

Supplemental Appendix

Ethnic Diversity, Historical Economic Exchange, and Development: Evidence from Andean Peru

Miriam Artiles

PUC-Chile

A Supplemental Appendix - Figures and Tables

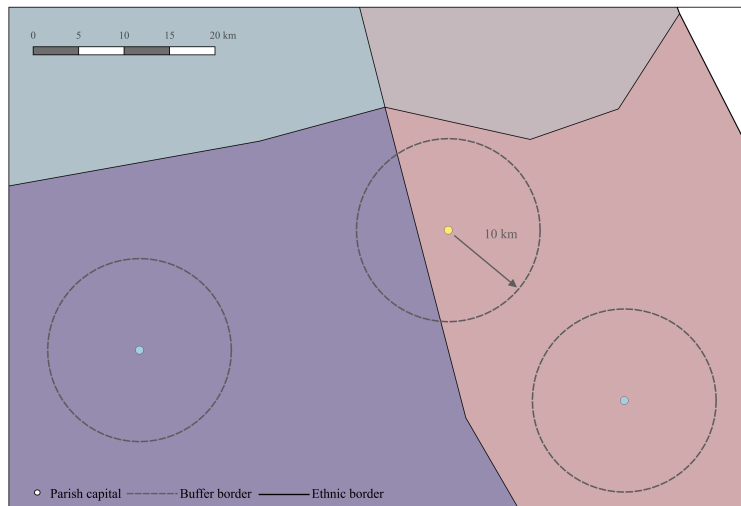


FIGURE A.1: Buffer Exercise

Notes. Construction of 10-km buffer around each parish capital. The map is displayed using a World Geodetic System projection (WGS 1984).

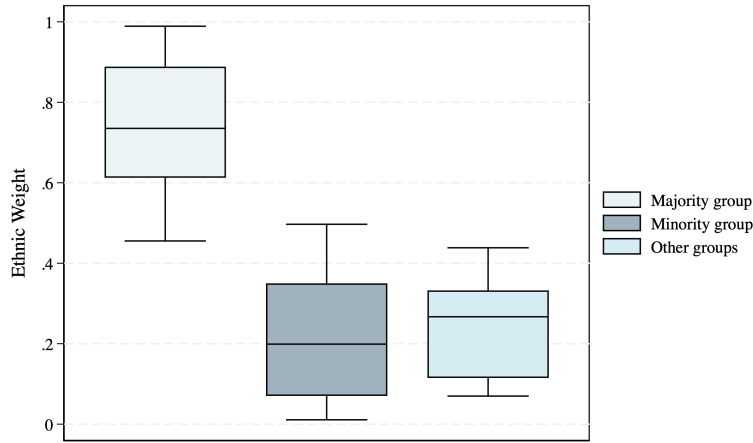


FIGURE A.2: Boxplots of Ethnic Weights

Notes. The ethnic weight (w_{ep}) is defined as the area share of ethnic group e within a 10-km buffer from the capital of parish p (using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers). The boxplots refer to the ethnic weights of the majority group (i.e., the group with the highest area share within the 10-km buffer), the minority group (i.e., the group the lowest area share), and other groups, considering parishes with ethnic diversity only. The category “Other groups” applies only for parishes that concentrated either three or four groups (15 percent of the parishes with ethnic diversity).

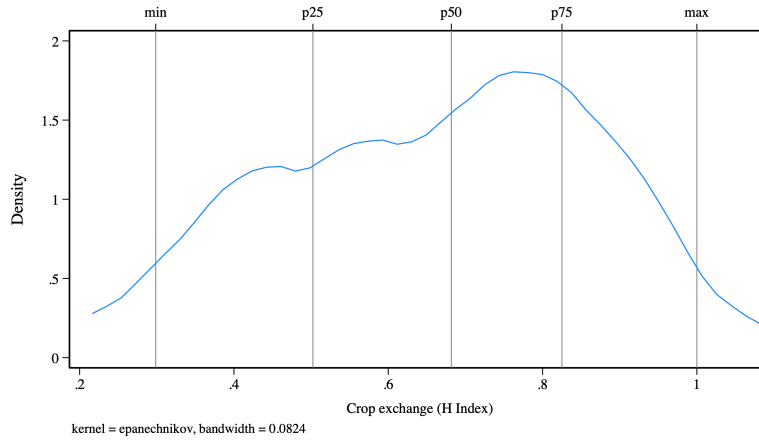


FIGURE A.3: H Index Across Ethnic Groups (Kernel Density)

Notes. Kernel density of the ethnic-level proxy for potential exposure to internal crop exchange (H index), defined as the reciprocal of the Herfindahl index: $H_e = 1/\sum_j s_{je}^2$, where s_{je} is the area share of zone j within the homeland of ethnic group e . The index is normalized to 1 for the group with the highest value.

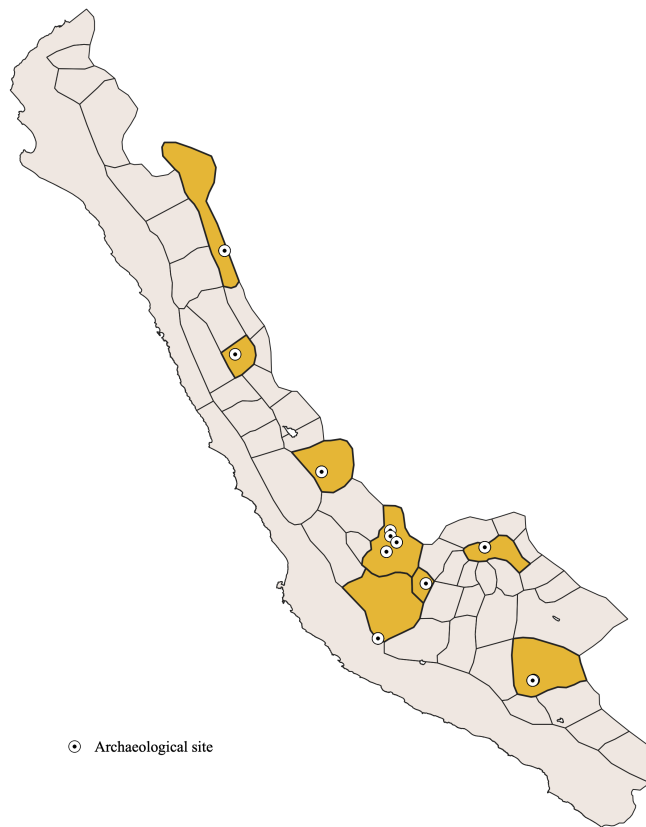


FIGURE A.4: Archaeological Sites with Information on Pre-Colonial Diets

Notes. Archaeological sites with information on pre-colonial individual diets. Geographic coordinates are from Wilson et al. (2022). Black lines represent the territorial extent of pre-colonial groups at the time of the Spanish conquest, based on Rowe (1946)'s map.

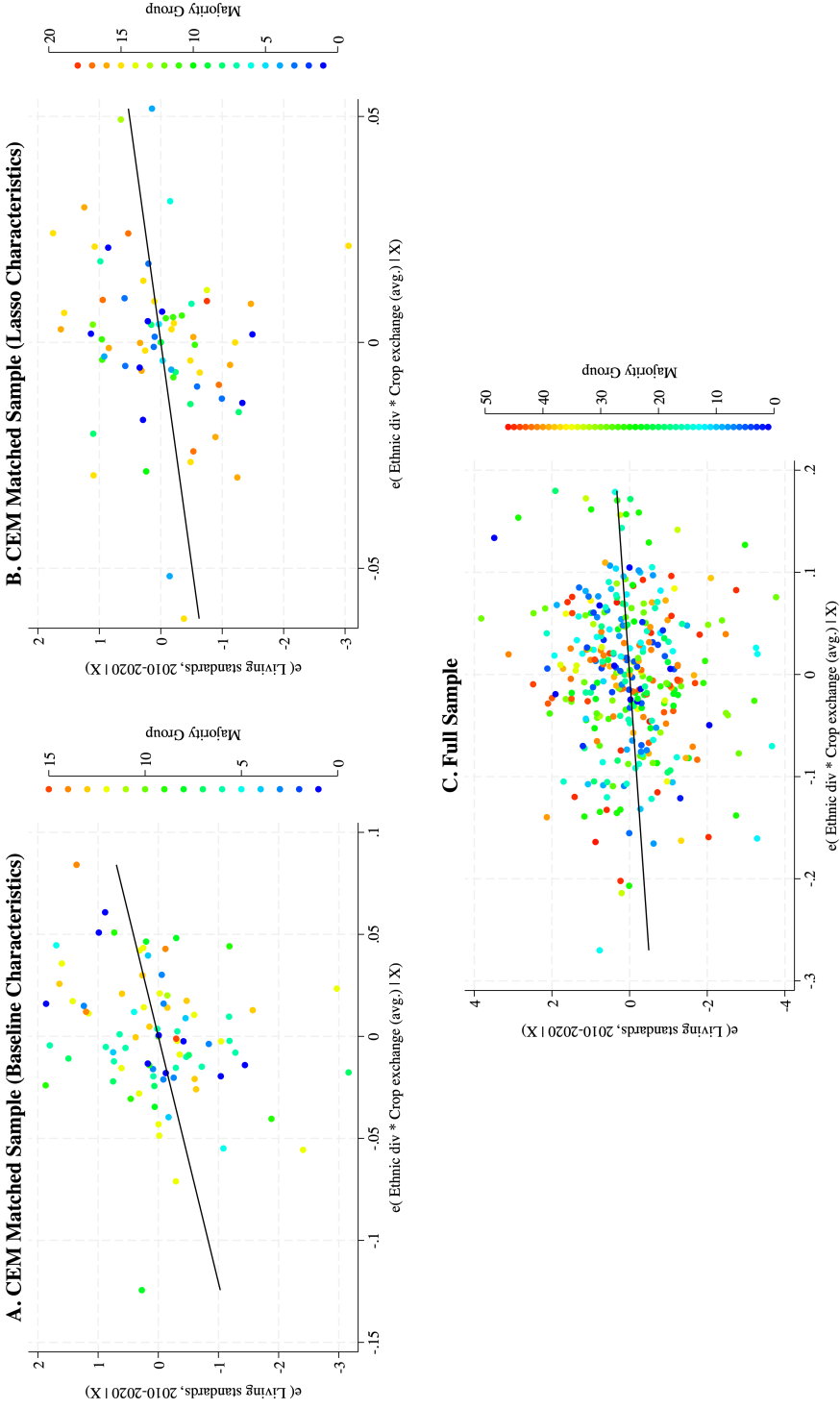


FIGURE A.5: Partial Correlation Plots

Notes. Partial correlation scatterplots. The unit of observation is the parish. Each color indicates a different majority group. The dependent variable is the score of the first principal component derived from the following variables: the average number of small- and medium-sized firms per 100 inhabitants (2010-2013, graph B), a dummy variable for whether the share of farmers practicing non-subsistence agriculture is above the sample median (2012, graph C), the log of 1 plus average light intensity per capita (2013, graph D), the share of dwellings with access to public sanitation (2017, graph E), and the share of dwellings with access to the public water network (2017, graph F). Ethnic diversity is a dummy variable taking value 1 if there is an ethnic border within a buffer of 10-km radius from the parish capital, using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers. Crop exchange (avg.) is the parish-level proxy for average exposure to internal crop exchange, computed using the area shares of ethnic groups within the 10-km parish buffer as weights. All regressions include baseline controls (parish-level mean and standard deviation of elevation, mean and standard deviation of land caloric suitability, longitude, latitude, log distance to perennial rivers, log distance to pre-colonial native shrines, log 16th-century expected tribute, and log distance to *mita* mines), weight of the majority group, and colonial province fixed effects, with robust standard errors clustered at the province level. The matched sample refers to the sample of parishes selected by the CEM algorithm of Iacus et al. (2012), using parishes with average exposure to internal crop exchange above the median as treated parishes (see Table 6).

