Working Paper in Economics No. 714

The Origins of Cultural Divergence: Evidence from a Developing Country

Hoang-Anh Ho, Peter Martinsson and Ola Olsson

Department of Economics, December 2017

ISSN 1403-2473 (Print)
ISSN 1403-2465 (Online)

UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

The Origins of Cultural Divergence: Evidence from a Developing Country*

Hoang-Anh Ho **Peter Martinsson** **Ola Olsson**

University of Gothenburg

December 2017

Abstract

Cultural norms diverge substantially across societies, often even within the same country. In the present paper, we study the voluntary settlement hypothesis, proposing that individualistic people tend to self-select into migrating out of reach from collectivist states towards the periphery and that such patterns of historical migration are reflected even in the contemporary distribution of norms. During most of the first millennium CE, the modern north of Vietnam was under an exogenously imposed Chinese rule. From the eleventh to the eighteenth centuries, historical Vietnam gradually expanded its territory to the Mekong River Delta through various waves of conquest and migration. In contrast to some recent research, we find very little support from historical sources for any major discontinuities in this territorial expansion. Combining archives with household survey and lab-in-the-field experiment, we demonstrate that areas being annexed earlier into historical Vietnam are nowadays more (less) prone to collectivist (individualist) culture. We argue that the southward out-migration of individualistic people was the main mechanism behind this finding, which is also in line with many historical accounts.

Keywords: Culture; Individualism-Collectivism; Voluntary Settlement.

JEL Classification: N45; O53; Z13.

*We are grateful for helpful comments from Fredrik Carlsson, Carl-Johan Dalgaard, Martin Dufwenberg, Masayuki Kudamatsu, Laura Mayoral, Khanh-Nam Pham, Pablo Selaya, Ardeshir Sepehri and seminar participants at University of Economics Ho Chi Minh City, University of Gothenburg, and the Sixteenth Nordic Conference in Development Economics. We would also like to thank EfD-Vietnam for kindly hosting and supporting our field research and data collection. All remaining errors are ours. Contact: Hoang-Anh Ho (hoang-anh.ho@economics.gu.se), Peter Martinsson (peter.martinsson@economics.gu.se), and Ola Olsson (ola.olsson@economics.gu.se).

1 Introduction

Economic research has uncovered strong associations between many cultural traits and various indicators of individual behavior, institutional and economic development (Guiso, Sapienza, and Zingales 2011; Fernández 2011; Algan and Cahuc 2014; Doepke and Zilibotti 2014; Alesina and Giuliano 2014, 2015).[1] Among them, the individualism-collectivism (IC) contrast has been found to be a powerful predictor of long-run democratic and economic development in a large sample of countries (Gorodnichenko and Roland 2011, 2015, 2017; Olsson and Paik 2016a).[2] The basic idea is that an individualist culture is conducive to technological and institutional innovations since it emphasizes personal initiative, value and achievement. Meanwhile a collectivist culture is favorable to collective action as it prioritizes individual embeddedness and conformity to group norms and values. In terms of innovation, collectivism only has a static effect, on the other hand individualism has a dynamic effect. Hence the latter dominates over longer horizons and individualist societies perform better with respect to economic and institutional development. The empirical findings of the IC contrast naturally lead us to an important question: why are some societies more (less) collectivist (individualist) than others?

In the present paper, we hypothesize that cultural differences along the IC dimension across modern societies can be traced back to territorial expansion and migration processes that happened in the historical times. We investigate a version of the voluntary settlement hypothesis, proposing that as collectivist societies emerged, they triggered the out-migration of individualistic members towards peripheral regions. This migration pattern then repeated over and over again as the individualistic immigrants inhabited and developed the new peripheral regions into less collectivist communities compared to the ones they left behind, which in turn induced more individualistic members to migrate towards more peripheral regions. Eventually, these migration processes gave rise to cultural differences along the IC dimension across communities, which have persisted through time and constituted the cultural landscape found in the present day.

[1] Some notable traits are generalized trust, individualism-collectivism, family ties, generalized morality, and attitudes toward work and the perception of poverty.

[2] The IC contrast has also been the single most relevant cultural dimension in cultural psychology (see the review in Gorodnichenko and Roland 2012). For an early literature survey on theory and measurement of the IC contrast in social psychology, see Triandis (1995).

We recognize the territorial expansion process in the history of Vietnam as an ideal setting to empirically examine the voluntary settlement hypothesis. During most of the first millennium, the modern north of Vietnam was under the colonization of the centralized imperial bureaucracy of historical China. After gaining independence in 939 CE, historical Vietnam initially governed the region surrounding the Red River plain with a rice-agriculture economy, a centralized bureaucratic system, and a collectivist social organization. Meanwhile, the territory in the south was sparsely populated by other indigenous people who were largely unaffected by Chinese influence. From the eleventh to the eighteenth centuries, historical Vietnam gradually expanded its territory along the coast from the Red River to the Mekong River plains through various waves of conquest and migration, which in turn displaced most of the indigenous people, to establish a country as today (see Figure 1).[3] We hypothesize that as historical Vietnam expanded its territory, individualistic people self-selected into migrating towards the annexed areas. This pattern repeated over and over again through eight centuries and eventually generated cultural differences along the IC dimension across regions in historical Vietnam whereby the most individualistic people inhabited the most recently annexed regions. Most importantly, these cultural differences have persisted to the modern day.

[Figure 1 is about here.]

To test the above hypothesis, ideally, one would first need historical data to show that people who migrated to and inhabited the annexed areas were more individualistic than those who stayed in the initial regions of historical Vietnam, which in turn generated cultural differences along the IC dimension across regions back in at least the eighteenth century. Then, one should collect present-day data and conduct a proper empirical analysis to provide convincing causal evidence that descendants in these annexed areas are nowadays also more individualistic than descendants in the initial regions.

To emulate this ideal setting, we first review historical studies showing that historical Vietnamese who migrated and inhabited in the annexed areas possessed typical characteristics of individualist culture, and that cultural differences along the IC dimension across regions in historical Vietnam had already been in place as early as the eighteenth

[3] The Vietnamese southern advance is well documented in typical historiographies written by Vietnamese (Le 1982) and Western scholars (Taylor 2013).

century. To capture the IC dimension in the present day, we focus on the societal ability to solve collective action problems, which is a critical feature in theoretical models linking the IC trait to economic and institutional development (Gorodnichenko and Roland 2015, 2017). We then conduct a reduced-form analysis linking the key explanatory variable, i.e. the time elapsed since being annexed into historical Vietnam, to the indicators of the IC dimension in the present day. Our hypothesis thus states that those areas being annexed into historical Vietnam earlier are nowadays more prone to collectivist norms.

Using the Vietnam Household Living Standard Survey, we aggregate data on household labor contributions to public goods production at district level to measure the societal ability to solve collective action problems. Three indicators are constructed based on labor contributions to public goods production: (i) the percentage of households and (ii) the number of members per household making contributions and (iii) the number of days contributed per household. We find that districts being annexed into historical Vietnam earlier nowadays have higher percentages of households and numbers of members per household making labor contributions as well as numbers of days contributed per household. The estimated effects are substantial and robust to various control variables as well as different empirical specifications including an instrumental variable estimation. Furthermore, we also distinguish the migration of individualistic people in historical times from other potential channels in explaining these cultural differences including differential living conditions in the annexed areas, Western influence in the Mekong River plain and important population movements in the modern history.

To investigate the voluntary settlement hypothesis further, we carry out a lab-in-the-field public goods experiment with high-school students from the north and the south of Vietnam. In doing so, we expect the public goods experiment to capture the prevailing cooperation attitudes in the north and the south. To ensure the credibility in terms of causal inference, we take a great care in selecting experimental sites in a way to separate the voluntary settlement hypothesis from other potential channels mentioned above. We find that students from the north contribute significantly more in the public goods experiment compared to their southern peers, and that the result is mainly driven by the belief about the contribution levels of other subjects. To sum up, this finding thus corroborates the survey data analysis and suggests that the cultural differences along the IC dimension across regions are also present in the youngest generation of the country.

Our research relates to a growing multidisciplinary literature on the origins of cultural differences along the IC dimension.[4] The pathogen hypothesis argues that societies with historically greater prevalence of disease-causing pathogens are nowadays more prone to collectivism because its emphasis on in-group/out-group distinction and conformity can serve as an anti-pathogen defense function (Fincher et al. 2008). The ecological context theory posits that some forms of production in subsistence economies (such as farming) require more functional interdependence than others (such as hunting) and hence give rise to collectivism as an adaptation mechanism (e.g. Vandello and Cohen 1999; Talhelm et al. 2014). Motivated by the history of settlement in the United States and its well-known individualist culture, Kitayama et al. (2006, 2009) put forward and provide experimental evidence for the voluntary settlement hypothesis, which postulates that voluntary settlers are likely to have a highly autonomous, independent, and goal-oriented mental set (see also Bazzi, Fiszbein, and Gebresilasse 2017). Olsson and Paik (2016b) argue and provide cross-country evidence that the advent of farming in the core regions during the Neolithic Revolution were accompanied with collectivism, which in turn induced individualistic farmers to out-migrate towards peripheral areas and eventually gave rise to cultural differences along the IC dimension between societies.

Our research builds on and adds to this literature in various ways. First, to the extent of our knowledge, no studies on the origins of cultural differences along the IC dimension have examined the societal ability to solve collective action problems. Second, our research combines survey and experimental data in a coherent manner to study the same type of indicator, i.e. contribution to public goods production. Finally, studies in this literature so far have mostly employed cross-country comparisons and concentrated on developed societies, and it is well-known that these countries have gone through the modernization process to a great extent. Developing countries, on the other hand, have still retained most of their traditional cultural landscapes. Using the setting within a developing country, our research thus corroborates Olsson and Paik (2016b) and Kitayama et al. (2006, 2009) in establishing the role of the migration of individualistic people over the long term in explaining cultural differences along the IC dimension across societies.

[4] The modernization hypothesis, arguing that societies become more individualistic as they reach higher levels of economic development, essentially focuses on the convergent tendency towards individualism, rather than pre-existing cultural differences, across modern societies (Inglehart and Baker 2000).

Within economics, our research joins a growing literature aimed at understanding the persistence of various cultural traits as an important channel through which historical events could exercise their impacts on current economic development.[5] Our paper is also closely related to a recent study involving Vietnamese history by Dell, Lane, and Querubin (2017). Using a regression discontinuity design to compare villages along a presumed fixed 150-year border between historical Vietnam and the territory in the Mekong Delta, then controlled by the Khmer empire, the authors show that today's living standards are higher in the more strongly centralized territories of the former. As one possible mechanism behind this finding, the authors argue that the strong state in historical Vietnam was better at instilling norms promoting collective action, which is reflected by a better ability to organize for public goods and redistribution through civil society and local government. Although the two papers reach a similar conclusion as regards the strong ability of historical Vietnam to promote collective action, our voluntary settlement interpretation of the finding and explicit focus on the IC dimension, differentiate our approaches. We further argue that our approach of studying the gradual southward expansion of historical Vietnam as a more or less continuous historical process, is strongly supported by the official chronicles of historical Vietnam, i.e. *Dai Viet Su Ky Toan Thu* (204 BCE-1675 CE) and *Dai Nam Thuc Luc* (1558-1888 CE). Indeed, our Vietnamese sources provide very weak support for alternative scenarios assuming an exploitable 150-year fixed border between the Khmer empire and historical Vietnam.

