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

Miriam Artiles†

December 26, 2024

Abstract

Is ethnic diversity good or bad for economic development? Most studies find corrosive effects. This paper shows that historical exposure to economic exchange can mitigate these effects in the long run. I collect data from a natural experiment of Peru's colonial history: the forced resettlement of native populations in the 16th century. Where the resettlement concentrated ethnically diverse populations with a history of internal crop exchange, contemporary populations perform better systematically. Additional evidence suggests that prior experience with mutually beneficial crop exchange shaped more open attitudes toward out-group members. Economic complementarities helped sustain long-run, market-oriented cooperation and local trade.

JEL Codes: J15, N36, N76, O12, Z10

Keywords: ethnic diversity, minorities, economic exchange, trade, economic complementarities, cooperative behavior, culture, long-run economic development

*I am grateful to Gianmarco León-Ciliotta, Luigi Pascali, Marta Reynal-Querol, and Joachim Voth for their guidance and support throughout this project. I am also grateful to Leah Boustan, Ekaterina Zhuravskaya, and anonymous referees for their suggestions and detailed advice. I thank Alberto Bisin, Emilio Depetris-Chauvin, Miguel Ángel Carpio, Javier A. Birchenall, Martín Fernández-Sánchez, Felipe González, Jenny Guardado, Jonas Hjort, Stelios Michalopoulos, Omer Moav, Ömer Özak, Elias Papaioannou, Paul Schaudt, Felipe Valencia Caicedo, David Weil, Yanos Zylberberg, and seminar participants for their comments and suggestions. I thank Gonzalo González Melo, Gshan Irigoin, and Diana Tadeo Ruesta for providing excellent research assistance. An earlier version of this paper circulated as "Within-Group Heterogeneity in a Multi-Ethnic Society." Financial support from FONDECYT Iniciación 11230265 is gratefully acknowledged.

†Pontificia Universidad Católica de Chile, Instituto de Economía, E-mail: miriam.artiles@uc.cl

1 Introduction

The effect of ethnic diversity on economic growth and development is a question of long-standing interest in economics. Following the initial work by Easterly and Levine (1997) and Alesina and Glaeser (2004), a rich literature has examined the costs and benefits of ethnic diversity.[1] Most empirical studies find corrosive effects. When ethnic groups differ in their preferences for policies or public goods, conflicting preferences can lead to inefficiencies in public good provision or to policy choices that may not benefit the entire society (e.g., Alesina, Baqir and Easterly 1999; Miguel and Gugerty 2005). Inter-group tensions can also result in civil conflicts or exacerbate mistrust and lack of cooperation (e.g., Alesina and La Ferrara 2000; Fearon and Laitin 2003). On the other hand, some studies find that when ethnic groups differ in their specializations or skills, economic complementarities can sustain mutually beneficial coexistence (Jha 2013; Becker and Pascali 2019; Jedwab et al. 2019; Grosfeld et al. 2020). While there is a general understanding that diversity brings opportunities and challenges, there is scarce evidence on which factors determine its positive or negative consequences. When is ethnic diversity good for economic development, and when is it bad?

This paper provides systematic evidence that historical exposure to economic exchange facilitates mutually beneficial interactions between ethnic groups, mitigating the negative effect of ethnic diversity on economic development in the long run. I analyze new data from a natural experiment of Peru's colonial history—the forced resettlement of native populations in the 16th century. Before colonization, the economy was primarily agricultural, with coethnic individuals settling at different altitudes and exchanging complementary crops to maximize the economic base:[2]

> "In a territory so broken up by altitude ..., we should expect wide differences between ecological or production zones ... Access to the productivity of contrasting zones becomes indispensable. This could have been achieved by maintaining a series of markets at different altitudes, run by the ethnic groups inhabiting each separate ecological niche. However, this was not the Andean solution. They opted for the simultaneous access of a given ethnic group to the productivity of many microclimates." (Murra 1995, p. 60-61)

The colonial resettlement created delimited, small-scale, jurisdictions (called parishes). Unintentionally on the part of the Spanish colonizers, some of these parishes forced together

[1] See Alesina and La Ferrara (2005) for a survey of the initial literature.

[2] Throughout the paper, I use the term "ethnic group" to refer to the societies that coexisted in the Andean highlands of Peru before the Spanish conquest. I refer to the issue of ethnic identity in Section 2.

various ethnic groups. At the same time, the resettlement disrupted the internal settlement pattern, creating variation in past exposure to internal crop exchange (a pre-treatment variable) across parishes. The colonial intervention accidentally resulted in a setting where historical exposure to internal exchange was arguably orthogonal to ethnic diversity. This unique scenario provides an opportunity to study the consequences of ethnic diversity across parishes with varying levels of past exposure to economic exchange.

Did exposure to economic exchange facilitate the functioning of multi-ethnic societies? The answer is not obvious. Throughout much of our evolutionary history, trust has generally been stronger among coethnics (Diamond 1997). However, prior experience with mutually beneficial exchange may shape more open attitudes toward out-group members (Enke 2023). Furthermore, recent research shows that the benefits of ethnic diversity tend to flourish at the local level (Montalvo and Reynal-Querol 2021). The literature has emphasized the positive role of local-level interactions (Desmet, Gomes and Ortuño-Ortín 2020) and complementarities (Jha 2013, 2018) between ethnic groups. The colonial resettlement concentrated native groups into small-scale jurisdictions. If ethnic groups with a history of internal crop exchange were more willing to engage with others after the resettlement, local inter-ethnic interactions and mutually beneficial exchange may have become more frequent.

The first result of the paper documents the direct effect of ethnic diversity in the long run. Guided by the historical narrative, I geolocated parishes that were accidentally created close to spatial boundaries between ethnic groups. This narrative assumes that colonial officials were not fully aware of the vertical distribution of coethnics across space (Murra 1975), making it unlikely for them to systematically consider ethnic boundaries when determining parish locations. As a result, geographic proximity to ethnic boundaries created quasi-random variation in ethnic diversity across parishes (Wachtel 1976; Pease 1989). The research design follows the historical narrative, assuming that the distance from parishes to pre-colonial ethnic boundaries was *as good as* random, conditional on baseline characteristics. I provide empirical support for this assumption and validate ethnic diversity using surnames from colonial baptism records available for a subset of parishes. The results show a robust pattern when examining contemporary living standards. On average, parishes created close to ethnic boundaries, whose initial populations were more likely to be ethnically diverse, tend to exhibit lower living standards in the long run, compared to parishes with an ethnically homogeneous founding population. The results highlight the persistent consequences of ethnic diversity—created during the formation of colonial jurisdictions—for local-level living standards today, as proxied by measures of local economic activity and access to public facilities.

I then explore whether the average effect of ethnic diversity depends on pre-colonial exposure to crop exchange. The main result of the paper is that the negative effect of ethnic diversity on long-run development is mitigated by historical exposure to internal crop exchange. To construct a parish-level proxy for average exposure to internal crop exchange, I took the following steps. First, I computed an ethnic-level measure of potential gains from internal crop exchange, based on the spatial distribution of complementary elevation zones within the pre-colonial homelands of ethnic groups, similar to the approach in Fenske (2014). Second, I validated this measure using data from paleodietary reconstructions available for a subset of the groups. Specifically, I show that the measure of potential gains helps explain pre-colonial carbon-enriched diets where such diets were unlikely in the absence of internal crop exchange. Finally, I averaged this ethnic-level measure across the groups concentrated in each parish to obtain a parish-level proxy for average exposure to crop exchange. The estimates show a negative coefficient on ethnic diversity and a positive coefficient on its interaction with average exposure to internal crop exchange. At the 10th percentile of average exposure, contemporary living standards are 0.3 standard deviations lower in parishes with ethnic diversity, compared to parishes with an ethnically homogeneous founding population; at the sample median of average exposure, the negative effect of ethnic diversity decreases to 0.1 standard deviations.

The estimated interaction effect persists when controlling for initial prosperity and geography. It also remains significant when considering administrative and ecclesiastical provinces, as well as proximity to the colonial mining *mita* (Dell 2010). Permutation-based p-values from randomization inference in the presence of heterogeneous treatment effects align with baseline results. Using a matching procedure to construct a sample of parishes with varying levels of exposure to internal exchange while ensuring similarity in other ethnic characteristics, such as pre-colonial political complexity and population density, shows consistent results.

I examine historical and modern data to explore potential mechanisms. The results are consistent with the idea that ethnic groups with a historical practice of internal crop exchange helped sustain a market-oriented agricultural economy in the long run. Specifically, where the resettlement concentrated ethnically diverse populations with a history of internal crop exchange, the results show a shift toward tertiary-sector activities, mainly driven by local trade, in a predominantly agricultural setting. I explore the potential drivers of this shift. First, the evidence is consistent with the hypothesis that prior experience with mutually beneficial crop exchange shaped more open attitudes toward out-group members, contributing to the likelihood of inter-group interaction after resettlement and fostering a trend toward a more integrated society. The analysis of colonial inter-group unions—commonly used as a proxy for societal integration—supports this interpretation. Contemporary data on self-reported identities and volunteering for military service are consistent with a more integrated, nation-oriented society. The data show a more integrated society, with a long-run tendency for engaging in associational activities, non-subsistence agriculture, and local trade, where the resettlement forced together ethnically diverse populations historically exposed to crop exchange. A broader analysis of U.S. immigrant descendants—individuals born in the U.S. with immigrant parents from various countries worldwide—provides consistent results. Second, contemporary populations tend to perform better where the historical ethnic minority complemented the majority group in terms of prior exposure to crop exchange. The results also support the hypothesis that these ethnic minorities specialized in intermediary services related to agriculture (e.g., retailing), reducing inter-group competition and contributing to long-run coexistence.

This paper contributes to the body of research on the economic factors that promote inter-ethnic coexistence. The results are in line with the theoretical framework developed by Jha (2013, 2018), which establishes that peaceful inter-ethnic coexistence can be sustained through the specialization of ethnic groups into complementary activities that are both costly to replicate and expropriate. Jha (2013) provides consistent empirical evidence on tolerance toward Muslims in South Asian medieval towns. Other studies have focused on anti-Semitism, with Jedwab, Johnson and Koyama (2019) presenting evidence from the Black Death in Western Europe and Becker and Pascali (2019) from the Protestant Reformation in Germany. In line with the previous results, Grosfeld, Sakalli and Zhuravskaya (2020) show that during times of political stability in the Russian Empire, anti-Jewish pogroms were not prevalent in places where Jews specialized in intermediary occupations related to agriculture, providing insurance against economic shocks. This paper examines a natural experiment of history to systematically study the long-run consequences of ethnic diversity. The analysis uses data from both ethnically homogeneous and diverse jurisdictions within a context with varying levels of historical exposure to economic exchange. I provide micro-level evidence that past exposure to economic exchange facilitates mutually beneficial inter-ethnic interactions, mitigating the negative effect of ethnic diversity on economic development in the long run.

The consequences of ethnic diversity for economic growth and development have inspired a rich literature in economics and political science. The initial studies tended to emphasize its negative effects at various levels of analysis (Alesina and La Ferrara 2005). Across countries and US localities, ethnic diversity has been associated with lower levels of economic growth, public good provision, and quality of government, as well as with greater political instability and civil conflict.[3] Using micro-level data, Miguel and Gugerty (2005) show that ethnic diversity is associated with lower public good provision in Kenya. Hjort (2014) focuses on the private sector, providing causal evidence for the effect of ethnic diversity on team productivity at a flower plant in Kenya. The results show that teams of ethnically diverse workers are, on average, less productive than homogeneous teams. Additional evidence points toward a taste for discrimination against coworkers of different ethnic origin.

More recent papers have focused on the role of local-level interactions, highlighting the importance of the level of analysis and the positive effects of local exposure to ethnic diversity. Specifically, Desmet, Gomes and Ortuño-Ortín (2020) emphasize that local inter-ethnic interactions can mitigate the negative effects of ethnic diversity in larger jurisdictions. They show that local inter-ethnic interactions help mitigate the adverse effects of ethnic diversity on the provision of public goods at the national level. Montalvo and Reynal-Querol (2021) focus on the size of the unit of analysis, documenting a positive, local-level, relationship between ethnic diversity and economic growth across grid cells of fixed size. The authors find that a potential explanation for this result in the context of Africa is the increase in trade close to spatial boundaries between ethnic groups, suggesting ethnic specialization into complementary activities and inter-ethnic contact. In this paper, I document that prior experience with internal economic exchange facilitates local-level interactions between ethnic groups, mitigating the local costs of ethnic diversity in the long run.

The evidence on potential mechanisms also contributes to the literature on the long-run effects of cultural traits (e.g., Nunn and Wantchekon 2011, Voigtländer and Voth 2012; Alesina,

[3]See Easterly and Levine (1997), Alesina, Baqir and Easterly (1999), La Porta et al. (1999), Alesina and La Ferrara (2000), and Fearon and Laitin (2003), among others.

Giuliano and Nunn 2013; Buggle and Durante 2021). First, the results are consistent with Enke (2023)'s finding that internal exposure to markets and economic exchange contributed to prosocial behavior, including cooperative behavior, during pre-industrial times. Contemporary data on neighborhood associations, used as a proxy for a culture of cooperation (e.g., Guiso, Sapienza and Zingales 2016), show consistent results. Second, the main results hold beyond the effect of local crop diversity, suggesting that it is not the availability of crops but rather the transmission of more open attitudes toward out-group members that facilitated agriculture-related market interactions in the study setting. This finding relates to a broader literature on the social norms that govern market interactions (Tabellini 2008; Guiso, Sapienza and Zingales 2009; Rohner, Thoenig and Zilibotti 2013).

Following Gennaioli and Rainer (2007) and Michalopoulos and Papaioannou (2013), pre-colonial ethnic traits have received increasing attention in the historical persistence literature (see Nunn 2020). While much of this literature has focused on the case of Africa, this paper provides evidence from pre-colonial ethnic groups in Latin America (see also Angeles and Elizalde 2017). I focus on the internal economic organization of the group, highlighting that exposure to internal economic exchange can help us understand potential gains from the interaction between groups.[4] The subsistence strategy of pre-colonial Andean groups, characterized by a single ethnic group managing various altitude-specific resources, contributed to shaping the long-run effects of ethnic diversity in the study setting. The Maya are also known for their investments in internal vertical organization, particularly in terraces (Rubio-Ramos, Isendahl and Olsson 2024). Similar subsistence strategies have been observed in other cultures, such as those of Bali and Polynesia.[5] The results suggest that further studying the internal economic organization of the groups can help us better understand comparative development.

Finally, this study contributes to the literature on the long-run effects of colonial interventions (see, e.g., Acemoglu, Johnson and Robinson 2001, Banerjee and Iyer 2005, Huillery 2009,

[4]Fenske (2014) has also examined internal exchange, providing evidence that it contributed to the emergence of states in pre-colonial Africa. Depetris-Chauvin and Özak (2020) show that population heterogeneity, as proxied by linguistic and genetic diversity, contributed to the emergence of within-group economic specialization in pre-colonial times; see Ashraf and Galor (2013) for evidence on the hump-shaped effect of within-group population heterogeneity on long-run development. Dippel (2014) has explored the internal political organization of the group instead. The author shows that when coethnic individuals form autonomous subpolities (i.e., they do not have a history of shared or centralized governance even though they are ethnically homogeneous), their forced coexistence in a centralized institutional system can negatively impact long-run development.

[5]I am grateful to an anonymous referee for this reference.

Nunn 2008, Dell 2010, Bruhn and Gallego 2012, Guardado 2018, Valencia Caicedo 2019, Montalvo and Reynal-Querol 2020). I add to the literature by exploring the role of ethnic identity in Latin American colonization, a topic that has been overlooked in previous studies. The results also provide insight into the lasting impact of the forced reorganization of indigenous populations as a result of colonization, a research topic with limited empirical evidence (see Valencia Caicedo (2019) for evidence from the Jesuit missions in South America).[6] Section 2 describes the historical context, Section 3 presents the empirical strategy, Section 4 explains the main results, Section 5 discusses the mechanisms, and Section 6 concludes.

2 Historical Background

Pre-colonial ethnic groups. By the time the Spanish conquerors arrived, the Andean civilization was composed of several groups that had been incorporated into the Inca empire (1438-1532), including the *Chocorvos*, *Lucanas*, *Soras*, and *Chankas* (Tello 1939; Rowe 1946; Dulanto 2008). The 47 groups in my study region coexisted approximately from the disintegration of the *Wari* civilization (ca. 1000) until the Spanish conquest (ca. 1532).

How ethnically distinct were these groups from one another? The term "ethnic group" is introduced by Murra (1975) to refer to the groups in my study region.[7] The prevailing view is that the society comprised various groups with diverse linguistic roots (e.g., Rowe 1946; Murra 1975; Isbell 2010) and differentiated material styles (e.g., architectural styles; see Stanish 1989) following the collapse of *Wari*. For some groups, there is anecdotal evidence that their languages coexisted with the official language of the Incas (*Quechua*) during a period of indirect rule (Garcilaso de la Vega (1960)[1609]). Group identity seems to have been reinforced by the absence of inter-group marriage. Specifically, the social unit is generally described as an endogamous group of several extended families with descent traced through

[6] See Becker (2022) for a summary of the literature on the long-run effects of forced displacements.

[7] See also Rostworowski (1990) on the term "macroetnia." The issue of ethnic identity has been vaguely discussed in the literature, possibly due to the European perception of the region as culturally homogeneous despite the existence of ethnolinguistic differences at the time of Spanish contact. Charney (1998) argues that the use of the Spanish term "Indio" to collectively refer to all native peoples in official documents played a role in obscuring ethnic distinctions in the eyes of Europeans—"Indio" was not merely a label but the imposed *new ethnicity* for native individuals in the colonial legal system. Stanish (2001) points toward the interest of Inca and Spanish powers in promoting cultural unity via state propaganda.

the male line (Rowe 1946). The group usually claimed descent from a mythical ancestor, such as an animal or element of nature. This mythical kin was worshipped and sometimes honored with rites and sacrifices; see Garcilaso de la Vega (1960)[1609].[8]

Vertical settlement pattern and crop exchange. In the human ecology literature, the mountain environment of the Andean highlands is described as a vertical resource system (Brush 1976). Differences in elevation give rise to various microclimates within short distances, and each microclimate is, in turn, suited to a different assortment of natural resources and crops. Following the pioneering ethnohistoric work of Murra (1975), studies across various disciplines documented a pre-colonial subsistence strategy characterized by individuals from the same ethnic group simultaneously controlling different altitudes.[9] Murra's model is often referred to as a zonal complementarity model (e.g., Stanish 1989; Aldenderfer 1993; Isbell and Silverman 2002b). The group tried to maximize its economic base by establishing permanent settlements in vertically arranged zones (Murra 1975, 1995, 2002a,b).[10] Since certain crops can only be grown at specific altitudes, these zones are interpreted as complements. By exchanging crops between populations settled in different zones, the group increased access to resources, thus maximizing total output at the group level.

