ONLINE APPENDIX FOR "CROP YIELD AND DEMOCRACY"

1. Distribution of Potential Crop Yield

Table A1: Bottom and Top Decile Countries by Potential Crop Yield

Continent	Bottom Decile		Top Decile	
	Country (Crop yield)	Polity2 (1961-2015)	Country (Crop yield)	Polity2 (1961-2015)
Africa	Egypt (36)	-6	Madagascar (11475)	1
	Djibouti (70)	-3	Swaziland (11661)	-8
	Libya (323)	-6	Benin (11914)	0
	Mauritania (425)	-6	Tanzania (12150)	-4
	Algeria (512)	-5	Mozambique (12223)	-1
			Malawi (12241)	-3
			Guinea-Bissau (12331)	-1
			Zambia (12339)	-1
Asia	Yemen (0)	-2	Korea, Republic (14018)	2
	Bahrain (0)	-9		
	Qatar (0)	-10		
	Saudi Arabia (6)	-10		
	Kuwait (23)	-8		
	Oman (59)	-9		
	UAE (63)	-8		
	Mongolia (597)	1		
	Jordon (880)	-6		
Europe	Norway (908)	10	Hungary (11743)	1
North America			Cuba (11848)	-7
			Dominican Republic (12071)	4
			Haiti (12339)	-4
South America			Paraguay (12856)	-1

Notes: provide and The index of potential crop yield in million kilocalories per hectare per year comes from Galor and Özak (2016). Democracy is measured by Polity2 from Marshall et al. (2015).

2. Summary Statistics and Correlations

Table A2: Summary Statistics

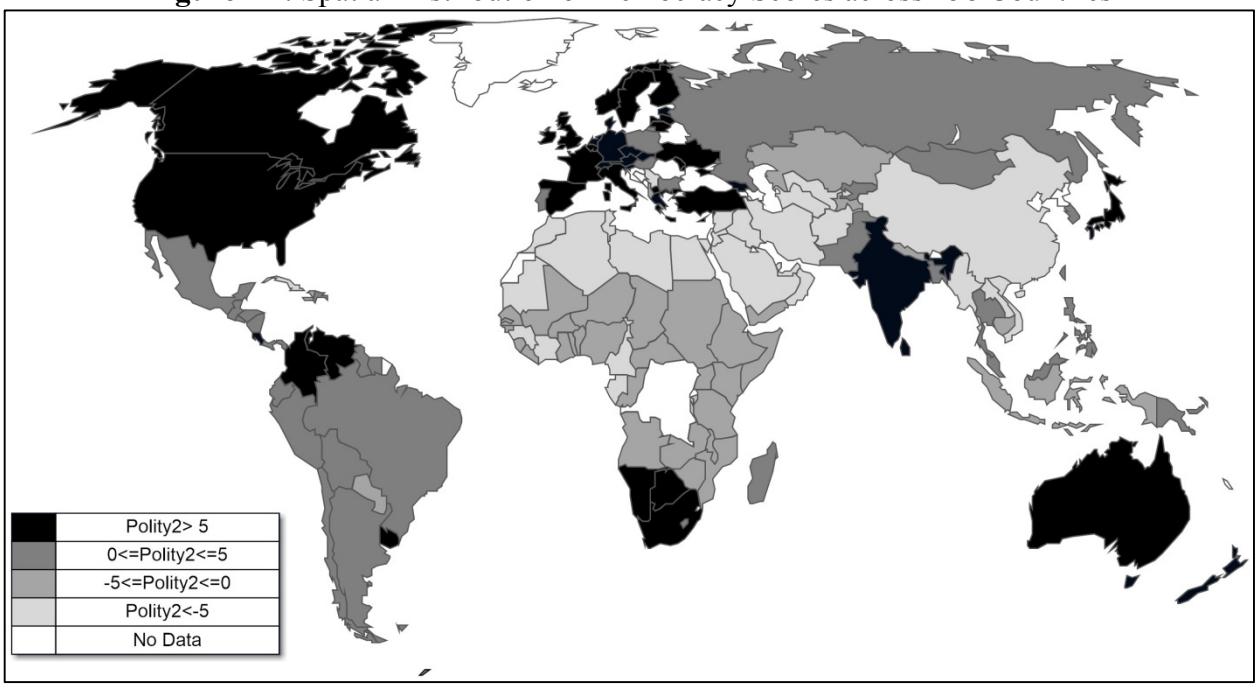
	Number of observations	Mean	Standard deviation	Minimum	Maximum
Polity2 (1961-2015)	147	0.77	6.15	-10.0	10.0
Crop yield	147	7262	3816	0.0	14017
Absolute latitude	147	26.51	16.94	1.0	64.0
Terrain ruggedness	147	1.32	1.25	0.03	6.74
Landlocked	147	0.24	0.43	0.0	1.0
Elevation (average)	147	0.57	0.50	0.02	2.67
Elevation (variation)	147	0.38	0.36	0.0	1.92
Distance to waterways	147	0.37	0.42	0.0	2.21

