-\beta_{k+1}}, q_{k+2}, ..., q_n, p_{k+2}, ..., p_n) \\
\textit{such that } & 0 \leq \beta_{k+1} \leq 1
\end{aligned}$$

which proves the induction hypothesis. As can be seen, the induction hypothesis implies that, necessarily:

$$\beta_k^* = \frac{(r_{k+1} - r_k)(w_{k,k+1} - w_{k-1,k})}{r_k(q_{k+1} - q_k)},$$

which proves the lemma. □

Lemma 3. *(1) If the shares of all assets are positive at the optimum, necessarily* $r_k < r_{k+1}$*, which implies* $q_k < q_{k+1}$*, for any* $1 \leq k \leq n-1$*.*
(2) With the notations of Lemma 2, assuming that the shares of all other assets are positive at the optimum:
(2.1) the share of asset 1 *is positive only if* $w_{0,1} < w_{1,2}$*,*
(2.2) the share of asset k*, for* $1 < k < n$*, is positive only if* $r_{k-1} < r_k$ *and* $w_{k-1,k} < w_{k,k+1}$*,*
(2.3) the share of asset n *is positive if and only if* $r_{n-1} < r_n$*.*

Proof. (1) $r_k < r_{k+1}$ is a direct implication of the concavity of the objective function and the statewise stochastic order on the probabilities of positive yields of assets. Since $r_k = (p_k + \sum_{i=k+1}^{n}) q_k < \sum_{i=k+1}^{n} q_k$, $r_k < r_{k+1}$ implies $q_k < q_{k+1}$, for any $1 \leq k \leq n-1$.
(2.1) From Lemma 2, assuming that the shares of all other assets are positive, the share of asset 1 is positive only if $\beta_1^* > 0$, which implies $w_{1,2} - w_{0,1} > 0$, since $r_2 > r_1$ from (1).
(2.2) From Lemma 2, assuming that the shares of all other assets are positive, the share of asset k is positive only if $\beta_k^* > 0$ and $\beta_{k-1}^* < 1$, i.e.:

$$(r_{k+1} - r_k)(w_{k,k+1} - w_{k-1,k}) > 0 \Leftrightarrow w_{k,k+1} - w_{k-1,k} > 0$$

since $r_k < r_{k+1}$ from (1), and

$$\frac{(r_k - r_{k-1})(w_{k-1,k} - w_{k-2,k-1})}{r_{k-1}(q_k - q_{k-1})} < 1$$
$$\Leftrightarrow (r_k - r_{k-1})(w_{k-1,k} - w_{k-2,k-1}) < r_{k-1}(q_k - q_{k-1}),$$

since $q_k > q_{k-1}$ from (1), and this last inequality is equivalent to the following ones:

$$\begin{aligned} r_{k-1}w_{k-2,k-1} - r_k w_{k-2,k-1} + p_{k-1}q_{k-1}q_k &< r_{k-1}q_k - r_{k-1}q_{k-1} \\ \Leftrightarrow r_{k-1}(w_{k-2,k-1} + q_{k-1}) &< r_k(w_{k-2,k-1} + q_{k-1}) \\ \Leftrightarrow r_{k-1} &< r_k \end{aligned}$$

since $q_k > q_{k-1}$ from (1).

(2.3) From Lemma 2, assuming that the shares of all other assets are positive, the share of asset n is positive if and only if $\beta^*_{n-1} < 1$, i.e.

$$\frac{(r_n - r_{n-1})(w_{n-1,n} - w_{n-2,n-1})}{r_{n-1}(q_n - q_{n-1})} < 1$$

which is equivalent to $r_{n-1} < r_n$, since $q_n > q_{n-1}$ from (1), using the same steps as in the proof of (2.2).

□

B.3 Farmer's problem

We consider a farmer who may allocate his total time, normalized to one, between non-farming labor (ℓ) and farming $(1-\ell)$. The total area that he may cultivate is normalized to one, and we interpret $(1-\ell)$ as the size of the area cultivated. There are n crops available, and the farmer allocates a share α_i of cultivated land to the cultivation of crop i, $i = 1...n$.

We assume that the farmer's utility is:

$$u(w, s, \ell, v_1, ..., v_n, \alpha_1, ..., \alpha_n) \equiv \log\left(w + s\ell + (1-\ell)\sum_{i=1}^{n}\alpha_i v_i\right), \tag{B.1}$$

where w is the farmer's wealth, s is the salary for non-farming labor, and v_i is the revenue from the production of crop i per unit of land cultivated with i.

The farmer's problem is:

$$\begin{cases} \underset{\{\alpha_i,\ i=1...n\}}{\text{Max}} \mathbf{E}\Big[\log\Big(w + s\ell + \sum_{i=1}^n \alpha_i v_i\Big)\Big] \\ \\ \text{such that } 0 \le \ell \le 1,\ 0 \le \alpha_i \le 1, \sum_{i=1}^n \alpha_i = 1 \end{cases}$$

To simplify the discussion, we assume that the variables v_i are stochastic, whereas s and w are deterministic.