According to the research of Pulgar Vidal (1941), a Peruvian geographer who integrated local knowledge of geography and native folklore to offer a detailed account of the mountain environment, the study region consists of five complementary elevation zones: Yunga (*warm valley*, 500–2,300 m), Quechua (*temperate land*, 2,300–3,500 m), Suni or Jalca (*high land*, 3,500–4,000 m), Puna (*cold land*, 4,000–4,800 m), and Janca (*white land*, 4,800–6,768 m), where figures in parentheses refer to elevation in meters above sea level. Each zone is traditionally known for specific crops. For example, the natural limit of maize cultivation is the Quechua zone, while grains such as quinua and kañiwa, as well as lupins such as tarwi, are best grown in the Jalca zone. Various potato varieties, which can provide more carbohydrates per hectare than maize at high altitudes, grow exceptionally well in the Puna zone (Burger and

[8]For example, the *Chankas* believed that they were closely connected to the Andean lion (*puma*). During festivities, they usually dressed in *puma* skins and adopted *puma* imagery. Bauer et al. (2010) provide anecdotal evidence of public support for this identity.

[9]See Brush (1976), Pease (1989), Stanish (1989), and Aldenderfer (1993), among others, for perspectives from human ecology, history, anthropology, and archaeology, respectively.

[10]These settelements are different from those established for the *mitmaqkuna* under Inca rule (Rowe 1946).

Merwe 1990; Sandweiss and Richardson 2008). This subsistence strategy has been particularly well-documented for groups of the central and southern highlands.[11]. This paper focuses on internal crop exchange as a subsistence strategy in highland Peru.

Continuity after the Inca expansion. According to Murra's research, the subsistence strategy based on vertical economic exchange existed during pre-Inca times (Murra 1956, 1975). The Inca expansion (1438–1525) was achieved through the gradual conquest of pre-existing groups. The dominant view is that this led to a dynamic process of state formation, where distinct regions or provinces were sequentially established based on ethnic identity (Rowe 1946). Ethnohistoric research suggests that the Inca government was indirect, with each region governed by local leaders of the corresponding ethnic group (Murra 1975, 2002b). This is a key characteristic of Inca rule that suggests the preservation of ethnic traits during this period.[12] The groups maintained "a jurisdictional base that defined an area as their homeland" (Mayer 2002, p.50), and ethnic rulers were pushed to maintain control over their respective vertical zones in order to sustain the empire (Murra 1956, 1975). Rowe (1946) mapped the approximate extent of the groups' homelands at the time of the Spanish conquest based on archaeological evidence and early ethnohistoric accounts. The map was published in the second volume of the *Smithsonian Handbook of South American Indians* (1948)—a georeferenced version is presented in Figure 1.

The colonial resettlement. The contemporary administrative division of Peru has its origins in the colonial period. When Spanish Viceroy *Francisco Toledo* arrived in Peru in 1569, native populations still lived scattered along mountain slopes.[13] This settlement pattern was perceived as an "obstacle" to both tribute collection and religious indoctrination. As Spanish official *J. Matienzo* noted, "the *indios*, for being isolated in *huaycos* and ravines, do not live in proper order, and this is the main obstacle to their indoctrination."[14] Between 1570 and 1575, the

[11]For evidence on genetic homogeneity, see, e.g., Nakatsuka et al. (2020).

[12]One example is the festivity of the *Chankas* in honor of their mythical connection to the *puma*, which, according to early chronicles, was still celebrated during the Inca period (Garcilaso de la Vega (1960)[1609]).

[13]At that time, most native populations were subject to the *encomienda* system, which granted Spanish conquerors and settlers a share of the product of native labor (Dougnac Rodríguez 1994). Historical literature suggests that this system did not disrupt pre-colonial ethnic divisions, as the *encomienda* was based on population rather than territory (Pease 1992; Murra 2002b).

[14]Translated from the Spanish version in Medina (1974a, p. 155).

colonial administration ordered the forced reorganization of native populations into residential jurisdictions (*reducciones*). In turn, several jurisdictions were assigned to a single *doctrina*, a parish served by either the regular or secular clergy (Medina 1974a,b, 1993).

The new model limited population movement, concentrating the populations already living around specific locations into small-scale, continuous, and delimited jurisdictions (Jiménez de la Espada 1881). This model, based on a horizontal conception of the world, pointed against the exchange of resources between different elevation zones, thus creating a new paradigm for native populations (Pease 1989). It is important to note that variation in the extent to which the forced concentration of populations was implemented could potentially affect long-run development. Historical studies have noted that, in practice, movement limitations were generally more effective at the parish level (Saignes 1991). This system was maintained throughout the entire colonial period, and at the time of independence from Spain, parishes were renamed as districts, forming the basis for the current third-level administrative division of the country.[15] Following the historical literature, I use the parish as the unit of analysis in the empirical design. Section 3.1 describes official recommendations for desirable locations.

3 Empirical Strategy

Sample. I collected information on the parishes involved in the natural experiment. These parishes satisfy two conditions: (i) they were created as a result of the colonial resettlement of the 16th century in highland Peru (i.e., they were located more than 500 meters above sea level), and (ii) they remained within the Viceroyalty of Peru for the entire colonial period.[16] The analysis excludes the two capital parishes of Cuzco and Arequipa, as well as six parishes now within Chilean territory. The resulting 336 parishes represent approximately 24 percent of the current districts in Peru, covering around 12 percent of the country's land area and 10 percent of its population, according to the 2017 population census.

[15] Throughout the text, I use "parish" and "district" interchangeably.

[16] The census conducted from 1791 to 1795 under the administration of *Gil de Taboada y Lemos* lists all parishes in this territory. The document was signed by *José Ignacio de Lequanda* and dated January 10, 1796; see Vollmer (1967). Using secondary historical sources, I checked for priests in charge of religious indoctrination; see, e.g., Lissón Chávez (1943) and de Armas Medina (1953).

3.1 Average Effect of Ethnic Diversity

The historical narrative suggests that geographic proximity to ethnic boundaries created quasi-random variation in ethnic diversity across parishes. This narrative is based on the assumption that colonial officials were not fully aware of the vertical distribution of coethnic individuals across space (Murra 1975), making it unlikely for them to systematically consider ethnic boundaries when determining parish locations. As a result, given the vertical settlement pattern, parishes created close to ethnic boundaries unintentionally concentrated populations from different ethnic origins (Wachtel 1976; Pease 1989, 1992). According to the historical literature, the tension between the vertical settlement pattern—a native response to geography—and the Spanish notion of jurisdiction—based on a horizontal conception of the world—potentially affected the ethnic composition of colonial parishes in the 16th century.

The first specification intends to explore whether the ethnic composition of colonial parishes influenced long-run comparative development:

$$y_p = \beta_0 + \beta_1 Ethnic\ div_p + X_p^{'}\gamma + \upsilon_p \tag{1}$$

where y_p is a contemporary development outcome for parish p, $Ethnic\ div_p$ is a dummy variable indicating whether parish p was accidentally created close to an ethnic boundary, X_p is a vector of parish-level control variables, and υ_p is an error term. This section provides empirical support for the historical narrative and validates the measure of ethnic diversity using surnames from colonial baptism records available for a subset of parishes.

Measuring ethnic diversity. In the absence of a map delineating parish borders, I follow the historical narrative and define a buffer of 10-km radius from the parish capital (see Figure A.1).[17] I then define $Ethnic\ div_p$ as a dummy variable for the presence of an ethnic boundary within the 10-km buffer. Parishes with an ethnic boundary within the buffer (*border* parishes)

[17]Colonial accounts describe geographic distances typically ranging between two and three *leguas* (approximately 10 km; see Jiménez de la Espada 1881). For the correspondence between *leguas* and kilometers during the 16th century, see Paz Soldán (1877). When the distance between the capitals of two parishes is less than 10 km, I use equidistant boundaries to ensure non-overlapping buffers. The resulting buffers have a mean and median area of 240.44 km^2 and 256.51 km^2, respectively. An ethnic group is considered as part of the buffer only if its homeland occupies at least one percent of the buffer's area, ensuring that the ethnic group has at least one grid cell of 1×1 km inside the buffer.

are labeled as parishes with ethnic diversity (35 percent of the sample; displayed in yellow in panel (a) of Figure 1), while those located further inside ethnic homelands (*interior* parishes; displayed in blue) are labeled as ethnically homogeneous parishes. I use Rowe (1946)'s mapping of pre-colonial ethnic boundaries for this exercise and validate that the dummy for *border* parishes captures ethnic diversity using surname data from colonial baptism records.[18] I consider a range of different radii and placebo ethnic boundaries in supplementary analyses.

Balance tests. I present empirical evidence supporting the historical narrative that colonial officials did not systematically consider ethnic boundaries when determining parish locations. Specifically, I show that factors that could have influenced locations and affected post-resettlement development did not vary significantly with proximity to ethnic boundaries.

Spanish officials may have followed recommendations or avoided locations where they suspected it would be easier for native populations to escape (e.g., plains or lower elevations). The colonial regulation of 1569–1570 described three desirable characteristics (Jiménez de la Espada 1881): (*i*) land quality and abundance, as enough land was needed for native families to work following their own rules of crop rotation; (*ii*) proximity to surface water, which was a key advantage for irrigation and sustaining populations that depended on subsistence agriculture; and (*iii*) distance to *huacas*, sacred native shrines that honored nature, to facilitate religious indoctrination. The extent to which Spanish officials applied these recommendations is unclear (Pease 1989). Nonetheless, the results in Table 1 show that, on average, *border* and *interior* parishes are statistically similar in the highlighted characteristics.

I start by exploring the mean and standard deviation of elevation, computed based on all 1 $\times$ 1 km grid cells within the 10-km buffer. There are no significant differences between *border* and *interior* parishes. Similarly, I find no significant differences in the mean and standard deviation of pre-colonial land caloric suitability (Galor and Özak 2016), alleviating the concern that colonial officials selected locations differently (i.e., in a way that resulted in systematic differences in proximity to ethnic boundaries) at different elevations or in plains as opposed to more rugged terrain. Log distance to the nearest perennial river from the parish capital is also balanced. I collected data on pre-colonial shrines to explore the third recommendation. On average, there are no significant differences between *border* and *interior* parishes in log

[18] Surname data have become increasingly used in economics and political science (e.g., Cruz et al. 2017).

distance to the nearest native shrine. The table also shows balance in log distance to *mita* mines (Dell 2010), considering both the Huancavelica and Potosí mines, and local prosperity at the time of the policy, as proxied by the value of expected tribute.[19] See Appendix D for data sources and definitions.

Conditional random assignment. The research design assumes that the distance from parish capitals to pre-colonial ethnic boundaries at the time of the resettlement was *as good as* random, conditional on baseline characteristics of the land and native populations. The vector of baseline controls (X_p) includes all the variables listed in Table 1 (mean and standard deviation of elevation, mean and standard deviation of land caloric suitability, log distance to perennial rivers, log distance to pre-colonial native shrines, log 16th-century expected tribute, log distance to *mita* mines), as well as longitude and latitude of the parish capital. The coefficient β_1 captures the average difference in contemporary development between *border* parishes, whose initial populations were more likely to be ethnically diverse, and *interior* parishes, which, located further inside ethnic homelands, concentrated ethnically homogeneous populations, conditional on baseline characteristics (Equation 1).

Validating ethnic diversity. The ethnicity of the individuals concentrated in each parish was not systematically registered. Appendix B provides supporting evidence that the dummy variable for *border* parishes captures diversity in ethnicity at the parish level. For this, I constructed a dictionary of native linguistic roots and compiled a dataset of 112,340 individuals with native paternal surname using colonial baptism records from digital genealogical sources. The dataset includes 65 parishes, of which 20 percent are *border* parishes, covering the period 1605–1780. I explore whether there was significantly higher surname diversity in *border*

[19] The colonial requirement to send native populations to *mita* mines started in 1573. Expected tribute was based on an official assessment of the number of native individuals at the time of the policy (Cook 1982). Table A.1 provides evidence on statistical balance for other pre-colonial characteristics. To proxy for the threat of native attack at the time of the policy, I georeferenced data on pre-colonial defensive sites (e.g., fortresses, walled sites, and *pukaras*). The table shows balance in log distance to these sites, as well as in log distance to pre-colonial socio-economic and institutional centers (urban sites, elite residences, and political sites). I also explore log distance to pre-colonial infrastructures (Inca roads, canals, and bridges), as colonial officials could have been interested in exploiting them. In each case, I compute the distance to the nearest site or type of infrastructure from the parish capital. Finally, colonial officials could also have been interested in specific crops, such as maize or potatoes (Brush 1976). The mean land caloric suitability for these crops is also balanced. Appendix D reports data sources and definitions.

parishes compared to those located in the interior of ethnic homelands. Two contextual features are worth noting: (*i*) pre-colonial endogamy traced through the male line, and (*ii*) the absence of a family name system, with only first names related to mythical ancestors, before the Spanish conquest. The Catholic Church introduced the Hispanic family name system for the purpose of religious indoctrination. Qualitative evidence suggests that early priests commonly selected a Spanish first name, imposing the mythical first names of the parents as surnames (see Carpio and Guerrero 2021). Table B.1 shows that *border* parishes tend to exhibit higher surname diversity (between 0.41 and 0.56 standard deviations, on average) compared to *interior* parishes (panels A-C); no such pattern is observed among individuals with non-native surnames (panel D). Throughout the rest of the paper, I refer to *border* parishes as ethnically diverse parishes and to *interior* parishes as parishes built on ethnically homogeneous populations.

3.2 Heterogeneous Effects of Ethnic Diversity

The main specification of interest explores heterogeneous effects ($\beta_3 \neq 0$):

$$\begin{aligned} y_p = \beta_0 + \beta_1 Ethnic\ div_p + \beta_2 \overline{Crop\ exchange}_p \\ + \beta_3 \left(Ethnic\ div_p \times \overline{Crop\ exchange}_p \right) + X_p^{'} \gamma + \epsilon_p \end{aligned} \tag{2}$$

where $\overline{Crop\ exchange}_p$ is a proxy for average exposure to internal crop exchange during pre-colonial times. In particular, I consider a weighted average among the ethnic groups concentrated in each parish: $\overline{Crop\ exchange}_p = \sum_e w_{ep} H_e$, where w_{ep} is the area share of ethnic group e within the buffer of parish p, and H_e is an ethnic-level measure of potential gains from internal crop exchange (a pre-resettlement variable). In the following paragraph, I describe the measure of potential gains from internal crop exchange. I then validate this measure as a proxy for historical exposure to internal crop exchange using data from paleodietary reconstructions (Table A.3). Figure A.2 characterizes ethnic weights (w_{ep}), which range from 0.013 to 0.986 among parishes with ethnic diversity, with similar mean (0.467) and median (0.452) values. Most parishes with ethnic diversity (85 percent) concentrated two ethnic groups, while the remaining 15 percent of parishes concentrated either three or four groups.

Measuring potential gains from crop exchange. Fenske (2014) uses an ethnic-level measure of ecological diversity to proxy for potential gains from internal exchange in pre-colonial times (see also Depetris-Chauvin and Özak 2020 and Enke 2023). I follow a similar approach. In my setting, pre-colonial subsistence relied on crop exchange between coethnics settled in complementary elevation zones (Murra 1975). I use the classification of complementary zones proposed by Pulgar Vidal (1941), excluding the Janca zone from the analysis (3.43 percent of the total territory in the study region) because it cannot be permanently inhabited due to oxygen constraints (e.g., Sandweiss and Richardson 2008).

I constructed spatial data on the distribution of these zones to compute a measure of potential gains from internal crop exchange: $H_e = 1/\sum_j s_{je}^2$, where s_{je} is the area share of elevation zone j within the homeland of ethnic group e.[20] Panel (b) of Figure 1 shows the spatial distribution of the zones using 1×1 km elevation data from FAO's Harmonized World Soil Database (Fischer et al. 2008).[21] The index is normalized to one for the group with the highest value, increasing as the composition of complementary zones becomes more diverse. Figure A.3 shows the density of H_e in the data. Approximately 23 percent of the groups have an index value below 0.5, while the index for the remaining 77 percent ranges from 0.5 to 1, with similar mean and median values (0.661 and 0.682, respectively). Appendix A.1 (Table A.2) provides correlational evidence that H_e helps explain crop variety in the data. On average, a one standard deviation increase in H_e is associated with a 0.5 standard deviation increase in log crop variety (considering native crops only).

Validating crop exchange: evidence from pre-colonial diets. I explore pre-colonial diets and provide suggestive evidence that H_e helps explain carbon-enriched diets where such diets were unlikely in the absence of internal crop exchange. Biochemical analyses of archaeological human remains can inform the role that carbon-enriched crops play in individuals' diets. In particular, stable isotope measures of carbon in bone and dentin collagen ($\delta^{13}C_{col}$) can provide information on the presence of certain plants in the protein component of the diet (Ambrose 1993; Ambrose and Norr 1993). Plants characterized by using the C_4 photosynthetic pathway for carbon fixation have particularly high $\delta^{13}C_{col}$ values compared to those using the C_3

[20] The reciprocal of the Herfindahl index is a common index to measure ecological diversity (Magurran 2004).

[21] Each grid cell of 1×1 km at the equator is assigned to an elevation zone based on its median elevation.

pathway. Maize, sorghum, and millet are well-known C_4 plants, while most plants, including tubers, use the C_3 pathway.

Wilson et al. (2022) provide comparable stable isotope values for 196 archaeological individuals in my study region, distributed across eight ethnic groups. Five groups are observed in the Quechua zone, the upper limit of maize cultivation, while the remaining three groups are observed in the Jalca zone.[22] The stable isotope values from bone collagen found in the Jalca zone of the *Soras* suggest carbon-enriched diets. While these isotope values might not be representative of the *Soras* population, it is noteworthy that they were found in a zone where growing maize, the main C_4 staple crop in the study region, was challenging due to geoclimatic conditions.[23] Tung and Knudson (2018) find similar evidence in southern Peru—the authors document carbon-enriched diets likely resulting from maize consumption in a non-suitable location, suggesting crop exchange between elevation zones.

In Table A.3, I regress individual-level $\delta^{13}C_{col}$ scores on H_e. The positive correlation in Column 1 suggests that individuals from ethnic groups with a higher potential for internal crop exchange tended to have more carbon-enriched diets, at least in their protein component. Column 2 includes zone fixed effects, thus comparing individuals settled in the same zone but from ethnic groups with different values of H_e. The positive correlation is also significant after controlling for differences in crop variety (Column 3) and average caloric suitability (Column 4) across ethnic groups' homelands. Column 5 shows consistent results when excluding children from the sample. Although the data do not allow direct testing of internal exchange, Column 6 provides evidence that the ethnic boundary matters in explaining carbon-enriched diets. Specifically, I replicate the analysis using 50 × 50 km grid cells instead of ethnic groups.[24] 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 line with the idea that ethnic boundaries matter, the results from the falsification exercise suggest no correlation between the grid-level H index

[22] The database covers 23 percent of the land area in the study region (Figure A.4). In the absence of individual-level data on ethnicity, I assign each individual to an ethnic group using Rowe (1946)'s ethnic boundaries and the geographic coordinates of the archaeological site where the human remains were found.