Table A3: Correlation Matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Polity2 (1961-2015)	1.00							
(2) Crop yield	0.20	1.00						
(3) Absolute latitude	0.43	-0.25	1.00					
(4) Terrain ruggedness	0.01	-0.00	0.15	1				
(5) Landlocked	-0.10	0.06	0.10	0.27	1			
(6) Elevation (average)	-0.15	-0.08	-0.02	0.71	0.48	1		
(7) Elevation (variation)	-0.06	-0.10	-0.01	0.57	0.14	0.74	1	
(8) Distance to waterways	-0.21	-0.26	0.11	-0.02	0.57	0.42	0.23	1

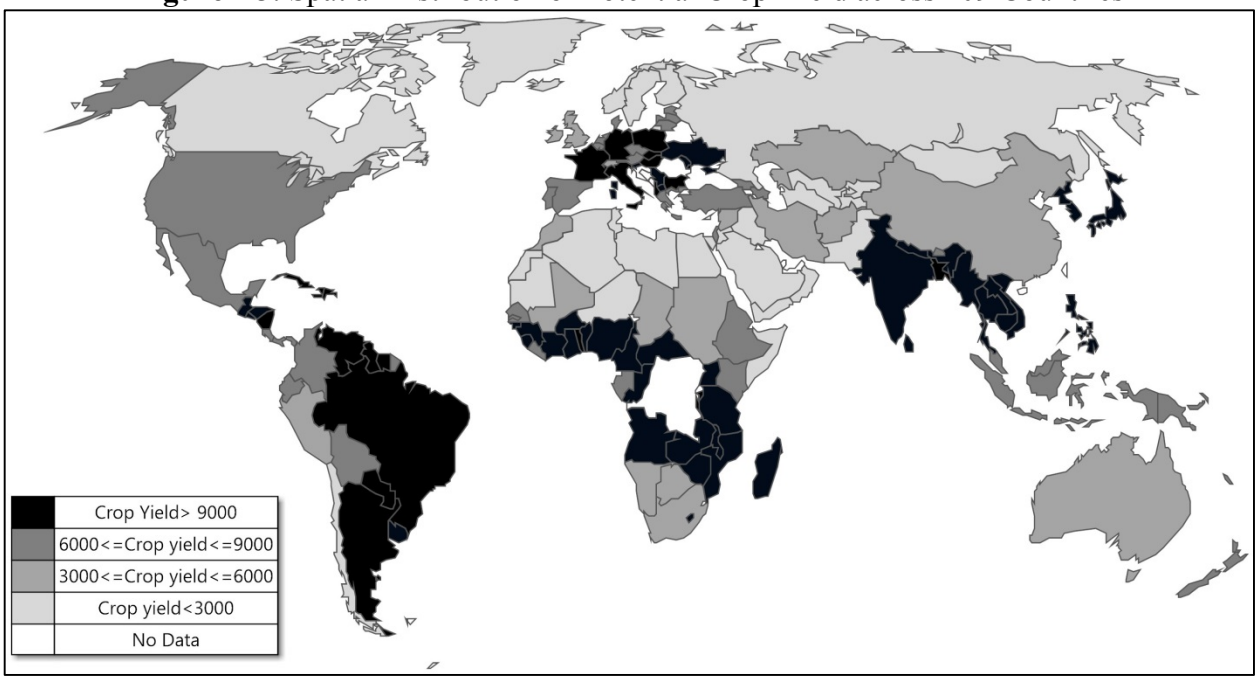
Notes: This table presents summary statistics and correlations of variables used in baseline estimation. The number of observations is 147.

Figure A2: Spatial Distribution of Democracy Scores across 158 Countries



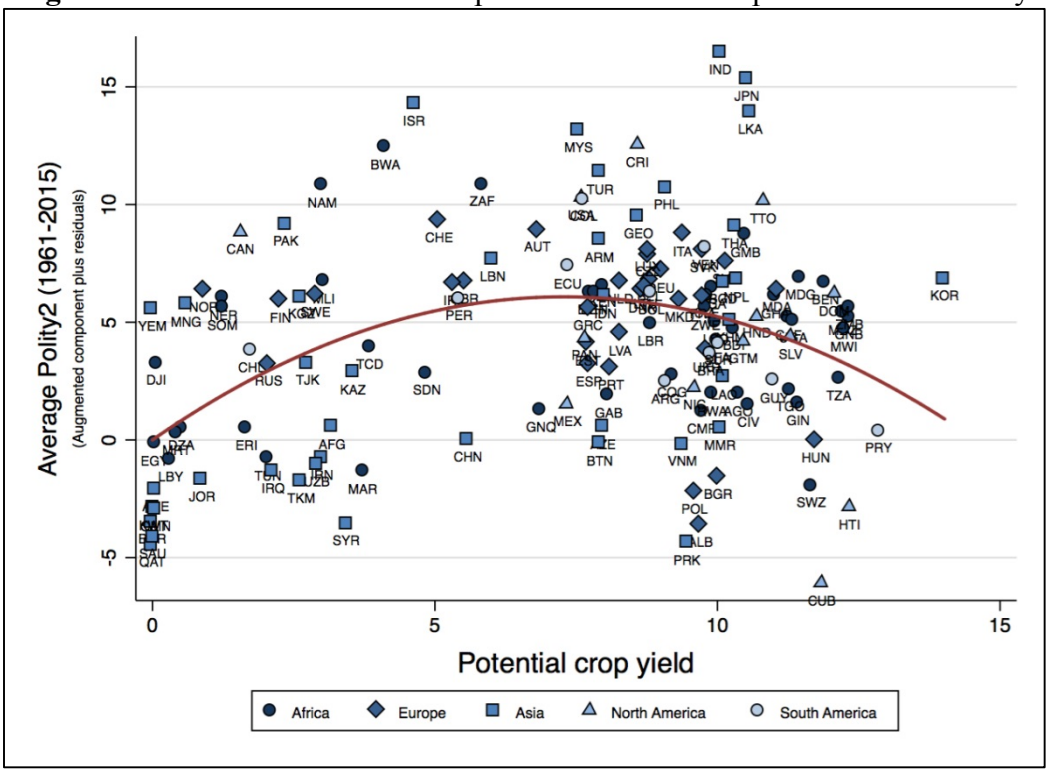
Notes: The democracy scores are measured as the average of Polity2 score over the period 1961-2015.