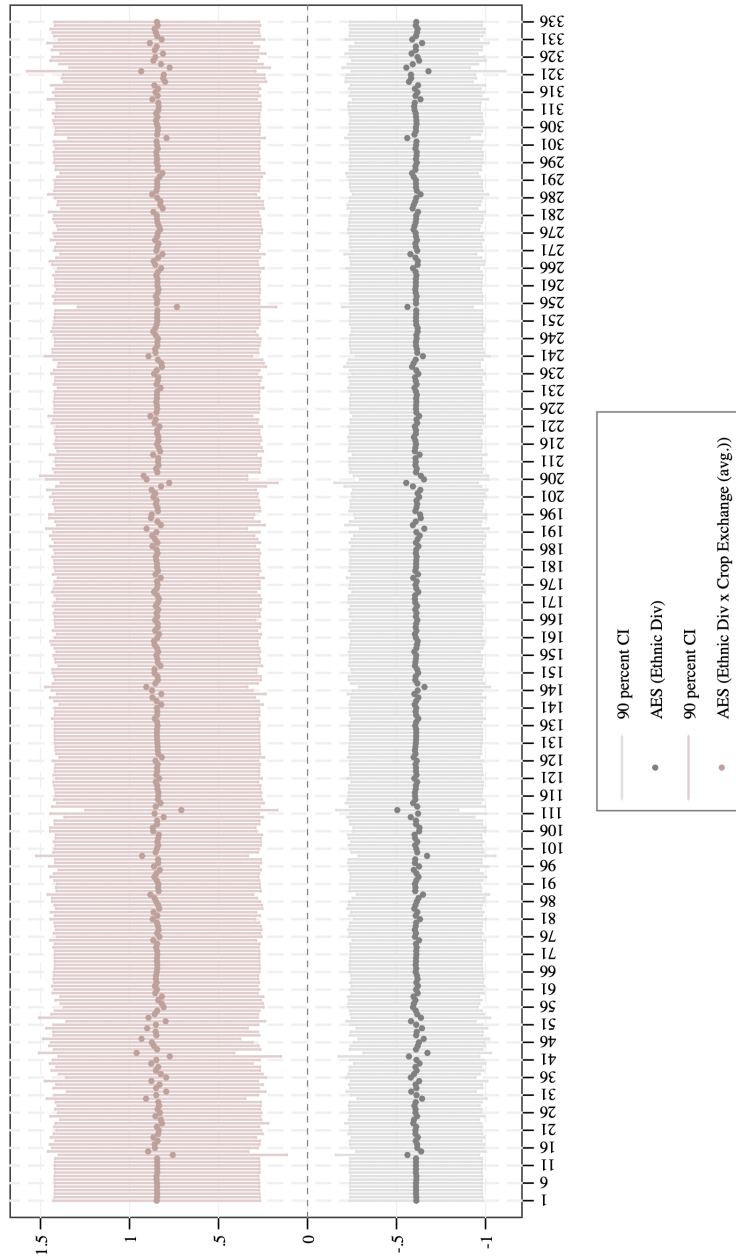


FIGURE A.6: Influential Observations

Notes. Standardized AES (Kling et al. 2004; Clingingsmith, Khwaja, and Kremer 2009) and 90 percent confidence intervals after baseline controls and colonial province fixed effects. The unit of observation is the parish. Each regression excludes one parish (indicated on the x-axis) at a time. Robust standard errors are clustered at the level of the colonial province. The standardized AES refers to the following outcomes: a dummy variable indicating whether the share of farmers practicing non-subsistence agriculture is above the sample median (2012), the log of 1 plus average light intensity per capita (2013, VIIRS), the average number of small- and medium-sized firms per 100 inhabitants (2010-2013), the share of dwellings with access to public sanitation (2017), and the share of dwellings with access to the public water network (2017). Ethnic diversity is a dummy variable taking value 1 if there is an ethnic border within a buffer of 10-km radius from the parish capital, using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers. Crop exchange (avg.) is the parish-level proxy for average exposure to internal crop exchange, computed using the area shares of ethnic groups within the 10-km parish buffer as weights. The vector of baseline controls includes parish-level mean and standard deviation of elevation, mean and standard deviation of land caloric suitability, longitude, latitude, log distance to perennial rivers, log distance to pre-colonial native shrines, log 16th-century expected tribute, and log distance to *mita* mines.

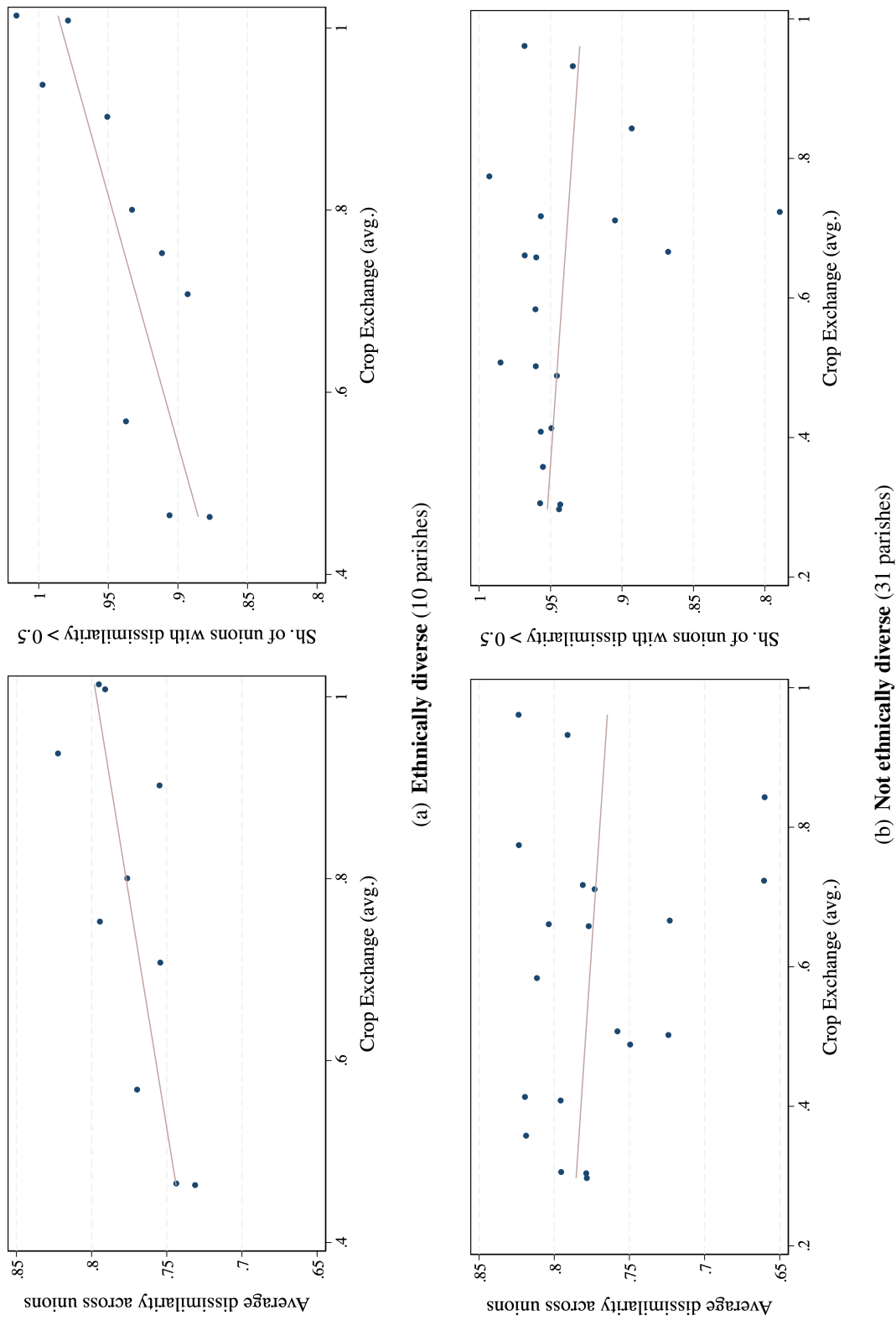


FIGURE A.7: Historical Crop Exchange and Inter-Group Unions (1605–1870)

Notes. Binned scatterplots. On the left-hand side, the y-axis variable is average dissimilarity, defined as the mean normalized Levenshtein distance across all unions. On the right-hand side, the y-axis variable refers to the share of unions with dissimilarity above 0.5. All graphs control for the log number of individuals found in the parish records, mitigating concerns about parish size influencing the results. Ethnically diverse parishes are those with an ethnic border within a buffer of 10-km radius from the parish capital, using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers. Crop exchange (avg.) is the parish-level proxy for average exposure to internal crop exchange, computed using the area shares of ethnic groups within the 10-km parish buffer as weights.

TABLE A.1: Balance Tests—Additional Pre-Colonial Characteristics

	Border Parishes		Interior Parishes		Diff.	p-value ^a	p-value ^b
	mean	sd	mean	sd			
(1) Ln dist. to defensive site	4.187	0.704	4.162	0.926	-0.025	[0.783]	[0.856]
(2) Ln dist. to urban site	4.211	0.742	4.129	0.984	-0.082	[0.391]	[0.567]
(3) Ln dist. to political site	4.087	0.671	3.931	0.984	-0.155	[0.089]*	[0.232]
(4) Ln dist. to elite residence	4.067	1.046	3.951	1.289	-0.116	[0.374]	[0.225]
(5) Ln dist. to road	1.409	2.597	0.983	2.458	-0.426	[0.146]	[0.188]
(6) Ln dist. to canal	3.901	0.810	3.913	0.908	0.012	[0.901]	[0.909]
(7) Ln dist. to bridge	4.249	0.760	4.249	0.783	0.001	[0.993]	[0.994]
(8) Caloric suitability for maize	357.869	1010.757	304.422	899.243	-53.447	[0.632]	[0.691]
(9) Caloric suitability for potato	588.712	750.783	626.101	801.952	37.389	[0.671]	[0.808]
Number of parishes	117	117	219	219	336	336	336

Notes. The unit of observation is the parish. Border parishes are parishes with an ethnic border within a buffer of 10-km radius from the parish capital (using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers), while interior parishes are parishes located further inside ethnic homelands. P-values from OLS regressions of each of the variables listed in the first column on a dummy variable indicating border parish; (^a) with robust standard errors, (^b) with standard errors corrected for spatial dependence using a distance cutoff of approximately one degree at the equator (Colella et al. 2019).