[23] Kiwicha, which is also a C_4 plant native to Peru, can present a carbon isotopic signature similar to that of maize (Turner et al. 2010). However, it is unclear whether its consumption became widespread in this region during pre-colonial times (e.g., Tung and Knudson 2018).

[24] The size of 50 × 50 km at the equator ensures that the number of grid cells coincides with the number of ethnic groups in Columns 1-5 of Table A.3.

and individual $\delta^{13}C_{col}$ scores. While the H index captures crop variety at different levels of geographical aggregation (i.e., grid cells of different sizes and ethnic groups; see Appendix A.1), it is associated with carbon-enriched diets only at the ethnic group level. This evidence is consistent with the narrative that crop exchange for subsistence was centralized at the level of ethnic group (Murra 2002b).

Identifying assumptions. A causal interpretation of β_3 requires two conditions. First, ethnic diversity should not be determined by average exposure to internal crop exchange ($\overline{Crop\ exchange}_p$). The 16th-century resettlement created a unique setting where $Ethnic\,div_p$ was likely orthogonal to $\overline{Crop\ exchange}_p$. However, there might be concerns that ethnic groups with a greater potential exposure to crop exchange (H_e) negotiated locations in the interior of ethnic homelands. This could result in ethnically homogeneous parishes systematically concentrating populations from these ethnic groups. Table A.4 provides supporting evidence that parish locations were not significantly influenced by groups with greater potential exposure to crop exchange. There is no significant correlation between H_e and the total number of parishes where an ethnic group was concentrated. The proportion of parishes located close to ethnic boundaries is also uncorrelated with H_e. Consistently, Figure 2 shows a similar distribution of $\overline{Crop\ exchange}_p$ among ethnically homogeneous parishes (left boxplot) and those concentrating populations from various ethnic groups (right boxplot)—$Ethnic\,div_p$ is not significantly correlated with $\overline{Crop\ exchange}_p$ in the data (Table A.5).

Second, β_3 should not capture any differential effects of ethnic diversity due to correlates of H_e that may have been relevant for post-resettlement economic development. Columns 1-3 of Table 2 show that mean elevation, land caloric suitability, and river density are not significantly correlated with H_e. There is no significant correlation between H_e and ethnic group size, as measured by land area (Column 4) and approximate population (Column 5) before colonization. However, H_e does exhibit a positive correlation with pre-colonial population density (Column 6), which could reflect economic prosperity (Ashraf and Galor 2011, 2013; Maloney and Valencia Caicedo 2016). Columns 7-9 explore pre-colonial socio-economic and institutional characteristics of the groups. In the absence of systematized ethnographic data, and following the approach of recent economic studies (see Matranga and Pascali 2021), I collect information from archaeological sources (Appendix D). In line with Column 6, the data suggest that H_e is positively correlated with urbanization, as measured by a dummy for the presence of towns and urban centers within ethnic homelands (Column 7). Column 8 shows evidence consistent with the idea that incentives for internal exchange may lead to political centralization (Fenske 2014). In particular, I create a dummy for any material indicator that could evince political complexity (i.e., administrative centers and monumental architecture—public buildings and communal spaces, including temples, palaces, and complex mound platforms, as defined in Stanish 2001) and find a positive correlation with H_e. Column 9 shows no correlation with the presence of elite residences, nonetheless. In robustness checks, I conduct various empirical exercises to mitigate the concern that pre-colonial correlates of H_e might confound the interaction effect.

In Table A.6, I explore the correlation of H_e with various pre-colonial infrastructures. Inca roads, for example, may have facilitated internal crop exchange. Column 1 shows no significant correlation between potential exposure to internal crop exchange (H_e) and the density of Inca roads inside the ethnic homeland. Columns 2-3 show no evidence of a correlation with other structures, such as water canals and bridges. In the case of roads, it is still possible that the road network facilitated crop exchange *across* ethnic homelands. Column 1 of Table A.7 explores whether H_e correlates with the distance to the nearest Inca road located in a different ethnic homeland.[25] While small-N results should be interpreted cautiously, the results suggest no statistically significant correlation. For completeness, the remaining columns of Table A.7 report the correlation of H_e with the distance to the nearest road inside the ethnic homeland (Column 2)[26] and to the nearest road (either inside or outside the ethnic homeland; Column 3). These correlations are not statistically significant, aligning with archaeologists' views that the Inca road network served various purposes (including political and military purposes) and may not have been primarily intended to promote crop exchange as a subsistence strategy for the *common* population (Hyslop 1990).[27]

[25]For each ethnic group, I compute the minimum distance across all 10×10 km grid cells inside the ethnic homeland to an Inca road located in a different ethnic homeland.

[26]This variable is not defined for seven groups for which there are no roads inside the ethnic homeland.

[27]Hyslop (1990, p. 276) writes: "Thus it appears that principal arteries [of the road network] occasionally avoided major population zones ... A particularly good example of this is found in the central Andes where the main north-south Inka road is located in the high, flat Puna region to the west of the more populated valleys on the eastern Andean slopes. There the north-south artery was constructed along a route with the least obstacles, and the main administrative centers such as *Huanuco Pampa* and *Pumpu* were placed on the road rather than in the areas where most of the subject populations lived."

4 Results

4.1 Main Results

Outcome variables. I examine contemporary living standards using various measures of local economic activity and access to public facilities for the decades 1990–2000 and 2010–2020. For each decade, the measures correspond to different years depending on data availability (see Appendix D for data sources and definitions). I first use luminosity data from satellite images at night to proxy for local economic activity (Michalopoulos and Papaioannou 2013). Specifically, I compute the log of 1 plus average light intensity per capita using data from satellite F15 of the DMSP-OLS nightlight series (years 2000-2003) and from the improved VIIRS nightlight series (year 2013).[28] As a second proxy for local economic activity, I use non-subsistence agriculture (Dell 2010). Specifically, I use data from the 1994 and 2012 agricultural censuses, conducted by the national institute of statistics (INEI). Since the definitions in these two censuses are not perfectly aligned, I created a dummy variable for whether the share of farmers practicing non-subsistence agriculture—defined as selling most of their harvest in local markets rather than using it for self-consumption—is above the sample median. For the 2010–2020 period, I also use the average number of small- and medium-sized firms per 100 inhabitants (years 2010 and 2013), as reported by the national tax administration (SUNAT). Although this variable only captures the formal sector of the economy, it serves as a third proxy for local economic activity.[29] Finally, I use data on access to public water and sanitation from the 1993 and 2017 population censuses (INEI). Local governments in Peru play an important role in ensuring access to public facilities, especially since the fiscal decentralization wave of 2002 (Pique 2019). I use the share of dwellings with access to public water and sanitation as measures of access to public facilities.

Graphical analysis. To explore potential systematic patterns in the data, Figure 3 plots mean development outcomes for the 2010–2020 period across four subgroups of parishes. These

[28] A previous version of this paper used nightlight data from satellite F18 of the DMSP-OLS nightlight series (years 2010-2013); see also Column 2 of Table 4. The improved VIIRS nightlight data are available after 2012 only. Using VIIRS data for seven countries of South America (including Peru), McCord and Rodriguez-Heredia (2022) show that VIIRS nighttime luminosity is positively correlated with GDP at the first subnational level.

[29] This information is not available for the 1990–2000 period.

subgroups are defined by ethnic diversity status (indicated by the dummy variable for whether there is an ethnic border within a 10-km buffer from the parish capital; $Ethnic\ div_p$) and a dummy variable for whether average exposure to pre-colonial crop exchange ($\overline{Crop\ exchange_p}$) is above the sample median.[30] Graph A shows the mean score of the first principal component derived from the five outcome variables, while graphs B to F plot the raw data for each variable separately. All variables, except for the dummy variable indicating non-subsistence agriculture, are standardized to have a mean of zero and a standard deviation of one. The data suggest a negative correlation between ethnic diversity and contemporary development where average exposure to crop exchange is below the sample median, with the magnitude of this correlation varying for the different outcome variables. However, in line with the hypothesis that historical exposure to crop exchange helps mitigate the local costs of ethnic diversity, this negative correlation does not appear in the data for parishes above the median level of crop exchange. In what follows, I examine this relationship using a regression framework.

Baseline regression analysis. To analyze the average effect of ethnic diversity on long-run development, I start by following the methodology in Kling et al. (2004) and Clingingsmith, Khwaja and Kremer (2009). Specifically, Table 3 reports the standardized average effect size (AES) across the different outcome variables, separately for each decade, thus accounting for the covariance across underlying individual effects.

In Column 1 of Table 3, I compare average living standards (1990–2000) between ethnically diverse parishes and those with an ethnically homogeneous founding population ($\hat{\beta}_1$, Equation 1), using heteroskedasticity-robust standard errors. On average, contemporary living standards are 0.2 standard deviations lower in ethnically diverse parishes (Column 1). This disparity remains statistically significant at the 5 percent level after considering parish-level baseline characteristics (X_p, Column 2). The estimated negative coefficient is consistent with the literature on the costs of ethnic diversity (Miguel and Gugerty 2005; Hjort 2014), highlighting the persistent consequences of forced ethnic diversity in the study setting. Column 3 presents the results from estimating the interaction effect of ethnic diversity ($Ethnic\ div_p$) and the average level of potential exposure to pre-colonial crop exchange ($\overline{Crop\ exchange_p}$), with

[30]Above the sample median of average exposure to crop exchange, there are 61 ethnically diverse parishes and 108 ethnically homogeneous parishes. Below the median, there are 56 ethnically diverse parishes and 111 ethnically homogeneous parishes.

baseline controls (X_p) and heteroskedasticity-robust standard errors; Equation 2. The estimated coefficients are consistent with a long-run comparative development pattern in which: (i) on average, ethnically diverse parishes exhibit lower living standards ($\hat{\beta}_1 < 0$); but (ii) those with initial populations exposed to higher levels of crop exchange perform relatively better in the long run ($\hat{\beta}_3 > 0$).

The coefficients remain statistically significant in Column 4, when using ecclesiastical jurisdiction fixed effects that account for potential differences across five colonial bishoprics, jointly with robust standard errors clustered at this level. The same pattern arises when I use fixed effects accounting for the colonial administrative province (44 provinces) instead of the ecclesiastical jurisdiction, jointly with robust standard errors clustered at the province level (Column 5). Columns 6 and 7 report the estimated direct effect of pre-colonial exposure to crop exchange ($\hat{\beta}_2$) for reference. The estimates suggest a positive correlation between average exposure to internal crop exchange and contemporary living standards, both after ecclesiastical jurisdiction (Column 6) and administrative province (Column 7) fixed effects.

Overall, the results for the 1990-2000 period are consistent with the hypothesis that prior experience with mutually beneficial crop exchange between coethnics helped mitigate the long-run costs of ethnic diversity—at the 10th percentile of average exposure to crop exchange ($\overline{Crop\ exchange}_p = 0.4$), contemporary living standards are 0.3 standard deviations lower in ethnically diverse parishes, compared to parishes with an ethnically homogeneous founding population; at the sample median of average exposure ($\overline{Crop\ exchange}_p = 0.6$), the negative effect of ethnic diversity decreases to 0.1 standard deviations (Column 5). The estimates in Column 8 show the same pattern for the 2010–2020 period. As a robustness exercise, the specification in Column 9 includes a measure of heterogeneity in complementary elevation zones at the parish level. Specifically, I control for $H_p = 1/\sum_j s_{jp}^2$, where s_{jp} represents the area share of elevation zone j within the 10-km buffer from the parish (p) capital.[31] The results are consistent with the baseline estimates (Column 8). Even in a saturated specification that additionally includes the set of parish-level pre-colonial characteristics used in Table A.1, the estimated coefficients are qualitatively and quantitatively consistent (Column 10).

Columns 1-7 of Table 4 report OLS estimates for each outcome variable of the 2010–2020

[31]The correlation between this variable and the standard deviation of elevation (also included as a baseline control variable) is 0.625, while the correlation with average exposure to internal crop exchange is 0.215.

period, separately.[32] All regressions include baseline controls (X_p) and administrative province fixed effects. Heteroskedasticity-robust standard errors in brackets are clustered at the province level. In parentheses, standard errors are adjusted for spatial autocorrelation with a distance cutoff of approximately 1 degree at the equator (Colella et al. 2019). The estimated coefficients are consistent with the long-run development pattern described previously.[33] The results also show that much of the documented average effect of ethnic diversity (AES, Table 3) is driven by differences in economic activity rather than access to public facilities.[34] Survey data on household consumption support this interpretation (Columns 8-10). I use the ENAHO annual surveys conducted by the national institute of statistics to compute the log of real household consumption per capita (2004–2017).[35] The surveys cover 280 parishes, of which 98 are ethnically diverse parishes. I use the household as the unit of analysis and include year fixed effects and household-level controls (the household head's gender, age, age squared, years of schooling, civil status, and language spoken at home) in the specification, with robust standard errors clustered at the parish level. Table A.8 reports wild cluster bootstrap t-statistics for all the outcomes in Table 4 (Cameron et al. 2008), showing consistent results.

Pre-colonial characteristics of ethnic groups. In Table 5, I show that the documented interaction effect is robust to controlling for the correlates of average exposure to pre-colonial crop exchange. I consider the set of pre-colonial ethnic characteristics analyzed in Table 2. In particular, I compute the weighted average of each characteristic among the ethnic groups concentrated in the parish and augment Equation 2 to control for the resulting average ($\overline{G}_p = \sum_e w_{pe} G_e$) and its interaction with ethnic diversity ($Ethnic\ div_p \times \overline{G}_p$). The first column shows the baseline specification (Column 8 of Table 3), Columns 2 to 8 introduce one characteristic at a time (only the coefficient on the interaction with ethnic diversity is reported), and Column 9 includes all characteristics. The results alleviate the concern that

[32] Additionally, the table reports the results from using the DMSP-OLS nightlight series (Column 2), instead of the VIIRS series (Column 1), to compute nightlight per capita. The table also reports the results from using the share of farmers practicing non-subsistence agriculture as outcome variable (Column 5), rather than a dummy for above-median share of farmers (Column 4).

[33] Due to space limitations, only the estimated coefficients on ethnic diversity and its interaction with average exposure to crop exchange are reported.

[34] Note, however, that night-time luminosity data may also capture public lighting (Hodler and Raschky 2014).

[35] I use the spatial deflators provided by the national institute of statistics (INEI) and follow Dell (2010) in subtracting public transfers received by the household. Real household consumption without transfers is divided by the number of household members to obtain a per capita measure.

relevant socio-economic and institutional characteristics of ethnic groups could be driving the entire result.[36] Column 10 shows that the positive interaction effect persists when using lasso methods to *select* the set of ethnic characteristics to be included in the specification (Belloni, Chernozhukov and Hansen 2014). The lasso routine considers all ethnic characteristics included in Column 9 of Table 5, land caloric suitability for maize, and ethnic infrastructures (Table A.6)—a total of 11 characteristics, of which 7 are *selected* by lasso.

Coarsened exact matching. In Table 6, I show the results from using a matching procedure to construct a counterfactual for parishes with high exposure to pre-colonial crop exchange (defined as $\overline{Crop\ exchange}_p$ above the sample median). Specifically, I use coarsened exact matching (Iacus, King and Porro 2012) to create a sample of parishes that present varying levels of exposure to crop exchange but are statistically similar in other pre-colonial characteristics ($\overline{G}_p$). In Columns 1-2, the pre-colonial ethnic characteristics to be balanced by the matching algorithm are the characteristics used in Column 9 of Table 5. In Columns 3-4, the procedure uses the set of lasso-*selected* characteristics.[37] Despite the reduced sample sizes, the results from the matched samples (Columns 1 and 3) are consistent with the documented pattern.

To alleviate the concern that certain ethnic groups could be driving the results, the remaining columns of Table 6 control for the weight (w_{pe}) of the majority group, defined as the group with the highest area share within the 10-km buffer. The estimated interaction effect remains positive and statistically significant in both the matched and full samples. Figure A.5 shows partial correlation scatterplots for the estimated interaction effect between ethnic diversity and average exposure to crop exchange. Different colors represent different majority groups. The lack of distinct color clusters suggests that no single majority group is predominantly influencing the main result.

Randomization inference. The empirical analysis is guided by the historical narrative that the colonial resettlement resulted in quasi-random variation in ethnic diversity across parishes.

[36] Table A.9 shows that the positive interaction effect is robust to controlling for the weighted average of log land area and its interaction with ethnic diversity, rather than controlling for log population density and log river density (which may affect the interpretation of the estimated coefficients). The table also shows robustness to controlling for variation in elevation and land caloric suitability across pre-colonial ethnic homelands.

[37] Table A.10 documents that, in the matched samples, ethnically diverse and non-diverse parishes continue to be balanced in terms of geographic and initial factors.

Table A.11 shows the results from permuting $Ethnic\ div_p$ within the regression sample. I use the procedure proposed by Young (2024) for randomization inference in the presence of heterogeneous treatment effects. In each iteration, $Ethnic\ div_p$ is permuted, and the interaction term with average exposure to pre-colonial crop exchange ($\overline{Crop\ exchange_p}$) is recalculated accordingly. The results from 1,000 iterations show stable permutation-based p-values.

4.2 Supplementary Analyses

Pre-colonial land occupation and transition zones. The estimates in tables 3 and 4 are likely affected by non-classical measurement errors. A potential source of error is the underlying assumption that pre-colonial individuals were uniformly distributed over space. In Table A.12, I consider alternative scenarios. In the absence of historical data on the spatial distribution of the population, I use archaeological site records as evidence of land occupation. Columns 1 and 2 show consistent estimates after restricting the analysis to 20 km and 10 km around pre-colonial archaeological sites, respectively.[38] Another underlying assumption is that all coethnic individuals were equally exposed to crop exchange (i.e., the baseline estimates capture the average effect of potential exposure). In Columns 3 and 4, I restrict the analysis to 20 km and 10 km around the transitions from one elevation zone to another. Although the results in Section 3.2 suggest that ethnic boundaries are relevant in explaining crop exchange, potential gains may have been higher around transition zones (Bates 2001). The estimates are similar to baseline results, suggesting no significant differences in exposure among coethnics.

Placebos for pre-colonial ethnic boundaries. Following Alesina, Michalopoulos and Papaioannou (2016), Column 5 of Table A.12 presents the results from using artificial ethnic boundaries (from Thiessen polygons) instead of historical ones.[39] Compared to baseline results, the estimates are smaller and not significant, suggesting that historical ethnic boundaries matter. This aligns with the evidence in Section 3.2, where I used grid cells instead of Thiessen polygons. In Column 6, I use the boundaries of the first administrative demarcations of the colonial period (*corregimientos*). Close correspondence between the spatial boundaries of

[38]For this exercise, I combined a public inventory of pre-colonial archaeological sites (*Catastro de Monumentos Arqueológicos Prehispánicos* (SIGDA), Ministerio de Cultura, Perú) with my own survey of published archaeological studies and handbooks (see Isbell and Silverman 2002a, 2008).