Figure A3: Spatial Distribution of Potential Crop Yield across 209 Countries



Notes: Potential crop yield is measured in millions of kilocalories per year per hectare.

Figure A4: The Non-linear Relationship between Potential Crop Yield and Democracy



Notes: This figure demonstrates a hump-shaped effect using a least-squares quadratic fit estimator, following the approach of Ashraf and Galor (2013). The estimation includes all control variables, following the baseline line model in the last column of Table 1. This figure is an augmented component-plus-residual plot rather than the typical added-variable plot of residuals against residuals. The vertical axis represents fitted values (as predicted by potential crop yield and its square) of average Polity2 score plus the residuals from the full regression model. The horizontal axis represents potential crop yield rather than the residuals obtained from regressing democracy on the control variables in the model. This approach enables us to illustrate the overall non-monotonic effect of potential crop yield on democracy.

4. Additional Regression Results

Table A4: Tobit estimations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Lower bound (-10)					Upper bound (+10)				
Dependent variable	Polity2 (1961-2015)									
Crop yield	0.35***	2.75***	2.43***	2.47***	1.74***	0.31**	2.90***	2.53***	2.63***	1.80***
	(2.64)	(5.76)	(5.77)	(5.92)	(4.48)	(2.10)	(5.40)	(5.41)	(5.66)	(4.20)
Crop yield square		-0.19***	-0.15***	-0.16***	-0.12***		-0.21***	-0.16***	-0.18***	-0.13***
		(-5.21)	(-4.62)	(-4.87)	(-3.93)		(-4.98)	(-4.37)	(-4.74)	(-3.78)
Absolute latitude			0.17***	0.17***	0.10**			0.20***	0.21***	0.13***
			(6.69)	(6.75)	(2.59)			(6.94)	(7.02)	(2.92)
Terrain ruggedness				-0.37	-0.22				-0.46	-0.30
				(-0.68)	(-0.44)				(-0.76)	(-0.55)
Elevation (Average)				-0.40	2.32				-0.40	2.55
				(-0.23)	(1.43)				(-0.20)	(1.42)
Elevation (std dev.)				-0.01	-2.91				-0.30	-3.28*
				(-0.00)	(-1.65)				(-0.15)	(-1.68)
Landlocked				-0.75	-1.17				-0.74	-1.16
				(-0.58)	(-1.03)				(-0.51)	(-0.92)
Distance to waterways				-2.10	-1.23				-2.73*	-1.73
				(-1.49)	(-0.97)				(-1.72)	(-1.23)
Continent FE	No	No	No	No	Yes	No	No	No	No	Yes
Observations	147	147	147	147	147	147	147	147	147	147

Notes: The table presents Tobit estimation of crop yield on average Polity2 score over 1961-2015. The continent dummies are Africa, Asia, Australia, Europe, North America, Oceania, and South America. Baseline controls used are absolute latitude, terrain ruggedness, elevation (average), elevation (variation), landlockedness, distance to waterways. Robust t-statistics are given in parentheses. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively.

Table A5: Establishing an Inverted U-shaped Relationship

	(1)	(2)	(3)
Dependent Variable	Polity2 (1961-2015)		
Crop yield	0.40***	-0.22***	1.64***
	(3.99)	(-2.76)	(2.66)
Crop yield square			-2.42*
			(-1.69)
Crop yield cube			0.94
			(1.10)
Baseline controls	Yes	Yes	Yes
Continent FE	No	No	Yes
Observations	52	95	147
Sample	Crop yield < 7,286	Crop yield > 7,286	Global
adj. R^2	0.40	0.35	0.54

Notes: The continent dummies are Africa, Asia, Australia, Europe, North America, Oceania, and South America. Baseline controls used are absolute latitude, terrain ruggedness, elevation (average), elevation (variation), landlockedness, distance to waterways. Robust t-statistics are given in parentheses. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively.