The remainder of the present paper is organized as follows. The next section provides the historical background of the southward territorial expansion of historical Vietnam and the accompanied migration process, together with details on historical sources to support our interpretation of the Vietnamese southern advance. Section 3 presents the analysis with survey data including empirical strategy, data, and results. Section 4 describes the experimental design, site selection, and results. Section 5 closes the paper with some concluding remarks.

[5] Some remarkable traits are gender equality (Alesina, Giuliano, and Nunn 2013; Hansen, Jensen, and Skovsgaard 2015), trust (Nunn and Wantchekon 2011; Becker et al. 2016; Bigoni et al. 2016; Guiso, Sapienza, and Zingales 2016; Litina 2016), preferences for redistribution (Alesina and Fuchs-Schündeln 2007; Basten and Betz 2013), and various others such as anti-semitic attitude (Voigtländer and Voth 2012), time preference (Galor and Özak 2016), and civic values (Lowes et al. 2017). See Nunn (2012, 2014) and Spolaore and Wacziarg (2013) for critical reviews of the literature.

2 Historical Background

Archaeological evidence indicates that ancient populations had settled down in the Red River plain with rice agriculture around 2000 BCE during the Neolithic Revolution (Nguyen, Pham, and Tong 2004). Oral history maintains that these people belonged to various tribes among whom was the Lac, ancestor of the Vietnamese (Taylor 2013). The whole region of what is now northern Vietnam was brought under the control of Han China already in 111 BCE. This external intervention initiated a long period of Chinese colonization during most of the first millennium CE. After the victory in 939 CE, the first stable state of historical Vietnam was founded in 968 CE and endowed with a centralized bureaucratic system from the colonizer. Subsequent dynasties governing this land continued to build stronger structures and orders into the society by adopting the Confucian doctrine, whose basic tenet emphasized the values of social groups above the needs and desires of its constituent members (Whitmore 1984, 1997). The lingering military threat from China in the north further necessitated a strong military organization with its associated collectivist norms. The strength of this collectivist society was best demonstrated in the warfare history of historical Vietnam, e.g. being the only land-based polity in Eurasia to defeat the Mongols in the thirteenth century.

Peripheral areas in the south of historical Vietnam were initially places where people belonging to marginalized social classes such as landless peasants, refugees, fugitives, and adventurers, often sought livings (Taylor 2013). This group "lacked standing in a social group, like the family or the village, was less than a full person and could hope for no better future in traditional village society" (Tana 1998, p. 111). Because they were not recognized in the collectivist society, we argue that they were more inclined towards an individualistic mindset typical of pioneer societies.[6] Gradually from the eleventh to the eighteenth centuries, this type of people was induced to migrate southward as historical Vietnam conquered and established its administration along the coast to the Mekong River Delta. While some of them had already inhabited the annexed areas long before the historical states reached there, most of them migrated to seek fortune in the south as the states granted private property rights on any new cleared croplands (Taylor 2013). They thus "had learned to rely on their own resources and initiative to a degree not

[6] One may recall the analogous image of European immigrants to the United States during the age of mass migration (e.g. Abramitzky, Boustan, and Eriksson 2012).

possible in the north" (Taylor 2013, p. 252). These resourceful immigrants together with their wet-rice technology later on transformed the Mekong River Delta from a wild marshy land into a prosperous agricultural economy (Tana 1998; Taylor 2013).

The Vietnamese southern advance took place gradually through various steps. Available historical sources appear to be in agreement that this process effectively ended in 1757 and they provide very little support for assuming any major discontinuities.[7] In each step, the state conquered and established its administration in the annexed territory, which induced a more individualistic population to migrate further south. During this territorial expansion, historical Vietnamese (the Kinh ethnicity) immigrants displaced indigenous people in the annexed areas, most of whom ran away and sought living in neighboring territories while a few others stayed and assimilated in Vietnamese culture (Taylor 2013).[8] The state capital was also shifted from modern Ha Noi to modern Thua Thien Hue (see Figure 1). The social environment in the annexed region was characterized by greater openness, mobility and autonomy relative to the north, "a new way of being Vietnamese" (Tana 1998). Respecting these cultural differences, Gia Long (1802-1820 CE), the first king of the Nguyen dynasty, governed from the capital in the central and delegated two viceroys with a high level of autonomy, one with literary and one with military backgrounds, to govern the northern and southern regions respectively (Taylor 2013). In contrast, his successor Minh Mang (1820-1841 CE) determined to impose by force more uniformity in government, religion, culture, and even apparel for the whole country. These policies had reduced the cultural diversity to a large extent and provoked many rebels in the southern region (Wook 2004; Taylor 2013).

The modern Vietnamese history started with the French Colonization in 1858 CE when the colonizers took control of the South East and Mekong River Delta (see Figure 1), and around 20 years later they did so with the whole country. Following the end

[7] For secondary sources, see Briggs (1947, p. 358), Nguyen (1999, p. 18), Engelbert (1994, p. 173), and Taylor (2013, p. 325-336). For a detailed chronological description in the primary historical sources of the Vietnamese southern advance, see the Supplementary Materials. In contrast, Dell, Lane, and Querubin (2017) claim that the southern region in modern Vietnam was divided by a fixed border between the Khmer empire and historical Vietnam from 1698 until 1833, i.e. just prior to the French Colonization in 1859.

[8] These groups nowadays only constitute minor fractions in the Vietnamese population, for example in 1979 the Cham 0.2% and the Khmer 1.4% compared to the Kinh 87% (Dien 2002).

of this colonization was the Second Indochina War (1954-1975 CE), during which the country was divided into the socialist North and the capitalist South. Whether this recent experience with Western capitalism in the southern region, or the more distant past, was the main driver of the cultural differences in Vietnam is an empirical question and examined further below. After the Reunification in 1975, the central planning economy was implemented with limited success and subsequent economic reforms during the last 30 years have brought prosperity and modernization to the whole country.

To summarize, cultural differences across regions in Vietnam, as in any other countries, are well expressed in various anecdotes, many of which are stereotypes or even prejudices. There are, however, some serious observations worth mentioning. Ending his Vietnamese historiography, Taylor (2013, p. 624) argues that cultural differences across regions in Vietnam have well persisted to the modern day, that "northerners are more disciplined to accept and to exercise government authority" and "southerners are more individualistic, egalitarian, entrepreneurial, interested in wealth more than in authority". In a practical guide to doing business and living in Vietnam, although regarding collectivism as the main cultural theme, Ashwill and Diep (2005) note that "northerners are considered to be more intelligent, conservative, austere, serious, and frugal, ..., are more apt to save for a rainy day" (p. 71) and "southerners are perceived as fun-loving, easy-going, open people who rarely think of saving for a rainy day" (p. 72). Scant empirical evidence indicates that southerners are more loss averse, but not different from northerners with respect to patience or present bias (Tanaka, Camerer, and Nguyen 2010).

3 Survey Data Analysis

3.1 Empirical Model

In this section, we use the Vietnam Household Living Standard Survey (VHLSS) as our main dataset to investigate the voluntary settlement hypothesis. Our baseline empirical model revolves around regressing various indicators of the IC dimension on an explanatory variable capturing the time elapsed since being annexed into historical Vietnam while controlling for other potential channels. Owing to data availability, we measure culture at district level through aggregation from individual data.[9] The model thus takes the

[9] There is certainly individual heterogeneity within a society, but in aggregation one can observe what social psychologists call "cultural syndrome" of each society (Triandis 1995). This is also a standard

following form:

$$IC_i = \beta TimeSinceAnnexed_i + \gamma X_i + \epsilon_i.$$

In this equation, IC_i is a measure of the average expression of cultural norms in district i, $TimeSinceAnnexed_i$ is the time since being annexed into historical Vietnam, X_i is a set of control variables, and ϵ_i is a random error term. Our hypothesis postulates that β is positive (negative) with respect to collectivist (individualist) norms. All variables are described in detail in the subsection below.

There are two sources of endogeneity that may threaten causal inference of the time since being annexed variable. The first one is reverse causality, i.e. the more collectivist a society was the earlier it was annexed by historical Vietnam. Historical evidence, however, does not suggest that this is a strong cause of concern. First, there were fundamental differences between the indigenous culture of historical Vietnam and the Sinic civilization when historical China colonized the region (Taylor 1983). And most importantly, later conquests of historical Vietnam into the south were either unintentional responses to attacks from southern enemies or primarily driven by economic calculation to appropriate land resource for agricultural production (Tana 1998). The second threat comes from various omitted variables that may influence both the time since being annexed into historical Vietnam and the IC dimension.

To examine the potential endogeneity bias problem, an instrumental variable (IV) estimation is employed. We argue that while many factors might play a role in the decision to annex an area, the north-south geographical order is impossible to manipulate and hence exogenous to the decision. In other words, from the Red River plain, one could not conquer the Mekong River plain without annexing all areas located in between.[10] Thus, in the subsample of the annexed areas, the north-south geographical order can serve as a valid instrument for the time since being annexed into historical Vietnam. We propose to use the walking distance along the coast from an annexed area to the military center of historical Vietnam (modern Nghe An) as such an instrument. From

exercise in economics, see Alesina and Giuliano (2015) for a review.

[10]Theoretically, one could do so by either passing the ocean in the east or the mountainous band in the west separating historical Vietnam from Laos and Cambodia. Both strategies were arguably impossible given the logistic and transportation technologies in historical Vietnam. We indeed do not find any attempt to do so in the historical narratives.

north to south, one needed to pass through modern Thua Thien Hue and Ninh Thuan (see Figure 1). The walking distance (instead of the geodesic, "bird-fly" distance) is therefore employed to capture the military route in the historical times.

3.2 Variables

The IC Trait

In the present paper, we adopt the conventional definition of culture in the multidisciplinary literature as "decision making heuristics or a rules of thumb that have evolved given our need to make decisions in complex and uncertain environments" (Nunn 2012, p. S109).[11] Using survey data, there are two strategies to measure culture. The first is to construct direct indicators extracted from value survey such as the WVS (Olsson and Paik 2016b). Unfortunately, we cannot rely on the WVS in Vietnam to conduct a causal analysis since it does not contain detailed locational information of the respondents.[12] The second is to employ indirect outcomes extracted from observational data such as marriage stability, inventiveness, family structure, or infrequent children names (Vandello and Cohen 1999; Talhelm et al. 2014; Bazzi, Fiszbein, and Gebresilasse 2017). To be a good measure of the IC trait, we argue that an outcome has to be both theoretically and practically relevant. In other words, it has to capture an important aspect of the IC trait and feature as a traditional practice in the society under study.