[39]Thiessen polygons are created using the centroids of historical ethnic homelands as input.

pre-colonial ethnic groups and *corregimientos* would suggest that the Spanish administration based the latter on prior knowledge of the spatial distribution of the groups.[40] The estimated coefficients are not statistically significant, suggesting that this was not the case.

Sensitivity analyses. Table A.13 reports the results from varying the size of the buffer used to measure ethnic diversity. In Columns 1-2 of Table A.14, I use an index of ethnic fractionalization ($1 - \sum_e w_{ep}^2$),[41] while in Columns 3-4, I use an alternative index of potential gains from crop exchange ($\widetilde{H}_e = 1 - \sum_j s_{je}^2$). Columns 5-6 consider the level of exchange exposure of the majority group in the parish instead of the weighted average of H_e, while Columns 7-8 focus on the level of the minority group. The results align with the documented pattern and support that both the majority and minority groups matter. Figure A.6 displays point estimates and confidence intervals after excluding one parish at a time, alleviating concerns about influential observations.

Additional robustness checks. It is important to consider that any potential effect of ethnic diversity after resettlement depends on the survival rates of the groups. The decline in native populations after European contact has been extensively documented by historical studies (e.g., Cook 1982). To the extent that all groups were similarly affected by disease and abuse, the estimates should be interpreted as the long-run effects of ethnic diversity among the descendants of survivors. Table A.15 shows that the positive coefficient on the interaction term is robust to controlling for the potential pre-resettlement spread of smallpox—an infectious disease caused by the variola virus that may have affected native populations during the conquest. The table also shows consistent results when considering factors related to the Inca period. Appendix A.2 and Table A.16 address potential selective migration.

[40] I am grateful to Jenny Guardado for this suggestion.

[41] The parish-level correlation between the ethnic dummy and the index of ethnic fractionalization in the data is 0.8522. Note, however, that, since w_{ep} represents the area share of ethnic group e within the buffer of parish p rather than the exact population share, and individuals were probably not uniformly distributed over space, this measure may arguably be more affected by measurement error than the dummy for ethnic diversity.

5 Mechanisms

In this section, I analyze historical and modern data to explore potential mechanisms. First, I examine whether prior experience with mutually beneficial crop exchange shaped more open attitudes toward out-group members, contributing to cooperative behavior and societal integration. Second, I explore the role of economic complementarities in sustaining market-oriented cooperation in the context of a predominantly agricultural economy. Finally, I examine whether these dynamics resulted in a shift toward tertiary-sector activities and local trade.

Societal integration and culture of cooperation. I first explore whether prior experience with mutually beneficial exchange facilitated integration with other ethnic groups. Did historical exposure to crop exchange contribute to the likelihood of inter-group interaction after resettlement? Comparing the two surnames of each individual in colonial baptism records (1605–1780) offers the opportunity to explore inter-group unions, commonly used as a proxy for societal integration (e.g., Bazzi, Gaduh, Rothenberg and Wong 2019).[42] The sample includes 17,411 individuals with native surnames across 41 parishes, of which 10 are ethnically diverse parishes. Since ethnic identities were not systematically registered, I use a measure of dissimilarity between the two surnames to *detect* parents who were potentially from different ethnic groups.[43] Given the small sample size, a graphical summary of the data is provided in Figure A.7. In line with the hypothesized channel, the left graph of Panel (a) suggests a positive correlation between average dissimilarity and historical exposure to crop exchange among parishes with ethnically diverse initial populations, pointing toward a more integrated society. The right graph shows a similar pattern when focusing on the share of unions with dissimilarity above 50 percent at the parish level. Panel (b) replicates this exercise for the subsample of ethnically homogeneous parishes, revealing no significant correlation, which is reassuring.[44]

[42]Each individual inherits two surnames in the Hispanic system of family names. The first surname corresponds to the paternal surname of the father, while the second corresponds to the paternal surname of the mother. Colonial marriage records from digital genealogical sources are limited in quantity and geographic coverage.

[43]I use the minimum number of spelling changes required to transform one surname into another (Levenshtein), normalized by the length of the longest surname to be interpreted as the percentage of dissimilarity. See Dickens (2022) for an application of this measure to compute distance between languages.

[44]It is worth noting that the ethnic identities represented in the sample can only be *detected* in historical data, primarily because native dialects have gradually been displaced by regional varieties of Quechua and Spanish,

In Appendix C, I explore this pattern more generally. Specifically, I present correlational evidence from a sample of U.S. immigrant descendants. I use the sample of descendants created by Giuliano and Nunn (2021) based on the *March Supplement of the Current Population Survey* (CPS) for the period 1994-2014. The sample includes all married individuals (born in the U.S.) with immigrant parents. Consistent with the previous results, the estimates suggest that individuals with greater ancestral exposure to economic exchange, based on the pre-industrial ancestors who lived in their parents' country of origin, are less likely to marry within their own group, on average. The sample covers 83 different countries of origin (see Table C.1). Nonetheless, it is important to note that this is a selected sample— the level of exposure to economic exchange at the origin may have influenced emigration choice.

The visual analysis based on colonial surnames sheds light on the long-run process of societal integration and assimilation in the study region (Figure A.7). Nonetheless, small-N results should be interpreted cautiously. Contemporary survey data on individuals' self-reported identities are consistent with a more integrated, nation-oriented society (Table 7). Individuals tend to exhibit stronger identification with the state where the resettlement forced together ethnically diverse populations that had historically experienced higher levels of crop exchange (Columns 1-3).[45] In line with a more nation-oriented society, these individuals are more likely to vote in presidential elections, a result that holds beyond trust in the state (Columns 4-5). Parish-level data on volunteers for military service exhibit a similar pattern (Columns 6-7).

The previous results suggest that being traditionally used to engage in mutually beneficial crop exchange shaped more open attitudes toward out-group members, fostering a trend toward a more integrated society. These results are in line with Enke (2023)'s result for the pre-industrial era, showing that a group's degree of exposure to markets and economic exchange contributed to prosocial behavior, which includes cooperative behavior as one of its dimensions. The estimates in Columns 8-9 of Table 7 are consistent with the transmission of cooperative behavior. Specifically, I use parish-level data on the presence of neighborhood associations to proxy for a culture of cooperation (Guiso, Sapienza and Zingales 2016). The results point

shaping today's linguistic landscape. The results in this paper should be interpreted as the long-run effects of ethnic diversity—created during the formation of 16th-century colonial jurisdictions.

[45]The ENAHO Peruvian survey (2004-2017) includes a question on whether individuals identify more strongly with their state administrative unit (including the administrative region, province, or district), as opposed to other groups (ethnic, religious, or other groups). All regressions include individual-level socio-demographic controls and survey-year fixed effects, with robust standard errors clustered at the parish level.

toward this culture facilitating inter-ethnic interactions where the colonial resettlement forced together populations with a more intense experience of internal crop exchange.[46] These results are also consistent with papers showing that strategies to cope with environmental risk and adverse geography can help sustain cooperation and more trusting attitudes over time (Nunn and Wantchekon 2011; Nunn and Puga 2012; Buggle and Durante 2021).

Economic complementarities. Being more willing to interact, economic complementarities may have sustained market-oriented cooperation and economic development in the long run. I examine whether, in line with Jha (2013), contemporary populations tend to perform better where the historical ethnic minority complemented the majority group in terms of prior exposure to crop exchange (i.e., where the historical ethnic minority experienced significant exposure to crop exchange, while the majority lacked such exposure). To test this hypothesis, I define the minority group as the group with the lowest area share within the 10-km buffer and run regressions of the following form for the subsample of parishes with ethnically diverse initial populations (Table 8, Panel A):

$$y_p = \delta_0 + \delta_1 High\ min_p + \delta_2 High\ maj_p + \delta_3 \left(High\ min_p \times High\ maj_p\right) + X_p^{'}\gamma + \varepsilon_p$$

where $High\ min_p$ is a dummy variable indicating whether the minority belonged to a group with a high degree of potential exposure to crop exchange (H_e above the 75th percentile) and $High\ maj_p$ is an analogous dummy variable for the majority group—the correlation between these two variables is 0.325. Consistent with the proposed interpretation, where the historical ethnic minority had a comparative advantage over the majority group, contemporary populations tend to perform better in the long run ($\hat{\delta}_1 > 0$, Columns 1-2).[47] In Table A.18, I

[46] These results align with anecdotal evidence suggesting the existence of reciprocity and food sharing practices in pre-colonial times. For example, Berezkin (2015)'s folklore catalog mentions the avaricious man motif—*A man does not share food with his wife or kinsfolk. He or his food is transformed (turns into a bird, into worms, etc.) in punishment*—in the *Conchucos*'s homeland, the same region that early chronicles describe as "very fertile and abundant, with many crops and resources that everyone has and sows" (translated from the Spanish version in Cieza de León 1962 [1553], p. 221). Stern (1995, p. 76) suggests that "Andean rules of reciprocity and redistribution served to govern exchanges ... Andean peoples sought self-sufficiency ... by engaging in reciprocities enabling the collective kin or ethnic group to directly produce diverse goods in scattered ecological zones." The results in Table 7 are robust to controlling for log population and consistent with individual-level data on participation in voluntary associations (Table A.17).

[47] In these parishes, the mean advantage of the minority over the majority group in terms of H_e is 12.93 percent, with a maximum of 34.17 percent.

use a continuous measure of minority group advantage, defined as the difference in potential crop exchange exposure between the minority and majority groups. In line with the idea that a greater minority group advantage increases the scope of trade between groups via economic complementarities, the minority group advantage is positively associated with contemporary development. The table also shows that the beneficial effects of this advantage tend to become more pronounced as the average level of potential exposure to crop exchange increases.

The 1876 population census—the earliest post-colonial census with information on occupations—reveals that most parishes remained primarily agricultural by the late 19th century, with an average of 70 percent of the population employed in the agricultural sector. In the remaining columns of Panel A (Table 8), I explore the hypothesis that ethnic minorities with a comparative advantage over the majority group specialized in intermediary services related to agriculture (e.g., retailing), thus reducing inter-ethnic competition (Grosfeld, Sakalli and Zhuravskaya 2020). In line with this hypothesis, Columns 3 and 4 show that contemporary local retail markets for selling agricultural products tend to be located where the historical ethnic minority had a comparative advantage. Moreover, in line with the idea of reduced competence over local resources, colonial data for the period 1700–1800 show that a larger share of the population had access to agricultural land in these parishes (Columns 5 and 6).[48]

Discussion and consequences: a market-oriented agricultural economy. The evidence on potential mechanisms suggests that prior experience with internal crop exchange contributed to more open attitudes toward out-group members, facilitating inter-group cooperation and fostering a trend toward a more integrated society (Table 7). In a region where subsistence farming has historically been prevalent (Mayer 2002), the results also show that non-subsistence agriculture flourished where the 16th-century resettlement brought together ethnically diverse populations with a history of internal crop exchange (Columns 4-5, Table 4). The emergence of intermediary services related to agriculture, primarily supported by ethnic minorities (Table 8, Panel A), aligns with the previous results.

In Panel B of Table 8, I explore the potential consequences of sustaining market-oriented cooperation related to agriculture. Specifically, I explore the share of the population employed in retail and services using the 1876 population census. In line with the documented pattern,

[48]This information is available for 44 percent of the parishes with ethnic diversity.

the estimates show a shift toward tertiary-sector activities in a predominantly agricultural setting.[49] As the degree of historical exposure to crop exchange increases, employment tends to become more oriented toward the tertiary sector where the resettlement concentrated ethnically diverse populations (Column 1, Panel B). Consistent with the previous results, decomposing tertiary-sector employment shows that local trade drives the overall effect (Column 3, Panel B). Census data for the 21st century show a similar pattern (Columns 2 and 4, Panel B).[50] Overall, the results align with the idea that ethnic groups with a historical practice of internal crop exchange helped sustain a market-oriented agricultural economy in the long run.

Furthermore, the results on non-subsistence agriculture remain significant when controlling for local (parish-level) crop diversity (Table A.19), suggesting that it is not crop availability but rather the transmission of more open attitudes governing agriculture-related market interactions that matters the most in this setting. Ethnic groups with greater exposure to crop exchange may have also played a role in developing stronger institutions in support of the market-oriented economy. In Table A.20, I use data on the share of the municipal budget executed as a proxy for local institutional performance (Artiles et al. 2021). The results for the 2002-2013 period are qualitatively consistent, though only statistically significant (at the 10 percent level) for the 2006-2009 electoral term. Finally, I find no significant evidence that pre-colonial crop diversity may have indirectly fostered secondary or tertiary-sector occupation through the potential increased availability of skills (Fiszbein 2022); see Table A.21.

6 Conclusion

The consequences of ethnic diversity have inspired a rich literature in the social sciences, with most papers documenting negative effects on economic development. This paper contributes to the literature on the economic drivers of inter-ethnic coexistence. I examine a natural experiment of history to learn about the consequences of ethnic diversity for long-run economic development, documenting that historical exposure to economic exchange helps mitigate the long-run costs of ethnic diversity at the local level. In contemporary societies where multiple

[49] Consistent with a predominantly agricultural context, I find qualitatively consistent but statistically insignificant results when exploring urbanization data (Column 7, Table A.19).

[50] I compute average tertiary-sector outcomes using data from the 2007 and 2017 population censuses.

ethnicities coexist (e.g., due to forced displacements, climate-induced migrations, or voluntary migrations in an increasingly globalized world), understanding the economic factors that help mitigate the costs of ethnic diversity is also important for policy debates.

I collect data on the 16-century resettlement of native populations in highland Peru, where the pre-colonial economy was primarily based on agriculture, with coethnic individuals engaging in crop exchange for subsistence. The colonial resettlement created new jurisdictions for native populations, unintentionally concentrating populations from different ethnic groups. This resettlement resulted in a unique scenario where historical exposure to internal crop exchange was arguably orthogonal to ethnic diversity. The results show that prior experience with internal crop exchange facilitated the functioning of multi-ethnic societies in the long run. Both historical and modern census data show a shift toward tertiary-sector activities, mainly driven by local trade, where ethnically diverse populations had prior experience with internal crop exchange. The results are consistent with the idea that ethnic groups with a historical practice of internal crop exchange helped develop a market-oriented agricultural economy after the resettlement. Additional evidence suggests that underlying this development process is the formation of more open attitudes toward out-group members, which contributed to the likelihood of inter-group interaction after resettlement. In a predominantly agricultural setting, the evidence on mechanisms also highlights the positive role of economic complementarities in sustaining long-run, market-oriented cooperation and local trade.

References

Acemoglu, Daron, Simon Johnson, and James A Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation," *The American Economic Review*, 2001, *91* (5), 1369–1401.

Aldenderfer, Mark S, *Domestic Architecture, Ethnicity, and Complementarity in the South-Central Andes*, University of Iowa Press, 1993.

Alesina, Alberto and Edward Glaeser, *Fighting Poverty in the US and Europe: A world of Difference*, Oxford University Press, 2004.

_ and Eliana La Ferrara, "Participation in Heterogeneous Communities," *The Quarterly Journal of Economics*, 2000, *115* (3), 847–904.

_ and _ , "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, 2005, *43* (3), 762–800.

— **, Paola Giuliano, and Nathan Nunn**, "On the Origins of Gender Roles: Women and the Plough," *The Quarterly Journal of Economics*, 2013, *128* (2), 469–530.

— **, Reza Baqir, and William Easterly**, "Public Goods and Ethnic Divisions," *The Quarterly Journal of Economics*, 1999, *114* (4), 1243–1284.

— **, Stelios Michalopoulos, and Elias Papaioannou**, "Ethnic Inequality," *Journal of Political Economy*, 2016, *124* (2), 428–488.

Ambrose, Stanley H, "Isotopic Analysis of Paleodiets: Methodological and Interpretive Considerations," in "Investigations of Ancient Human Tissue: Chemical Analysis in Anthropology," Philadelphia, USA: Gordon and Breach Science Publishers, 1993, pp. 59–130.

— **and Lynette Norr**, "Experimental Evidence for the Relationship of the Carbon Isotope Ratios of Whole Diet and Dietary Protein to Those of Bone Collagen and Carbonate," in "Prehistoric Human Bone: Archaeology at the Molecular Level," Berlin: Springer, 1993, pp. 1–37.

Angeles, Luis and Aldo Elizalde, "Pre-Colonial Institutions and Socioeconomic Development: The Case of Latin America," *Journal of Development Economics*, 2017, *124*, 22–40.

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.

Ashraf, Quamrul and Oded Galor, "Dynamics and Stagnation in the Malthusian Epoch," *American Economic Review*, 2011, *101* (5), 2003–41.

— **and** — , "The "Out of Africa" Hypothesis, Human Genetic Diversity, and Comparative Economic Development," *American Economic Review*, 2013, *103* (1), 1–46.

Banerjee, Abhijit and Lakshmi Iyer, "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India," *The American Economic Review*, 2005, *95* (4), 1190–1213.

Bates, Robert H., *Prosperity and Violence: The Political Economy of Development*, W. W. Norton & Company, Jan 1, 2001 2001.

Bauer, Brian S, LC Kellett, and Aráoz Silva, *The Chanka: Archaeological Research in Andahuaylas (Apurimac)*, Cotsen Institute of Archaeology, UCLA, 2010. With contributions by Sabine Hyland and and Carlo Socualaya Dávila.

Bazzi, Samuel, Arya Gaduh, Alexander D Rothenberg, and Maisy Wong, "Unity in Diversity? How Intergroup Contact Can Foster Nation Building," *American Economic Review*, 2019, *109* (11), 3978–4025.

Becker, Sascha O, "Forced Displacement in History: Some Recent Research," *Australian Economic History Review*, 2022, *62* (1), 2–25.

_ **and Luigi Pascali**, "Religion, Division of Labor, and Conflict: Anti-Semitism in Germany over 600 Years," *American Economic Review*, 2019, *109* (5), 1764–1804.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "Inference on Treatment Effects After Selection Among High-Dimensional controls," *The Review of Economic Studies*, 2014, *81* (2), 608–650.

Berezkin, Yu E, "Folklore and Mythology Catalogue: Its Lay-out and Potential for Research," *The Retrospective Methods Network*, 2015, *1* (S10), 58–70.

Bruhn, Miriam and Francisco A Gallego, "Good, Bad, and Ugly Colonial Activities: Do They Matter for Economic Development?," *The Review of Economics and Statistics*, 2012, *94* (2), 433–461.

Brush, Stephen B, "Man's Use of an Andean Ecosystem," *Human Ecology*, 1976, *4* (2), 147–166.

Buggle, Johannes C and Ruben Durante, "Climate Risk, Cooperation and the Co-Evolution of Culture and Institutions," *The Economic Journal*, 2021, *131* (637), 1947–1987.

Burger, Richard L and Nikolaas J Van der Merwe, "Maize and the Origin of Highland Chavín Civilization: An Isotopic Perspective," *American Anthropologist*, 1990, *92* (1), 85–95.