Table A6: Accounting for the Quadratic Effects of Early Development

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	Polity2 (1961-2015)						
Crop yield	1.05***	0.86***	1.02***	0.94***	0.97***	1.09***	0.96***
	(3.82)	(3.09)	(3.82)	(3.57)	(3.61)	(4.02)	(3.69)
Crop yield square	-0.83***	-0.87***	-0.96***	-0.86***	-0.99***	-0.94***	-0.87***
	(-3.18)	(-3.07)	(-3.65)	(-3.28)	(-3.69)	(-3.50)	(-3.33)
Pathogen stress	-0.23						
	(-1.60)						
Pathogen stress square	0.03						
	(0.39)						
State antiquity (1500)		-0.54**					
		(-2.05)					
State antiquity (1500) square		0.45*					
		(1.80)					
Years since agricultural transition			-0.41				
			(-1.13)				
Years since agricultural transition square			0.18				
			(0.50)				
Logged population density (1500)				-0.24**			
				(-2.25)			
Logged population density (1500) square				0.31***			
				(4.49)			
Biogeography					1.33		
					(1.36)		
Biogeography square					-1.58		
					(-1.66)		
Human settlement						-0.24	
						(-0.70)	
Human settlement square						0.31	
						(0.93)	
Genetic diversity (predicted)							2.94
							(0.99)
Genetic diversity (predicted) square							-3.08
							(-1.09)
Observations	132	131	144	137	129	139	147
Adj. R-squared	0.585	0.520	0.545	0.570	0.547	0.541	0.538
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Inverted U-shape test	2.45	2.83	3.28	2.85	3.55	2.88	2.86
	[p=0.01]	[p=0.00]	[p=0.00]	[p=0.00]	[p=0.00]	[p=0.00]	[p=0.00]

Notes: This table presents standardized coefficients for the effect of (mean) crop yield (measured in millions of kilocalories per hectare per year) on Polity2 during 1961-2014. All specifications use an intercept term but it is not reported for brevity. The continent dummies are Africa, Asia, Australia, Europe, North America, Oceania, and South America. Baseline controls used are absolute latitude, terrain ruggedness, elevation (average), elevation (variation), landlockedness, distance to waterways. Robust t-statistics are given in parentheses. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively. The inverted U-shape test is performed using the approach of Lind and Mehlum (2010).

Table A7: Further Robustness Tests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Dependent Variable	Polity2 (1961-2015)										
Crop yield	1.28***	0.78**	0.82**	0.97***	0.98***	1.07***	0.92**	0.74**	1.14***	1.22***	1.03***
	(4.84)	(2.34)	(2.26)	(4.07)	(3.95)	(3.78)	(2.56)	(2.20)	(2.95)	(4.91)	(3.87)
Crop yield square	-1.15***	-0.73**	-0.76**	-0.91***	-0.89***	-0.94***	-0.70*	-0.61*	-0.87**	-1.03***	-0.91***
	(-4.52)	(-2.24)	(-2.17)	(-3.65)	(-3.49)	(-3.39)	(-1.95)	(-1.82)	(-2.17)	(-4.05)	(-3.34)
Daily Newspaper circulation	0.36***										
	(4.46)										
Human capital stock – Lee and Lee (2016)		0.51***									
		(4.41)									
Alternative human capital stock - Lee and Lee (2016)			0.43***								
			(3.66)								
Oil export >1/3 of total export				-0.21***							
				(-3.51)							
OPEC dummy					-0.13***						
					(-2.75)						
Presence of oil/diamond						-0.06					
						(-0.88)					
Risky behavior - Importance of secure surroundings							-0.08				
							(-0.79)				
Risky behavior - Importance of taking risk								-0.21			
								(-1.45)			
Risky behavior - Preference for security									0.16*		
									(1.81)		
F_{ST} distance										-0.02	
										(-0.22)	
Blood distance											-0.04
											(-0.49)
Observations	126	96	96	142	147	129	68	68	51	140	141
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	0.58	0.57	0.54	0.55	0.55	0.54	0.47	0.49	0.55	0.55	0.52
Inverted U-shape test	3.98	2.03	1.96	3.14	2.94	2.88	1.37	1.37	1.52	3.14	2.76
	[p=0.00]	[p=0.02]	[0.02]	[p=0.00]	[p=0.00]	[p=0.00]	[p=0.08]	[p=0.08]	[p=0.06]	[p=0.00]	[p=0.00]

Notes: This table presents standardized coefficients for the effect of potential crop yield (measured in millions of kilocalories per hectare per year) on various measures of democracy using OLS regression. Baseline controls used are absolute latitude, terrain ruggedness, elevation (average), elevation (variation), landlockedness, distance to waterways. Neo–European countries are USA, Canada, Australia, and New Zealand. All specifications use an intercept term but not reported for brevity. Robust t-statistics are given in parentheses. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively. The inverted U-shape test is performed using the approach of Lind and Mehlum (2010).