For our main analysis, we employ labor contribution to public goods production (e.g building or repairing infrastructure roads, wells, irrigation, schools or health clinics) to capture the strength of collectivism based on two underpinnings. From a theoretical perspective, the ability to solve collective action problems is a critical feature of collectivist societies (Gorodnichenko and Roland 2015, 2017). Since collectivist societies are considered to be better in this aspect, one should observe their members to contribute more resources to public goods production. From a practical perspective, labor contribution to public goods production (such as irrigation facilities, roads, and communal buildings) has also been a typical activity among people in historical Vietnam (e.g. Adams and Hancock 1970). This tradition is still prevalent in the modern day, especially in rural areas, as labor

[11] This definition is closely related to another prominent one used in Guiso, Sapienza, and Zingales (2006, p. 23): "customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation."

[12] A descriptive analysis using the WVS is presented in the Supplementary Materials.

is probably the most abundant resource that people possess. Exploiting the VHLSS, we construct three related variables at district level based on labor contributions to public goods production: (i) the percentage of households and (ii) the number of members per household making contributions, and (iii) the number of days contributed per household.

The VHLSS covers all provinces in Vietnam with information on geographical location at district level that enables one to conduct a nuanced empirical analysis. Although the basic structure is the same, many contents of the survey have changed over time since the first round in 1992. Thus, owing to data availability, we only employ the 2002 survey round in the empirical investigation. The 2002 sample was drawn from the 1999 Population Census and covered 607 districts across 61 provinces. Appendix A1 shows that around 26 percent of households contributing labor to public goods production, whereas 0.44 and 3.24 are respectively the numbers of members and days contributed per household. To avoid blank areas for non-surveyed districts on the map in Figure 2, we aggregate all three variables at province level and depict them. A visual comparison with Figure 1 gives an impression that provinces being annexed into historical Vietnam earlier nowadays have higher percentages of households and numbers of members per household making labor contributions as well as numbers of days contributed per household.

[Figure 2 is about here.]

The Time since Being Annexed into Historical Vietnam

As already mentioned above, our main explanatory variable is the time elapsed since being annexed into historical Vietnam, measured in number of 100 years. Following the historical background discussed above, we choose the first unified state of historical Vietnam in 968 CE as the beginning year to code this variable. For any year chosen, however, it only affects areas in the initial territory. And as shown in the following empirical analysis, the result remains qualitatively unchanged when the estimation is restricted to the subsample covering only the annexed areas. The terminal year is 1990.[13]

To construct this variable for each modern district, two dimensions are needed: (i) its corresponding area in historical Vietnam and (ii) the year that this area was annexed into historical Vietnam.[14] For the time dimension, we rely on two official chronicles

[13]This is an arbitrary choice. The results are robust to other choices and available upon request.

[14]All details on the coding procedure are presented in the Supplementary Materials.

of historical Vietnam, *Dai Viet Su Ky Toan Thu* (Cao 2012) and *Dai Nam Thuc Luc* (Nguyen 2002), which record events from the beginning to 1675 CE and in the periods 1558-1888 CE respectively. These chronicles were written by top scholars of the historical states of Vietnam to keep track of historical events for future generation, and constituted the primary sources for all Vietnamese historiographies that we are aware of. We consider an area to be annexed when there is a record in the chronicles saying that this area either a tribute to or was conquered by historical Vietnam, which means the state could start establishing its administration and collecting tax. To link areas in historical Vietnam to their modern counterparts, we rely on two seminal works of Vietnamese historians: Dao (2005) and Phan et al. (2011). Given that the administrative boundaries had changed many times throughout history, matching historical areas with modern units is apparently not a perfect mapping. This is more of a problem at district level, which potentially drives the estimated coefficient towards zero. As shown in the following empirical analysis, however, the time since being annexed into historical Vietnam does carry enough signal to deliver a satisfactory statistical performance.

In our context, the time since being annexed into historical Vietnam captures the migration channel whereby the most individualistic migrants occupied the most recently annexed areas. However, one might argue that it also reflects the length of time that the historical states of Vietnam had in instituting collectivist norms and values into the people. In other words, the longer an area experienced the control of historical Vietnam, the stronger was its collectivism. Although we cannot rule out this possibility, the above historical discussion indicates that a dominant weight should be assigned to the voluntary settlement hypothesis since most of the Vietnamese immigrants only moved to the annexed areas after the states had conquered them. Thus, in terms of state experience, these immigrants generally had the same duration as those who stayed in the old areas.

Control Variables

We prepare two sets of control variables to be included in the regression analysis whose descriptive statistics can be found in Appendix A1. The first is intended to capture district demographics, which is constructed by aggregating standard household characteristics including the number of members, residing in urban area, and head ethnicity, gender, age and schooling years. We use dummy variables for the Kinh ethnicity (historical Vietnamese) and male household head.

The second set of control variables consists of various potentially confounding factors normally found in the literature. The first one is per capita expenditure accounting for the influence of modernization (Inglehart and Baker 2000). The percentage of cropland in total land area is used as a control for the prevalence of agriculture (Talhelm et al. 2014), which is mainly rice in the context of Vietnam. The third variable is natural land productivity, aimed at capturing the incentive to cooperate, as motivated in Litina (2016). Finally, Triandis (1995) hypothesizes that isolated and densely populated areas are conducive to the development of collectivist culture. To account for these factors, we control for elevation, terrain ruggedness and population density. Finally, we also control for latitude following the popular practice in empirical studies although there is no theoretical underpinning for its impact on the IC trait.

Information on per capita expenditure is taken directly from the VHLSS. The percentage of cropland is taken from the Global Land Cover Characteristics Database, which is constructed at 30 arc-second resolution based on satellite inputs in the period of 1992-1993 (Loveland et al. 2000). Natural land productivity is measured by the caloric suitability index constructed at 5 arc-minute resolution by Galor and Özak (2016), who make their calculation based on data from the Global Agro-Ecological Zones. This index measures the average potential yield (million kilo calories per squared kilometer per year) attainable in each grid cell given the set of crops that are suitable for cultivation in the post-1500 period. To capture the natural component of productivity, the production conditions are set at low level of inputs and rain-fed agriculture with the climatic conditions (precipitation, temperature, etc.) are taken from the baseline period 1961-1990.

Elevation is taken from the GTOPO30 dataset provided by the Earth Resources Observation and Science Center. The terrain ruggedness index is originally devised by Riley, DeGloria, and Elliot (1999) to quantify topographic heterogeneity in wildlife habitats providing concealment for preys and lookout posts. This index is calculated at 30 arc-second resolution across the entire surface of the Earth by Nunn and Puga (2012) based on the GTOPO30 dataset. Finally, population density is taken from the Gridded Population of the World Version 3 provided by the Center for International Earth Science Information Network (CIESIN and CIAT 2005). The data are from 1990, which is the earliest year available. Latitude and distance are measure at the district centroids using data taken from the Global Administrative Unit Layers.

Owing to data availability, we cannot control for an important factor found in previous studies, i.e. historical pathogen prevalence. For several reasons, we think that this omission does not seriously impact our analysis. First, because Vietnam is a small country located within the tropical climate zone, the geographical variation in historical pathogen prevalence is arguably not substantial. Indeed, the history of disease ecology do not indicate any significant north-south difference in disease environment (Bamber 1993; Taylor 1993). Second, pathogens have been shown to be more prevalent in hot regions (Guernier, Hochberg, and Guégan 2004), which are also more favorable to growing rice. Indeed, Talhelm et al. (2014) shows that, across provinces in China, pathogen prevalence is insignificantly related to the IC trait when rice agriculture (percentage of cultivated land) is controlled for. Finally, the instrumental variable strategy mentioned above is adopted exactly to give us a clue on this issue.

3.3 Baseline Results

In the following analysis, we focus on the number of days contributed per household to public goods production and report the corresponding results regarding the percentage of households and the number of members per household making labor contributions in the Appendix. Table 1 presents the result of regressing the number of days contributed per household on the time since being annexed into historical Vietnam and standard demographic characteristics. As can be seen, the estimated coefficient of the time since being annexed is significantly positive whether demographic controls are included or not. Districts being annexed into historical Vietnam earlier nowadays on average have a higher number of days contributed per household. Relative to the mean number of days contributed, the marginal effect is reasonably large and stable through various regressions. In the full specification, for example, one 100-year increase in the time since being annexed is associated with roughly 0.52 more days contributed per household, which is roughly 13 percent of the mean value. The time since being annexed also accounts for approximately 11 percent of the total variation in the number of days contributed, which is one half the part explained by all the demographic controls. Districts having higher average head age and percentage of Kinh heads are significantly associated with less days contributed, while higher average household size and percentage of male heads are significantly associated with more days contributed.

[Table 1 is about here.]

Table 2 reports the result of including additional control variables. The estimated coefficient of the time since being annexed remains relatively unchanged in sign, absolute value, and statistical significance. The estimated coefficient of per capita expenditure has a positive sign but is insignificant. The estimated coefficient of the percentage of cropland is significant and negative when entering alone but turns insignificant and positive in the full specification. The significant and negative estimated coefficient of the caloric suitability index indicates that higher natural land productivity is associated with lower number of days contributed, which is in line with the finding in Litina (2016). While the positive estimated coefficient of ruggedness turns insignificant in the full specification, the significant estimated coefficients of the average elevation and population density suggest that more highly elevated and densely populated areas are respectively associated with higher and lower numbers of days contributed. These additional control variables all together only account for roughly another five percent of the total variation in the number of days contributed per household. While latitude has an insignificant estimated coefficient, adding it results in a substantially inflated standard error of the estimated coefficient of the time since being annexed and even reduces the adjusted R^2. This is not a surprise given the high correlation between the two variables (Pearson's coefficient = 0.97). We thus exclude latitude in the rest of the empirical analysis.

[Table 2 is about here.]

3.4 Robustness Analysis

Although many control variables have been introduced, it is hardly enough to claim that there is no omitted variable bias left in the regression model. As mentioned above, a stronger test is to restrict the analysis to the sub-sample of annexed areas and employ the walking distance to modern Nghe An as an instrumental variable for the time since being annexed into historical Vietnam. Table 3 reports the results of the instrumental variable estimation using 2SLS estimator. The estimated coefficient of the time since being annexed remains relatively unchanged in sign and statistical significance whether or not all control variables are included. Relative to the mean number of days contributed, the marginal effect increases substantially, which indicates the presence of measurement errors. In addition, the first-stage estimations show that the distance to Nghe An is

a strong predictor of the time since being annexed, i.e. it has significant and negative estimated coefficients. Finally, we reject the exogeneity assumption of the time since annexed when it enters the regression alone, but we cannot do so when demographic characteristics and additional controls are included. Conditioning on our control variables, the time since being annexed is arguably exogenous and hence the above result from the OLS estimator is valid.

[Table 3 is about here.]