Caicedo, Felipe Valencia, "The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America," *The Quarterly Journal of Economics*, 2019, *134* (1), 507–556.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller, "Bootstrap-Based Improvements for Inference with Clustered Errors," *The Review of Economics and Statistics*, 2008, *90* (3), 414–427.

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.

Charney, Paul, "A Sense of Belonging: Colonial Indian Cofradias and Ethnicity in the Valley of Lima, Peru," *The Americas*, 1998, *54* (3), 379–407.

Chávez, Emilio Lissón, *La Iglesia de España en el Perú. Colección de Documentos para la Historia de la Iglesia en el Perú.*, Vol. 3, 4, Sevilla: Editorial Católica Española, 1943.

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.

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

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

Cruz, Cesi, Julien Labonne, and Pablo Querubin, "Politician Family Networks and Electoral Outcomes: Evidence from the Philippines," *American Economic Review*, 2017, *107* (10), 3006–3037.

de Armas Medina, Fernando, *Cristianización del Perú (1532-1600)*, Sevilla: Escuela de Estudios Hispano-Americanos, 1953.

de Cieza de León, Pedro, *La Crónica del Perú*, Madrid: Espasa–Calpe, 1962. [1553].

de la Espada, Marcos Jiménez, *Relaciones Geográficas de Indias. Perú*, Vol. 1, Madrid: Ministerio de Fomento, 1881.

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

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

Depetris-Chauvin, Emilio and Ömer Özak, "The Origins of the Division of Labor in Pre-Modern Times," *Journal of Economic Growth*, 2020, *25* (3), 297–340.

Desmet, Klaus, Joseph Flavian Gomes, and Ignacio Ortuño-Ortín, "The Geography of Linguistic Diversity and the Provision of Public Goods," *Journal of Development Economics*, 2020, *143*, 102–384.

Diamond, Jared, *Guns, Germs, and Steel: The Fates of Human Societies*, New York: Norton, 1997.

Dickens, Andrew, "Understanding Ethnolinguistic Differences: The Roles of Geography and Trade," *The Economic Journal*, 2022, *132* (643), 953–980.

Dippel, Christian, "Forced Coexistence and Economic Development: Evidence from Native American Reservations," *Econometrica*, 2014, *82* (6), 2131–2165.

Dulanto, Jalh, "Between Horizons: Diverse Configurations of Society and Power in the Late pre-Hispanic Central Andes," in "The Handbook of South American Archaeology," Springer, 2008, pp. 761–782.

Easterly, William and Ross Levine, "Africa's Growth Tragedy: Policies and Ethnic Divisions," *The Quarterly Journal of Economics*, 1997, *112* (4), 1203–1250.

Enke, Benjamin, "Market Exposure and Human Morality," *Nature Human Behaviour*, 2023, *7* (1), 134–141.

Fearon, James D and David D Laitin, "Ethnicity, Insurgency, and Civil War," *American Political Science Review*, 2003, *97* (1), 75–90.

Fenske, James, "Ecology, Trade, and States in Pre-Colonial Africa," *Journal of the European Economic Association*, 2014, *12* (3), 612–640.

Fischer, Guenther, Freddy Nachtergaele, Sylvia Prieler, HT Van Velthuizen, Luc Verelst, and David Wiberg, "Global Agro-Ecological Zones Assessment for Agriculture (GAEZ 2008)," 2008. IIASA, Laxenburg, Austria and FAO, Rome, Italy.

Fiszbein, Martin, "Agricultural Diversity, Structural Change, and Long-Run Development: Evidence from the United States," *American Economic Journal: Macroeconomics*, 2022, *14* (2), 1–43.

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

Gennaioli, Nicola and Ilia Rainer, "The Modern Impact of Precolonial Centralization in Africa," *Journal of Economic Growth*, 2007, *12* (3), 185–234.

Giuliano, Paola and Nathan Nunn, "Understanding Cultural Persistence and Change," *The Review of Economic Studies*, 2021, *88* (4), 1541–1581.

Grosfeld, Irena, Seyhun Orcan Sakalli, and Ekaterina Zhuravskaya, "Middleman Minorities and Ethnic Violence: Anti-Jewish Pogroms in the Russian Empire," *The Review of Economic Studies*, 2020, *87* (1), 289–342.

Guardado, Jenny, "Office-Selling, Corruption, and Long-Term Development in Peru," *American Political Science Review*, 2018, *112* (4), 971–995.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "Cultural Biases in Economic Exchange?," *The Quarterly Journal of Economics*, 2009, *124* (3), 1095–1131.

_ , _ , and _ , "Long-Term Persistence," *Journal of the European Economic Association*, 2016, *14* (6), 1401–1436.

Hjort, Jonas, "Ethnic Divisions and Production in Firms," *The Quarterly Journal of Economics*, 2014, *129* (4), 1899–1946.

Hodler, Roland and Paul A Raschky, "Regional Favoritism," *The Quarterly Journal of Economics*, 2014, *129* (2), 995–1033.

Huillery, Elise, "History Matters: The Long-Term Impact of Colonial Public Investments in French West Africa," *American Economic Journal: Applied Economics*, 2009, *1* (2), 176–215.

Hyslop, John, *Inka Settlement Planning*, University of Texas Press, 1990.

Iacus, Stefano M, Gary King, and Giuseppe Porro, "Causal Inference Without Balance Checking: Coarsened Exact Matching," *Political Analysis*, 2012, *20* (1), 1–24.

Isbell, William H, "La Arqueología Wari y la Dispersión del Quechua," *Boletín de Arqueología PUCP*, 2010, (14), 199–220.

_ **and Helaine Silverman**, *Andean Archaeology I*, Springer, 2002.

_ **and** _ , "Theorizing Variations in Andean Sociopolitical Organization," in "Andean Archaeology I," Springer, 2002, pp. 3–11.

_ **and** _ , *The Handbook of South American Archaeology*, Springer, 2008.

Jedwab, Remi, Noel D Johnson, and Mark Koyama, "Negative Shocks and Mass Persecutions: Evidence from the Black Death," *Journal of Economic Growth*, 2019, *24* (4), 345–395.

Jha, Saumitra, "Trade, Institutions, and Ethnic Tolerance: Evidence from South Asia," *American Political Science Review*, 2013, *107* (4), 806–832.

_ , "Trading for Peace," *Economic Policy*, 2018, *33* (95), 485–526.

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 No. RWP04-035.

Magurran, Anne E, *Measuring Biological Diversity*, Oxford, U.K.: Blackwell Science, 2004.

Maloney, William F and Felipe Valencia Caicedo, "The Persistence of (Subnational) Fortune," *The Economic Journal*, 2016, *126* (598), 2363–2401.

Matranga, Andrea and Luigi Pascali, "The Use of Archaeological Data in Economics," in "The Handbook of Historical Economics," Elsevier, 2021, pp. 125–145.

Mayer, Enrique, *The Articulated Peasant: Household Economies in the Andes*, Westview Press, 2002.

McCord, Gordon Carlos and Mario Rodriguez-Heredia, "Nightlights and Subnational Economic Activity: Estimating Departmental GDP in Paraguay," *Remote Sensing*, 2022, *14* (5), 1150.

Medina, Alejandro Málaga, "Las Reducciones en el Perú (1532-1600)," *Historia y Cultura*, 1974, *8*, 141–172. Museo Nacional de Arqueología e Historia del Perú.

__ , “Las Reducciones en el Perú durante el Gobierno del Virrey Francisco de Toledo,” *Anuario de Estudios Americanos*, 1974, *31*, 819–842. Escuela de Estudios Hispanoamericanos de Sevilla, Consejo Superior de Investigaciones Científicas.

__ , “Las Reducciones Toledanas en el Perú,” in “Pueblos de Indios: Otro Urbanismo en la Región Andina,” Quito: Ediciones ABYA-YALA, 1993, pp. 263–316.

Michalopoulos, Stelios and Elias Papaioannou, “Pre-Colonial Ethnic Institutions and Contemporary African Development,” *Econometrica*, 2013, *81* (1), 113–152.

Miguel, Edward and Mary Kay Gugerty, “Ethnic Diversity, Social Sanctions, and Public Goods in Kenya,” *Journal of Public Economics*, 2005, *89* (11-12), 2325–2368.

Montalvo, José G and Marta Reynal-Querol, “Colonization, Early Settlers and Development: The Case of Latin America,” 2020. BSE Working Paper: 1189.

Montalvo, Jose G and Marta Reynal-Querol, “Ethnic Diversity and Growth: Revisiting the Evidence,” *Review of Economics and Statistics*, 2021.

Murra, John V., “The Economic Organization of the Inca State,” Technical Report, The University of Chicago 1956.

__ , *Formaciones Económicas y Políticas del Mundo Andino*, Lima: Instituto de Estudios Peruanos, 1975.

__ , “Did Tribute and Markets Prevail in the Andes before the European Invasion?,” in “Ethnicity, Markets, and Migration in the Andes: At the Crossroads of History and Anthropology,” Duke University Press, 1995, pp. 57–72.

__ , *El Mundo Andino: Población, Medio Ambiente y Economía*, Vol. 24, Lima: Instituto de Estudios Peruanos, 2002.

__ , “Las Sociedades Andinas antes de 1532,” in “América Latina en la Época Colonial,” Vol. 1, Barcelona: Crítica, 2002, pp. 56–82.

Nakatsuka, Nathan, Iosif Lazaridis, Chiara Barbieri, Pontus Skoglund, Nadin Rohland, Swapan Mallick, Cosimo Posth, Kelly Harkins-Kinkaid, Matthew Ferry, Éadaoin Harney et al., “A Paleogenomic Reconstruction of the Deep Population History of the Andes,” *Cell*, 2020, *181* (5), 1131–1145.

Nunn, Nathan, “The Long Term Effects of Africa’s Slave Trades,” *Quarterly Journal of Economics*, 2008, *123* (1), 139–176.

__ , “The Historical Roots of Economic Development,” *Science*, 2020, *367* (6485), eaaz9986.

__ **and Diego Puga**, “Ruggedness: The Blessing of Bad Geography in Africa,” *Review of Economics and Statistics*, 2012, *94* (1), 20–36.

— **and Leonard Wantchekon**, "The Slave Trade and the Origins of Mistrust in Africa," *American Economic Review*, 2011, *101* (7), 3221–52.

Pease, Franklin, *Del Tawantinsuyu a la Historia del Perú*, Lima: Fondo Editorial de la Pontificia Universidad Católica del Perú, 1989.

— , *Perú: Hombre e Historia. Entre el siglo XVI y el XVIII*, Lima: Edubanco, Fundación del Banco Continental para el Fomento de la Educación y la Cultura, 1992.

Pique, Ricardo, "Higher Pay, Worse Outcomes? The Impact of Mayoral Wages on Local Government Quality in Peru," *Journal of Public Economics*, 2019, *173*, 1–20.

Porta, Rafael La, Florencio Lopez de Silanes, Andrei Shleifer, and Robert Vishny, "The Quality of Government," *The Journal of Law, Economics, and Organization*, 1999, *15* (1), 222–279.

Rodríguez, Antonio Dougnac, *Manual de Historia del Derecho Indiano*, Universidad Nacional Autónoma de México, Instituto de Investigaciones Jurídicas, 1994.

Rohner, Dominic, Mathias Thoenig, and Fabrizio Zilibotti, "War Signals: A Theory of Trade, Trust, and Conflict," *Review of Economic Studies*, 2013, *80* (3), 1114–1147.

Rostworowski, María, "Las Macroetnias en el Ámbito Andino," *Allpanchis*, 1990, *22* (35/36), 3–28.

Rowe, John Howland, "Inca Culture at the Time of the Spanish Conquest," in "Handbook of South American Indians," Vol. 2, US Government Printing Office, 1946, pp. 183–330.

Rubio-Ramos, Melissa, Christian Isendahl, and Ola Olsson, "The Political Economy of Bread and Circuses: Weather Shocks and Classic Maya Monument Construction," 2024. Working Paper.

Saignes, Thierry, "Lobos y Ovejas. Formación y Desarrollo de los Pueblos y Comunidades en el Sur Andino (Siglos XVI-XX)," in "Reproducción y Transformación de las Sociedades Andinas, Siglos XVI-XX," Quito: Ediciones ABYA-YALA, 1991, pp. 91–137.

Sandweiss, Daniel H and James B Richardson, "Central Andean Environments," in "The Handbook of South American Archaeology," Springer, 2008, pp. 93–104.

Soldán, Mariano Felipe Paz, *Diccionario Geográfico Estadístico del Perú*, Lima: Imprenta del Estado, 1877.

Stanish, Charles, "Household Archeology: Testing Models of Zonal Complementarity in the South Central Andes," *American Anthropologist*, 1989, *91* (1), 7–24.

— , "The Origin of State Societies in South America," *Annual Review of Anthropology*, 2001, *30* (1), 41–64.

Stern, Steve J., "The Variety and Ambiguity of Native Andean Intervention in European Colonial Markets," in "Ethnicity, Markets, and Migration in the Andes: At the Crossroads of History and Anthropology," Duke University Press, 1995, pp. 73–100.

Tabellini, Guido, "The Scope of Cooperation: Values and Incentives," *The Quarterly Journal of Economics*, 2008, *123* (3), 905–950.

Tello, Julio César, "Origen y Desarrollo de las Civilizaciones Prehistóricas Andinas," in "Actas y Trabajos. XXVII Congreso Internacional de Americanistas," Vol. 1, Lima: Librería e Imprenta Gil, 1939, pp. 589–720.

Tung, Tiffiny A and Kelly J Knudson, "Stable Isotope Analysis of a Pre-Hispanic Andean Community: Reconstructing Pre-Wari and Wari era Diets in the Hinterland of the Wari empire, Peru," *American Journal of Physical Anthropology*, 2018, *165* (1), 149–172.

Turner, Bethany L, John D Kingston, and George J Armelagos, "Variation in Dietary Histories Among the Immigrants of Machu Picchu: Carbon and Nitrogen Isotope Evidence," *Chungara, Revista de Antropología Chilena*, 2010, *42* (2), 515–534.

Vidal, Javier Pulgar, "Las Ocho Regiones Naturales del Perú," in "Boletín del Museo de Historia Natural "Javier Prado"" number 17, Lima: Imprenta D. Miranda, 1941, pp. 145–160.

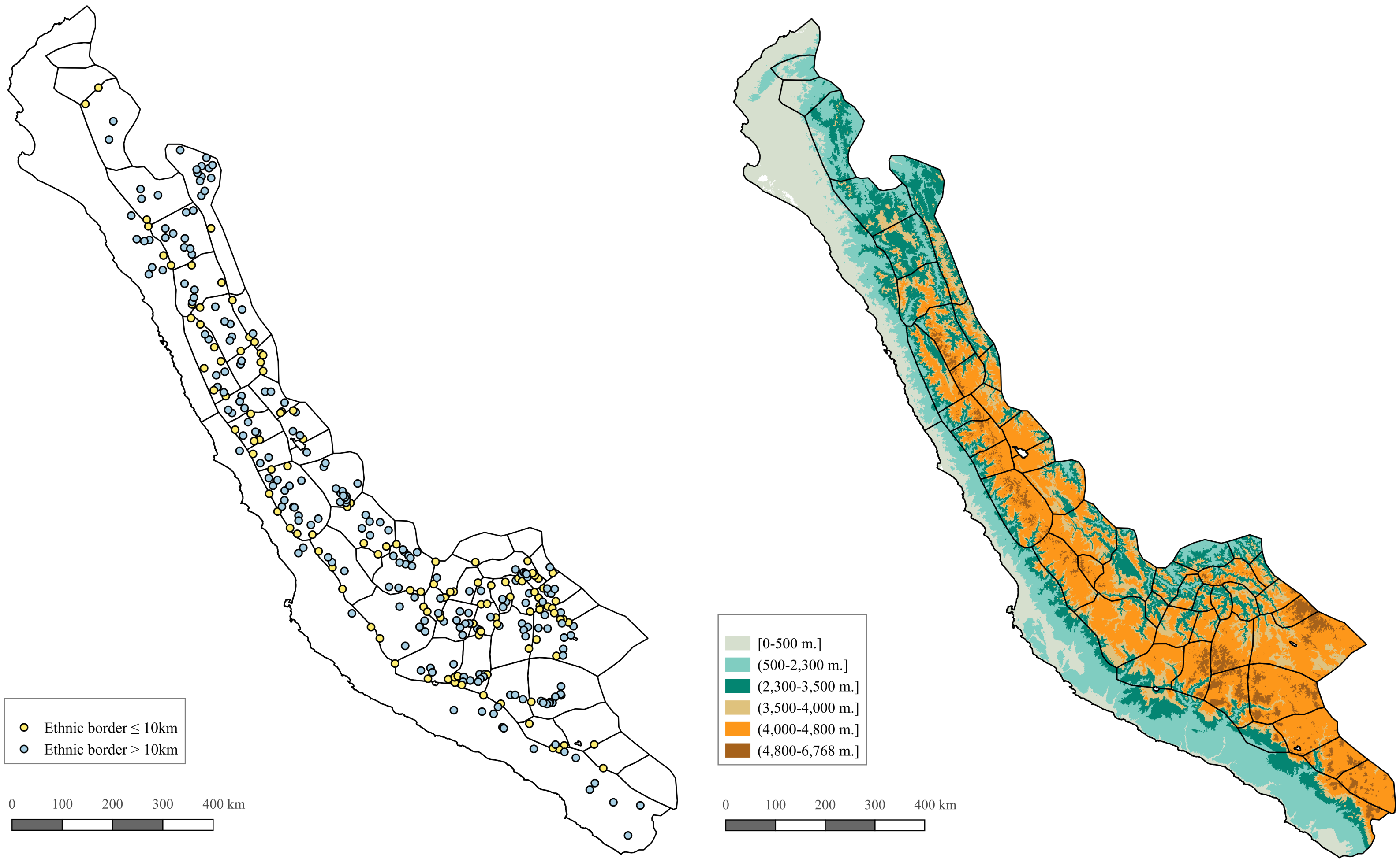
Voigtländer, Nico and Hans-Joachim Voth, "Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany," *The Quarterly Journal of Economics*, 2012, *127* (3), 1339–1392.

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

Wachtel, Nathan, *Los Vencidos. Los Indios del Perú frente a la Conquista Española (1530-1570)*, Madrid: Alianza Editorial, 1976.

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 Codding, "Climate and Demography Drive 7000 Years of Dietary Change in the Central Andes," *Scientific Reports*, 2022, *12* (1), 1–16.

Young, Alwyn, "Asymptotically Robust Permutation-based Randomization Confidence Intervals for Parametric OLS Regression," *European Economic Review*, 2024, *163*, 104644.



(a) Colonial Parishes

(b) Complementary Elevation Zones

FIGURE 1: Colonial Parishes and Elevation Zones

Notes. Black lines represent the territorial extent of pre-colonial groups at the time of the Spanish conquest, based on Rowe (1946)'s mapping. The dots in Panel (a) represent the capitals of colonial parishes: those with an ethnic border within a buffer of 10-km radius are displayed in yellow (using equidistant boundaries when the distance between parish capitals is less than 10 km to ensure non-overlapping buffers); the rest are displayed in blue. Panel (b) shows the distribution of complementary elevation zones (Pulgar Vidal 1941). Elevation intervals refer to meters above sea level. For elevation data, I use version 1.2 of the Harmonized World Soil Database (Fischer et al. 2008). It provides 30 arc-second raster data with median elevation constructed based on information from the NASA Shuttle Radar Topographic Mission. The maps are displayed using a World Geodetic System projection (WGS 1984).