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A Complementary Elevation Zones and Crop Variety

This section presents correlational evidence that the measure of heterogeneity in complementary elevation zones (H index) helps explain crop variety in the data. For this exercise, one would ideally use land suitability data for all crops available before 1500. In the absence of these data, I use information on native crops from modern sources. I rely on the 2012 agricultural census, which provides geo-referenced data for an extensive set of native crops. I explore the determinants of crop variety across grid cells of different sizes, as well as across ethnic groups, based on the number of native crops reported by farmers at the time of the census.⁵¹

I start by computing the H index for $25 \text{ km} \times 25 \text{ km}$ grid cells covering the entire study region (i.e., located more than 500 meters above sea level). Specifically, I compute the grid-level H index as $H_g = 1/\sum_j s_{jg}^2$, where s_{jg} is the area share of elevation zone j within grid cell g . In Column 1 of Table A.2, I regress the log of the number of crops on this measure (H_g), mean elevation, log land area, latitude, and longitude. The estimated beta coefficient is positive (0.504) and statistically significant. The coefficient remains stable in magnitude and statistical significance when I control for another potential predictor of crop variety—variation in elevation (Column 3). According to OLS estimates, a one standard deviation increase in H_g is associated with a 0.504 standard deviation increase in log crop variety (Column 1, statistically significant at the 1 percent level). In the case of variation in elevation, the associated standard deviation increase in log crop variety is 0.418 (Column 2, statistically significant at the 1 percent level). The coefficient on variation in elevation becomes small (0.062) and statistically insignificant after controlling for H_g (Column 3). The same pattern arises when including fixed effects for the 87 hydrographic basins in the study region (Column 4). In Column 5, I substitute the grid-level H index with dummy variables indicating the number of elevation zones within the grid cell. The magnitudes of the estimated coefficients increase with the number of zones (relative to the coefficient for grid cells with one elevation zone only). Columns 6 and 7 report similar results across $50 \text{ km} \times 50 \text{ km}$ grid cells and ethnic groups (H_e), respectively.

⁵¹I follow the classification of native crops in Tapia (2013). Farmers were only required to report the list of crops harvested at the time of the census—crop rotation practices may have affected the list of reported crops.

TABLE A.2: Validating Elevation Zones

	Dependent Variable: Ln (# Native Crops)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
H index	0.504 [0.036]*** (0.045)***		0.461 [0.044]*** (0.058)***	0.448 [0.057]*** (0.057)***		0.445 [0.075]*** (0.090)***	0.527 [0.106]*** (0.086)***
SD of elevation		0.418 [0.040]*** (0.063)***	0.062 [0.052] (0.081)	0.074 [0.057] (0.062)	0.021 [0.060] (0.059)	0.060 [0.120] (0.124)	0.135 [0.141] (0.157)
Ln land area	0.186 [0.033]*** (0.027)***	0.183 [0.034]*** (0.036)***	0.181 [0.033]*** (0.029)***	0.133 [0.024]*** (0.022)***	0.136 [0.033]*** (0.031)***	0.399 [0.070]*** (0.062)***	0.387 [0.126]*** (0.124)***
Mean elevation	-0.052 [0.044] (0.073)	0.146 [0.044]*** (0.092)	-0.030 [0.052] (0.085)	-0.100 [0.063] (0.068)	-0.304 [0.062]*** (0.069)***	-0.002 [0.106] (0.117)	0.134 [0.130] (0.106)
Dummy (number of zones=2)					0.961 [0.130]*** (0.162)***		
Dummy (number of zones=3)					1.691 [0.150]*** (0.190)***		
Dummy (number of zones=4)					1.848 [0.182]*** (0.216)***		
Hydrographic basin FE	No	No	No	Yes	Yes	No	No
Observations	665	665	665	665	665	200	47

Notes. The unit of observation is the 25 km $\times$ 25 km grid cell in Columns 1-5, the 50 km $\times$ 50 km grid cell in Column 6, and the ethnic group in Column 7. The table reports OLS estimates. Robust standard errors in brackets; clustered at the basin level in Columns 4-5. In parentheses, standard errors are corrected for spatial dependence with a distance cutoff of approximately one degree at the equator (Colella et al. 2019). All variables except dummies are standardized to have zero mean and standard deviation equal to one. All regressions control for longitude and latitude.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.3: Historical Crop Exchange — Pre-Colonial Diets

	Dependent Variable: Carbon Isotope Score ($\delta^{13}C_{col}$)					
	(1)	(2)	(3)	(4)	(5)	(6)
H index (ethnic group level)	0.501 [0.153]**	0.279 [0.089]**	0.851 [0.145]***	0.261 [0.071]***	0.269 [0.085]**	
	(0.145)***	(0.094)***	(0.107)***	(0.069)***	(0.081)***	
H index (grid cell level)						-0.216 [0.756] (0.420)
Number of individuals	196	196	196	196	182	196
Zone FE	No	Yes	Yes	Yes	Yes	Yes
Ln (# native crops)	No	No	Yes	No	No	No
Land caloric suitability	No	No	No	Yes	Yes	Yes
Excluding children	No	No	No	No	Yes	No
Number of ethnic groups	8	8	8	8	8	—
Number of grid cells (50×50 km)	—	—	—	—	—	8

Notes. The unit of observation is the individual. The table reports OLS estimates. Robust standard errors clustered at the ethnic group (columns 1-5) or grid cell (column 6) level in brackets. In parentheses, standard errors are corrected for spatial dependence with a distance cutoff of approximately one degree at the equator (Colella et al. 2019). In column 3, *Ln (# native crops)* refers to the log number of different native crops in the ethnic homeland (see Appendix A). *Land caloric suitability* refers to the average caloric suitability of the ethnic homeland (columns 4-5) or grid cell (column 6). All regressions control for the longitude and latitude of the archaeological site where the individual's remains were found. All variables are standardized to have zero mean and standard deviation equal to one. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.4: Colonial Parishes and Historical Crop Exchange

	Dependent Variable:	
	Number of Parishes	% Parishes with Ethnic Div.
	(1)	(2)
Crop exchange (H index)	0.154 [0.145]	-0.114 [0.169]
	(0.174)	(0.165)
Number of ethnic groups	47	47

Notes. The unit of observation is the ethnic group. The table reports OLS estimates. Robust standard errors in brackets. In parentheses, standard errors are corrected for spatial dependence with a distance cutoff of approximately one degree at the equator (Colella et al. 2019). Crop exchange is the ethnic-level proxy for potential exposure to internal crop exchange (H index). All variables are standardized to have zero mean and standard deviation equal to one.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.5: Ethnic Diversity and Historical Crop Exchange

	Ethnic Diversity = 1		Ethnic Diversity = 0		Diff.	p-value ^a	p-value ^b
	mean	sd	mean	sd			
Crop exchange	0.674	0.173	0.652	0.179	-0.023	[0.260]	[0.319]
Number of parishes	117	117	219	219	336	336	336

Notes. The unit of observation is the parish. Ethnic diversity is a dummy variable taking value 1 if there is an ethnic border within a buffer of 10-km radius from the parish capital, using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers. Crop exchange is the parish-level proxy for average exposure to internal crop exchange, computed using the area shares of ethnic groups within the 10-km parish buffer as weights. P-values from OLS regressions of average exposure to crop exchange on ethnic diversity; (^a) with robust standard errors, (^b) with standard errors corrected for spatial dependence using a distance cutoff of approximately one degree at the equator (Colella et al. 2019).

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.6: Historical Crop Exchange — Pre-Colonial Infrastructure

	Dependent Variable:		
	Ln Road Density	Dummy Canals	Dummy Bridges
	(1)	(2)	(3)
Crop exchange (H index)	0.125 [0.131] (0.143)	-0.008 [0.062] (0.046)	-0.038 [0.056] (0.045)
Number of ethnic groups	47	47	47

Notes. The unit of observation is the ethnic group. The table reports OLS estimates. Robust standard errors in brackets. In parentheses, standard errors are corrected for spatial dependence with a distance cutoff of approximately one degree at the equator (Colella et al. 2019). Crop exchange is the ethnic-level proxy for potential exposure to internal crop exchange (H index). All variables except dummies are standardized to have zero mean and standard deviation equal to one. The dummy variables for canals and bridges take value 1 for

27.66 and 23.40 percent of the groups, respectively.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.7: Historical Crop Exchange — Inca Road

	Dependent Variable:		
	Ln Dist Road (Outside)	Ln Dist Road (Inside)	Ln Dist Road (Any)
	(1)	(2)	(3)
Crop exchange (H index)	-0.120 [0.118] (0.122)	-0.101 [0.196] (0.217)	-0.074 [0.138] (0.149)
Number of ethnic groups	47	40	47

Notes. The unit of observation is the ethnic group. The table reports OLS estimates. Robust standard errors in brackets. In parentheses, standard errors are corrected for spatial dependence with a distance cutoff of approximately one degree at the equator (Colella et al. 2019). Crop exchange is the ethnic-level proxy for potential exposure to internal crop exchange (H index). The dependent variables refer to the minimum distance across all 10×10 km grid cells inside the ethnic homeland to an Inca road located (1) in a different ethnic homeland, (2) inside the ethnic homeland, or (3) any. All variables are standardized to have zero mean and standard deviation equal to one.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.8: Wild Cluster Bootstrap (WCB) Inference

	(1)	(2)	(3)	(4)	(5)
	Log(1+light intensity per capita), 2013, VIIRS	Log(1+light intensity per capita), 2010-2013, DMSP	Firms per 100 inh., 2010-2013	Non-subs. farming (share of farmers > median), 2012	Non-subs. farming (share of farmers), 2012
Ethnic diversity	-0.036** [-2.080]	-0.107*** [-2.966]	-11.991 [-1.585]	-0.416** [-2.368]	-0.131* [-1.765]
Ethnic div. × Crop exchange	0.042* [1.909]	0.119** [2.665]	19.147* [1.710]	0.593** [2.372]	0.151 [1.405]
Baseline controls	Yes	Yes	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes	Yes	Yes
Number of parishes	336	336	336	336	336
	(6)	(7)	(8)	(9)	(10)
	Access to public water (share of dwellings), 2017	Access to public sanitation (share of dwellings), 2017	Log(hh. consumption per capita), 2004-2017	Log(hh. consumption per capita), 2004-2010	Log(hh. consumption per capita), 2011-2017
Ethnic diversity	-0.088 [-0.954]	-0.096 [-0.847]	-0.301* [-1.866]	-0.386* [-1.914]	-0.238 [-1.415]
Ethnic div. × Crop exchange	0.122 [0.913]	0.134 [0.713]	0.491** [2.156]	0.601** [2.129]	0.420* [1.784]
Baseline controls	Yes	Yes	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes	Yes	Yes
Year FE	–	–	Yes	Yes	Yes
Household controls	–	–	Yes	Yes	Yes
Number of parishes	336	336	280	220	262
Number of households	–	–	53,361	21,258	32,103