Table A8: Geographic Region Exclusion Analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	Polity2 (1961-2015)								
Crop yield	1.04***	1.27***	0.83***	1.08***	0.95***	1.00***	1.05***	0.97***	1.52***
	(4.05)	(3.74)	(3.19)	(3.44)	(3.64)	(3.86)	(4.03)	(3.53)	(3.51)
Crop yield square	-0.91***	-1.04***	-0.91***	-0.92***	-0.80***	-0.84***	-0.91***	-0.80***	-1.31***
	(-3.46)	(-3.00)	(-3.39)	(-2.83)	(-3.03)	(-3.22)	(-3.45)	(-2.90)	(-2.87)
Observations	147	116	105	100	134	136	144	123	57
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	0.537	0.342	0.668	0.496	0.565	0.537	0.525	0.537	0.533
Omitted continent	None	Europe	Asia	Africa	N. America	S. America	Oceania	None	None
Sample	Global	Global	Global	Global	Global	Global	Global	Old world	Industrial
Inverted U-shape test	2.83 [p=0.00]	2.28 [p=0.01]	3.19 [p=0.00]	2.25 [p=0.01]	2.40 [p=0.01]	2.55 [p=0.01]	2.82 [p=0.00]	2.25 [p=0.01]	2.28 [p=0.01]

Notes: This table presents standardized coefficients for the effect of potential crop yield (measured in millions of kilocalories per hectare per year) on various measures of democracy using OLS regression. Baseline controls used are absolute latitude, terrain ruggedness, elevation (average), elevation (variation), landlockedness, distance to waterways. All specifications use an intercept term but not reported for brevity. Robust t-statistics are given in parentheses. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively. The inverted U-shape test is performed using the approach of Lind and Mehlum (2010).

Table A9: Controlling for Crop-based Inequality Measures, Precipitation and Temperature

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	Polity2 (1961-2015)						
Crop yield	1.08***	1.13***	1.08***	0.80***	0.72**	0.61**	0.90***
	(4.31)	(4.43)	(4.32)	(2.92)	(2.52)	(2.23)	(3.75)
Crop yield square	-0.95***	-0.99***	-0.95***	-0.75***	-0.75***	-0.60**	-0.82***
	(-3.69)	(-3.83)	(-3.72)	(-2.84)	(-2.81)	(-2.33)	(-3.32)
Wheat/sugar	0.12						
	(1.26)						
(Wheat+maize) / (sugar+rice)		0.16**					
		(2.10)					
(Wheat+maize) / (sugar+cotton)			0.13				
			(1.40)				
Precipitation				0.20*	0.54*		
				(1.94)	(1.69)		
Precipitation square					-0.32		
					(-1.11)		
Precipitation spline 1						0.30*	
						(1.92)	
Precipitation spline 2						-0.07	
						(-0.43)	
Temperature							-0.37
							(-1.65)
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	142	142	142	147	147	139	147
Adj. R-squared	0.54	0.55	0.54	0.55	0.55	0.55	0.55
Inverted U-shape test	3.25	3.40	3.01	2.60	2.29	1.83	2.58
	[p=0.00]	[p=0.00]	[p=0.00]	[p=0.00]	[p=0.01]	[p=0.03]	[p=0.00]

Notes: This table measures the standardized coefficients for the effect of potential crop yield (measured in millions of kilocalories per hectare per year) on Polity2 during 1961-2015. The continent dummies are Africa, Asia, Australia, Europe, North America, Oceania, and South America. Baseline controls used are absolute latitude, terrain ruggedness, elevation (average), elevation (variation), landlockedness, distance to waterways. All specifications use an intercept term but it is not reported for brevity. Robust t-statistics are given in parentheses. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively. The inverted U-shape test is performed using the approach of Lind and Mehlum (2010).

5. Data Description and Sources

A. Outcome variable

Democ: Institutional democracy is an additive eleven-point scale (0-10). The operational indicator of democracy is derived from the coding of the competitiveness of political participation, the regulation of participation, the openness and competitiveness of executive recruitment, and constraints on the chief executive (source: Marshall et al., 2015).

Autoc: Institutional autocracy is an additive eleven-point scale (0-10). The operational indicator of autocracy is derived from the coding of the competitiveness of political participation, the regulation of participation, the openness and competitiveness of executive recruitment, and constraints on the chief executive (source: Marshall et al., 2015).

Polity2: A composite index from the Polity IV project that is computed by subtracting the *Autoc* score from the *Democ* score. The variable has possible values of -10 to +10, where 10 indicates the highest degree of democracy (source: Marshall et al., 2015).

DemocAve: the average *Democ* value for years 1961-2015.

AutocAve: the average *Autoc* value for years 1961-2015.

Voter turnout: This is defined as the number of votes as a percent of total eligible voters registered in an area. The data are obtained from the following sources:
http://www.electproject.org/home/voter-turnout/voter-turnout-data and
http://eci.nic.in/eci_main1/ElectionStatistics.aspx, for the US and India, respectively.

Democracy (*civil liberties and political rights*): An index developed by Freedom House to indicate the quality of political rights and civil liberties in various countries (source: Freedom House, 2016).

B. Crop yield (main explanatory variable)

The following description is based largely on Fischer et al. (2002). The crop yield index captures maximum potential crop yield (measured in millions of kilocalories per hectare per year) across different regions in a country. Crop yield index is constructed using crop yield (measured in tons per hectare per year) from the Global Agro-Ecological Zones (GAEZ) project of the Food and Agriculture Organization (FAO) and caloric content of various crops from the US Department of Agriculture Nutrient Database for Standard Reference (source: Galor and Özak, 2016).