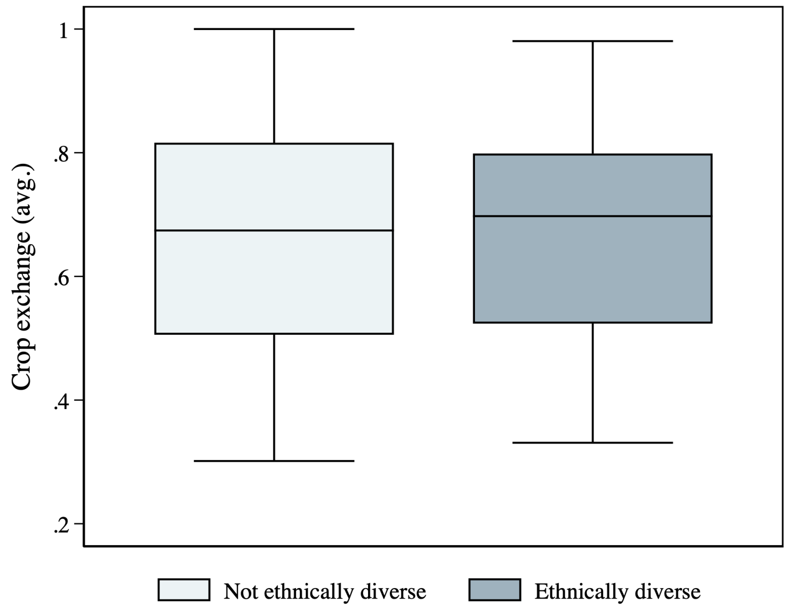


FIGURE 2: Ethnic Diversity and Historical Crop Exchange

Notes. Boxplots of 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. 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.

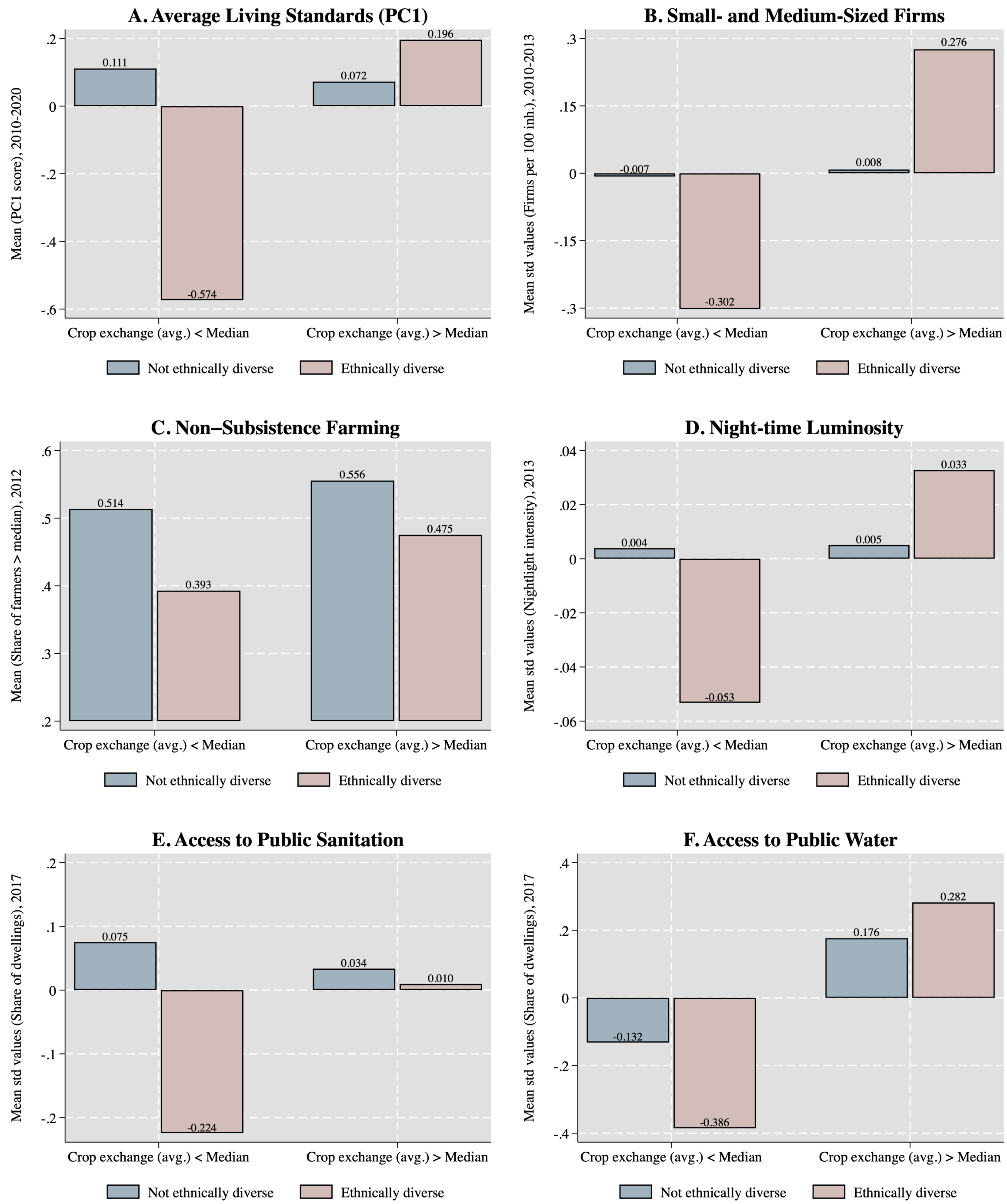


FIGURE 3: Ethnic Diversity, Historical Crop Exchange, and Contemporary Development

Notes. Mean development outcomes by ethnic diversity status and a dummy for whether the parish-level proxy for average exposure to internal crop exchange is above the sample median. Graph A shows the mean score of the first principal component 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), averaged across parishes. All variables, except for the dummy variable indicating non-subsistence agriculture, are standardized to have a mean of zero and a standard deviation of one. Graphs B to F plot the raw data for each variable, separately.

Table 1: Balance Tests for Ethnic Diversity

	Border Parishes		Interior Parishes				
	mean	sd	mean	sd	Diff.	p-value[a]	p-value[b]
(1) Mean elevation	3480.341	529.292	3406.865	734.882	-73.476	[0.293]	[0.487]
(2) SD of elevation	479.734	188.314	448.043	178.185	-31.691	[0.136]	[0.215]
(3) Mean caloric suitability	126.104	275.867	117.969	262.657	-8.135	[0.794]	[0.844]
(4) SD of caloric suitability	138.934	241.901	122.805	219.042	-16.128	[0.548]	[0.639]
(5) Ln dist. to perennial river	0.673	1.054	0.744	1.085	0.070	[0.564]	[0.646]
(6) Ln dist. to native shrine	4.160	0.999	4.310	1.090	0.150	[0.205]	[0.384]
(7) Ln expected tribute (16th c.)	6.510	0.723	6.492	0.641	-0.017	[0.830]	[0.777]
(8) Ln dist. to *mita* mine	5.667	0.744	5.702	0.727	0.035	[0.679]	[0.735]
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). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2: Historical Crop Exchange — Pre-Colonial Correlates

	Dependent Variable:								
	Mean Elevation	Mean Caloric Suit.	Ln River Density	Ln Land Area	Ln Population	Ln Population Density	Dummy Urbanization	Dummy Political Complexity	Dummy Elite Residences
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Crop exchange (H index)	-0.268	0.203	0.041	-0.205	0.119	0.304	0.104	0.155	0.036
	[0.193]	[0.126]	[0.120]	[0.148]	[0.138]	[0.121]**	[0.045]**	[0.050]***	[0.053]
	(0.234)	(0.150)	(0.123)	(0.198)	(0.140)	(0.108)***	(0.048)**	(0.064)**	(0.059)
Number of ethnic groups	47	47	47	47	46	46	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 for dummies are standardized to have zero mean and standard deviation equal to one. Population is available for 46 (out of 47) groups; see Appendix D. The dummy variables for urbanization, political complexity, and elite residences take value 1 for 12.77, 21.28, and 21.28 percent of the groups, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE 3: Ethnic Diversity, Historical Crop Exchange, and Contemporary Development I

	Living Standards (AES, 1990 – 2000)							Living Standards (AES, 2010 – 2020)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ethnic diversity	-0.200***	-0.165**	-0.689***	-0.556**	-0.604***			-0.610***	-0.627***	-0.561***
	[0.070]	[0.066]	[0.243]	[0.247]	[0.206]			[0.225]	[0.224]	[0.209]
$\overline{\text{Crop exchange}}$			0.080	0.253	0.226	0.468**	0.514*	0.299	0.245	-0.074
			[0.213]	[0.235]	[0.226]	[0.183]	[0.277]	[0.318]	[0.322]	[0.331]
Ethnic div. $\times$ $\overline{\text{Crop exchange}}$			0.781**	0.659**	0.789***			0.846**	0.870**	0.836**
			[0.349]	[0.321]	[0.303]			[0.350]	[0.352]	[0.343]
Heterogeneity in elevation zones									0.277	0.403
									[0.323]	[0.294]
Joint significance (p-value)			0.003	0.020	0.008			0.018	0.012	0.021
Baseline controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ecclesiastical jurisd. FE	No	No	No	Yes	No	Yes	No	No	No	No
Colonial province FE	No	No	No	No	Yes	No	Yes	Yes	Yes	Yes
Extended set of controls	No	No	No	No	No	No	No	No	No	Yes
Number of parishes	336	336	336	336	336	336	336	336	336	336

Notes. The unit of observation is the parish. Columns 1-7 report the standardized AES (Kling et al. 2004; Clingingsmith, Khwaja and Kremer 2009) across four outcomes measured for the decade 1990–2000: a dummy variable indicating whether the share of farmers practicing non-subsistence agriculture is above the sample median, the log of 1 plus average light intensity per capita, the share of dwellings with access to public sanitation, and the share of dwellings with access to the public water network. Columns 8-10 report the standardized AES for the decade 2010–2020, using the same four variables along with the average number of small- and medium-sized firms per 100 inhabitants. Robust standard errors in columns 1-3. Robust standard errors clustered at the ecclesiastical jurisdiction (columns 4 and 6) or province (columns 5 and 7-10) level 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 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 columns 9-10, the specification includes the H index of heterogeneity in elevation zones within the buffer of 10-km radius from the parish capital as a control variable. In column 10, the specification also includes the set of parish-level pre-colonial characteristics used in Table A.1 as additional control variables. The p-value refers to the joint significance of ethnic diversity terms. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Ethnic Diversity, Historical Crop Exchange, and Contemporary Development II

	(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	-0.107	-11.991	-0.416	-0.131
	[0.016]**	[0.033]***	[7.041]*	[0.163]**	[0.069]*
	(0.018)**	(0.037)***	(6.664)*	(0.127)***	(0.049)***
Ethnic div. × $\overline{\text{Crop exchange}}$	0.042	0.119	19.147	0.593	0.151
	[0.021]**	[0.042]***	[10.425]*	[0.233]**	[0.100]
	(0.023)*	(0.041)***	(9.650)**	(0.178)***	(0.065)**
Mean Dep. Var.	0.023	0.056	23.992	0.500	0.369
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	-0.096	-0.301	-0.386	-0.238
	[0.086]	[0.106]	[0.161]*	[0.202]*	[0.168]
	(0.057)	(0.060)	(0.156)*	(0.173)**	(0.190)
Ethnic div. × $\overline{\text{Crop exchange}}$	0.122	0.134	0.491	0.601	0.420
	[0.124]	[0.175]	[0.228]**	[0.282]**	[0.235]*
	(0.096)	(0.084)	(0.219)**	(0.198)***	(0.272)
Mean Dep. Var.	0.763	0.468	8.098	7.918	8.217
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. Robust standard errors in brackets are clustered at the province level in columns 1-7 and at the parish level in columns 8-10. 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 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 5: Pre-Colonial Characteristics of Ethnic Groups

	Living Standards (AES, 2010 – 2020)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ethnic diversity	-0.610***	-0.973***	-0.708***	-1.249*	-0.523**	-0.598***	-0.625***	-0.592***	-1.303	-1.318
	[0.225]	[0.309]	[0.237]	[0.751]	[0.236]	[0.225]	[0.225]	[0.219]	[1.083]	[1.296]
Ethnic div. × $\overline{\text{Crop exchange}}$	0.846**	0.839***	1.099***	0.706**	0.801**	0.863**	0.937**	0.842**	1.175***	1.034**
	[0.350]	[0.325]	[0.376]	[0.321]	[0.347]	[0.355]	[0.375]	[0.347]	[0.440]	[0.420]
Ethnic div. × $\overline{\text{Elevation}}$		0.475*							-0.161	-0.348
		[0.286]							[0.450]	[0.459]
Ethnic div. × $\overline{\text{Caloric suitability}}$			-0.951						-1.328	-1.310
			[0.614]						[0.926]	[0.906]
Ethnic div. × $\overline{\text{Ln river density}}$				0.165					0.174	0.227
				[0.168]					[0.256]	[0.269]
Ethnic div. × $\overline{\text{Ln population density}}$					0.078				0.047	0.026
					[0.095]				[0.123]	[0.136]
Ethnic div. × $\overline{\text{Urbanization}}$						-0.115			0.002	
						[0.170]			[0.319]	
Ethnic div. × $\overline{\text{Political complexity}}$							-0.111		-0.184	-0.147
							[0.143]		[0.273]	[0.208]
Ethnic div. × $\overline{\text{Elite residences}}$								-0.039	0.113	
								[0.122]	[0.231]	
Ethnic div. × $\overline{\text{Water canals}}$										-0.034
										[0.149]
Ethnic div. × $\overline{\text{Ln 1+ road density}}$										0.005
										[0.070]
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lasso characteristics	No	No	No	No	No	No	No	No	No	Yes
Number of parishes	336	336	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. 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; 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 et al. 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 6: Coarsened Exact Matching

	Living Standards (AES, 2010 – 2020)					
	CEM Matched Sample (Baseline Characteristics)		CEM Matched Sample (Lasso Characteristics)		Full Sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Ethnic diversity	-0.156	-0.554*	-0.565**	-0.769**	-0.228***	-0.326***
	[0.126]	[0.283]	[0.233]	[0.391]	[0.074]	[0.095]
Ethnic div. × High $\overline{\text{Crop exchange}}$	0.433***	0.524***	0.883***	0.906***	0.354***	0.360***
	[0.154]	[0.194]	[0.280]	[0.304]	[0.133]	[0.135]
Number of parishes	90	90	74	74	336	336
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes	Yes	Yes	Yes
Majority group weight	No	Yes	No	Yes	No	Yes

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 matched sample refers to the sample of parishes selected by the CEM algorithm of Iacus et al. (2012), using parishes with average 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. For the weighted averages of log population density, log river density, and log road density, the CEM algorithm uses the sample median as cutoff point to define categories. For the weighted averages of urbanization, political complexity, and elite residences, the algorithm uses {0.2, 0.4, 0.6, 0.8} as cutoff points. 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. Columns 2, 4 and 6 control for the area share of the ethnic group with the highest share of area within the 10-km buffer. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: Mechanisms — Societal Integration and Cooperative Behavior

	Dependent Variable:								
	Which Group Do You Identify With The Most? (2004-2017)			Voting in Presidential Elections (2007-2011)		Military Service (2008-2014)		Dummy Neighborhood Associations	
	State	Ethnicity or Race	Religion	Dummy (Yes, I Voted in the 2006 Election)		Ln (1 + Av. Volunteers)	Dummy Volunteers	2002-2004	2012-2014
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ethnic diversity	-0.151	0.020	0.001	-0.057	-0.057	-0.122	-0.231	-0.440	-0.534
	[0.047]***	[0.013]	[0.027]	[0.035]	[0.035]	[0.062]*	[0.104]**	[0.208]**	[0.162]***
	(0.054)***	(0.018)	(0.030)	(0.043)	(0.043)	(0.048)**	(0.074)***	(0.226)*	(0.228)**
Ethnic div × $\overline{\text{Crop exchange}}$	0.202	-0.026	0.006	0.110	0.110	0.235	0.367	0.698	0.718
	[0.067]***	[0.018]	[0.038]	[0.046]**	[0.046]**	[0.108]**	[0.145]**	[0.315]**	[0.234]***
	(0.083)**	(0.024)	(0.048)	(0.054)**	(0.054)**	(0.093)**	(0.122)***	(0.345)**	(0.326)**
Trust in the state					0.011				
					[0.012]				
					(0.012)				
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	–	–	–	–
Year FE	Yes	Yes	Yes	Yes	Yes	–	–	–	–
Number of parishes	280	280	280	186	186	336	336	336	336
Number of individuals	52,875	52,875	52,875	17,422	17,422	–	–	–	–
Mean Dep. Var.	0.511	0.037	0.169	0.831	0.831	0.064	0.134	0.372	0.354

Notes. The unit of observation is the individual in columns 1-5 (yearly data from the ENAHO Peruvian household surveys; 2004-2017 waves) and the parish in columns 6-9. Robust standard errors clustered at the parish (columns 1-5) or province (columns 6-9) levels 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 columns 1-3, the dependent variable is a dummy for whether 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; column 1), ethnicity or race (column 2), and religion (column 3). The question on having voted in the 2006 presidential election (columns 4-5) was only asked in the 2007-2011 waves of the survey, and it was not repeated for other presidential elections. 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 1-5, the vector of individual-level controls includes gender, age, age squared, years of schooling, civil status, and language spoken at home. 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 8: Mechanisms — A Market-Oriented Agricultural Economy