Notes. The unit of observation is the parish in columns 1-7 and the household in columns 8-10. Wild cluster bootstrap t-statistics in brackets (Cameron et al. 2008), with clusters at the province level in columns 1-7 and at the parish level in columns 8-10. Ethnic diversity is a dummy variable taking value 1 if there is an ethnic border within a buffer of 10-km radius from the parish capital, using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers. Crop exchange is the parish-level proxy for average exposure to internal crop exchange, computed using the area shares of ethnic groups within the 10-km parish buffer as weights. The vector of baseline controls includes parish-level mean and standard deviation of elevation, mean and standard deviation of land caloric suitability, longitude, latitude, log distance to perennial rivers, log distance to pre-colonial native shrines, log 16th-century expected tribute, and log distance to *mita* mines. The vector of household controls includes gender, age, age squared, years of schooling, civil status, and language spoken at home of the household head. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.9: Robustness—Pre-Colonial Characteristics of Ethnic Groups

Living Standards (AES, 2010 – 2020)						
	(1)	(2)	(3)	(4)	(5)	(6)
Ethnic diversity	-0.122 [0.762]	-1.449 [1.495]	-3.333 [2.120]	-0.445* [0.240]	-0.706*** [0.225]	-3.087 [1.913]
Ethnic div. × <u>Crop exchange</u>	0.768** [0.348]	1.332** [0.604]	1.423*** [0.547]	0.934*** [0.332]	1.117*** [0.365]	1.302** [0.662]
Ethnic div. × <u>Ln land area</u>	-0.049 [0.075]	0.040 [0.123]	0.132 [0.140]			0.199 [0.162]
Ethnic div. × <u>Elevation</u>		0.319 [0.458]	0.165 [0.457]			0.623 [0.615]
Ethnic div. × <u>Caloric suitability</u>		-0.968 [0.744]	-1.036 [0.868]			-2.601 [1.824]
Ethnic div. × <u>Ln river density</u>			0.265 [0.265]			
Ethnic div. × <u>Ln population density</u>			0.060 [0.124]			
Ethnic div. × <u>Urbanization</u>		0.044 [0.298]	0.056 [0.334]			-0.224 [0.410]
Ethnic div. × <u>Political complexity</u>		-0.173 [0.278]	-0.275 [0.299]			-0.160 [0.264]
Ethnic div. × <u>Elite residences</u>		0.099 [0.229]	0.132 [0.234]			0.077 [0.236]
Ethnic div. × <u>SD of elevation</u>				-0.391 [0.327]		-0.208 [0.586]
Ethnic div. × <u>SD of caloric suitability</u>					-0.398	1.369
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of parishes	336	336	336	336	336	336

Notes. The unit of observation is the parish. Robust standard errors clustered at the level of the colonial province in brackets. For each pre-colonial characteristic, the specification includes the weighted average of the characteristic (computed using the area shares of ethnic groups within the 10-km parish buffer as weights) and its interaction with ethnic diversity. Only the estimated coefficient on the interaction term is reported.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.10: Balance Tests for Ethnic Diversity—CEM Matched Samples

Dependent Variable:									
Mean	SD of	Mean	SD of	Ln Dist	Ln Dist	Ln Dist	Ln Expected	Ln Dist	Ln Dist
Elevation	Elevation	Caloric Suit.	Caloric Suit.	Perennial River	Native Shrine	Tribute (16th c.)	Mita Mine		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
Panel A: Matched Sample Based on CEM for Baseline Pre-Colonial Characteristics									
Ethnic diversity	92.649 [0.212]	-17.039 [0.100]	-25.414 [0.121]	-0.186 [0.574]	-0.190 [0.235]	-0.014 [0.922]	0.256 [0.322]		
Observations	90	90	90	90	90	90	90		
Panel B: Matched Sample Based on CEM for Lasso-Selected Pre-Colonial Characteristics									
Ethnic diversity	-33.329 [0.710]	-15.592 [0.283]	-27.970 [0.225]	-0.394 [0.450]	-0.028 [0.941]	0.304** [0.033]	-0.086 [0.834]		
Observations	74	74	74	74	74	74	74		

Notes. The unit of observation is the parish. Robust standard errors in brackets. Ethnic diversity is a dummy variable taking value 1 if there is an ethnic border within a buffer of 10-km radius from the parish capital, using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers. The matched sample refers to the sample of parishes selected by the CEM algorithm of Iacus, King, and Porro (2012), using parishes with average exposure to internal crop exchange above the median as treated parishes. The pre-colonial characteristics to be balanced by the algorithm are the baseline characteristics (Column 9 of Table 5) in Panel A and the lasso-selected characteristics (Column 10 of Table 5) in Panel B. OLS estimates with CEM matching weights.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.11: Randomization Inference

Firms per 100 inh.									
Non-subs. farming			Light intensity per capita			Randomization-t p -value			
Randomization-t p -value			Randomization-t p -value			Randomization-t p -value			
Minimum	Maximum	Randomized	Minimum	Maximum	Randomized	Minimum	Maximum	Randomized	Randomized
Ethnic diversity	0.020	0.021	0.020	0.046	0.046	0.119	0.120	0.119	
Ethnic div. $\times$ Crop exchange	0.025	0.026	0.025	0.068	0.068	0.115	0.116	0.115	

Notes. The unit of observation is the parish. The regressions include baseline controls and colonial province fixed effects, with robust standard errors clustered at the level of the colonial province. Randomization-t p -values result from permuting ethnic diversity (1,000 iterations) and recalculating its interaction with average exposure to crop exchange accordingly, following the procedure proposed by Young (2024). Robust standard errors are clustered at the level of the colonial province. The outcome variables are: a dummy variable indicating whether the share of farmers practicing non-subsistence agriculture is above the sample median (2012), the log of 1 plus average light intensity per capita (2013, VIIRS), and the average number of small- and medium-sized firms per 100 inhabitants (2010-2013).

TABLE A.12: Robustness—Pre-Colonial Land Occupation, Transition Zones, and Placebos

	Living Standards (AES, 2010 – 2020)					
	Pre-Colonial Land Occupation			Transition Zones		Placebos
	(1)	(2)	(3)	(4)	(5)	(6)
Ethnic diversity	-0.641*** [0.233]	-0.742*** [0.273]	-0.645*** [0.247]	-0.669** [0.305]		
Ethnic div. $\times$ $\overline{\text{Crop exchange}}$ (20km correction)	0.886** [0.360]					
Ethnic div. $\times$ $\overline{\text{Crop exchange}}$ (10km correction)		0.985** [0.394]				
Ethnic div. $\times$ $\overline{\text{Crop exchange}}$ (20km transition-zone buffer)			0.877** [0.374]			
Ethnic div. $\times$ $\overline{\text{Crop exchange}}$ (10km transition-zone buffer)				0.864** [0.435]		
Dummy (Artificial ethnic border within parish buffer)					-0.260 [0.395]	
Dummy $\times$ $\overline{\text{Crop exchange}}$ (Artificial)					0.399 [0.498]	
Dummy (<i>Corregimiento</i> border within parish buffer)						0.165 [0.246]
Dummy $\times$ $\overline{\text{Crop exchange}}$ (<i>Corregimiento</i>)						-0.288 [0.359]
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of parishes	336	336	336	336	336	336

Notes. The unit of observation is the parish. Robust standard errors clustered at the level of the colonial province in brackets. Ethnic diversity is a dummy variable taking value 1 if there is an ethnic border within a buffer of 10-km radius from the parish capital, using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers. The table reports the standardized AES (Kling et al. 2004; Clingingsmith, Khwaja and Kremer 2009) across five outcomes: a dummy variable indicating whether the share of farmers practicing non-subsistence agriculture is above the sample median (2012), the log of 1 plus average light intensity per capita (2013, VIIRS), the average number of small- and medium-sized firms per 100 inhabitants (2010-2013), the share of dwellings with access to public sanitation (2017), and the share of dwellings with access to the public water network (2017). The vector of baseline controls includes parish-level mean and standard deviation of elevation, mean and standard deviation of land caloric suitability, longitude, latitude, log distance to perennial rivers, log distance to pre-colonial native shrines, log 16th-century expected tribute, and log distance to *mita* mines. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.13: Robustness—Varying the Buffer Size

	Living Standards (AES, 2010 – 2020)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	7km	8km	9km	10km	11km	12km	13km	14km	15km
Ethnic diversity	-0.308 [0.229]	-0.341 [0.212]	-0.435** [0.189]	-0.610*** [0.225]	-0.394** [0.181]	-0.454*** [0.170]	-0.567*** [0.168]	-0.580*** [0.194]	-0.533*** [0.211]
Ethnic div. $\times$ Crop exchange	0.417 [0.365]	0.449 [0.333]	0.598** [0.304]	0.846** [0.350]	0.525* [0.296]	0.673** [0.269]	0.890*** [0.286]	0.905*** [0.317]	0.793*** [0.360]
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of parishes	336	336	336	336	336	336	336	336	336
% parishes intersecting ethnic border	27.38	30.36	32.14	34.82	37.20	40.18	42.26	45.24	46.43

Notes. The unit of observation is the parish. Robust standard errors clustered at the level of the colonial province in brackets. Ethnic diversity takes value 1 if there is an ethnic border within a non-overlapping buffer of r -km radius from the parish capital, for $r \in \{7, 8, 9, 10, 11, 12, 13, 14, 15\}$ (indicated at the top of each column), and 0 otherwise. Crop exchange is the parish-level proxy for average exposure to internal crop exchange, computed using the area shares of ethnic groups within the 10-km parish buffer as weights. The table reports the standardized AES (Kling et al. 2004; Clingsmith, Khwaja and Kremer 2009) across five outcomes: a dummy variable indicating whether the share of farmers practicing non-subsistence agriculture is above the sample median (2012), the log of 1 plus average light intensity per capita (2013, VIIRS), the average number of small- and medium-sized firms per 100 inhabitants (2010-2013), the share of dwellings with access to public sanitation (2017), and the share of dwellings with access to the public water network (2017). The vector of baseline controls includes parish-level mean and standard deviation of elevation, mean and standard deviation of land caloric suitability, longitude, latitude, log distance to perennial rivers, log distance to pre-colonial native shrines, log 16th-century expected tribute, and log distance to *mita* mines. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.14: Robustness—Alternative Diversity Indices