In their natural state, many soils, in particular in the tropics, cannot be continuously cultivated without undergoing degradation. Such degradation is marked by a decrease in crop yields and a deterioration of soil structure; nutrient status; and other physical, chemical, and biological attributes. Under traditional low input farming systems, this deterioration is kept in check by alternating some years of cultivation with periods of fallow. Regeneration of nutrients and maintenance of soil fertility under low input cultivation is achieved through natural bush or grass fallow. At somewhat higher inputs to soils, soil fertility is maintained through fallow that may include for a portion of time grass, grass-legume ley, or a green-manure crop. The high level assumes the best farming technology, soil nutrient inputs, and management known today.

With balanced fertilizer applications and proper pest and disease management (which is most likely at high level of inputs), only limited fallow will be required to maintain soil fertility and to keep pest and disease outbreaks in check. At low level of inputs (assuming virtually no application of chemicals and only limited organic fertilizer, and very limited or no application of biocides), there is a need for considerable fallow periods in the crop rotations to restore soil nutrient status and to break pest and disease cycles. The expected long-term yields as estimated by the GAEZ procedures assume the proper crop/fallow program.

As a rule of thumb, for low level input/management conditions, fallow period requirements may vary between 30–90% of the cultivation fallow cycle. At high levels of inputs and management, fallow requirements are uniformly set at 10%. At intermediate level of inputs, the fallow requirements are set at one third of the levels required under low level of inputs. On average, long-term achievable yields are 10%, 20%, and 55% lower than maximum attainable yields at high, intermediate, and low levels of inputs, respectively.

The 48 crops are alfalfa, banana, barley, buckwheat, cabbage, cacao, carrot, cassava, chickpea, citrus, coconut, coffee, cotton, cowpea, dry pea, flax, foxtail millet, green gram, groundnuts, indigo rice, maize, oat, palm oil, olive, onion, palm heart, pearl millet, phaseolus bean, pigeon pea, rye, sorghum, soybean, sunflower, sweet potato, tea, tomato, wetland rice, wheat, spring wheat, winter wheat, white potato, yams, giant yams, subtropical sorghum, tropical highland sorghum, tropical lowland, sorghum, and white yams.

C. Geographic controls

Distance to waterways: The distance, in thousands of km, from a GIS grid cell to the nearest ice-free coastline or sea-navigable river, averaged across the grid cells of a country (source: Gallup et al., 2010).

Elevation: The average elevation of a country, in thousands of kilometers above sea level, calculated using geospatial data at a 1-degree resolution (source: Geographically based Economic data (G-ECON) project (2008).

Latitude: The absolute value of latitude of a country's centroid (source: http://gothos.info/resources).

Precipitation: Average precipitation in 1000 millimeters between 1961 and 1990 (source: Geographically based Economic data (G-ECON) project (2008).

Ruggedness: The calculation for ruggedness takes a point on the earth's surface and measures the difference in elevation between this point and each of the eight major directions of the compass (north, northeast, east, southeast, south, southwest, west, and northwest). The index at the central point is given by the square root of the sum of the squared differences in elevation between the central point and the eight adjacent points (source: Nunn and Puga, 2012).

Temperature: Average temperature in Celsius between 1961 and 1990 (source: Geographically based Economic data (G-ECON) project (2008).

D. Other control variables

Age dependency: The ratio is calculated by dividing the sum of young-aged populations (0–14) and old-aged populations (65+) by the populations aged between 15 and 64 in 1980 (source: National Center for Health Statistics, 1984; Minnesota population Center, 2016).

Birthplace diversity: This index measures the probability that two individuals drawn randomly from the entire population have two different countries of birth (source: Alesina et al., 2016).

Colonial history: A classification of a country's colonial origin into British, French, Portuguese, Spanish, or other European origins (i.e., Dutch, Belgian, and Italian) since 1700. For countries ruled by several colonial powers, the most recent one is taken provided that it was ruled for ten years or longer (source: Nunn and Puga, 2012).

Duration of colonization: This variable indicates the number of years a country was under colonial rule of European imperial powers (source: Olsson, 2009).

Emigration rate: This variable is created by dividing the total number of emigrants by the average population of the source country (source: IAB brain-drain dataset, Brücker et al., 2013).

Ethnic fractionalization: This is defined as the probability that two randomly selected individuals in a country will be from different ethnic groups (source: Fearon, 2003).

Genetic diversity: An index that incorporates the expected heterozygosities of the pre-colonial ancestral populations of contemporary subnational groups as well as the pairwise genetic distances between these ancestral populations (source: Ashraf and Galor, 2013).

European population (%): The fraction of a country's population having European ancestors during colonization (source: Easterly and Levine, 2016).

European language (%): The fraction of a country's population speaking one of the five primary Western European languages (i.e., English, French, German, Portuguese, and Spanish) as their first language (source: Hall and Jones, 1999).

Family ties: Data from the following three responses are combined to obtain an overall measure of strength of family ties for each state in the U.S.: (1) the percentage of people living alone (reversed); (2) the percentage of elderly people (65+) living alone (reversed); and (3) and the percentage of households with grandchildren (source: Ang and Fredriksson, 2017).

Foreign aid (% of GDP): This variable is constructed using data on 'net official development assistance and official aid received' and GDP for recipient countries (source: WDI, 2015).

GDP per capita: GDP per capita data at 2005 US dollar (source: Feenstra et al., 2015).