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Economic Complementarities					
	Living Standards (AES)		Dummy Agricultural Retail Market (1993)		Access to Agricultural Land (Mean Share of Pop, 1700 – 1800)	
	(1990 – 2000)	(2010 – 2020)				
High min	0.926***	0.546**	0.340**	0.293*	0.430***	0.372***
	[0.337]	[0.215]	[0.165]	[0.153]	[0.141]	[0.109]
High maj	0.339	0.280	-0.065	-0.099	0.149	0.086
	[0.245]	[0.198]	[0.072]	[0.080]	[0.168]	[0.121]
High min × High maj	-0.443	-0.064	-0.185	-0.196	-0.227	-0.211
	[0.638]	[0.383]	[0.154]	[0.147]	[0.180]	[0.164]
Ln population (1993)				0.077*		
				[0.038]		
Ln mean population (1700 – 1800)						-0.042
						[0.032]
Number of parishes	117	117	117	117	51	51
Mean Dep. Var.	–	–	0.077	0.077	0.793	0.793
	Panel B: Tertiary-Sector Occupation (Share of Pop)					
	Tertiary Sector		Tertiary Sector: Local Trade		Tertiary Sector: Other Services	
	(1876)	(2007 – 2017)	(1876)	(2007 – 2017)	(1876)	(2007 – 2017)
Ethnic diversity	-0.113***	-0.249**	-0.112***	-0.108**	-0.001	-0.076*
	[0.041]	[0.103]	[0.040]	[0.044]	[0.003]	[0.038]
Ethnic div. × $\overline{\text{Crop exchange}}$	0.153**	0.339*	0.154**	0.147**	-0.001	0.105
	[0.067]	[0.169]	[0.066]	[0.066]	[0.004]	[0.066]
Number of parishes	282	336	282	336	282	336
Mean Dep. Var.	0.073	0.336	0.064	0.124	0.009	0.105
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The unit of observation is the parish. Robust standard errors clustered at the level of the colonial province in brackets. *High min* is a dummy variable indicating whether the minority belonged to a group with a high degree of exposure to crop exchange (H_e above the 75th percentile). Variable *High maj* is an analogous dummy for the majority group. 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. In Panel A, the sample includes parishes with ethnic diversity only. In Panel B, 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A Supplemental Appendix - Figures and Tables

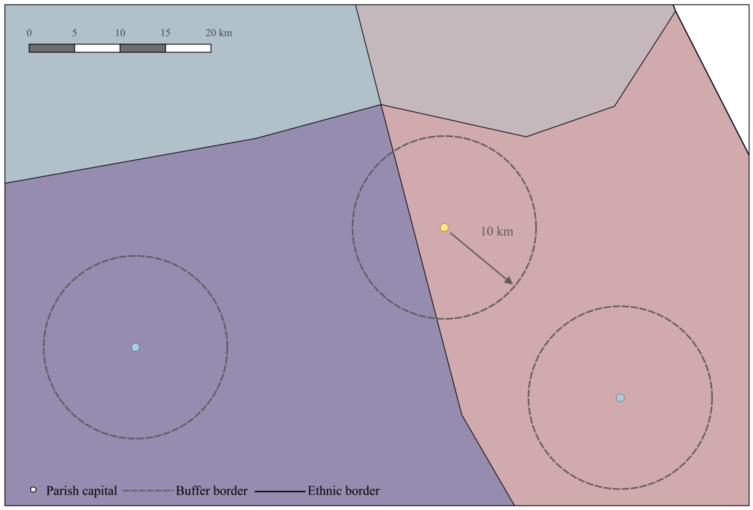


FIGURE A.1: Buffer Exercise

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

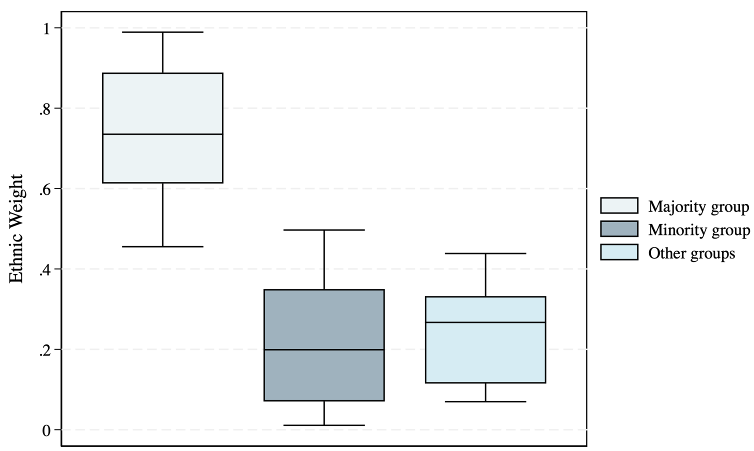


FIGURE A.2: Boxplots of Ethnic Weights

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).

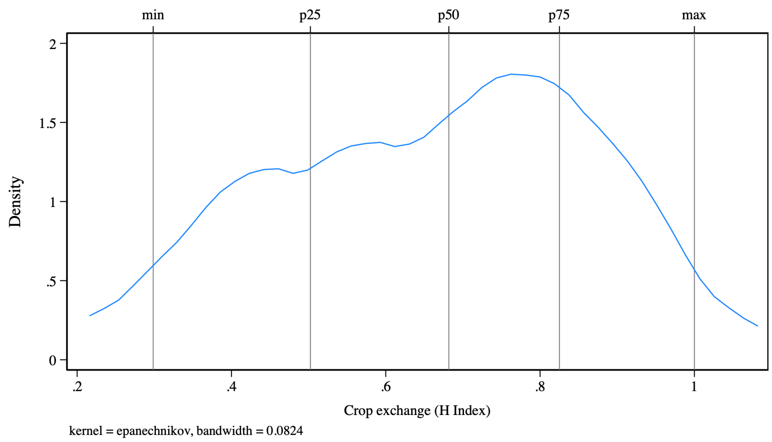


FIGURE A.3: H Index Across Ethnic Groups (Kernel Density)

Notes. Kernel density of the ethnic-level proxy for potential exposure to internal crop exchange (H index), defined as the 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.

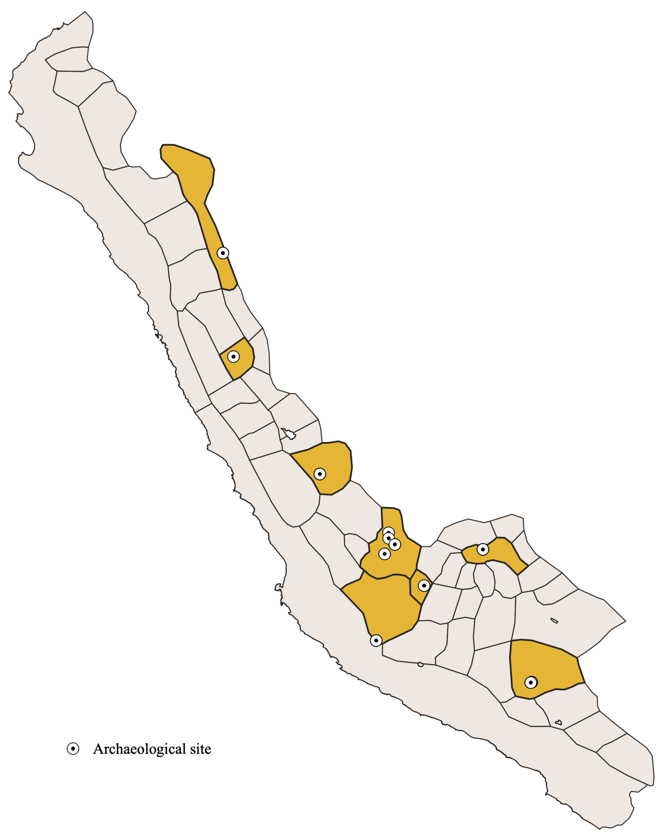


FIGURE A.4: Archaeological Sites with Information on Pre-Colonial Diets

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.

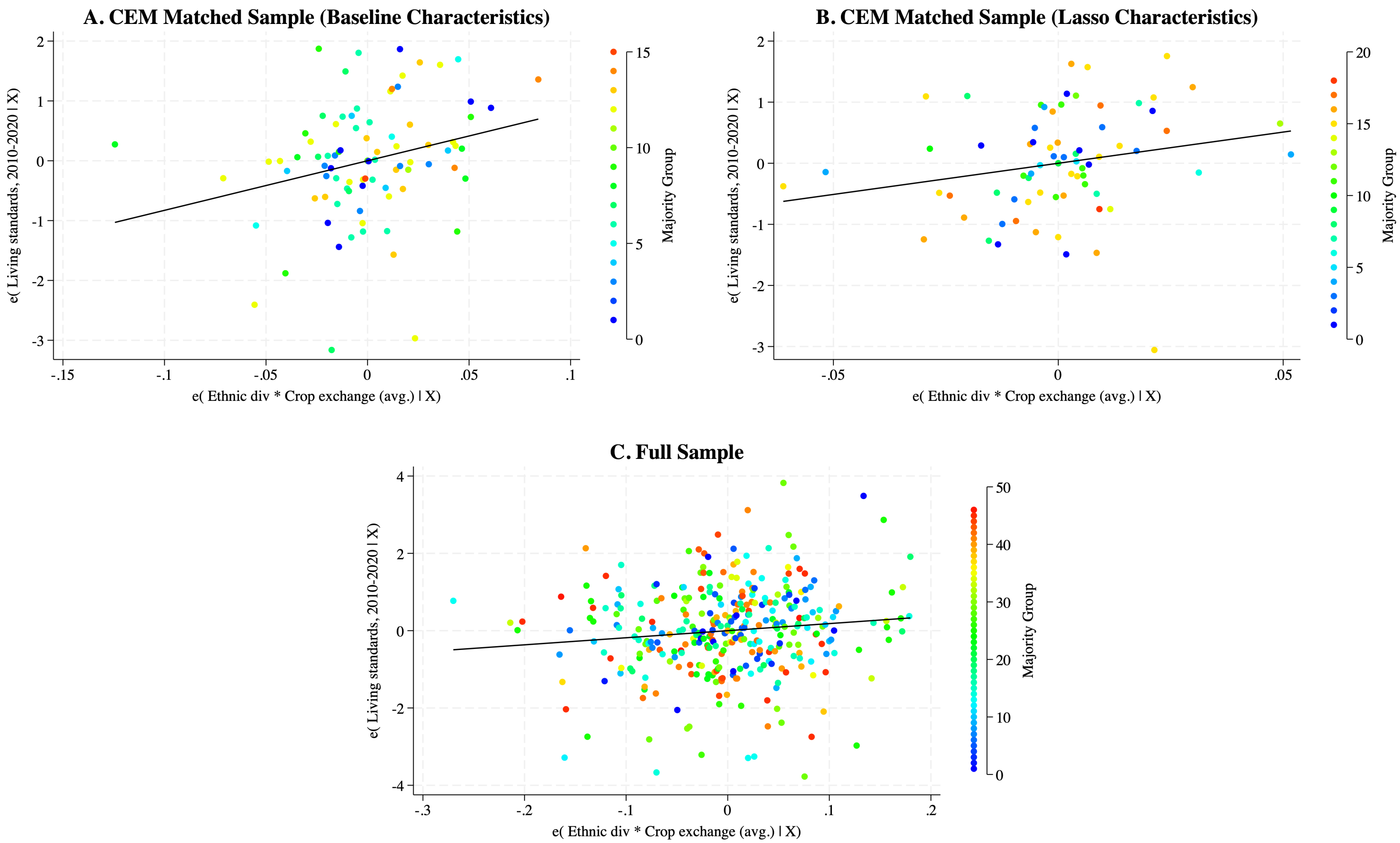


FIGURE A.5: Partial Correlation Plots

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).

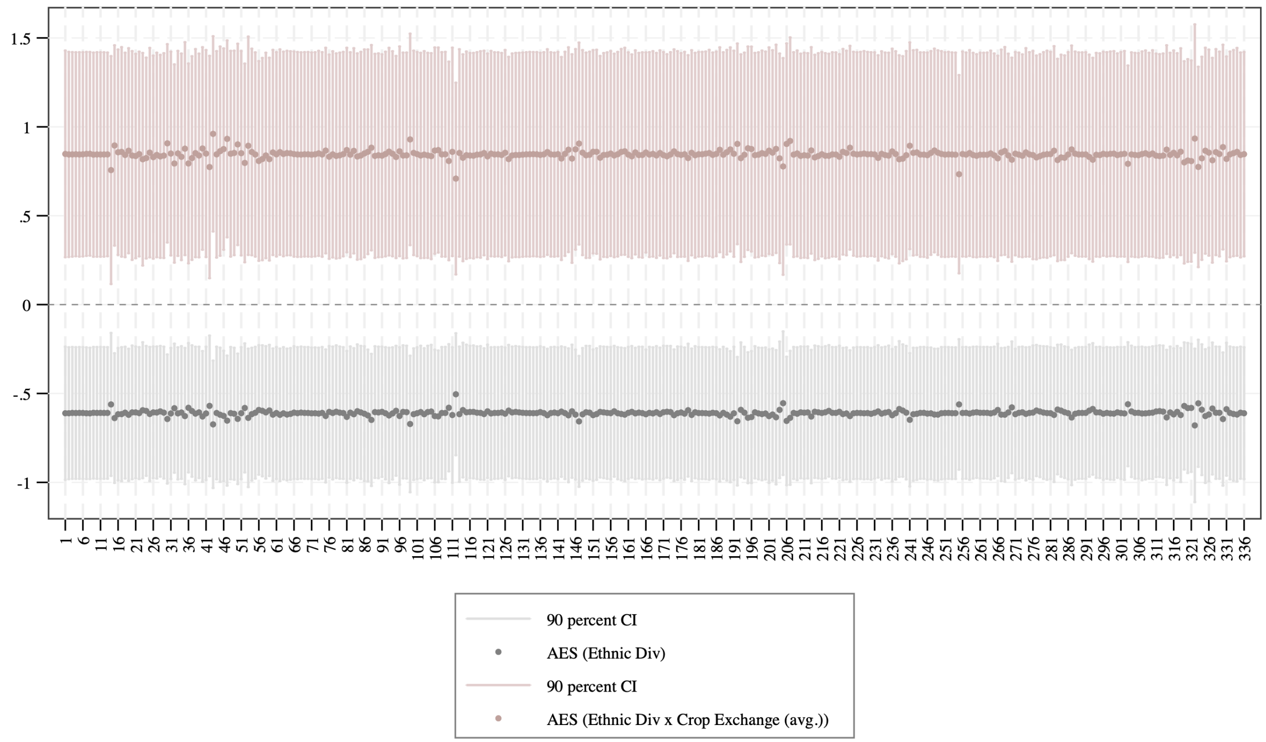
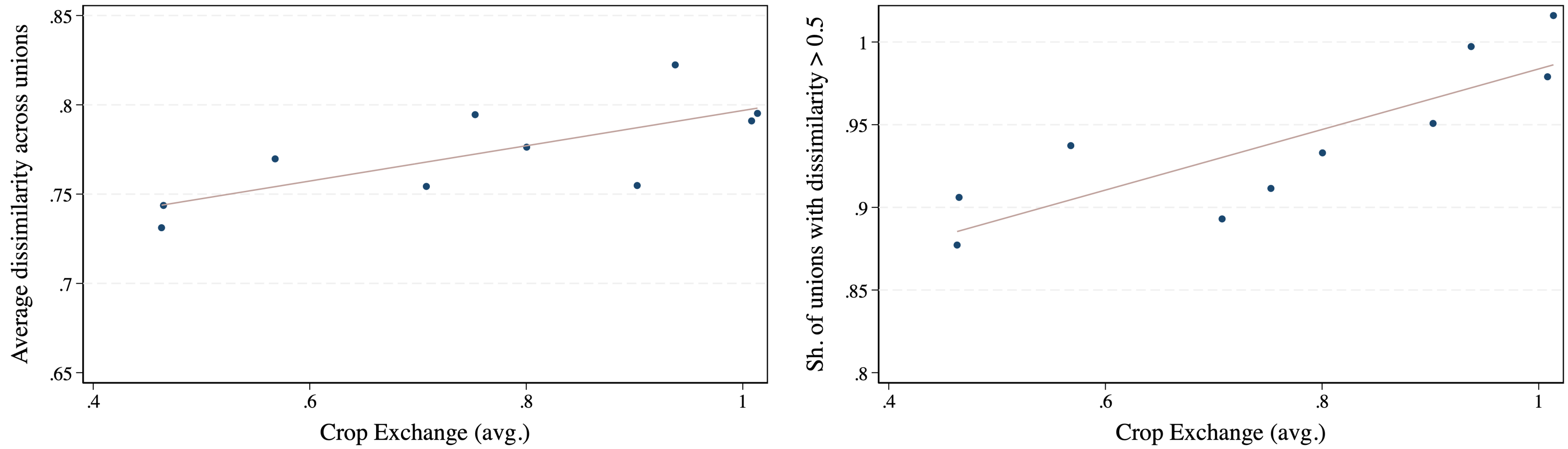
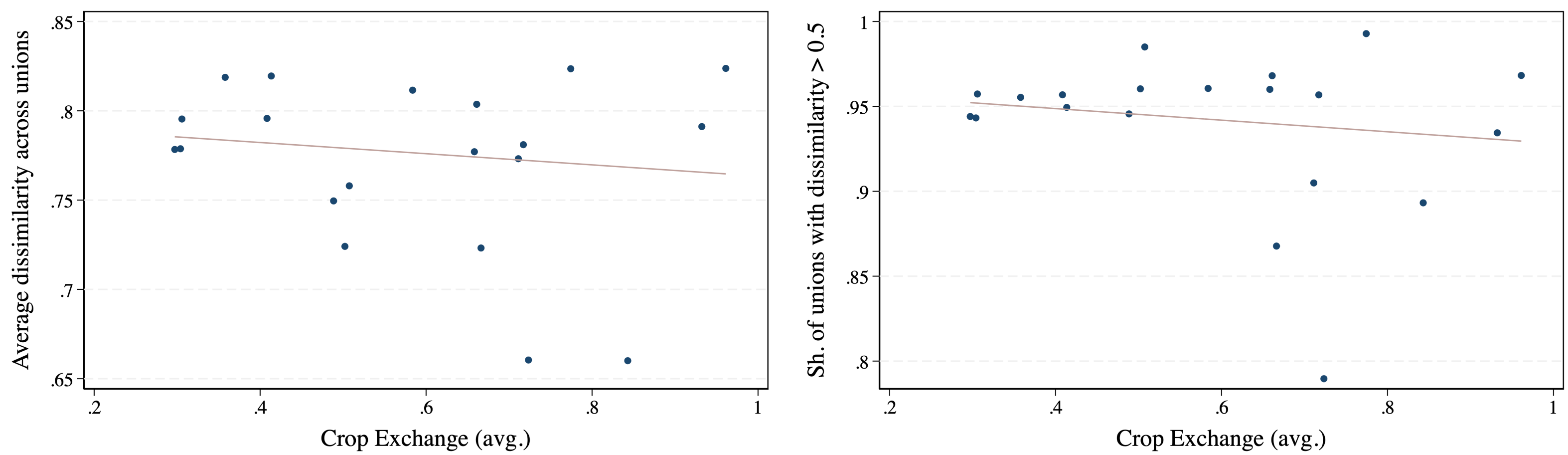


Figure A.6: Influential Observations

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.



(a) **Ethnically diverse** (10 parishes)



(b) **Not ethnically diverse** (31 parishes)

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				
	mean	sd	mean	sd	Diff.	p-value[a]	p-value[b]
(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.1 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 km × 25 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 km × 50 km grid cells and ethnic groups (H_e), respectively.

[51] I 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.