	Living Standards (AES, 2010 – 2020)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ethnic fractionalization	-1.461*** [0.557]	-0.999 [0.641]						
Ethnic frac. $\times$ $\overline{\text{Crop exchange}}$	1.741** [0.795]	1.476 [0.982]						
Ethnic diversity (dummy)			-0.959*** [0.221]	-0.612** [0.269]	-0.681*** [0.250]	-0.496** [0.214]	-0.904*** [0.241]	-0.596*** [0.223]
Ethnic div. $\times$ $\overline{\text{Crop exchange}}$ ($\tilde{H}$ index)			1.475*** [0.403]	1.014** [0.495]				
Ethnic div. $\times$ $\overline{\text{Crop exchange}}$ (majority group)					0.812** [0.371]	0.660** [0.324]		
Ethnic div. $\times$ $\overline{\text{Crop exchange}}$ (minority group)							1.218*** [0.340]	0.887*** [0.344]
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
Colonial province FE	No	Yes	No	Yes	No	Yes	No	Yes
Number of parishes	336	336	336	336	336	336	336	336

Notes. The unit of observation is the parish. Robust standard errors clustered at the level of the colonial province in brackets. Columns 1-2 use ethnic fractionalization, computed as $ethnic\frac{frac_p}{1 - \sum_e w_{ep}^2}$. Columns 3-8 use the ethnic diversity dummy taking value 1 if there is an ethnic border within a buffer of 10-km radius from the parish capital (using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers). As a measure of average exposure to pre-colonial crop exchange, columns 1-2 use the weighted average defined in the main text, $Crop\ exchange_p = \sum_e w_{ep} H_e$, where $H_e = 1 / \sum_j s_{je}^2$; columns 3-4 use a Herfindahl index to measure H_e instead, specifically $\tilde{H}_e = 1 - \sum_j s_{je}^2$; columns 5-6 use the value (H_e) of the majority group in the parish, defined as the group with the highest area share within the 10-km buffer, instead of a weighted average; while columns 7-8 use the value (H_e) of the minority group. The table reports the standardized AES (Kling et al. 2004; Clingsmith, Khwaja and Kremer 2009) across five outcomes: a dummy variable indicating whether the share of farmers practicing non-subsistence agriculture is above the sample median (2012), the log of 1 plus average light intensity per capita (2013, VIIRS), the average number of small- and medium-sized firms per 100 inhabitants (2010-2013), the share of dwellings with access to public sanitation (2017), and the share of dwellings with access to the public water network (2017). The vector of baseline controls includes parish-level mean and standard deviation of elevation, mean and standard deviation of land caloric suitability, longitude, latitude, log distance to perennial rivers, log distance to pre-colonial native shrines, log 16th-century expected tribute, and log distance to *mita* mines.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.15: Robustness—Inca Period and Pre-Resettlement Spread of Smallpox

	Living Standards (AES, 2010 – 2020)			
	(1)	(2)	(3)	(4)
Ethnic diversity	-0.610*** [0.225]	-0.504 [0.393]	-0.615*** [0.229]	-0.539*** [0.263]
Ethnic div. $\times$ <u>Crop exchange</u>	0.846** [0.350]	0.865** [0.351]	0.845** [0.354]	0.786** [0.399]
Ethnic div. $\times$ <u>Ln dist. smallpox outbreak</u>		-0.022 [0.065]		
Baseline controls	Yes	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes	Yes
Inca region (<i>suyu</i>) FE	No	No	Yes	No
Excluding groups potentially affected by Inca resettlements	No	No	No	Yes
Number of parishes	336	336	336	241

Notes. The unit of observation is the parish. Robust standard errors clustered at the level of the colonial province in brackets. Ethnic diversity is a dummy variable taking value 1 if there is an ethnic border within a buffer of 10-km radius from the parish capital, using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers. Crop exchange is the parish-level proxy for average exposure to internal crop exchange, computed using the area shares of ethnic groups within the 10-km parish buffer as weights. The table reports the standardized AES (Kling et al. 2004; Clingsmith, Khwaja and Kremer 2009) across five outcomes: a dummy variable indicating whether the share of farmers practicing non-subsistence agriculture is above the sample median (2012), the log of 1 plus average light intensity per capita (2013, VIIRS), the average number of small- and medium-sized firms per 100 inhabitants (2010–2013), the share of dwellings with access to public sanitation (2017), and the share of dwellings with access to the public water network (2017). In Column 2, I control for the average distance from each ethnic group's centroid to the closest outbreak of smallpox, either Tomebamba or Cuzco, after the Spanish conquest (1524–1526). The last two columns present point estimates after including fixed effects for the four major Inca regions or *suyus* (Column 3) and after excluding parishes that concentrated groups potentially affected by Inca resettlements (Column 4) according to Rowe (1946). The vector of baseline controls includes parish-level mean and standard deviation of elevation, mean and standard deviation of land caloric suitability, longitude, latitude, log distance to perennial rivers, log distance to pre-colonial native shrines, log 16th-century expected tribute, and log distance to *mita* mines.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B Household Consumption and Selective Migration

Individuals may have managed to escape from their respective parishes after the resettlement. For example, one possibility is that the most *capable* individuals from ethnically diverse parishes tried to join their coethnics located in parishes with an ethnically homogeneous founding population. This type of selective migration may help explain the overall negative impacts of ethnic diversity if the skills and knowledge that these individuals (the most *capable* ones) brought with them were (i) heritable and (ii) relevant for economic development; Dell (2010); Lowes and Montero (2021).

In Table A.16, I use survey data on household consumption to explore this possibility. The outcome variable is the log of real household consumption per capita (as in Column 8 of Table 4). Although the surveys are not representative at the local (parish) level, they allow me to trim the sample. Specifically, I trim the sample by dropping the top x percent of households (in terms of real household consumption per capita) from not ethnically diverse parishes where average exposure to crop exchange was above the sample median, i.e. the better-off parishes. I present the results for different bandwidths ($x \in \{2.2, 4.2, 6.2, 8.2, 10.2\}$ percent). Reassuringly, the results for the trimmed samples are consistent with the main result of the paper.

TABLE A.16: Robustness—Household Consumption and Selective Migration

	Dependent Variable: Log Real Household Consumption Per Capita (2004 – 2017)				
	(1)	(2)	(3)	(4)	(5)
Ethnic diversity	-0.287 [0.149]*	-0.279 [0.144]*	-0.269 [0.142]*	-0.268 [0.140]*	-0.266 [0.139]*
	(0.147)*	(0.145)*	(0.143)*	(0.143)*	(0.142)*
Ethnic div. $\times$ Crop exchange	0.498	0.507	0.514	0.532	0.548
	[0.214]**	[0.209]**	[0.206]**	[0.205]**	[0.203]**
	(0.213)**	(0.212)**	(0.211)**	(0.211)**	(0.210)**
Baseline controls	Yes	Yes	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes	Yes	Yes
Household controls	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Number of parishes	280	280	280	280	280
Number of households	52,493	51,707	50,916	50,133	49,335
% Trimmed	2.2 %	4.2 %	6.2 %	8.2 %	10.2 %

Notes. The unit of observation is the household. The table reports OLS estimates. Robust standard errors are clustered at the parish level in brackets. In parentheses, standard errors are corrected for spatial dependence with a distance cutoff of approximately one degree at the equator (Colella et al. 2019). Ethnic diversity is a dummy variable taking value 1 if there is an ethnic border within a buffer of 10-km radius from the parish capital, using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers. Crop exchange is the parish-level proxy for average exposure to internal crop exchange, computed using the area shares of ethnic groups within the 10-km parish buffer as weights. The sample is trimmed by dropping the top x percent of households (in terms of real household consumption per capita) from interior parishes (ethnic diversity=0) where average exposure to crop exchange is above the sample median, for $x \in \{2.2, 4.2, 6.2, 8.2, 10.2\}$. The vector of household controls includes gender, age, age squared, years of schooling, civil status, and language spoken at home of the household head. The vector of baseline controls includes parish-level mean and standard deviation of elevation, mean and standard deviation of land caloric suitability, longitude, latitude, log distance to perennial rivers, log distance to pre-colonial native shrines, log 16th-century expected tribute, and log distance to *mita* mines. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.17: Mechanisms—Societal Integration and Cooperative Behavior

Panel A:	Dependent Variable:			
	Military Service (2008-2014)		Dummy Neighborhood Associations	
	Ln (1 + Av. Volunteers)	Dummy Volunteers	2002-2004	2012-2014
	(1)	(2)	(3)	(4)
Ethnic diversity	-0.130 [0.046]***	-0.242 [0.090]**	-0.440 [0.210]**	-0.526 [0.158]***
Ethnic div. × <u>Crop exchange</u>	(0.030)*** 0.244 [0.079]***	(0.070)*** 0.379 [0.124]***	(0.227)* 0.698 [0.316]**	(0.221)** 0.708 [0.228]***
Ln population	(0.072)*** Yes	(0.124)*** Yes	(0.344)** Yes	(0.318)** Yes
Number of parishes	336	336	336	336
Panel B:	Participation Dummy (2004-2017)			
	Neighborhood Associations	Professional Associations	Labor Associations	Some Association
	(1)	(2)	(3)	(4)
Ethnic diversity	-0.053 [0.037]	-0.023 [0.009]**	-0.050 [0.027]*	-0.122 [0.045]***
Ethnic div. × <u>Crop exchange</u>	(0.052) 0.085 [0.047]*	(0.012)* 0.035 [0.014]**	(0.030)* 0.060 [0.034]*	(0.056)** 0.175 [0.059]***
Individual controls and year FE	(0.070) Yes	(0.021)* Yes	(0.036)* Yes	(0.072)** Yes
Number of parishes	280	280	280	280
Number of individuals	54,676	54,676	54,676	54,676
Baseline controls	Yes	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes	Yes

Notes. The unit of observation is the parish in Panel A and the individual in Panel B (individual-level yearly data from the ENAHO Peruvian household surveys; 2004-2017 waves). Robust standard errors clustered at the province level (Panel A) or parish level (Panel B) in brackets. In parentheses, standard errors are corrected for spatial dependence with a distance cutoff of approximately one degree at the equator (Colella et al. 2019). The table reports OLS estimates. Ethnic diversity is a dummy variable taking value 1 if there is an ethnic border within a buffer of 10-km radius from the parish capital, using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers. Crop exchange is the parish-level proxy for average exposure to internal crop exchange, computed using the area shares of ethnic groups within the 10-km parish buffer as weights. In panels A and B, the vector of baseline controls includes parish-level mean and standard deviation of elevation, mean and standard deviation of land caloric suitability, longitude, latitude, log distance to perennial rivers, log distance to pre-colonial native shrines, log 16th-century expected tribute, and log distance to *mita* mines. In Panel B, the vector of individual-level controls includes gender, age, age squared, years of schooling, civil status, and language spoken at home. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.18: Mechanisms—Minority Group Advantage

Living Standards (AES, 2010 – 2020)			
	(1)	(2)	(3)
Crop exchange	1.055* [0.594]	2.011** [0.791]	1.955*** [0.719]
Minority group advantage		0.931** [0.449]	-1.510 [1.305]
Crop exchange $\times$ Minority group advantage			3.499* [2.022]
Number of parishes	117	117	117
Baseline controls	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes

Notes. The unit of observation is the parish. The sample includes parishes with ethnic diversity only. Robust standard errors clustered at the level of the colonial province in brackets. Crop exchange represents the average level of exposure to internal crop exchange among the minority and majority groups in the parish. This average is computed using the area shares of these ethnic groups within the 10-km parish buffer as weights. The *Minority group advantage* is defined as the difference in crop exchange exposure between the minority and majority groups, computed as: *Crop exchange (minority group) - Crop exchange (majority group)*. The table reports the standardized AES (Kling et al. 2004; Clingingsmith, Khwaja and Kremer 2009) across five outcomes: a dummy variable indicating whether the share of farmers practicing non-subsistence agriculture is above the sample median (2012), the log of 1 plus average light intensity per capita (2013, VIIRS), the average number of small- and medium-sized firms per 100 inhabitants (2010-2013), the share of dwellings with access to public sanitation (2017), and the share of dwellings with access to the public water network (2017). Column 10 shows the coefficients of lasso-selected characteristics (Belloni, Chernozhukov and Hansen 2014). The lasso algorithm uses the score of the first principal component derived from the five outcomes as dependent variable. The vector of baseline controls includes parish-level mean and standard deviation of elevation, mean and standard deviation of land caloric suitability, longitude, latitude, log distance to perennial rivers, log distance to pre-colonial native shrines, log 16th-century expected tribute, and log distance to *mita* mines.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.19: Mechanisms—Non-Subsistence Agriculture and Local Crop Availability

	Dependent Variable:						
	Non-Subsistence Agriculture (1994)			Non-Subsistence Agriculture (2012)			Urban Status (2017)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ethnic diversity	-0.291 [0.162]* (0.134)**	-0.298 [0.168]* (0.136)**	-0.292 [0.162]* (0.135)**	-0.416 [0.163]** (0.127)**	-0.444 [0.170]** (0.123)**	-0.392 [0.172]** (0.144)**	-0.138 [0.183] (0.074)*
Ethnic div. $\times$ Crop exchange	0.516 [0.265]* (0.220)**	0.524 [0.272]* (0.222)**	0.518 [0.265]* (0.221)**	0.593 [0.233]** (0.178)**	0.624 [0.238]** (0.173)**	0.556 [0.240]** (0.197)**	0.191 [0.268] (0.139)
Ln (1 + # native crops)		0.048 [0.066] (0.055)			0.198 [0.092]** (0.086)**		
Native crop frac.			-0.013 [0.174] (0.195)			0.365 [0.203]* (0.141)**	
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of parishes	336	336	336	336	336	336	336

Notes. The unit of observation is the parish. Robust standard errors clustered at the level of the colonial province in brackets. In parentheses, standard errors are corrected for spatial dependence with a distance cutoff of approximately one degree at the equator (Colella et al. 2019). The table reports OLS estimates. Ethnic diversity is a dummy variable taking value 1 if there is an ethnic border within a buffer of 10-km radius from the parish capital, using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers. Crop exchange is the parish-level proxy for average exposure to internal crop exchange, computed using the area shares of ethnic groups within the 10-km parish buffer as weights. In columns 1-6, the dependent variable is a dummy variable indicating whether the share of farmers practicing non-subsistence agriculture is above the sample median (1994 in columns 1-3 and 2012 in columns 4-6). In column 7, the dependent variable is a dummy variable indicating whether the share of population living in urban areas (defined as areas where at least 100 dwellings are grouped contiguously by INEI) is above the sample median. The vector of baseline controls includes parish-level mean and standard deviation of elevation, mean and standard deviation of land caloric suitability, longitude, latitude, log distance to perennial rivers, log distance to pre-colonial native shrines, log 16th-century expected tribute, and log distance to *mita* mines.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.20: Mechanisms—Institutional Performance

	Dependent Variable: Share of Budget Executed		
	(2002-2005)	(2006-2009)	(2010-2013)
	(1)	(2)	(3)
Ethnic diversity	-0.070 [0.060]	-0.144* [0.085]	-0.036 [0.083]
	(0.062)	(0.079)	(0.045)
Ethnic div. $\times$ <u>Crop exchange</u>	0.109 [0.090]	0.220* [0.127]	0.033 [0.120]
	(0.084)	(0.114)	(0.076)
Mean Dep. Var.	0.771	0.625	0.568
Baseline controls	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes
Number of parishes	336	336	336

Notes. The unit of observation is the parish. Robust standard errors clustered at the level of the colonial province in brackets. In parentheses, standard errors are corrected for spatial dependence with a distance cutoff of approximately one degree at the equator (Colella et al. 2019). The table reports OLS estimates. Ethnic diversity is a dummy variable taking value 1 if there is an ethnic border within a buffer of 10-km radius from the parish capital, using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers. Crop exchange is the parish-level proxy for average exposure to internal crop exchange, computed using the area shares of ethnic groups within the 10-km parish buffer as weights. The dependent variables refer to the average share of the municipal budget executed (Ministry of Economy and Finance, MEF). The vector of baseline controls includes parish-level mean and standard deviation of elevation, mean and standard deviation of land caloric suitability, longitude, latitude, log distance to perennial rivers, log distance to pre-colonial native shrines, log 16th-century expected tribute, and log distance to *mita* mines.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A.21: Mechanisms—Pre-Colonial Crop Diversity and Structural Transformation

	Dependent Variable: Share of Population					
	Tertiary Sector		Primary Sector		Secondary Sector	
	(1876)	(2007-2017)	(1876)	(2007-2017)	(1876)	(2007-2017)
	(1)	(2)	(3)	(4)	(5)	(6)
Ethnic diversity	-0.024 [0.086]	-0.214 [0.243]	0.084 [0.269]	0.309 [0.308]	-0.060 [0.271]	-0.095 [0.134]
Ethnic div. $\times$ $\overline{\text{Crop frac.}}$	0.041 [0.186]	0.388 [0.456]	-0.163 [0.505]	-0.560 [0.568]	0.123 [0.512]	0.172 [0.251]
Mean Dep. Var.	0.073	0.336	0.687	0.517	0.240	0.147
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of parishes	282	336	282	336	282	336

Notes. The unit of observation is the parish. Robust standard errors clustered at the level of the colonial province in brackets. The outcomes refer to the share of the population employed in the tertiary (1-2), primary (3-4), and secondary (5-6) sectors. Ethnic diversity is a dummy variable taking value 1 if there is an ethnic border within a buffer of 10-km radius from the parish capital, using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers. Crop frac is a measure of average crop fractionalization, computed using the area shares of ethnic groups within the 10-km parish buffer as weights and considering native crops only. To detect native crops within the pre-colonial homelands of ethnic groups, I rely on the 2012 agricultural census, which provides geo-referenced data for an extensive set of crops reported by farmers at the time of the census. I then follow the classification of native crops in Tapia (2013). Note that farmers were only required to report the list of crops harvested at the time of the census—crop rotation and fallow practices can affect the list of reported crops. Regressions are weighted by the square root of the total population. The vector of baseline controls includes parish-level mean and standard deviation of elevation, mean and standard deviation of land caloric suitability, longitude, latitude, log distance to perennial rivers, log distance to pre-colonial native shrines, log 16th-century expected tribute, and log distance to *mita* mines.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B Supplemental Appendix - Colonial Surnames

In certain contexts, measures based on the frequency distribution of surnames can shed light on the biological relationships between human populations. Provided that surnames are inherited, the underlying premise of this approach is that surname commonality between individuals (isonymy) can be used to trace common ancestry (Lasker 1980, 1985; Colantonio 2003). Two main diversity indices have been applied to surnames:

$$D = 1 - \sum_{k=1}^K p_k^2, \quad S = - \sum_{k=1}^K p_k \ln(p_k)$$

where p_k represents the proportion of individuals with surname k in the population and K is the total number of different surnames. The first index, $D \in [0, 1]$, is a standard measure of diversity based on the Simpson or Herfindahl index. The second index, $S \in [0, \ln(K)]$, takes its theoretical basis from information theory (Shannon 1948) and can be interpreted in terms of uncertainty in predicting ancestry, as long as any two individuals with the same surname inherited the surname from a common ancestor.

I created a dataset of colonial individuals with native paternal surnames using baptism records from digital genealogical sources ([FamilySearch.org](https://www.familysearch.org)). Each record reports the full name of the individual, parish, and date of baptism (1605–1780). The dataset provides information for 112,340 individuals.⁵² Table B.1 presents OLS estimates from regressing surname diversity (either the S index or the D index) on the dummy for *border* parishes. Since it can be reasonably assumed that not all historical records have been preserved, the results should be interpreted with caution. Panel A shows the baseline results. For each surname diversity index, the first column shows the unconditional correlation; the second column controls for the log number of individuals found in the records of the parish and for the share of individuals with non-native surnames; and the last column includes ecclesiastical jurisdiction fixed effects, accounting for potential differences in the administration of baptism across five colonial bishoprics. Panel B shows the estimates obtained after dropping individuals whose surnames occur only once in the dataset, which results in a sample size

⁵²I use information from the collection “Perú, bautismos, 1556-1930,” accessed December 2018. The number of parishes with information varies by year. The mean parish comprises 1,726 individuals with native paternal surnames, of whom 857 are men, relative to a sample mean of 1,627 individuals per parish according to the census of 1791–1795 (of which 769 are men).

of 106,124 individuals. In Panel C, I show the estimates obtained from using groups of similar surnames (instead of raw surnames) to compute surname diversity indices.⁵³ While not perfect, this approach helps account for potential changes in the written spelling of surnames over time. Panel D restricts the analysis to individuals with non-native paternal surnames.

Identification of native surnames. In order to identify native surnames, I constructed a dictionary of linguistic roots from the Quechuan and Aymaran language families. There is no unique source for the identification of surnames from these families. For Quechua, the main sources are the classic dictionary by González Holguín (1952)[1608] and a recent dictionary compiled by the Academia Mayor de la Lengua Quechua (2005). I also include the list of names provided by the Peruvian *Registro Nacional de Identificación y Estado Civil* (RENIEC 2012). For Aymara, the main sources are the classic dictionary by Bertonio (2011)[1612], the list of surnames provided by De Lucca (1983), and a recent dictionary compiled by CONADI (2011). I complement the analysis using two additional sources: (1) *Vocabulario Políglota Incaico*, originally compiled by Franciscan missionaries in Peru, which provides an extensive list of words in four dialects of Quechua (varieties of Cuzco, Ayacucho, Junín and Ancash) and Aymara, see Fide (1998)[1905]; and (2) the *An Crúbadán-Corpus Building for Minority Languages* project (Scannell 2007), which provides downloadable text datasets for different dialects of Quechua and Aymara based on online text resources, including translations of the Bible and the Universal Declaration of Human Rights (accessed May 2019).

⁵³Specifically, I group surnames if the deletion, insertion, or substitution of only one character is required to transform one surname into another (i.e., the surnames have a Levenshtein distance equal to one).