There are various complications of the historical process that may influence our interpretation of the above result as supporting the voluntary settlement hypothesis. First, historical Vietnamese immigrants (the Kinh ethnicity) often inhabited along the coastal plain with their traditional rice agriculture, while the highland areas were mainly inhabited by other ethnic tribes. After the Reunification in 1975, the Kinh started to migrate more to the highland areas in large scale through a state-sponsored program under the central planning economy to establish new production zones. This later migration, therefore, might be different from the one that happened in the historical times. We thus exclude from the analysis all districts in the highland areas, i.e. where average elevation are above 500 meters.[15] Furthermore, we also exclude two provinces of Ha Noi and Ho Chi Minh City, which are two biggest venues for immigrants in the modern time. Table 4 shows that the estimated coefficient of the time since being annexed remains relatively unchanged in sign, absolute value, and statistical significance (Columns 1-3).

[Table 4 is about here.]

Second, we choose the first stable state after gaining independence from historical China as the starting point in constructing the time since being annexed into historical Vietnam. One may, however, argue that this choice is just arbitrary. In addition, these initial areas were different from the annexed areas in many important aspects. For example, as mentioned above, historical Vietnamese immigrants had to face harsh living conditions and inhabited along various indigenous groups as they moved into the annexed areas. Finally, historical studies have also argued that the annexed areas in the south were

[15]The result is robust to other values such as 400 and 600 meters, and is available upon request.

governed by a more flexible and open administrative system (e.g. Tana 1998). These environmental and institutional factors might be the causes of cultural differences in the annexed areas other than the immigration of individualistic people. To purify the voluntary settlement hypothesis further, we thus continue to exclude all districts in the initial territory of the first historical state from the analysis. Table 4 shows that the estimated coefficient of the time since being annexed remains relatively unchanged in sign, absolute value, and statistical significance in this subsample (Columns 4-6).

Finally, around 20 years before taking the whole Vietnam, the French colonizers had taken control of the furthest south (Cochinchina) in 1858 CE, the place where they inhabited the most. During 20 years of the Second Indochina War, these areas were also subjected to a capitalist regime influenced by the United States. One therefore may argue that these more recent experiences with the Western world had created the more individualistic culture in the Mekong River plain other than the immigration of individualistic people in the further past. To examine this possibility, we include a dummy variable for the former Cochinchina, i.e. the modern South East and Mekong River Delta regions, in the regression. Although this dummy variable apparently captures many other things besides earlier Western contact, this exercise can be taken as the simplest check of Western influence. Table 4 shows that the estimated coefficient of the time since being annexed remains relatively unchanged in sign, absolute value, and statistical significance (Columns 7-9). Meanwhile, the dummy variable for Cochinchina has negative estimated coefficients, but highly insignificant.

Tables A2 and A3 in the Appendix also report the results of replicating the above analysis to the percentage of households and the number of members per household making labor contributions to public goods production. The estimated coefficients of the time since being annexed are also significant and positive as regards these two indicators. Areas being annexed earlier nowadays on average have a higher percentage of households and number of members per household making labor contributions. The marginal effects are also reasonably large and stable in various specifications. These findings altogether thus indicate that areas being annexed earlier into historical Vietnam nowadays have higher levels of labor contribution to public goods production in terms of not only intensity but also prevalence. In the Supplementary Materials, we also conduct additional robustness checks including aggregating variables from the VHLSS with sampling weights and us-

ing standard errors adjusted for potential spatial autocorrelation followed the method developed by Conley (1999). The results in general remain relatively intact.

3.5 Discussion

The analysis so far has shown that labor contribution to public goods production is both more intensive and prevalent in areas being annexed earlier into historical Vietnam. More importantly, this result is not driven by population movements in the modern day, differential living environment in the annexed areas or past contacts with the Western world in the south. These findings altogether lend support to the voluntary settlement hypothesis as the main explanation behind the cultural differences along the IC dimension across regions in Vietnam.

There nonetheless remains some limitations. First, our empirical strategy may not fully account for other institutional channels through which the time since being annexed into historical Vietnam might influence labor contribution to public goods production. For example, the government might spend less on infrastructure in areas being annexed earlier, which in turn might induce people in these areas to contribute more. In this case, the time since being annexed does not have an influence on labor contribution through the cultural channel, but instead through the government spending allocation. Second, the survey data we have used only have information until 2002, which naturally leads to the question if our finding can be carried over today. Third, money is certainly another form of contribution to public goods besides labor in the modern day. Hence, we should also be able to observe similar variation across societies when money contributions are involved rather than labor. Finally, the survey data reflect a situation where people may observe the labor contributions of others from the same district, and one may be curious if this cultural difference presents in an impersonal setting where decision is anonymous. We thus conduct a lab-in-the-field experiment to examine these questions in details.

4 Experimental Evidence

4.1 Site Selection

Before describing the public goods experiment, we explain our strategy in choosing experimental sites in ways to differentiate the voluntary settlement hypothesis from other potential channels as discussed above. First, we focus on the annexed areas to rule out differences in historical living environment. Second, we restrict to the territory of the

former south during the Second Indochina War (1954-1975) to eliminate the confounding effects of different regimes in the north and the south during this period. Third, we pick provinces located along the coast, which was the route that historical Vietnamese often migrated and settled down. In these provinces, we also pick rural districts locating along the coast. Fourth, we choose provinces, and rural districts in them, that were historically inhabited mainly by the Kinh ethnicity (historical Vietnamese) and whose populations have been living there for many generations. In this subsample, we randomly pick one district with the longest time since being annexed into historical Vietnam and one district in the youngest region, i.e. the Mekong Delta. Finally, we use the VHLSS 2014, which is the latest round, to find districts having similar levels of economic development based on income per capita. This process narrows our selection to one rural district in Thua Thien Hue Province and one rural district in Ben Tre Province (see Figure 1).

Each district in Vietnam has around 3 to 5 high schools, and we choose one in each. To maximize the sample size and representativeness, given the resource constraint, we pick schools that have more than 200 students and locate in the centers of the districts. The participation rates of students in the experiment are relatively similar, i.e. 70 percent in the north (140 out of 200) and 73 percent (235 out of 320) in the south. In accordance with our expectation, 98 percent of the subjects were born in the chosen districts, while 95 percent their parents were so. All subjects are from the same cohort, i.e. 97 percent of them were born in 1999. All subjects from the north, and 97 percent from the south, say that their households have contributed in at least one among many public building activities including road, school, health facilities, dam, and commune houses.

4.2 The Experiment

We conduct a public goods experiment with high-school students from the north and the south of Vietnam to analyze attitudes towards cooperation.[16] By using an experimental approach, we can keep the institutional setting constant and carefully select our sites to enable a causal investigation of the voluntary settlement hypothesis. We use high-school students as our subjects in the experiment since they are old enough to embody the cultural environments of the places where they grew up, but still have not been affected

[16]For a general discussion on public goods experiment see Zelmer (2003) and Chaudhuri (2011).

by living or working outside their community.[17] The experimental design thus allows a direct test of the voluntary settlement hypothesis by comparing the contribution levels to the public goods between the selected sites. In other words, we expect subjects in the north to share a stronger cultural norm of cooperation and hence on average contribute at a higher level compared to subjects from the south.

We build our design on the one-shot linear public goods experiment developed by Fischbacher, Gächter, and Fehr (2001). For other experiments using this design, see, e.g. Kocher et al. (2008), Herrmann and Thöni (2009), Fischbacher and Gächter (2010), Fischbacher, Gächter, and Quercia (2012), and Martinsson, Villegas-Palacio, and Wollbrant (2015). We begin by describing the public goods experiment before discussing the details of Fischbacher, Gächter, and Fehr (2001)'s design. The subjects are randomly assigned to groups of three, each member comes from a different class at the high school. This feature of the design is chosen to avoid that subjects are allocated to groups consisting of classmates. All subjects receive an endowment of 20 tokens and have to decide simultaneously how much of their endowments to invest in a public good. The marginal per capita return ($MPCR$) from the public good is 0.5, which means that each token given to the public goods results in 0.5 token to all group members including the one who gives the token. If a subject is rational and selfish, then a $MPCR$ below one leads to a dominant strategy to free ride, i.e. to contribute zero to the public good. Nevertheless, it is socially optimal to contribute the whole endowment if $MPCR \times n > 1$, where n is the number of group members. Our choice of the $MPCR$ of 0.5 thus generates the conflict between private and social optima that characterizes a public good. The payoff for subject i is given by two components, (i) the amount of the endowment that is not invested in the public goods $(20 - c_i)$ and (ii) the return from the public good:

$$\pi_i = (20 - c_i) + (0.5 \sum_{j=1}^{3} c_j).$$

Each token earned is exchanged to money at the rate of one token equals 3000 Vietnamese Dong. The experiment is calibrated such that each student on average receives a monetary payoff worth roughly three daily meals.

The specific design feature of Fischbacher, Gächter, and Fehr (2001) is that they use the strategy method. Each subject makes two types of contribution decisions to the

[17]This strategy has also been used earlier in the literature, e.g. Kocher, Martinsson, and Visser (2012).

public good: (i) unconditional contribution and (ii) conditional contribution. In the unconditional decision, which is similar to a standard public goods experiment, each subject states how much she would like to contribute to the public goods from her endowment of 20 tokens. In the conditional decision, a subject states how much she would like to contribute conditional on the average contribution of the other two group members. We round off the possible average contributions of the other two group members to integer numbers ranging from 0 to 20. A subject thus makes 21 conditional decisions in total. All decisions are incentive compatible by using the following strategy. For two randomly selected group members, it is the unconditional contribution to the public good that is pay-off relevant. For the third member, the average unconditional contribution of the other two group members is calculated, and the contribution of the third member is then determined from her conditional contribution given the average contribution of the other two group members. Thus, when making her decisions, a subject does not know which decision would be pay-off relevant. The experimental procedure was carefully described to the subjects before the experiment began, and various examples were given to ensure that the subjects understood.[18]

The strength of the strategy method is that subjects can be categorized into different types based on their conditional contributions to the public good. We use the same classification proposed in Fischbacher, Gächter, and Fehr (2001). A subject is classified as a "conditional cooperator" if her conditional contribution increases weakly monotonically with the average contribution of the other two group members or the relationship between her conditional contribution and the others' average contribution is positive and significant at 1 percent significance level using a Spearman rank correlation coefficient. A "free rider" is a subject who contributes zero to the public good for all levels of the other members' average contributions. A "hump-shaped" contributor is characterized by a subject who shows weakly monotonically increasing contributions with a Spearman rank correlation coefficient at 1 percent significance level, which is similar to a conditional contributor, but only up to an inflection point. For average contribution levels by others above this level, her own conditional contributions decrease weakly monotonically or show a significant and

[18]Details on the experimental instruction can be found in the Supplementary Materials. The public goods experiment is the first experiment in a bigger project, in which the second experiment is pay-off independent from the first one.

negative Spearman rank correlation coefficient at the 1 percent significance level. Those who cannot be categorized based on above criteria are referred to as "others".