Table A.2: Validating Elevation Zones

	Dependent Variable: Ln (# Native Crops)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
H index	0.504		0.461	0.448		0.445	0.527
	[0.036]***		[0.044]***	[0.057]***		[0.075]***	[0.106]***
	(0.045)***		(0.058)***	(0.057)***		(0.090)***	(0.086)***
SD of elevation		0.418	0.062	0.074	0.021	0.060	0.135
		[0.040]***	[0.052]	[0.057]	[0.060]	[0.120]	[0.141]
		(0.063)***	(0.081)	(0.062)	(0.059)	(0.124)	(0.157)
Ln land area	0.186	0.183	0.181	0.133	0.136	0.399	0.387
	[0.033]***	[0.034]***	[0.033]***	[0.024]***	[0.033]***	[0.070]***	[0.126]***
	(0.027)***	(0.036)***	(0.029)***	(0.022)***	(0.031)***	(0.062)***	(0.124)***
Mean elevation	-0.052	0.146	-0.030	-0.100	-0.304	-0.002	0.134
	[0.044]	[0.044]***	[0.052]	[0.063]	[0.062]***	[0.106]	[0.130]
	(0.073)	(0.092)	(0.085)	(0.068)	(0.069)***	(0.117)	(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 × 25 km grid cell in Columns 1-5, the 50 km × 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	0.279	0.851	0.261	0.269	
	[0.153]**	[0.089]**	[0.145]***	[0.071]***	[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.1). *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	-0.114
	[0.145]	[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				
	mean	sd	mean	sd	Diff.	p-value^a	p-value^b
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	-0.008	-0.038
	[0.131]	[0.062]	[0.056]
	(0.143)	(0.046)	(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	-0.101	-0.074
	[0.118]	[0.196]	[0.138]
	(0.122)	(0.217)	(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**	-0.107***	-11.991	-0.416**	-0.131*
	[-2.080]	[-2.966]	[-1.585]	[-2.368]	[-1.765]
Ethnic div. × $\overline{\text{Crop exchange}}$	0.042*	0.119**	19.147*	0.593**	0.151
	[1.909]	[2.665]	[1.710]	[2.372]	[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	-0.096	-0.301*	-0.386*	-0.238
	[-0.954]	[-0.847]	[-1.866]	[-1.914]	[-1.415]
Ethnic div. × $\overline{\text{Crop exchange}}$	0.122	0.134	0.491**	0.601**	0.420*
	[0.913]	[0.713]	[2.156]	[2.129]	[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	-1.449	-3.333	-0.445*	-0.706***	-3.087
	[0.762]	[1.495]	[2.120]	[0.240]	[0.225]	[1.913]
Ethnic div. × $\overline{\text{Crop exchange}}$	0.768**	1.332**	1.423***	0.934***	1.117***	1.302**
	[0.348]	[0.604]	[0.547]	[0.332]	[0.365]	[0.662]
Ethnic div. × $\overline{\text{Ln land area}}$	-0.049	0.040	0.132			0.199
	[0.075]	[0.123]	[0.140]			[0.162]
Ethnic div. × $\overline{\text{Elevation}}$		0.319	0.165			0.623
		[0.458]	[0.457]			[0.615]
Ethnic div. × $\overline{\text{Caloric suitability}}$		-0.968	-1.036			-2.601
		[0.744]	[0.868]			[1.824]
Ethnic div. × $\overline{\text{Ln river density}}$			0.265			
			[0.265]			
Ethnic div. × $\overline{\text{Ln population density}}$			0.060			
			[0.124]			
Ethnic div. × $\overline{\text{Urbanization}}$		0.044	0.056			-0.224
		[0.298]	[0.334]			[0.410]
Ethnic div. × $\overline{\text{Political complexity}}$		-0.173	-0.275			-0.160
		[0.278]	[0.299]			[0.264]
Ethnic div. × $\overline{\text{Elite residences}}$		0.099	0.132			0.077
		[0.229]	[0.234]			[0.236]
Ethnic div. × $\overline{\text{SD of elevation}}$				-0.391		-0.208
				[0.327]		[0.586]
Ethnic div. × $\overline{\text{SD of caloric suitability}}$					-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. 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).
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

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

	Dependent Variable:							
	Mean Elevation	SD of Elevation	Mean Caloric Suit.	SD of Caloric Suit.	Ln Dist Perennial River	Ln Dist Native Shrine	Ln Expected Tribute (16th c.)	Ln Dist *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	17.070	-17.039	-25.414	-0.186	-0.190	-0.014	0.256
	[0.212]	[0.783]	[0.100]	[0.121]	[0.574]	[0.235]	[0.922]	[0.322]
Observations	90	90	90	90	90	90	90	90
Panel B: Matched Sample Based on CEM for Lasso-*Selected* Pre-Colonial Characteristics								
Ethnic diversity	-33.329	-10.190	-15.592	-27.970	-0.394	-0.028	0.304**	-0.086
	[0.710]	[0.852]	[0.283]	[0.225]	[0.450]	[0.941]	[0.033]	[0.834]
Observations	74	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

	Non-subs. farming			Light intensity per capita			Firms per 100 inh.		
	Randomization-t p-value			Randomization-t p-value			Randomization-t p-value		
	Minimum	Maximum	Randomized	Minimum	Maximum	Randomized	Minimum	Maximum	Randomized
Ethnic diversity	0.020	0.021	0.020	0.046	0.047	0.046	0.119	0.120	0.119
Ethnic div. $\times \overline{\text{Crop exchange}}$	0.025	0.026	0.025	0.068	0.069	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***	-0.742***	-0.645***	-0.669**		
	[0.233]	[0.273]	[0.247]	[0.305]		
Ethnic div. × $\overline{\text{Crop exchange}}$ (20km correction)	0.886**					
	[0.360]					
Ethnic div. × $\overline{\text{Crop exchange}}$ (10km correction)		0.985**				
		[0.394]				
Ethnic div. × $\overline{\text{Crop exchange}}$ (20km transition-zone buffer)			0.877**			
			[0.374]			
Ethnic div. × $\overline{\text{Crop exchange}}$ (10km transition-zone buffer)				0.864**		
				[0.435]		
Dummy (Artificial ethnic border within parish buffer)					-0.260	
					[0.395]	
Dummy × $\overline{\text{Crop exchange}}$ (Artificial)					0.399	
					[0.498]	
Dummy (*Corregimiento* border within parish buffer)						0.165
						[0.246]
Dummy × $\overline{\text{Crop exchange}}$ (*Corregimiento*)						-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	-0.341	-0.435**	-0.610***	-0.394**	-0.454***	-0.567***	-0.580***	-0.533**
	[0.229]	[0.212]	[0.189]	[0.225]	[0.181]	[0.170]	[0.168]	[0.194]	[0.211]
Ethnic div. $\times$ $\overline{\text{Crop exchange}}$	0.417	0.449	0.598**	0.846**	0.525*	0.673**	0.890***	0.905***	0.793**
	[0.365]	[0.333]	[0.304]	[0.350]	[0.296]	[0.269]	[0.286]	[0.317]	[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; 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.14: Robustness—Alternative Diversity Indices

	Living Standards (AES, 2010 – 2020)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ethnic fractionalization	-1.461***	-0.999						
	[0.557]	[0.641]						
Ethnic frac. × $\overline{\text{Crop exchange}}$	1.741**	1.476						
	[0.795]	[0.982]						
Ethnic diversity (dummy)			-0.959***	-0.612**	-0.681***	-0.496**	-0.904***	-0.596***
			[0.221]	[0.269]	[0.250]	[0.214]	[0.241]	[0.223]
Ethnic div. × $\overline{\text{Crop exchange}}$ ($\widetilde{H}$ index)			1.475***	1.014**				
			[0.403]	[0.495]				
Ethnic div. × $\overline{\text{Crop exchange}}$ (majority group)					0.812**	0.660**		
					[0.371]	[0.324]		
Ethnic div. × $\overline{\text{Crop exchange}}$ (minority group)							1.218***	0.887***
							[0.340]	[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 $\textit{ethnic frac}_p = 1 - \sum_e w_{ep}^2$. 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, $\textit{Crop exchange}_p = \sum_e w_{ep} H_e$, where $H_e = 1/\sum_j s_{je}^2$; columns 3-4 use a Herfindahl index to measure H_e instead, specifically $\widetilde{H}_e = 1 - \sum_j s_{je}^2$; 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; 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.15: Robustness—Inca Period and Pre-Resettlement Spread of Smallpox

	Living Standards (AES, 2010 – 2020)			
	(1)	(2)	(3)	(4)
Ethnic diversity	-0.610***	-0.504	-0.615***	-0.539**
	[0.225]	[0.393]	[0.229]	[0.263]
Ethnic div. × $\overline{\text{Crop exchange}}$	0.846**	0.865**	0.845**	0.786**
	[0.350]	[0.351]	[0.354]	[0.399]
Ethnic div. × $\overline{\text{Ln dist. smallpox outbreak}}$		-0.022		
		[0.065]		
Baseline controls	Yes	Yes	Yes	Yes
Colonial province FE	Yes	Yes	Yes	Yes
Inca region (*suyu*) 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; 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). 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.

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

A.2 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	-0.279	-0.269	-0.268	-0.266
	[0.149]*	[0.144]*	[0.142]*	[0.140]*	[0.139]*
	(0.147)*	(0.145)*	(0.143)*	(0.143)*	(0.142)*
Ethnic div. × $\overline{\text{Crop exchange}}$	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

	Dependent Variable:			
Panel A:	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	-0.242	-0.440	-0.526
	[0.046]***	[0.090]**	[0.210]**	[0.158]***
	(0.030)***	(0.070)***	(0.227)*	(0.221)**
Ethnic div. $\times \overline{\text{Crop exchange}}$	0.244	0.379	0.698	0.708
	[0.079]***	[0.124]***	[0.316]**	[0.228]***
	(0.072)***	(0.124)***	(0.344)**	(0.318)**
Ln population	Yes	Yes	Yes	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	-0.023	-0.050	-0.122
	[0.037]	[0.009]**	[0.027]*	[0.045]***
	(0.052)	(0.012)*	(0.030)*	(0.056)**
Ethnic div. $\times \overline{\text{Crop exchange}}$	0.085	0.035	0.060	0.175
	[0.047]*	[0.014]**	[0.034]*	[0.059]***
	(0.070)	(0.021)*	(0.036)*	(0.072)**
Individual controls and year FE	Yes	Yes	Yes	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)
$\widetilde{\text{Crop exchange}}$	1.055*	2.011**	1.955***
	[0.594]	[0.791]	[0.719]
Minority group advantage		0.931**	-1.510
		[0.449]	[1.305]
$\widetilde{\text{Crop exchange}}$ × 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 et al. 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	-0.298	-0.292	-0.416	-0.444	-0.392	-0.138
	[0.162]*	[0.168]*	[0.162]*	[0.163]**	[0.170]**	[0.172]**	[0.183]
	(0.134)**	(0.136)**	(0.135)**	(0.127)***	(0.123)***	(0.144)***	(0.074)*
Ethnic div. $\times \overline{\text{Crop exchange}}$	0.516	0.524	0.518	0.593	0.624	0.556	0.191
	[0.265]*	[0.272]*	[0.265]*	[0.233]**	[0.238]**	[0.240]**	[0.268]
	(0.220)**	(0.222)**	(0.221)**	(0.178)***	(0.173)***	(0.197)***	(0.139)
Ln (1 + # native crops)		0.048			0.198		
		[0.066]			[0.092]**		
		(0.055)			(0.086)**		
Native crop frac.			-0.013			0.365	
			[0.174]			[0.203]*	
			(0.195)			(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	-0.144*	-0.036
	[0.060]	[0.085]	[0.083]
	(0.062)	(0.079)	(0.045)
Ethnic div. $\times \overline{\text{Crop exchange}}$	0.109	0.220*	0.033
	[0.090]	[0.127]	[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	-0.214	0.084	0.309	-0.060	-0.095
	[0.086]	[0.243]	[0.269]	[0.308]	[0.271]	[0.134]
Ethnic div. $\times$ $\overline{\text{Crop frac.}}$	0.041	0.388	-0.163	-0.560	0.123	0.172
	[0.186]	[0.456]	[0.505]	[0.568]	[0.512]	[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:

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

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). 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

[52] I 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).

surnames occur only once in the dataset, which results in a sample size 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).

[53] Specifically, 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	0.447	0.558	0.481	0.448	0.608
	[0.212]**	[0.182]**	[0.191]***	[0.170]***	[0.182]**	[0.218]***
	(0.213)**	(0.207)**	(0.203)***	(0.217)**	(0.201)**	(0.261)**
	Panel B: Non-Unique Surnames (Native Surnames)					
Border parish (dummy)	0.462	0.408	0.513	0.454	0.428	0.581
	[0.209]**	[0.159]**	[0.180]***	[0.173]**	[0.178]**	[0.214]***
	(0.208)**	(0.180)**	(0.204)**	(0.218)**	(0.196)**	(0.260)**
	Panel C: Grouped Surnames (Native Surnames)					
Border parish (dummy)	0.482	0.416	0.523	0.461	0.427	0.582
	[0.215]**	[0.184]**	[0.189]***	[0.172]***	[0.181]**	[0.216]***
	(0.213)**	(0.205)**	(0.190)***	(0.216)**	(0.199)**	(0.253)**
	Panel D: Non-Native Surnames					
Border parish (dummy)	0.284	0.116	0.005	0.333	0.185	0.112
	[0.257]	[0.101]	[0.098]	[0.209]	[0.116]	[0.136]
	(0.223)	(0.097)	(0.083)	(0.195)*	(0.126)	(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 - Descendants of U.S. Immigrants

I explore whether the descendants of societies with greater historical exposure to economic exchange are less likely to marry within their own group. For this analysis, I use the sample of U.S. immigrant descendants created by Giuliano and Nunn (2021) based on the *March Supplement of the Current Population Survey* (CPS) for the period 1994-2014. Specifically, the sample includes all married individuals, born in the U.S., with at least one immigrant parent (i.e., a parent born outside the U.S.). For each individual, I identify the pre-industrial ancestors who lived in the parent's country of origin. Based on this country of origin, I then assign a (country-level) value of ancestral exposure to economic exchange to each individual. I conduct the analysis separately for the father's and the mother's country of origin.

To measure ancestral exposure to economic exchange, I use the measure of 'markets' importance across pre-industrial societies constructed by Enke (2023). Specifically, the author constructs a continuous proxy for the presence of "intercommunity trade and/or money" for each society, based on Michalopoulos and Xue (2021)'s historical folklore data. This measure is then cross-validated by the author using a subsample of societies from the Standard Cross-Cultural Sample (SCCS). I aggregate this measure at the country level by considering the minimum level of 'markets' importance across the pre-industrial societies in a country. The measure is available for 83 different countries of origin in the CPS sample.

In Table C.1, the outcome is a dummy variable indicating whether the individual (born in the U.S.) is married with someone from the same group.[54] I regress this dummy variable on the country-level measure of ancestral exposure to economic exchange, based on the country of origin of the father (Columns 1 and 2) or the mother (Columns 3 and 4). The results are reported separately for married women (Panel A) and married men (Panel B). Robust standard errors are clustered at the level of country of origin.

Following Giuliano and Nunn (2021), the specification includes: (*i*) area of residence (i.e., metropolitan-area) fixed effects, (*ii*) survey-year fixed effects, (*iii*) a vector of individual-level characteristics (age, age squared, fixed effects for the level of educational attainment, metropolitan city status, and the share of first- or second-generation immigrants from the same

[54] I use the same definition as in Giuliano and Nunn (2021). Specifically, the individual's spouse is considered to be from the same group if the spouse was born in the same country as the individual's father (Columns 1 and 2) or mother (Columns 3 and 4) or if the spouse's father or mother was born in the same country.

country of origin that live within the individual's metropolitan area), and (iv) a vector of country-level characteristics for the parent's country of origin—either the father's (Columns 1 and 2) or the mother's (Columns 3 and 4)—including log GDP per capita, distance from the equator, genetic distance with the U.S., and separate variables for pre-industrial political and economic complexity. Finally, I also include a quadratic polynomial of the country of origin's average number of folklore motifs. The results are consistent with the idea that individuals with greater ancestral exposure to economic exchange, based on the pre-industrial ancestors who lived in their parents' country of origin, are less likely to marry within their own group, on average. Columns 2 and 4 show consistent (economically modest) results when controlling for Giuliano and Nunn (2021)'s country-level measure of cross-generational climatic instability for the father's country of origin (Column 2) or the mother's (Column 4).

Table C.1: Within-Group Marriage—Evidence from Descendants of U.S. Immigrants

	Dependent variable: Dummy variable for spouse being from the same group							
	Panel A. Married Women				**Panel B.** Married Men			
	Origin country identified from father		Origin country identified from mother		Origin country identified from father		Origin country identified from mother	
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Ancestral economic exchange	-0.030***	-0.029**	-0.032**	-0.030**	-0.003	-0.003	-0.005***	-0.005***
	[0.011]	[0.011]	[0.013]	[0.013]	[0.004]	[0.004]	[0.002]	[0.002]
Climatic instability		-0.344*		-0.426**		0.009		0.014
		[0.179]		[0.202]		[0.022]		[0.019]
Individual-level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Average motifs (quad. polynomial)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Metropolitan-area FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var.	0.425	0.425	0.425	0.425	0.980	0.980	0.981	0.981
Number of individuals	19,391	19,391	19,403	19,403	19,391	19,391	19,403	19,403
Number of countries	83	83	83	83	83	83	83	83

Notes. The unit of observation is the individual. The table reports OLS estimates. Robust standard errors are clustered at the level of the father's (Columns 1 and 2) or mother's (Columns 3 and 4) country of origin. The outcome variable is a dummy indicating whether the individual is married with someone from the same group. The individual's spouse is considered to be from the same group if the spouse was born in the same country as the individual's father (Columns 1-2) or mother (Columns 3-4) or if the spouse's father or mother was born in the same country. The vector of individual-level characteristics includes age, age squared, fixed effects for the level of educational attainment, metropolitan city status, and the share of first- or second-generation immigrants from the same country of origin that live within the individual's metropolitan area). The vector of country-level characteristics for the parent's country of origin—either the father's (Columns 1-2) or the mother's (Columns 3-4)—includes log GDP per capita, distance from the equator, genetic distance with the U.S., and separate variables for pre-industrial political and economic complexity.

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

D 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* (ENAHO), 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* (ENAHO), 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* (ENAHO), 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).

E 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 Taqe*, 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.

Carraffa, Alberto García and Arturo García Carraffa, *Diccionario Heráldico y Genealógico de Apellidos Españoles y Americanos*,1920–1963, Madrid: Imprenta Antonio Marzo.

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 Velthuizen, 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.

Giuliano, Paola, and Nathan Nunn, "Understanding Cultural Persistence and Change," *The Review of Economic Studies*, 2021, 88 (4): 1541–1581.

González Holguín, Diego, *Vocabvlario de la Lengva General de todo el Perv llamada Lengua Qquichua, 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.

Michalopoulos, Stelios, and Melanie Meng Xue, "Folklore," *The Quarterly Journal of Economics*, 2021, 136 (4): 1993–2046.

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.

Platt, Lyman D, *Hispanic Surnames and Family History*, 1996, Baltimore, MD: Genealogical Publishing Company.

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 Codding, "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.