

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, Revised June 2018 (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

May 2018

Abstract

Cultural norms diverge substantially across societies, even within the same country. The present paper examines the voluntary settlement hypothesis, proposing that individualistic people tend to self-select into migrating out from collectivist societies toward the frontier areas, and that such patterns of historical migration are reflected even in the current distribution of norms. Gaining independence in 939 CE after 1000 years of Chinese colonization, historical Vietnam occupied the region that is now north Vietnam with a collectivist social organization. From the 11th to the 18th centuries, historical Vietnam gradually expanded its territory southward to the Mekong River Delta through various waves of conquest and migration. Combining findings from household survey and lab-in-the-field experiment, we demonstrate that areas annexed earlier to historical Vietnam are currently more prone to a collectivist culture. Relying on many historical accounts, together with various checks and tests, we show that the dominant mechanism behind this finding is the southward out-migration of individualistic people during the territorial expansion.

Keywords: Culture; Voluntary Settlement; Vietnam.

JEL Classification: N45; O53; Z1.

*We are grateful for valuable comments from Fredrik Carlsson, Carl-Johan Dalgaard, Quoc-Anh Do, Martin Dufwenberg, Masayuki Kudamatsu, Laura Mayoral, Khanh-Nam Pham, Pablo Selaya, Ardeshir Sepehri, Fabrizio Zilibotti, and seminar participants at the University of Gothenburg and the 16th Nordic Conference in Development Economics. We would like to thank EfD-Vietnam for kindly hosting and supporting our field research and data collection. All remaining errors are our own. Research grant from the Swedish Research Council (ref 348 2013-6348) is gratefully acknowledged. Contact information: 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). Among the cultural traits, the individualism-collectivism contrast has been found to be a powerful predictor of democratic and economic development across many countries (Gorodnichenko and Roland 2011, 2015, 2017).¹ These empirical findings lead us to an important question: why are some societies more or less collectivistic or individualistic than others?

In the present paper, we hypothesize that cultural differences along the individualism-collectivism dimension across modern societies can be traced back to territorial expansion and migration processes that happened in historical times. In particular, we investigate the *voluntary settlement hypothesis*, consisting of three building blocks. First, collectivist societies emerged in early agricultural regions. Second, these societies triggered the out-migration of individualistic members toward peripheral areas. This pattern then repeated itself as the individualistic immigrants inhabited and developed these peripheral areas into less collectivist societies compared to the ones they left behind, which in turn induced more individualistic members to migrate toward more peripheral areas. Eventually, these migration processes gave rise to cultural differences along the individualism-collectivism dimension across societies. Third, these cultural differences have persisted over time and constituted the cultural landscape exhibited in modern societies.

Testing the voluntary settlement hypothesis requires a historical setting where there was a large out-migration of people from a collectivist society that settled down in new regions, and that descendants of people in both the collectivist society and the new regions can be traced to the present day. We recognize the territorial expansion process in the history of Vietnam as an ideal setting to test the voluntary settlement hypothesis. The modern north Vietnam (Figure 1) was under the colonization of the centralized imperial bureaucracy of historical China over a period of 1000 years. After gaining independence in

¹ Some other notable traits include trust, family ties, generalized morality, and attitudes toward work and the perception of poverty. The individualism-collectivism contrast is also the single most relevant dimension in cultural psychology (Gorodnichenko and Roland 2012). For a survey on the theory and measurement of the individualism-collectivism contrast in social psychology, see Triandis (1995).

939 CE, historical Vietnam initially governed this region with a centralized bureaucratic system. At the same time, the territory in the south of historical Vietnam was sparsely populated by many ethnic tribes that did not have a centralized government. From the 11th to the 18th centuries, historical Vietnam gradually expanded its territory southward to the Mekong River Delta to establish the country as it is today (Figure 1) through various waves of state conquest and civil migration, which resulted in the displacement of most of the ethnic tribes.

[Figure 1 is about here.]

To test the voluntary settlement hypothesis, ideally, one would need to demonstrate that some core societies are characterized by collectivism, and that people who migrated to the new regions were more individualistic than those who stayed. Second, historical data are required to prove that these selective migrations gave rise to cultural differences along the individualism-collectivism dimension across historical societies. Finally, present-day data should be collected to conduct an in-depth empirical analysis to provide causal evidence that descendants of the old societies are currently more prone to collectivist norms than descendants of the new societies.