Irrigation (1900): This variable is created by dividing the area irrigated in 1900 AD by land area of the country. (sources: irrigated area and land area retrieved from Siebert et al., 2015, and WDI, 2015, respectively).

Language fractionalization: This index is constructed as: $F = \sum_{i=1}^{m} n_i(1 - n_i)$, where n_i is the population share of group *i* and *m* is the number of groups (source: Esteban et al., 2012).

Legal origins: A classification of a country's legal tradition into British common law, French civil law, German civil law, Scandinavian law or Socialist law (source: La Porta et al., 1998).

Oil: A dummy variable indicating whether the proportion of export revenues from oil is larger than one third (source: Fearon, 2003).

Pathogen stress: A standardized index of historical pathogen prevalence for nine diseases, including leishmanias, schistosomes, trypanosomes, leprosy, malaria, typhus, filariae, dengue, and tuberculosis (source: Murray and Schaller, 2010).

Per capita income: The average household income over the years 1984–1989. It is measured in 2015 CPI-U-RS adjusted U.S. dollars (source: U.S. Census Bureau, 2017).

Population density (1500): This variable is constructed by dividing total population in 1500 AD by land area (sources: population in 1500 AD and land area retrieved from McEvedy and Jones, 1978, and WDI, 2015, respectively).

Religion (% of population): This variable indicates the fraction of population following major religions in each country in the year 2000 (source: ARDA, 2017).

Religious adherence: This variable indicates the fraction of people who attend their respective place of worship (source: ARDA, 2017).

Religious fractionalization: This variable is computed as one minus Herfindahl index of religious group shares. It reflects the probability that two randomly selected individuals from the population belong to different religion groups (source: Alesina et al., 2003).

School enrollment: The estimated public elementary and secondary enrollments as a percentage of the population aged between 5 and 17 in 1980 (source: U.S. Department of Education, 1998).

State Antiquity (1500): An index of state history covering the period between 1 AD and 1500 AD (source: Putterman, 2012).

Years since transition to agriculture: The number of years before 2000 AD when the people of a country, today's territories, started getting more than half of their calories from cultivated foods and domesticated animals (source: Putterman, 2006).

Years of schooling: Average number of years spent at school by people 15 years or older in 1960 (source: Barro and Lee, 2013).

Newspaper per capita: Total average circulation of daily newspapers per 1000 population in 1996 (source: UIS, 2018).

Human capital stock of Lee and Lee (2016): This variable measures human capital stock per person for population aged 15 to 64 in 1960. It is the sum of the share of workers (fraction of population group *a* with education level *j*), weighted by relative wage rates, across all education categories. Observations for the total population aged 15–64

are divided into 10 five-year age groups ranging from 15–19 years old (a=1) to 60–64 years old (a =10). Educational attainment is classified into seven categories, namely no formal education (j=1), incomplete primary (j=2), complete primary (j=3), lower secondary (j=4), upper secondary (j=5), incomplete tertiary (j=6), and complete tertiary (j=7). Unlike the widely used "years of schooling" measure, which assumes uneducated workers to have no human capital contribution and human capital increases proportionally with the duration of schooling, this measure assumes a Mincerian log-linear relationship between them and the contribution by uneducated labor force equals to one. This measure also assumes perfect substitutability across different skill types (source: Lee and Lee, 2016).

Alternative human capital stock of Lee and Lee (2016): This variable provides an alternative measure for human capital per worker for population aged 15 to 64 in 1960. Unlike the above which assumes perfect substitutability across different skill types, the substitutability is assumed to be imperfect here. The aggregate labor input is measured as a constant-elasticity-of-substitution aggregate of two types of labor - unskilled and skilled labor. The first four levels (j=1 to *4*) of educational attainment are classified as unskilled and the remaining ones (j=5 to 7) are classified as skilled labor force (source: Lee and Lee, 2016).

Importance of secure surroundings: A proxy of preference for risky behavior. This risk avoidance index is constructed using the following question "living in secure surroundings is important to this person (A191)" from the World Values Survey, waves 5 and 6 (source: WVS Database, 2015).

Importance of taking risk: A proxy of preference for risky behavior. This risk avoidance index is constructed using the following question "adventure and taking risks are important to this person (A195)" from the World Values Survey, waves 5 and 6 (source: WVS Database, 2015).

Preference for security: A proxy of preference for risky behavior. This risk avoidance index is constructed using the following question on "things done for reasons of security (H003), including "didn't carry much money",

"preferred not to go out at night", and "carried a knife, gun, or other weapon" from the World Values Survey, wave 6 (source: WVS Database, 2015).

Precipitation spline measures: A spline is a piecewise polynomial function where a continuous variable can be divided into separate segments using knots. Separate regression lines are then fitted for each segment to form a smooth and continuous fitted curve. A restricted cubic spline is a spline of the third order polynomial that is restricted to being linear at its end knots. The advantage of this spline is that it involves the non-linear transformation of the explanatory variable prior to running regressions, and hence is not influenced by the dependent variable. Following the approach of Haber and Menaldo (2011b), we construct a restricted cubic spline for the logs of precipitation. It consists of three knots which lie in the 0.10^{th}, 0.50^{th} and 0.90^{th} quantiles (108.28 mm, 819.81 mm and 2600.35 mm, respectively). Precipitation is linear before the first and after the last knot, but forms a cubic piecewise polynomial form between these knots.