4.3 Result

The average contribution level in our sample is 6.92, which is fairly similar to what have been found in the literature (Zelmer 2003; Chaudhuri 2011). Table 5 shows that subjects from the north unconditionally contributed 7.50 tokens on average while their peers from the south did so at 6.58 tokens, and the difference is significant (Mann-Whitney U test). Subjects from the north on average also gave a significantly higher guess about the average contribution of other members compared to their peers in the south, i.e. 8.25 versus 7.60 tokens. Except for free riders, other types in the north on average have higher levels of unconditional contribution and guess about the average contribution of other members compared to their counterparts in the south. The distributions of types are similar across the north and the south, i.e. the percentages of free riders are 3.62 in the north and 5.53 in the south while the corresponding numbers for conditional cooperators are 52.17 and 54.04 percent. We indeed cannot reject the null hypothesis that the compositions of types in the north and the south are drawn from the same distribution (Kolmogorov-Smirnov test). This finding suggests that the north-south difference in contribution behaviors is driven by belief rather than the composition of types.

[Table 5 is about here.]

To examine the result further, we use regress analysis to control for measures of risk attitude, trust perception, and social preferences as well as standard socioeconomic variables. The basic idea is to examine if these variables influence the contribution decision rather than the belief about the average contribution of the other two group members. Risk attitude is measured using a question validated by Dohmen et al. (2011) in which respondents are asked to rate their willingness to take risks on a scale from 0 to 10. Subjects also answer a question followed the format in the WVS as regards the extent to which they trust their friends.[19] We code trust "completely" and "somewhat" as one, and do not trust very much and do not trust at all as zero. Social preferences are captured through six hypothetical situations of allocating money between the subject and an anonymous

[19]We did not get the approval of the authority to ask about trust in schoolmates and villagers.

person proposed by Murphy, Ackermann, and Handgraaf (2011), based on which subjects are classified as altruistic, prosocial, individualistic, and competitive. We code altruists and prosocials as one, and the rest as zero. The socioeconomic variables include gender, household size, and a wealth index.[20]

Table 6 shows that the estimated coefficient of the north dummy is significantly positive with a marginal effect slightly smaller than the mean difference when standard socioeconomic characteristics are controlled for. The role of belief about the average contribution of the other two group members in driving the unconditional contribution is confirmed in the regression where it has a significant and positive estimated coefficient. Conditioning on this belief, the estimated coefficient of the north dummy is reduced substantially in both size and significance. Furthermore, the influence of belief about the average contribution of the other two group members on one's contribution behavior is found to be stronger in the north, which is reflected through the significant and positive estimated coefficient of the interaction term between belief and the north dummy.

[Table 6 is about here.]

The dummy for conditional cooperators has an insignificant estimated coefficient. The measure of risk attitude has a significant and positive estimated coefficient, which indicates that more risk-taking subjects are willing to contribute more. Trust in friends has an insignificant estimated coefficient. The socially oriented dummy (altruists and prosocials) also has an insignificant estimated coefficient. Finally, all socioeconomic variables have insignificant estimated coefficients (not shown). When all control variables are included separately or together, the estimated coefficients of the north dummy and its interaction with belief about the average contribution of the other two group members remain qualitatively intact.[21] To sum up, the regression analysis thus confirms that belief is the main driver behind the north-south difference in contribution behaviors. We take this finding as evidence for a stronger cultural norm of cooperation in the north.

[20] The wealth index is constructed by extracting the first principal component of six variables measuring the numbers of mobile phone, computer, motorbike, refrigerator, gas cooker, and air conditioner that the households possess.

[21] We also include the interaction terms between the north dummy and four control variables and the result remains qualitatively intact. Details are available upon request.

5 Concluding Remarks

The IC contrast has been found to be a powerful predictor of long-run democratic and economic development in a large sample of countries. Thus, why some societies have become more collectivist or individualistic than others is a crucial question in understanding long-run comparative development. In the present paper, we study the voluntary settlement hypothesis, postulating that cultural differences along the IC dimension are driven by the out-migration of individualistic people from collectivist societies to settled down in less (more) collectivist (individualistic) peripheries. We recognize the grand territorial expansion from the eleventh to the eighteenth centuries in historical Vietnam as an ideal empirical setting. During this period, historical Vietnam gradually expanded its territory southward along the coast from the Red River Delta to the Mekong River Delta through various waves of conquest and migration to form a country as today.

We examine the ability to solve collective action problems, which is a critical feature of collectivist societies, by using data on contribution to public goods. Combining archives with household survey, we find that areas being annexed earlier into historical Vietnam nowadays have higher levels of labor contribution to public goods production in terms of not only intensity but also prevalence. We eliminate many potential interpretations for this finding including differential living conditions in the annexed areas, Western influence in the Mekong River plain and important population movements in the modern history. To strengthen our analysis, we also conduct a public goods experiment with high-school students from the north and the south in which we carefully select experimental sites in ways to differentiate the voluntary settlement hypothesis from other potential channels. We find that students from the north contribute significantly more in the public goods experiment compared to their southern peers, and that the result is mainly driven by the belief about the contribution levels of other subjects. We argue that the southward out-migration of individualistic people was the main mechanism behind these cultural differences, which is also in line with many historical accounts.

We believe that the present paper provides a valuable input for understanding long-run comparative development. First and foremost, the migration patterns in the distant past played a crucial role in explaining cultural divergence across modern societies. As time goes on, similar processes may continue to enhance cultural differences across societies, which in turn may influence comparative development in the coming times.

References

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2012. "Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration". *American Economic Review* 102 (5): 1832–1856.

Adams, John, and Nancy Hancock. 1970. "Land and Economy in Traditional Vietnam". *Journal of Southeast Asian Studies* 1 (2): 90–98.

Alesina, Alberto, and Nicola Fuchs-Schündeln. 2007. "Goodbye Lenin (or Not?): The Effect of Communism on People". *American Economic Review* 97 (4): 1507–1528.

Alesina, Alberto, and Paola Giuliano. 2015. "Culture and Institutions". *Journal of Economic Literature* 53 (4): 898–944.

— . 2014. "Family Ties". In *Handbook of Economic Growth*, ed. by Philippe Aghion and Steven N. Durlauf, vol. 2A, 177–215. North-Holland: Elsevier.

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

Algan, Yann, and Pierre Cahuc. 2014. "Trust, Growth, and Well-Being: New Evidence and Policy Implications". In *Handbook of Economic Growth*, ed. by Philippe Aghion and Steven N. Durlauf, vol. 2A, 49–120. North-Holland: Elsevier.

Ashwill, Mark A., and Thai Ngoc Diep. 2005. *Vietnam Today: A Guide to a Nation at a Crossroads*. Maine: Intercultural Press.

Bamber, Scott. 1993. "Diseases of Antiquity and the Premodern Period in Southeast Asia". In *The Cambridge World History of Human Disease*, ed. by Kenneth F. Kiple, 425–440. Cambridge: Cambridge University Press.

Basten, Christoph, and Frank Betz. 2013. "Beyond Work Ethic: Religion, Individual, and Political Preferences". *American Economic Journal: Economic Policy* 5 (3): 67–91.

Bazzi, Samuel, Martin Fiszbein, and Mesay Gebresilasse. 2017. "Frontier Culture: The Roots and Persistence of "Rugged Individualism" in the United States". *NBER Working Paper No. 23997*.

Becker, Sascha O., et al. 2016. "The Empire Is Dead, Long Live the Empire! Long-Run Persistence of Trust and Corruption in the Bureaucracy". *Economic Journal* 126 (590): 40–74.

Bigoni, Maria, et al. 2016. "Amoral Familism, Social Capital, or Trust? The Behavioural Foundations of the Italian North-South Divide". *Economic Journal* 126 (594): 1318–1341.

Briggs, Lawrence Palmer. 1947. "A Sketch of Cambodian History". *The Far Eastern Quarterly* 6 (4): 345–363.

Cao, Huy-Giu (Translator). 2012. *Dai Viet Su Ky Toan Thu*. Ha Noi: Thoi Dai.

Chaudhuri, Ananish. 2011. "Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature". *Experimental Economics* 14 (1): 47–83.

CIESIN and CIAT. 2005. *Gridded Population of the World Version 3 (GPWv3): Population Density Grids*. New York: NASA Socioeconomic Data & Applications Center.

Conley, Timothy G. 1999. "GMM Estimation with Cross Sectional Dependence". *Journal of Econometrics* 92 (1): 1–45.

Dao, Duy-Anh. 2005. *Dat Nuoc Viet Nam Qua Cac Doi*. Ha Noi: Van Hoa Thong Tin.

Dell, Melissa, Nathan Lane, and Pablo Querubin. 2017. "The Historical State, Local Collective Action, and Economic Development in Vietnam". *NBER Working Paper No. 23208*.

Dien, Khong. 2002. *Population and Ethno-Demography in Vietnam*. Chiang Mai: Silkworm Books.

Doepke, Matthias, and Fabrizio Zilibotti. 2014. "Culture, Entrepreneurship, and Growth". In *Handbook of Economic Growth*, ed. by Philippe Aghion and Steven N. Durlauf, vol. 2A, 1–48. North-Holland: Elsevier.

Dohmen, Thomas, et al. 2011. "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences". *Journal of the European Economic Association* 9 (3): 522–550.

Engelbert, Thomas. 1994. "The Khmer in Southern Vietnam-Cambodians or Vietnamese?" In *Nationalism and Ethnicity in Southeast Asia*, ed. by Ingrid Wessel, 1:155–196. Berlin: Humboldt University.

Fernández, Raquel. 2011. "Does Culture Matter?" In *Handbook of Social Economics*, ed. by Jess Benhabib, Alberto Bisin, and Matthew Jackson, vol. 1A, 481–510. North-Holland: Elsevier.

Fincher, Corey L., et al. 2008. "Pathogen Prevalence Predicts Human Cross-Cultural Variability in Individualism/Collectivism". *Proceedings of the Royal Society B: Biological Sciences* 275 (1640): 1279–1285.

Fischbacher, Urs, and Simon Gächter. 2010. "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments". *American Economic Review* 100 (1): 541–56.

Fischbacher, Urs, Simon Gächter, and Ernst Fehr. 2001. "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment". *Economics Letters* 71 (3): 397–404.

Fischbacher, Urs, Simon Gächter, and Simone Quercia. 2012. "The Behavioral Validity of the Strategy Method in Public Good Experiments". *Journal of Economic Psychology* 33 (4): 897–913.

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

Gorodnichenko, Yuriy, and Gérard Roland. 2015. "Culture, Institutions and Democratization". *NBER Working Paper No. 21117*.

— . 2017. "Culture, Institutions and the Wealth of Nations". *Review of Economics and Statistics* 99 (3): 402–416.

— . 2012. "Understanding the Individualism-Collectivism Cleavage and its Effects: Lessons from Cultural Psychology". In *Institutions and Comparative Economic Development*, ed. by Masahiko Aoki, Timur Kuran, and Gérard Roland, 213–236. London: Palgrave.