Proposition 1. *(1) Assume that the random variables v_i are non-negative and bounded, and let $m \equiv argmax_{\{i=1...n\}}\boldsymbol{E}[v_i]$.*
(1.1) For any j, crop j is not grown if $s \geq \boldsymbol{E}[v_j]$.
(1.2) Given s, for any $j \neq m$ there exists w_j such that if $w \geq w_j$, crop j is not grown at the optimum. In addition, if $\boldsymbol{E}[v_m] > s$, there exists w_s such that if $w \geq w_s$, $\ell^* = 0$.
(2) Assume in addition that, for any $k < l$, the probability of positive yield of crop k statewise stochastically dominates the probability of positive yield of crop l, and that, when positive, the value of a crop is constant.
Then, the crop with the smallest number of positive yields among cultivated crops is necessarily the crop with the largest expected value.

Proof. (1.1) This is an implication of the concavity of the utility function and the deterministic s.
(1.2) Let $j \neq m$. We denote $f(v_1, ..., v_n)$ the density function of $(v_1, ..., v_n)$, and $\overline{v} = \text{Max}(s, \underset{\{k=1...n\}}{\text{Max}} \underset{\{v_k\}}{\text{Max}}\, v_k)$, the maximum of s and of the maximal value of all crops, which is finite by assumption, and

$$\mathbf{E}\left[\log\left(w + s\ell + \sum_{i=1}^{n} \alpha_i v_i\right)\right] = \int \cdots \int \log\left(w + s\ell + \sum_{i=1}^{n} \alpha_i v_i\right) f(v_1, ..., v_n) dv_1 ... dv_n \quad \text{(B.2)}$$

Let k be the crop with the smallest expected value larger than the expected value of j; k exists because $j \neq m$. Suppose that j and k are grown at the optimum, then necessarily, the solution β^* of the following problem:

$$\underset{\beta}{\text{Max}}\, \mathbf{E}\left[\log\left(w + s\ell^* + \sum_{i=1, i\neq j, i\neq k}^{n} \alpha_i^* v_i + \alpha_j^*(\beta v_j + (1-\beta)v_k)\right)\right]$$
$$\textit{such that} \quad 0 \leq \beta \leq 1,$$

should be strictly between zero and one, where ℓ^* and α_i^* are the optimal time of non-farming labor and share of crop i.

The derivative of the objective function is at $\beta = 0$ is:

$$\int \cdots \int \frac{v_k - v_j}{w + s\ell + \sum_{i=1, i\neq j, i\neq k}^{n} \alpha_i^* v_i + \alpha_j^* v_j} f(v_1, ..., v_n) dv_1 ... dv_n \quad \text{(B.3)}$$

Given that $v_i \geq 0$, this term is smaller than:

$$\int \cdots \int \frac{v_j}{w} f(v_1, ..., v_n) dv_1 ... dv_n - \int \cdots \int \frac{v_k}{w + \overline{v}} f(v_1, ..., v_n) dv_1 ... dv_n = \frac{\mathbf{E}[v_j]}{w} - \frac{\mathbf{E}[v_k]}{w + \overline{v}} \quad \text{(B.4)}$$

Let

$$w_j \equiv \frac{\overline{v}\mathbf{E}[v_j]}{\mathbf{E}[v_k] - \mathbf{E}[v_j]} \quad \text{(B.5)}$$

This term is positive and finite, and, if $w > w_j$, the derivative of the objective function at $\beta = 0$ is negative, so $\beta^* = 0$: crop j is not grown.
(2) This is a corollary of Lemma 3 - (2.3) above. □

The proposition presents two series of results. Part (1) derives results with minimal assumptions on crops revenues (they are non-negative and bounded). Part (2) derives additional results under two stronger assumptions on these revenues. The first assumption – for any $k < l$, the probability of positive yield of crop k statewise stochastically dominates the probability of positive yield of crop l – is consistent with historical crop yields (see Figure 1). The second assumption – when positive, the value of a crop is constant – is also consistent with the data on historical crop yields under the assumption that prices are constant. This assumption is aimed at formalizing the effect that changes in wealth or salary may have on subsistence farming, to which a substantial share of land may be devoted. In subsistence farming, the value of a unit of production of a crop is some function of nutritional value and taste, and hence, may not vary with demand variations, prices fluctuations or agro-climatic conditions.
We emphasize that, even if prices are constant, it is not possible in general to infer which crop can be considered high-risk/high-expected revenue from the distribution of yields only. Part (2) shows here that, given the specific properties of the distribution of yields, we do not need to observe crop prices to infer that a crop with a smaller number of positive yields than another one, among cultivated crops, necessarily has lower expected revenue, and can be considered low-risk/low-revenue. Crops with different expected revenues may be grown because the farmer is risk-averse, and crops with lower expected revenue may also have lower risk.
Remark. Perfect insurance on farming income is equivalent to an increase in w large enough to make the cultivation of any crop except perhaps the one with the largest expected revenue profitable.