REFERENCES (Online Appendix)

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S. & Wacziarg, R. 2003. "Fractionalization." *Journal of Economic Growth* 8: 155-194.

Alesina, A., Harnoss, J. & Rapoport, H. 2016. "Birthplace Diversity and Economic Prosperity." *Journal of Economic Growth* 21: 101-138.

ARDA 2017. "Association of Religion Data Archives." Online Data retrieved from www.thearda.com.

Ashraf, Q. & Galor, O. 2013. "The "Out of Africa" Hypothesis, Human Genetic Diversity, and Comparative Economic Development." *American Economic Review* 103: 1-46.

Barro, R.J. & Lee, J.W. 2013. "A New Data Set of Educational Attainment in the World, 1950–2010." *Journal of Development Economics* 104: 184-198.

Brücker, H., Capuano, S. & Marfouk, A. 2013. "Education, Gender and International Migration: Insights from a Panel-dataset 1980-2010." *mimeo*.

Easterly, W. & Levine, R. 2016. "The European Origins of Economic Development." *Journal of Economic Growth* 21: 225-257.

Esteban, J., Mayoral, L. & Ray, D. 2012. "Ethnicity and Conflict: An Empirical Study." *American Economic Review* 102: 1310-1342.

Fearon, J.D. 2003. "Ethnic and Cultural Diversity by Country." *Journal of Economic Growth* 8: 195-222.

Feenstra, R.C., Inklaar, R. & Timmer, M.P. 2015. "The Next Generation of the Penn World Table." *American Economic Review* 105: 3150-3182.

Fischer, G., Van Velthuizen, H.T., Shah, M.M. & Nachtergaele, F.O. 2002. *Global Agro-ecological Assessment for Agriculture in the 21st Century: Methodology and Results.* Laxenburg, Austria: International Institute for Applied Systems Analysis.

Freedom House 2016. "Freedom in the World Comparative and Historical Data." Ney York, NY and Washington, DC. Available at https://freedomhouse.org (accessed April 30, 2016).

G-Econ 2008. "Geographicallybased Economic data." Yale University. Available at http://gecon.yale.edu.

Galor, O. & Özak, Ö. 2016. "The Agricultural Origins of Time Preference." *American Economic Review* 106: 3064-3103.

Haber, S. & Menaldo, V. 2011. "Rainfall, Human Capital, and Democracy." *SSNR*: http://dx.doi.org/10.2139/ssrn.1667332.

Hall, R.E. & Jones, C.I. 1999. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics* 114: 83-116.

La Porta, R., Lopez-De-Silanes, F., Shleifer, A. & Vishny, R.W. 1998. "Law and Finance." *Journal of Political Economy* 106: 1113-1155.

Lee, J.W. & Lee, H. 2016. "Human Capital in the Long Run." *Journal of Development Economics* 122: 147-169.

Lind, J.T. & Mehlum, H. 2010. "With or Without U? The Appropriate Test for a U shaped Relationship." *Oxford Bulletin of Economics and Statistics* 72: 109-118.

Marshall, M., Gurr, T. & Jaggers, K. 2015. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2015 Dataset."

Mcevedy, C. & Jones, R. 1978. *Atlas of World Population History.* New York: Facts on File.

Murray, D.R. & Schaller, M. 2010. "Historical Prevalence of Infectious Diseases within 230 Geopolitical Regions: A Tool for Investigating Origins of Culture." *Journal of Cross-Cultural Psychology* 41: 99-108.

Nunn, N. 2008. "Slavery, Inequality, and Economic Development in the Americas: An Examination of the Engerman-Sokoloff Hypothesis" In *Institutions and Economic Performance,* ed. E. Helpman, 48-180. Cambridge: Harvard University Press.

Nunn, N. & Puga, D. 2012. "Ruggedness: The Blessing of Bad Geography in Africa." *Review of Economics and Statistics* 94: 20-36.

Olsson, O. 2009. "On the Democratic Legacy of Colonialism." *Journal of Comparative Economics* 37: 534-551.

Putterman, L. 2006. "Agricultural Transition Year Country Data Set." *Brown University*.

_____ 2012. "State Antiquity Index (Statehist) Version 3.1 (http://www.econ.brown.edu/fac/louis_putterman/)."

Siebert, S., Kummu, M., Porkka, M., Döll, P., Ramankutty, N. & Scanlon, B.R. 2015. "A Global Data Set of the Extent of Irrigated Land from 1900 to 2005." *Hydrology and Earth System Sciences* 19: 1521-1545.

UIS 2018. "UNESCO Institute for Statistics (UIS) Database (http://data.un.org/Data.aspx?q=newspaper&d=UNESCO&f=series%3aC_N_DNTPI), Retrived on September 27, 2018, UIS."

WDI. 2015. *World Development Indicators 2015 database.* Philadelphia: World Bank.

WVS Database. 2015. *World Value Survey 1981-2014 Longitudinal Aggregate v.20150418, 2015. World Values Survey Association (www.worldvaluessurvey.org). Madrid, Spain.*