TABLE B.1: Validating Ethnic Diversity

Dependent Variable: Surname Diversity (1605 – 1780)						
	S Index	S Index	S Index	D Index	D Index	D Index
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Baseline (Native Surnames)						
Border parish (dummy)	0.512 [0.212]** (0.213)**	0.447 [0.182]** (0.207)**	0.558 [0.191]**** (0.203)****	0.481 [0.170]**** (0.217)**	0.448 [0.182]** (0.201)**	0.608 [0.218]**** (0.261)**
Panel B: Non-Unique Surnames (Native Surnames)						
Border parish (dummy)	0.462 [0.209]** (0.208)**	0.408 [0.159]** (0.180)**	0.513 [0.180]**** (0.204)**	0.454 [0.173]** (0.218)**	0.428 [0.178]** (0.196)**	0.581 [0.214]**** (0.260)**
Panel C: Grouped Surnames (Native Surnames)						
Border parish (dummy)	0.482 [0.215]** (0.213)**	0.416 [0.184]** (0.205)**	0.523 [0.189]**** (0.190)****	0.461 [0.172]**** (0.216)**	0.427 [0.181]** (0.199)**	0.582 [0.216]**** (0.253)**
Panel D: Non-Native Surnames						
Border parish (dummy)	0.284 [0.257] (0.223)	0.116 [0.101] (0.097)	0.005 [0.098] (0.083)	0.333 [0.209] (0.195)*	0.185 [0.116] (0.126)	0.112 [0.136] (0.125)
Number of parishes	65	65	65	65	65	65
Ln total individuals (1605–1780)	No	Yes	Yes	No	Yes	Yes
% Non-native surnames (1605–1780)	No	Yes	Yes	No	Yes	Yes
Ecclesiastical jurisd. FE	No	No	Yes	No	No	Yes

Notes. The unit of observation is the parish. The table reports OLS estimates. Robust standard errors in brackets. In parentheses, standard errors are corrected for spatial dependence with a distance cutoff of approximately one degree at the equator (Colella et al. 2019). Border parishes are parishes with an ethnic border within a buffer of 10-km radius from the parish capital (using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers). All variables except dummies are standardized to have zero mean and standard deviation equal to one.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

C Supplemental Appendix - Data Sources and Definitions

Mean elevation [parish level]. Average elevation within a buffer of 10-km radius from the parish capital (using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers). Source: author's computation using version 1.2 of the [Harmonized World Soil Database](#) (Fischer et al. 2008); 30 arc-second raster data with median elevation based on information from the NASA Shuttle Radar Topographic Mission (SRTM).

Variation in elevation [parish level]. Standard deviation of elevation within a buffer of 10-km radius from the parish capital (using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers). Source: see *Mean elevation*.

Mean caloric suitability [parish level]. Average pre-1500 land caloric suitability within a buffer of 10-km radius from the parish capital (using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers). Source: author's computation using the [Caloric Suitability Index](#) constructed by Galor and Özak (2016); 5 arc-minute raster data on potential crop yield given the set of available crops before 1500CE.

Variation in caloric suitability [parish level]. Standard deviation of pre-1500 land caloric suitability within a buffer of 10-km radius from the parish capital (using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers). Source: see *Mean caloric suitability*.

Mean caloric suitability for maize/potato [parish level]. Average pre-1500 land caloric suitability for maize/potato within a buffer of 10-km radius from the parish capital (using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers). Source: See *Mean caloric suitability*.

Ln distance to perennial river [parish level]. Natural log of the geodesic distance (km) from the parish capital to the closest perennial river. Source: author's computation using water area features from version 10.0 of the [Seamless Digital Chart of the World](#).

Ln expected tribute [parish level]. Natural log of the total tribute (*pesos ensayados*) in the 16th century. Source: Cook (1982) and Puente Brunke (1991). The information exists for 116 parishes; for the remaining parishes, it is imputed using the average of the province.

Ln distance to mita mine [parish level]. Natural log of the geodesic distance (km) from the parish

capital to the closest *mita* mine, considering both the Huancavelica and Potosí mines. Source: author's computation using data from Dell (2010).

Ecclesiastical jurisdiction [parish level]. Categorical variable indicating the colonial bishopric (Lima, Arequipa, Huamanga, Trujillo, or Cuzco). Source: Unanue (1797).

Administrative province [parish level]. Categorical variable indicating the colonial administrative province (*partido*). Source: Vollmer (1967).

Ln distance to native shrine, defensive site, canal, bridge [parish level]. Natural log of the geodesic distance (km) from the parish capital to the closest pre-colonial shrine/defensive site/canal/bridge. Source: author's computation using pre-colonial archaeological data; see *Dummy urbanization [ethnic level]*.

Ln distance to road [parish level]. Natural log of the geodesic distance (km) from the parish capital to the closest pre-colonial road. Source: author's computation using the Inca road map (Qhapaq Ñan) produced by SIGDA (Ministerio de Cultura, Perú, accessed March 2021).

Ln light intensity per capita [parish level]. Natural log of 1 plus average light intensity per capita. The average sum of light intensity values within the 10-km parish buffer is divided by total population within the same buffer (using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers). Source: average cloud free coverages of the **DMSP-OLS Nighttime Lights Time Series**, produced by the NOAA's National Geophysical Data Center, which provide 30 arc-second yearly raster data (Elvidge et al. 1997). The paper uses data from satellites F15 and F18 for the periods 2000-2003 and 2010-2013, respectively. The paper also uses nightlight 15 arc-second raster data from the **Visible Infrared Imaging Radiometer Suite (VIIRS)** for the year 2013, produced by the Earth Observation Group (EOG); see Elvidge et al. (2021). Finally, version 4.10 of the **Gridded Population of the World** (Center for International Earth Science Information Network–CIESIN) provides 30 arc-second raster data with population counts for the years 2000 and 2010. Population counts are developed through the uniform areal-weighting method using census data adjusted to match the United Nation's population counts at the country level.

Non-subsistence agriculture [parish level]. Dummy variable taking value 1 if the share of agricultural producers devoting most of the harvest to sell in local markets is above the median. Source: *Censo Nacional Agropecuario* (CENAGRO), 1994, 2012; National Institute of Statistics (INEI).

Firms per 100 inhabitants [parish level]. Average number of small- and medium-sized firms per 100 inhabitants (2010-2013). Source: National Tax Administration (SUNAT).

Access to public sanitation [parish level]. Share of occupied dwellings with access to the public sewer system (inside or outside the dwelling unit). Source: *Censo Nacional de Población y Vivienda*, 1993, 2017; National Institute of Statistics (INEI).

Access to public water [parish level]. Share of occupied dwellings with access to the public network of water supply (inside or outside the dwelling unit). Source: *Censo Nacional de Población y Vivienda*, 1993, 2017, National Institute of Statistics (INEI).

Municipal budget executed [parish level]. Share of the municipal budget executed (2002-2013). Source: Artiles et al. (2021). The underlying data come from the Ministry of Economy and Finance (MEF).

Identification with the state, ethnicity or race, and religion [individual level]. Dummy variables taking value 1 if the individual reports to identify more strongly with a certain group (separate variables for the state administrative unit (including the administrative region, province, or district), ethnicity or race, and religion). Source: *Encuesta Nacional de Hogares* (ENAHOG), 2004-2017; National Institute of Statistics (INEI).

Dummy voted in the 2006 presidential election [individual level]. Dummy variable taking value 1 if the individual reports to have voted in the 2006 presidential election. Source: *Encuesta Nacional de Hogares* (ENAHOG), 2007-2011; National Institute of Statistics (INEI).

Participation in voluntary associations [individual level]. Dummy variables taking value 1 if the individual reports to participate in a voluntary association (separate variables for neighborhood, professional, and labor associations). Source: *Encuesta Nacional de Hogares* (ENAHOG), 2004-2017; National Institute of Statistics (INEI).

Ln volunteers for military service [parish level]. Natural log of 1 plus the average number of volunteers for military service (data for years 2008, 2011, and 2014). Source: Ministry of National Defense (MINDEF).

Dummy volunteers for military service [parish level]. Dummy variable taking value 1 if there is at least one volunteer for military service (data for years 2008, 2011, and 2014). Source: Ministry of National Defense (MINDEF).

Dummy neighborhood association [parish level]. Dummy variables taking value 1 for the presence

of neighborhood associations during the periods 2002-2004 and 2012-2014. Source: *Registro Nacional de Municipalidades* (RENAMU), 2004, 2012, 2013, 2014; National Institute of Statistics (INEI).

Dummy agricultural retail market [parish level]. Dummy variable taking value 1 for the presence of agricultural retail markets (focused on fruits and vegetables) established in 1993 or earlier. Source: *Censo Nacional de Mercados de Abastos* (CENAMA), 2016; National Institute of Statistics (INEI).

Access to agricultural land [parish level]. Mean share of the population with access to agricultural land between 1706 and 1800. Source: author's computation using data from Macera (1972).

Tertiary-sector occupation [parish level]. Share of population employed in the tertiary sector. Source: author's computation using data from the 1876, 2007 and 2017 population and housing censuses; National Institute of Statistics (INEI). The outcomes for the period 2007-2017 refer to the average outcomes from the 2007 and 2017 population censuses.

Mean elevation [ethnic level]. Average elevation within the ethnic homeland. Source: author's computation using Rowe (1946)'s ethnic boundaries; see *Mean elevation*.

Mean caloric suitability [ethnic level]. Average pre-1500 land caloric suitability within the ethnic homeland. Source: author's computation using Rowe (1946)'s ethnic boundaries; see *Mean caloric suitability*.

Ln river density [ethnic level]. Natural log of total river length (*km*, only perennial rivers) divided by total land area (km^2). Source: author's computation using Rowe (1946)'s ethnic boundaries; see *Ln distance to perennial river*.

Ln land area [ethnic level]. Natural log of total land area (km^2) of the ethnic homeland. Source: author's computation using Rowe (1946)'s ethnic boundaries.

Ln population [ethnic level]. Natural log of approximate population by the time of the Spanish conquest. Source: author's computation using Rowe (1946)'s ethnic boundaries and tributary population data (1532–1575) from Cook (1982). All population centers with centroid within the ethnic homeland are considered.

Ln population density [ethnic level]. Natural log of total population divided by land area. Source: see *Ln population [ethnic level]*.

Dummy urbanization [ethnic level]. Dummy variable taking value 1 for the presence of pre-colonial towns or urban centers. Source: author's computation using Rowe (1946)'s ethnic boundaries and

information on pre-colonial archaeological sites in Ravines Sánchez (1985), Ramos Giraldo (2001), Isbell and Silverman (2002, 2008), and the inventory of pre-colonial archaeological sites (*Catastro de Monumentos Arqueológicos Prehispánicos*) developed by SIGDA (Ministerio de Cultura, Perú, accessed March 2021).

Dummy political complexity [ethnic level]. Dummy variable taking value 1 for the presence of pre-colonial administrative centers and monumental architecture—public buildings and communal spaces, including temples, palaces, and complex mound platforms. Source: see *Dummy urbanization [ethnic level]*.