— . 2011. "Which Dimensions of Culture Matter for Long-Run Growth?" *American Economic Review: Papers and Proceedings* 101 (3): 492–498.

Guernier, Vanina, Michael Hochberg, and Jean-Francois Guégan. 2004. "Ecology Drives the Worldwide Distribution of Human Diseases". *PLoS Biology* 2 (6): 740–746.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2011. "Civic Capital as the Missing Link". In *Handbook of Social Economics*, ed. by Jess Benhabib, Alberto Bisin, and Matthew Jackson, vol. 1A, 417–480. North-Holland: Elsevier.

— . 2006. "Does Culture Affect Economic Outcomes?" *Journal of Economic Perspectives* 20 (2): 23–48.

— . 2016. "Long-term Persistence". *Journal of the European Economic Association* 14 (6): 1401–1436.

Hansen, Casper Worm, Peter Sandholt Jensen, and Christian Volmar Skovsgaard. 2015. "Modern Gender Roles and Agricultural History: The Neolithic Inheritance". *Journal of Economic Growth* 20 (4): 365–404.

Herrmann, Benedikt, and Christian Thöni. 2009. "Measuring Conditional Cooperation: A Replication Study in Russia". *Experimental Economics* 12 (1): 87–92.

Inglehart, Ronald, and Wayne E. Baker. 2000. "Modernization, Cultural Change, and the Persistence of Traditional Values". *American Sociological Review* 65 (1): 19–51.

Kitayama, Shinobu, et al. 2009. "A Cultural Task Analysis of Implicit Independence: Comparing North America, Western Europe, and East Asia". *Journal of Personality and Social Psychology* 97 (2): 236–255.

Kitayama, Shinobu, et al. 2006. "Voluntary Settlement and the Spirit of Independence: Evidence from Japan's 'Northern Frontier'". *Journal of Personality and Social Psychology* 91, no. 3 (): 369–384.

Kocher, Martin, Peter Martinsson, and Martine Visser. 2012. "Social Background, Cooperative Behavior, and Norm Enforcement". *Journal of Economic Behavior and Organization* 81 (2): 341–354.

Kocher, Martin G., et al. 2008. "Conditional Cooperation on Three Continents". *Economics Letters* 101 (3): 175–178.

Le, Thanh-Khoi. 1982. *Histoire du Vietnam des Origines à 1858*. Paris: Sudestasie.

Litina, Anastasia. 2016. "Natural Land Productivity, Cooperation and Comparative Development". *Journal of Economic Growth* 21 (4): 351–408.

Loveland, T. R., et al. 2000. "Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1 km AVHRR Data". *International Journal of Remote Sensing* 21 (6-7): 1303–1330.

Lowes, Sara, et al. 2017. "The Evolution of Culture and Institutions: Evidence from the Kuba Kingdom". *Econometrica* (forthcoming).

Martinsson, Peter, Clara Villegas-Palacio, and Conny Wollbrant. 2015. "Cooperation and Social Classes: Evidence from Colombia". *Social Choice and Welfare* 45 (4): 829–848.

Murphy, Ryan O., Kurt A. Ackermann, and Michel J. J. Handgraaf. 2011. "Measuring Social Value Orientation". *Judgment and Decision Making* 6 (8): 771–781.

Nguyen, Khac-Su, Minh-Huyen Pham, and Trung-Tin Tong. 2004. "Northern Vietnam: From the Neolithic to the Han Period". In *Southeast Asia: From Prehistory to History*, ed. by Ian Glover and Peter Bellwood, 177–201. New York: Routledge Curzon.

Nguyen, Thi-Dieu. 1999. *The Mekong River and the Struggle for Indochina: Water, War, and Peace*. Westport: Praeger.

Nguyen, Ngoc-Tinh (Translator). 2002. *Dai Nam Thuc Luc*. Ha Noi: Giao Duc.

Nunn, Nathan. 2012. "Culture and the Historical Process". *Economic History of Developing Regions* 27 (sup1): S108–S126.

— . 2014. “Historical Development”. In *Handbook of Economic Growth*, ed. by Philippe Aghion and Steven N. Durlauf, vol. 2A, 347–402. North-Holland: Elsevier.

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

Nunn, Nathan, and Leonard Wantchekon. 2011. “The Slave Trade and the Origins of Mistrust in Africa”. *American Economic Review* 101 (7): 3221–52.

Olsson, Ola, and Christopher Paik. 2016a. “A Western Reversal Since the Neolithic? The Long-Run Impact of Early Agriculture”. *Working Paper*.

— . 2016b. “Long-Run Cultural Divergence: Evidence From the Neolithic Revolution”. *Journal of Development Economics* 122:197–213.

Phan, Huy-Le, et al. 2011. *Qua Trinh Hinh Thanh Va Phat Trien Vung Dat Nam Bo*. Ha Noi: Hoi Khoa Hoc Lich Su Viet Nam.

Riley, Shawn J., Stephen D. DeGloria, and Robert Elliot. 1999. “A Terrain Ruggedness Index That Quantifies Topographic Heterogeneity”. *Intermountain Journal of Sciences* 5 (1-4): 23–27.

Spolaore, Enrico, and Romain Wacziarg. 2013. “How Deep Are the Roots of Economic Development?” *Journal of Economic Literature* 51 (2): 325–369.

Talhelm, Thomas, et al. 2014. “Large-Scale Psychological Differences Within China Explained by Rice Versus Wheat Agriculture”. *Science* 344 (6184): 603–608.

Tana, Li. 1998. *Nguyen Cochinchina: Southern Vietnam in the Seventeenth and Eighteenth Centuries*. Ithaca: Cornell University Press.

Tanaka, Tomomi, Colin F. Camerer, and Quang Nguyen. 2010. “Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam”. *American Economic Review* 100 (1): 557–571.

Taylor, Keith W. 2013. *A History of the Vietnamese*. Cambridge: Cambridge University Press.

— . 1993. “Diseases and Disease Ecology of the Modern Period in Southeast Asia”. In *The Cambridge World History of Human Disease*, ed. by Kenneth F. Kiple, 440–444. Cambridge: Cambridge University Press.

— . 1983. *The Birth of Vietnam*. Berkeley: University of California Press.

Triandis, Harry C. 1995. *Individualism and Collectivism*. Boulder: Westview Press.

Vandello, Joseph A., and Dov Cohen. 1999. "Patterns of Individualism and Collectivism Across the United States". *Journal of Personality and Social Psychology* 77 (2): 279–292.

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

Whitmore, John K. 1997. "Literati Culture and Integration in Dai Viet, c. 1430-c. 1840". *Modern Asian Studies* 31 (3): 665–687.

— . 1984. "Social Organization and Confucian Thought in Vietnam". *Journal of Southeast Asian Studies* 15 (2): 296–306.

Wook, Choi Byung. 2004. *Southern Vietnam under the Reign of Minh Mang (1820–1841)*. Ithaca: Cornell University Press.

Zelmer, Jennifer. 2003. "Linear Public Goods Experiments: A Meta-Analysis". *Experimental Economics* 6 (3): 299–310.

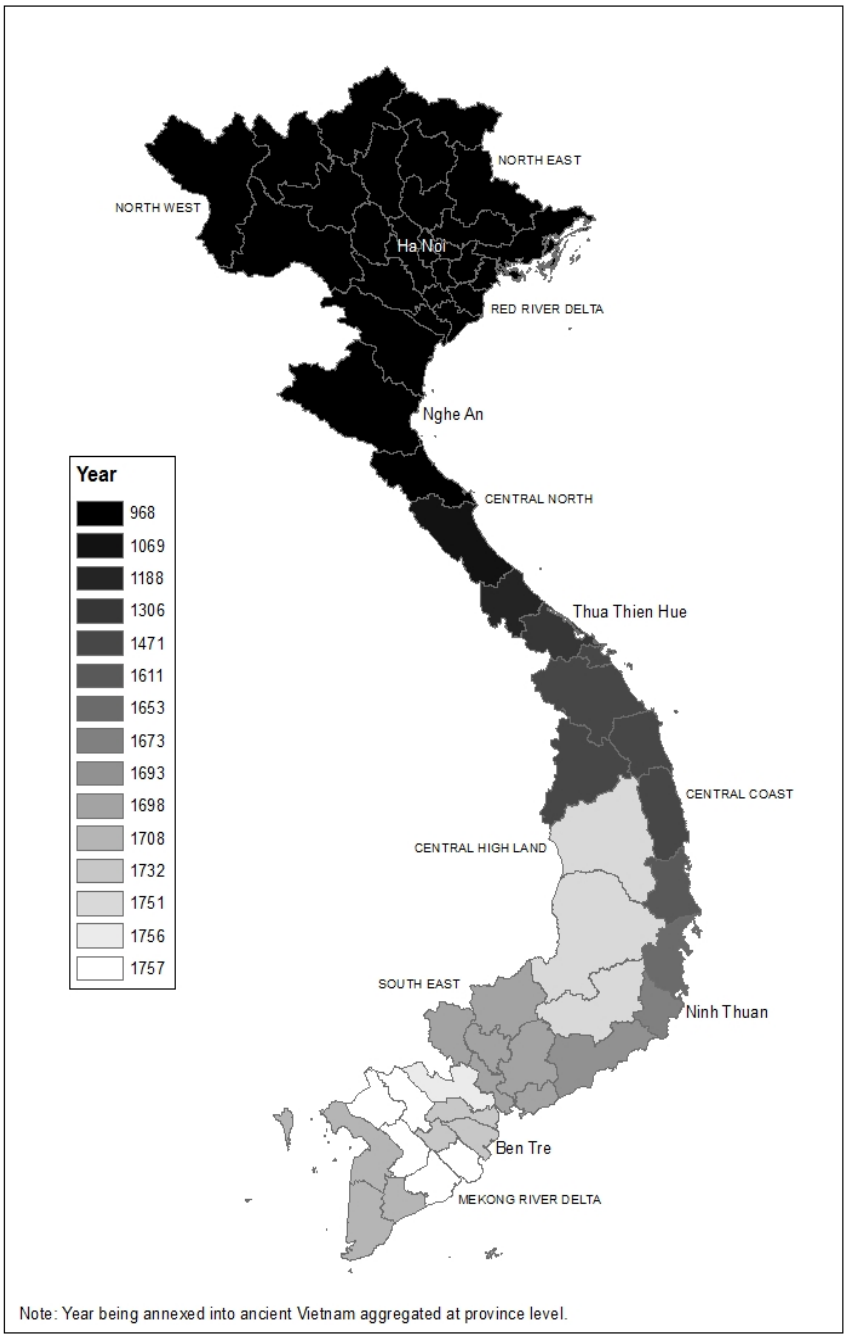
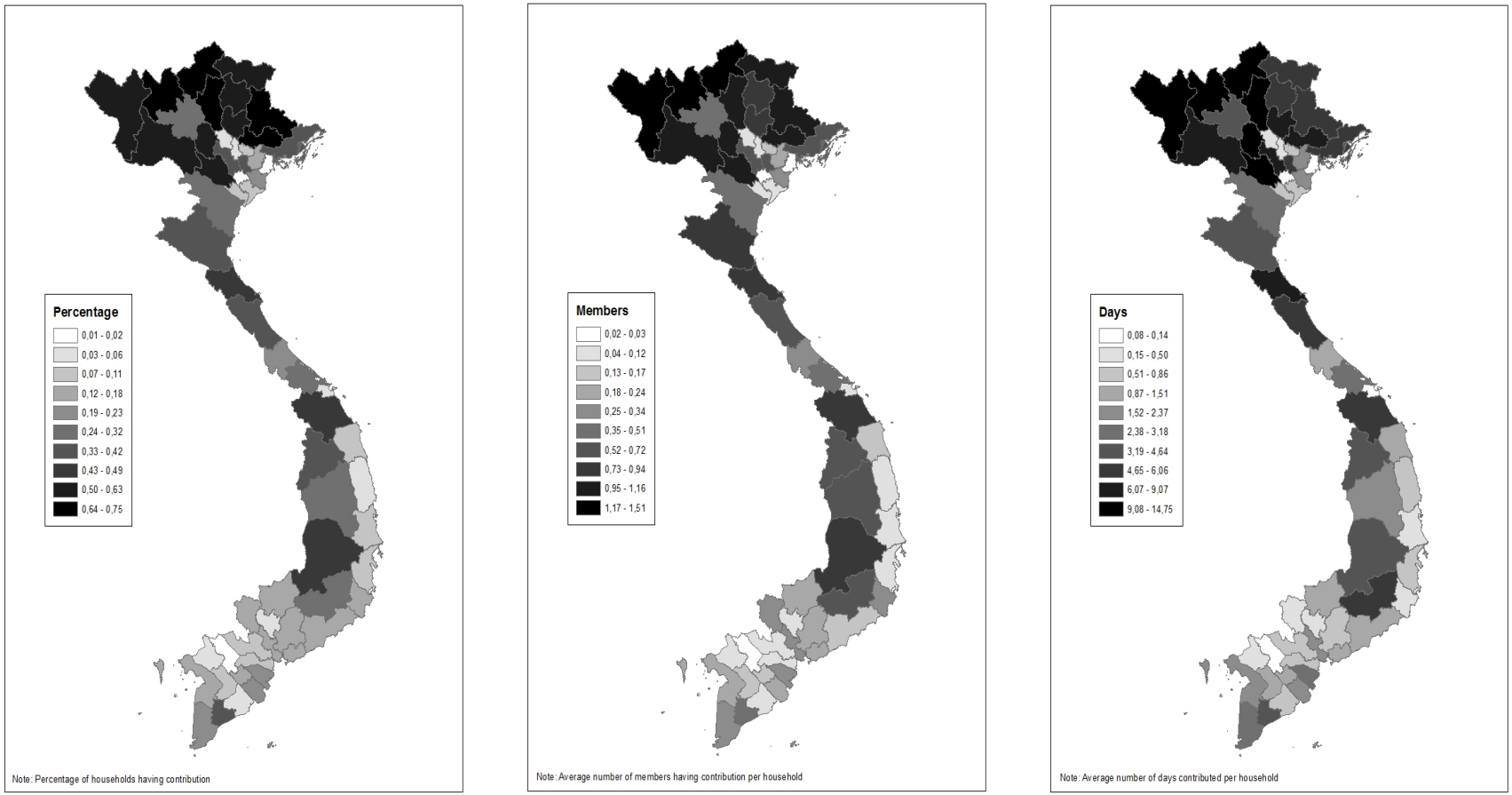