Dummy elite residences [ethnic level]. Dummy variable taking value 1 for the presence of elite residences. Source: see *Dummy urbanization [ethnic level]*.

Dummies for different types of infrastructure [ethnic level]. Separate dummy variables taking value 1 for the presence of canal or bridges structures. Source: see *Dummy urbanization [ethnic level]*.

Ln road density [ethnic level]. Natural log of 1 plus total road length (*km*) divided by total land area (*km*²) of the ethnic homeland. Source: author's computation using Rowe (1946)'s ethnic boundaries and the Inca road map (Qhapaq Ñan) produced by SIGDA (Ministerio de Cultura, Perú, accessed March 2021).

D Supplemental Appendix - References

Artiles, Miriam, Lukas Kleine-Rueschkamp, and Gianmarco León-Ciliotta, “Accountability, Political Capture, and Selection into Politics: Evidence from Peruvian Municipalities,” *Review of Economics and Statistics*, 2021, 103 (2), 397–411.

Academia Mayor de la Lengua Quechua, *Diccionario Quechua – Español – Quechua, Qheswa – Español – Qheswa: Simi Taque*, 2005, Cusco, Perú: Gobierno Regional Cusco, 2nd ed.

Bertonio, Ludovico, *Vocabulario de la Lengua Aymara*, 2011, La Paz, Bolivia: Instituto de Lenguas y Literaturas Andinas–Amazónicas (ILLA–A). [1612].

Carpio, Miguel Angel and María Eugenia Guerrero, “Did the Colonial mita Cause a Population Collapse? What Current Surnames Reveal in Peru,” *The Journal of Economic History*, 2021, 81 (4):1015–1051.

CIESIN, *Gridded Population of the World, Version 4 (GPWv4): Population Count Adjusted to Match 2015 Revision of UN WPP Country Totals, Revision 10*, 2017, Center for International Earth Science Information Network (CIESIN), NASA Socioeconomic Data and Applications Center (SEDAC).

Clingingsmith, David, Asim Ijaz Khwaja, and Michael Kremer, “Estimating the Impact of the Hajj:

Religion and Tolerance in Islam's Global Gathering," *The Quarterly Journal of Economics*, 2009, 124 (3):1133–1170.

Colantonio, Sonia E, Gabriel W Lasker, Bernice A Kaplan, and Vicente Fuster, "Use of Surname Models in Human Population Biology: A Review of Recent Developments," *Human Biology*, 2003, 75 (6):785–807.

Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig, "Inference with Arbitrary Clustering," 2019, IZA Discussion Paper n. 12584.

CONADI, *Diccionario Trilingüe: Aymara–Castellano–Inglés*, 2011, Chile: Corporación Nacional de Desarrollo Indígena (CONADI), Dirección Regional Arica y Parinacota.

Cook, Noble David, "Population Data for Indian Peru: Sixteenth and Seventeenth Centuries," *The Hispanic American Historical Review*, 1982, 62 (1):73–120.

Dell, Melissa, "The Persistent Effects of Peru's Mining Mita," *Econometrica*, 2010, 78 (6):1863–1903.

De la Vega, Garcilaso, *Obras Completas del Inca Garcilaso de la Vega*, 1960, Vol. 133. Ediciones Atlas. [1609].

De Lucca, Manuel, *Diccionario Aymara—Castellano, Castellano—Aymara*, 1983, La Paz, Bolivia: Comisión de Alfabetización y Literatura en Aymara (CALA).

Elvidge, C. D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., and Davis, E. R, "Mapping City Lights with Nighttime Data from the DMSP Operational Linescan System," *Photogrammetric Engineering and Remote Sensing*, 1997, 63(6), 727–734.

Elvidge, C. D., Zhizhin, M., Ghosh, T., Hsu, F. C., and Taneja, J., "Annual Time Series of Global VIIRS Nighttime Lights Derived from Monthly Averages: 2012 to 2019," *Remote Sensing*, 2021, 13(5), 922.

FamilySearch, *Perú, Bautismos, 1556-1930*, Database, 2018, <http://FamilySearch.org>: December 2018. Index based upon data collected by the Genealogical Society of Utah, Salt Lake City.

FAO/IIASA/ISRIC/ISS-CAS/JRC, *Harmonized World Soil Database (Version 1.2)*, 2012, FAO, Rome, Italy and IIASA, Laxenburg, Austria.

FIDE. Religiosos Franciscanos Misioneros de los Colegios de Propaganda Fide del Perú, *Vocabulario Políglota Incaico: Comprende más de 12,000 Voces Castellanas y 100,000 de Keshua del Cuzco, Ayacucho, Junín, Ancash y Aymará*, 1998, Lima, Perú: Ministerio de Educación. [1905] Editor: Cerrón-Palomino, R.

Fischer, G., F. Nachtergaele, S. Prieler, H.T. van Velthuisen, L. Verelst, D. Wiber, *Global Agro-ecological Zones Assessment for Agriculture (GAEZ)*, 2008, IIASA, Laxenburg, Austria and FAO, Rome, Italy.

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

González Holguín, Diego, *Vocabulario de la Lengua General de todo el Perv llamada Lengua*

Quichua, o del Inca, 1952, Lima: Universidad Nacional Mayor de San Marcos. [1608].

INEI, *Censos Nacionales IX de Población IV de Vivienda, 11 de julio de 1993, Perú: Resultados Definitivos*, 1994, Instituto Nacional de Estadística e Informática (INEI), Perú: Lima.

———, *III Censo Nacional Agropecuario (CENAGRO), 1994, Perú: Resultados Definitivos*, 1995, Instituto Nacional de Estadística e Informática (INEI), Perú: Lima.

———, *Censos Nacionales XI de Población y VI de Vivienda, 21 de octubre del 2007, Perú: Resultados Definitivos*, 2008, Instituto Nacional de Estadística e Informática (INEI), Perú: Lima.

———, *IV Censo Nacional Agropecuario (CENAGRO), 15 de noviembre del 2012, Perú: Resultados Definitivos*, 2013, Instituto Nacional de Estadística e Informática (INEI), Perú: Lima.

———, *Censo Nacional de Mercados de Abastos (CENAMA) 2016, 2017*, Instituto Nacional de Estadística e Informática (INEI), Perú: Lima.

———, *Censos Nacionales XII de Población y VII de Vivienda, 22 de octubre del 2017, Perú: Resultados Definitivos*, 2018, Instituto Nacional de Estadística e Informática (INEI), Perú: Lima.

———, *Encuesta Nacional de Hogares, 2004-2017*, Instituto Nacional de Estadística e Informática (INEI), Perú: Lima.

———, *Registro Nacional de Municipalidades (RENAMU), 2004*, Instituto Nacional de Estadística e Informática (INEI), Perú: Lima.

———, *Registro Nacional de Municipalidades (RENAMU), 2012-2014*, Instituto Nacional de Estadística e Informática (INEI), Perú: Lima.

Isbell, William H and Helaine Silverman, *Andean Archaeology I*, 2002, Springer.

———, *The Handbook of South American Archaeology*, 2008, Springer.

Kling, Jeffrey R, Jeffrey B Liebman, Lawrence F Katz, and Lisa Sanbonmatsu, “Moving to Opportunity and Tranquility: Neighborhood Effects on Adult Economic Self-Sufficiency and Health from a Randomized Housing Voucher Experiment,” 2004, KSG Working Paper.

Lasker, Gabriel W, “Surnames in the Study of Human Biology,” *American Anthropologist*, 1980, 82 (3):525–538.

———, *Surnames and Genetic Structure*, 1985, Cambridge University Press.

Lowes, Sara and Eduardo Montero, “Concessions, Violence, and Indirect Rule: Evidence from the Congo Free State,” *The Quarterly Journal of Economics*, 2021, 136 (4):2047–2091.

Macera, Pablo, *Tierra y Población en el Perú, SS XVIII-XIX*, 1972, Lima.

Perú, Dirección de Estadística, *Censo General de la República del Perú Formado en 1876*, 1878, Lima.

Puente Brunke, José de la., “Un Documento de Interés en Torno al Tributo Indígena en el Siglo XVI,” *Historica*, 1991, 15 (2):265–313.

Pulgar Vidal, Javier, “Las Ocho Regiones Naturales del Perú,” 1941, In *Boletín del Museo de Historia Natural “Javier Prado”*, 17, Lima: Imprenta D. Miranda, 145–160.

Ramos Giraldo, Jesús, *Contribución para un Primer Inventario General de Sitios Arqueológicos del Perú*, 2001, Lima: Instituto Nacional de Cultura.

Ravines Sánchez, Rogger H, *Inventario de Monumentos Arqueológicos del Perú*, 1985, Lima: Instituto Nacional de Cultura.

RENIEC, *Introducción a un Tesoro de Nombres Quechuas en Apurímac*, 2012, Lima, Perú: Registro Nacional de Identificación y Estado Civil, jointly with Terra Nuova and Apurimac ONLUS.

Rowe, John Howland, “Inca Culture at the Time of the Spanish Conquest,” 1946, In *Handbook of South American Indians*, Vol. 2, US Government Printing Office, 183–330.

Scannell, Kevin P, “The Crúbadán Project: Corpus Building for Under-Resourced Languages,” 2007, In *Building and Exploring Web Corpora (WAC3-2007): Proceedings of the 3rd Web as Corpus Workshop, Incorporating CleanEval*. Vol. 4.

Shannon, Claude Elwood, “A Mathematical Theory of Communication,” *Bell System Technical Journal*, 1948, 27 (3):379–423.

SIGDA, *Catastro de Monumentos Arqueológicos Prehispánicos*, 2021, Sistema de Información Geográfica de Arqueología (SIGDA, Ministerio de Cultura, Perú), accessed March 2021.

Tapia, Mario E, “Diagnóstico de los Ecosistemas de Montañas en el Perú,” 2013, Tech. rep., FAO–MINAM (Ministerio del Ambiente, Perú).

Unanue, José Hipólito, *Guía Política, Eclesiástica y Militar del Virreinato del Perú para el Año de 1793*, 1793, Imprenta Real de los Huérfanos.

Vollmer, Günter, *Bevölkerungspolitik und Bevölkerungsstruktur im Vizekönigreich Peru zu Ende der Kolonialzeit, 1741-1821*, 1967, Berlin: Gehlen.

Wilson, Kurt M, Weston C McCool, Simon C Brewer, Nicole Zamora-Wilson, Percy J Schryver, Roxanne Lois F Lamson, Ashlyn M Huggard, Joan Brenner Coltrain, Daniel A Contreras, and Brian F Coddig, “Climate and Demography Drive 7000 Years of Dietary Change in the Central Andes,” *Scientific Reports*, 2022, 12 (1), 1–16.

Zuidema, Reiner Tom and Deborah Poole, “Los Límites de los Cuatro Suyus Incaicos en el Cuzco,” *Bulletin de l’Institut Français d’Études Andines*, 1982, 11 (1-2):83–89.