Figure 1. The Vietnamese Southern Advance

Source: Authors' construction.



A. Percentage of Households B. Number of Members C. Number of Days

Figure 2. Labor Contribution to Public Goods Production

Source: VHLSS 2002, aggregation without sampling weights.

Table 1. Baseline Regression

	Number of days contributed per household							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Time since annexed	0.535***	0.596***	0.510***	0.453***	0.409***	0.418***	0.763***	0.520***
	(0.058)	(0.058)	(0.056)	(0.056)	(0.052)	(0.053)	(0.073)	(0.081)
Average household size		4.018***						2.617***
		(0.694)						(0.828)
Percentage of urban households			-3.099***					-0.522
			(0.537)					(0.684)
Average head age				-0.282***				-0.106**
				(0.050)				(0.053)
Percentage of male heads					13.098***			3.676**
					(1.676)			(1.714)
Percentage of Kinh heads						-5.872***		-2.267**
						(0.782)		(0.883)
Average head schooling years							-0.904***	-0.242
							(0.133)	(0.166)
Constant	0.427	-15.123***	1.276***	14.339***	-8.950***	5.848***	4.875***	-3.661
	(0.317)	(2.738)	(0.327)	(2.449)	(1.291)	(0.773)	(0.685)	(4.178)
Mean value of dep. var.	4.05	4.05	4.05	4.05	4.05	4.05	4.05	4.05
Adjusted R^2	0.113	0.235	0.143	0.159	0.202	0.237	0.209	0.308
Observations	607	607	607	607	607	607	607	607

Note: OLS estimator with robust standard errors in parentheses.
* p<0.1, ** p<0.05, *** p<0.01

Table 2. Additional Control Variables

	Number of days contributed per household							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Time since annexed	0.530***	0.542***	0.337***	0.508***	0.421***	0.530***	0.398***	0.575*
	(0.084)	(0.081)	(0.104)	(0.080)	(0.081)	(0.083)	(0.117)	(0.308)
Per capita expenditure	0.088						0.195	0.159
	(0.113)						(0.221)	(0.224)
Percentage of cropland area		-2.263***					-0.714	-0.890
		(0.564)					(0.782)	(0.815)
Caloric suitability			-0.461***				-0.385***	-0.396***
			(0.123)				(0.126)	(0.127)
Elevation				0.004***			0.003*	0.003**
				(0.001)			(0.002)	(0.002)
Ruggedness					0.851***		0.113	0.076
					(0.219)		(0.376)	(0.379)
Population density						0.010	-0.725**	-0.647**
						(0.069)	(0.317)	(0.320)
Latitude								-0.146
								(0.237)
Constant	-3.892	-4.377	4.544	-7.780*	-6.695	-3.832	-1.443	0.354
	(4.177)	(4.135)	(5.453)	(4.157)	(4.090)	(4.288)	(5.828)	(6.120)
Demographic controls	YES	YES	YES	YES	YES	YES	YES	YES
Mean value of dep. var.	4.05	4.05	4.27	4.05	4.05	4.11	4.27	4.27
Adjusted R^2	0.308	0.322	0.332	0.330	0.329	0.312	0.358	0.358
Observations	607	607	555	607	607	591	552	552

Note: OLS estimator with robust standard errors in parentheses. Demographic controls are household characteristics aggregated at district level including the number of members, residing in urban area, and head ethnicity, gender, age and schooling years.
* p<0.1, ** p<0.05, *** p<0.01

Table 3. IV Estimation

Second stage	Number of days contributed per household					
	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexed	0.409***	0.377**	0.432***	0.684***	0.535***	0.413**
	(0.152)	(0.149)	(0.162)	(0.172)	(0.163)	(0.195)
First stage	Time since being annexed					
	(1)	(2)	(3)	(4)	(5)	(6)
Distance to Nghe An	na	na	na	-0.376***	-0.404***	-0.485***
	na	na	na	(0.022)	(0.022)	(0.026)
Exogeneity (p-value)	na	na	na	0.007	0.111	0.887
Kleibergen-Paap F	na	na	na	289	331	359
Adjusted R^2	na	na	na	0.645	0.691	0.819
Demographic controls	NO	YES	YES	NO	YES	YES
Additional controls	NO	NO	YES	NO	NO	YES
Mean value of dep. var.	2.24	2.24	2.27	2.24	2.24	2.27
Adjusted R^2	0.031	0.098	0.158	na	na	na
Observations	311	311	287	311	311	287

Note: Two stages least squares estimator with robust standard errors in parentheses, constant terms in both stages are suppressed to save space. Demographic controls are household characteristics aggregated at district level including the number of members, residing in urban area, and head ethnicity, gender, age and schooling years. Additional controls include per capita expenditure, percentage of cropland area, caloric suitability index, elevation, terrain ruggedness and population density. All regressions only include districts in the annexed areas.

* p<0.1, ** p<0.05, *** p<0.01

Table 4. Alternative Mechanisms

	Number of days contributed per household								
	Population Movement			Annexed Areas			Western Influence		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Time since annexed	0.469***	0.421***	0.271**	0.467***	0.439**	0.430***	0.371*	0.414*	0.351*
	(0.054)	(0.079)	(0.107)	(0.159)	(0.170)	(0.160)	(0.206)	(0.215)	(0.207)
Cochinchina							-0.518	-0.143	-0.634
							(0.557)	(0.561)	(1.043)
Demographic controls	NO	YES	YES	NO	YES	YES	NO	YES	YES
Additional controls	NO	NO	YES	NO	NO	YES	NO	NO	YES
Mean value of dep. var.	3.37	3.37	3.50	1.87	1.87	1.95	1.87	1.87	1.95
Adjusted R^2	0.122	0.223	0.272	0.060	0.153	0.172	0.060	0.150	0.170
Observations	478	478	445	249	249	238	249	249	238

Note: OLS estimator with robust standard errors in parentheses, constant terms are suppressed to save space. Demographic controls are household characteristics aggregated at district level including the number of members, residing in urban area, and head ethnicity, gender, age and schooling years. Additional controls include per capita expenditure, percentage of cropland area, caloric suitability index, elevation, terrain ruggedness and population density. Columns 1-3 exclude all districts whose elevation are above 500 meters and belong to two municipalities, Ha Noi and Ho Chi Minh City. Columns 4-9 continue to exclude districts in the initial territory of historical Vietnam.

* p<0.1, ** p<0.05, *** p<0.01

Table 5. Distribution of Types, Unconditional Contribution, and Belief

	NORTH ($n = 138$)			SOUTH ($n = 235$)		
	Distribution (%)	Av. Un. Con. (Tokens)	Belief (Tokens)	Distribution (%)	Av. Un. Con. (Tokens)	Belief (Tokens)
All subjects	100	7.50	8.25	100	6.58	7.60
		(3.92)	(3.46)		(4.07)	(4.11)
Conditional cooperators	57.12	7.29	8.19	54.04	6.39	7.88
		(3.20)	(3.72)		(3.35)	(4.07)
Free riders	3.62	0.60	6.20	5.53	2.69	6.54
		(0.55)	(4.55)		(3.88)	(3.55)
Hump-shaped cooperators	5.80	10.50	8.38	4.26	6.90	6.70
		(5.45)	(3.42)		(5.92)	(5.19)
Others	38.41	7.98	8.49	36.17	7.42	7.44
		(4.03)	(2.98)		(4.51)	(4.15)

Note: Standard errors are in parentheses. Av. Un. Con. = average unconditional contribution. Belief = belief about the average unconditional contribution of other two groups members.

Table 6. Unconditional Contribution: Regression Analysis

	Unconditional contribution							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
North	0.892**	0.789*	-1.409	-1.370	-1.345	-1.292	-1.375	-1.183
	(0.435)	(0.410)	(0.966)	(0.976)	(0.976)	(0.970)	(0.973)	(0.999)
Belief		0.294***	0.215***	0.220***	0.228***	0.216***	0.217***	0.235***
		(0.059)	(0.072)	(0.072)	(0.073)	(0.073)	(0.072)	(0.073)
North×Belief			0.276**	0.271**	0.260**	0.275**	0.269**	0.249**
			(0.123)	(0.123)	(0.124)	(0.123)	(0.124)	(0.125)
Conditional cooperator				-0.500				-0.422
				(0.410)				(0.409)
Risk attitude					0.107*			0.104
					(0.065)			(0.066)
Trust in friends						-0.677		-0.572
						(0.627)		(0.614)
Socially oriented							0.340	0.407
							(0.405)	(0.407)
Constant	5.533***	3.618***	4.237***	4.484***	3.632***	4.724***	3.942***	3.938***
	(0.680)	(0.805)	(0.822)	(0.848)	(0.976)	(0.953)	(0.857)	(1.111)
Socioeconomic controls	YES	YES	YES	YES	YES	YES	YES	YES
Adjusted R^2	0.009	0.085	0.097	0.099	0.102	0.099	0.098	0.105
Observations	373	373	373	373	368	373	372	367

Note: OLS estimator with robust standard errors in parentheses. The sample includes 138 students from the north and 235 students from the south. Socioeconomic controls include a male dummy, household size and a wealth index.
* p<0.1, ** p<0.05, *** p<0.01

Appendix

Table A1. Variable Description

Variable	Description	Mean	SD	Min	Max	N	Source
Main Variables							
Having contribution	Percentage of household making labor contribution	0.31	0.33	0	1	607	VHLSS
Members contributed	Average number of members per household making labor contribution	0.55	0.65	0	4.10	607	VHLSS
Days contributed	Average number of days contributed per household	4.05	5.63	0	41	607	VHLSS
Time since annexed	Number of 100 years since being annexed into historical Vietnam	6.77	3.57	2.33	10.22	607	Archives
Demographic Characteristics							
Household size	Average number of members per household	3.77	0.49	1.80	7.50	607	VHLSS
Urban	Percentage of households residing in urban areas	0.22	0.32	0	1	607	VHLSS
Head age	Average age of household head	47.41	4.43	26.60	66.40	607	VHLSS
Male head	Percentage of household head being male	0.78	0.13	0.27	1	607	VHLSS
Kinh head	Percentage of household head with Kinh ethnicity	0.79	0.35	0	1	607	VHLSS
Head schooling years	Average schooling years of household head	6.63	2.14	0.60	13.59	607	VHLSS
Other Control Variables							
Per capita expenditure	Average real per capita expenditure (million dong)	3.51	2.29	0.89	21.96	607	VHLSS
Cropland area (%)	Percentage of cropland in total area (1992-1993)	0.45	0.36	0	1	607	GLCCv2 (EROS)
Caloric suitability[a]	Average yield (million kilo calories per km^2 per year)	25.10	2.49	15.88	31.74	555	Galor & Özak (2016)
Elevation	Average height above sea level (km)	0.20	0.30	0.001	1.56	607	GTOPO30 (EROS)
Terrain ruggedness	Average topographic heterogeneity (100 m)	1.14	1.43	0	6.27	607	Nunn & Puga (2012)
Population density[a]	Thousand persons per km^2 in 1990	0.75	2.47	0.01	35.33	591	GPWv3 (CIESIN)
Latitude	Latitude (degree)	16.44	4.79	8.65	23.24	607	GAUL (FAO)
Distance to Nghe An	Mainland distance to modern Nghe An (100 km)	6.12	4.47	0	14.70	607	GAUL (FAO)

Note: All variables are constructed at district level. Data from the VHLSS 2002 are aggregated without sampling weights. Missing data in *a* are owning to spatial operation.
VHLSS: Vietnam Household Living Standard Survey.
GLCC: Global Land Cover Characteristics Database Version 2.0.
GTOPO30: Global 30 Arc-Second Elevation Dataset.
EROS: Center for Earth Resources Observation and Science, United States Geological Survey.
GPWv3: Gridded Population of the World Version 3.
CIESIN: Center for International Earth Science Information Network.
GAUL: Global Administrative Unit Layers.
FAO: Food and Agriculture Organization, United Nations.

Table A2a. Percentage of Households Making Labor Contribution

	Percentage of Households								
	Baseline			Alternative Mechanisms					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Time since annexed	0.028***	0.017**	0.018**	0.021***	0.024***	0.029**	0.028*	0.016	0.022
	(0.008)	(0.007)	(0.007)	(0.006)	(0.008)	(0.013)	(0.015)	(0.017)	(0.018)
Cochinchina								-0.072	-0.045
								(0.048)	(0.073)
Constant	0.121**	0.517*	0.508	0.289	0.357	-0.294	0.261	-0.185	0.486
	(0.051)	(0.279)	(0.356)	(0.285)	(0.332)	(0.434)	(0.637)	(0.433)	(0.728)
Demographic controls	NO	YES	YES	YES	YES	YES	YES	YES	YES
Additional controls	NO	NO	YES	NO	YES	NO	YES	NO	YES
Mean value of dep. var.	0.31	0.31	0.33	0.28	0.29	0.19	0.20	0.19	0.20
Adjusted R^2	0.093	0.348	0.375	0.281	0.319	0.156	0.193	0.162	0.191
Observations	607	607	552	478	445	249	238	249	238

Note: OLS estimator with robust standard errors in parentheses. Demographic controls are household characteristics aggregated at district level including the number of members, residing in urban area, and head ethnicity, gender, age and schooling years. Additional controls include per capita expenditure, percentage of cropland area, caloric suitability index, elevation, terrain ruggedness and population density. Columns 4-5 exclude all districts whose elevation are above 500 meters and belong to two municipalities, Ha Noi and Ho Chi Minh City. Columns 6-9 continue to exclude districts in the initial territory of historical Vietnam.
* p<0.1, ** p<0.05, *** p<0.01

Table A2b. Percentage of Households Making Labor Contribution

Second stage	Percentage of households					
	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexed	0.030***	0.029***	0.032**	0.065***	0.048***	0.041**
	(0.011)	(0.011)	(0.014)	(0.013)	(0.012)	(0.016)
First stage	Time since being annexed					
	(1)	(2)	(3)	(4)	(5)	(6)
Distance to Nghe An	na	na	na	-0.376***	-0.404***	-0.485***
	na	na	na	(0.022)	(0.022)	(0.026)
Exogeneity (p-value)	na	na	na	0.000	0.011	0.333
Kleibergen-Paap F	na	na	na	289	331	359
Adjusted R^2	na	na	na	0.645	0.691	0.819
Demographic controls	NO	YES	YES	NO	YES	YES
Additional controls	NO	NO	YES	NO	NO	YES
Mean value of dep. var.	0.22	0.22	0.23	0.22	0.22	0.23
Adjusted R^2	0.028	0.183	0.224	na	na	na
Observations	311	311	287	311	311	287

Note: Two stages least squares estimator with robust standard errors in parentheses, constant terms in both stages are suppressed to save space. Demographic controls are household characteristics aggregated at district level including the number of members, residing in urban area, and head ethnicity, gender, age and schooling years. Additional controls include per capita expenditure, percentage of cropland area, caloric suitability index, elevation, terrain ruggedness and population density. All regressions only include districts in the annexed areas.

* p<0.1, ** p<0.05, *** p<0.01

Table A3a. Number of Members per Household Making Labor Contribution

	Number of members per household								
	Baseline			Alternative Mechanisms					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Time since annexed	0.056***	0.043***	0.045***	0.046***	0.050***	0.075**	0.067**	0.051	0.052
	(0.014)	(0.013)	(0.014)	(0.012)	(0.015)	(0.030)	(0.031)	(0.037)	(0.036)
Cochinchina								-0.142	-0.125
								(0.091)	(0.142)
Constant	0.172*	0.329	0.046	0.023	0.077	-0.634	1.219	-0.418	1.846
	(0.095)	(0.535)	(0.699)	(0.532)	(0.620)	(0.799)	(1.385)	(0.776)	(1.543)
Demographic controls	NO	YES	YES	YES	YES	YES	YES	YES	YES
Additional controls	NO	NO	YES	NO	YES	NO	YES	NO	YES
Mean value of dep. var.	0.55	0.55	0.58	0.46	0.48	0.31	0.32	0.31	0.32
Adjusted R^2	0.093	0.370	0.406	0.294	0.336	0.186	0.239	0.194	0.239
Observations	607	607	552	478	445	249	238	249	238

Note: OLS estimator with robust standard errors in parentheses. Demographic controls are household characteristics aggregated at district level including the number of members, residing in urban area, and head ethnicity, gender, age and schooling years. Additional controls include per capita expenditure, percentage of cropland area, caloric suitability index, elevation, terrain ruggedness and population density. Columns 4-5 exclude all districts whose elevation are above 500 meters and belong to two municipalities, Ha Noi and Ho Chi Minh City. Columns 6-9 continue to exclude districts in the initial territory of historical Vietnam.
* p<0.1, ** p<0.05, *** p<0.01

Table A3b. Number of Members per Household Making Labor Contribution

Second stage	Number of members per household					
	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexed	0.068***	0.069***	0.069**	0.141***	0.113***	0.090***
	(0.023)	(0.025)	(0.028)	(0.026)	(0.025)	(0.031)
First stage	Time since being annexed					
	(1)	(2)	(3)	(4)	(5)	(6)
Distance to Nghe An	na	na	na	-0.376***	-0.404***	-0.485***
	na	na	na	(0.022)	(0.022)	(0.026)
Exogeneity (p-value)	na	na	na	0.000	0.002	0.185
Kleibergen-Paap F	na	na	na	289	331	359
Adjusted R^2	na	na	na	0.645	0.691	0.819
Demographic controls	NO	YES	YES	NO	YES	YES
Additional controls	NO	NO	YES	NO	NO	YES
Mean value of dep. var.	0.37	0.37	0.39	0.37	0.37	0.39
Adjusted R^2	0.042	0.185	0.249	na	na	na
Observations	311	311	287	311	311	287

Note: Two stages least squares estimator with robust standard errors in parentheses, constant terms in both stages are suppressed to save space. Demographic controls are household characteristics aggregated at district level including the number of members, residing in urban area, and head ethnicity, gender, age and schooling years. Additional controls include per capita expenditure, percentage of cropland area, caloric suitability index, elevation, terrain ruggedness and population density. All regressions only include districts in the annexed areas.

* p<0.1, ** p<0.05, *** p<0.01