To emulate this ideal setting, we first present qualitative evidence demonstrating that the initial society of historical Vietnam was strongly characterized by collectivist norms. Second, we provide various accounts and analyses showing that people who migrated and inhabited the annexed areas were likely to be more individualistic than those who stayed. We also provide accounts and statistics showing evidence that cultural differences along the individualism-collectivism dimension across regions were already in place in 17th and 18th centuries. Third, to show that these cultural differences have persisted to the present day, we conduct a reduced-form analysis of the relationship between the key explanatory variable, i.e., the time elapsed since an area was annexed to historical Vietnam, and various indicators of collectivism in the present day. Our hypothesis states that areas annexed earlier to historical Vietnam are currently more prone to collectivist norms.

To capture the strength of collectivism, we focus on the societal ability to solve collective action problems, which is the main feature of collectivism studied in related economic models (Gorodnichenko and Roland 2015, 2017). What constitutes a collective action, of course, varies significantly across societies. In Vietnam, labor contribution to public goods production such as building or repairing public infrastructure (e.g., roads, wells,

irrigation, schools or health clinics) is a typical collective action (Adams and Hancock 1970). Using data from the Vietnam Household Living Standard Survey, we construct three related indicators at the district level for almost all districts in Vietnam: (i) the percentage of households contributing labor, (ii) the average number of persons per household making labor contributions, and (iii) the average number of labor days contributed per household. We find that districts annexed earlier to historical Vietnam currently have higher percentages of households contributing labor, more members per household making labor contributions, and more labor days contributed per household. The estimated effects are economically and statistically significant, and robust to the inclusion of standard control variables. By using various checks and tests, we further distinguish the migration of individualistic people in historical times from other potential channels in explaining these cultural differences, including different living conditions in the annexed region, omitted-variable bias, the Western influence in south Vietnam in the 19th and 20th centuries, and important population movements in modern day as a result of state planning and urbanization.

To investigate the voluntary settlement hypothesis further, we conduct public goods experiments with high school students from an earlier-annexed district and a later-annexed district, which is a subject pool old enough to be aware of the cooperative norms in their communities, but have not yet been significantly exposed to other external influences. By comparing the cooperation levels between the two studied districts, we test the voluntary settlement hypothesis. To ensure that the voluntary settlement hypothesis is examined without confounding effects, we took a great care in selecting experimental sites to separate the selective migration of individualistic people from other potential channels mentioned above. We find that subjects from the earlier-annexed district contribute substantially more in the public goods experiment compared to subjects from the later-annexed district, and that the result is mainly driven by the belief about the contribution levels of the other subjects, which is in line with conditional cooperation motivation. Therefore, the experimental findings corroborate the survey data analysis and further suggest that the cultural differences along the individualism-collectivism dimension across regions are still present among teenagers of the country.

Our research relates to a growing multidisciplinary literature examining the origins of cultural differences along the individualism-collectivism dimension. Theories based on

ecological context posit that some forms of production in subsistence economies (e.g., farming) require more functional interdependence than others (e.g., hunting), which gave rise to collectivism as an adaptation mechanism (Vandello and Cohen 1999; Talhelm et al. 2014). Fincher et al. (2008) argue that societies with historically greater prevalence of disease-causing pathogens are currently more prone to collectivist norms, because the emphasis on the in-group/out-group distinction could serve as an anti-pathogen defense function. Motivated by the history of settlement in the United States and its highly individualist culture, Kitayama et al. (2006, 2009) put forward the voluntary settlement hypothesis, asserting that settlers in frontier areas are likely to have highly autonomous, independent, and goal-oriented mindsets; see also Olsson and Paik (2016) and Bazzi, Fiszbein, and Gebresilas (2017).² Our research builds on and adds to this literature in various ways. To the best of our knowledge, no studies on the origins of cultural differences along the individualism-collectivism dimension have examined the societal ability to solve collective action problems, especially using a combination of survey and experimental data. Furthermore, most studies so far have employed cross-country comparisons or concentrated on developed societies, which have gone through the modernization process to a great extent, resulting in a reduction of the traditional cultural landscapes. Comparing different regions within a developing country, our research is thus able to overcome the limitations found in previous studies.

Our research also fits into a literature in economics examining the persistence of various cultural traits as an important channel through which historical events could influence on contemporary economic development (Nunn 2012, 2014; Spolaore and Wacziarg 2013). Some notable traits are gender equality (Alesina, Giuliano, and Nunn 2013; Hansen, Jensen, and Skovsgaard 2015), trust and cooperation (Nunn and Wantchekon 2011; Becker et al. 2016; Bigoni et al. 2016, 2018; Guiso, Sapienza, and Zingales 2016; Litina 2016), anti-Semitic attitude (Voigtländer and Voth 2012), time preference (Galor and Özak 2016), and civic values (Lowe et al. 2017). In a related study, Dell, Lane, and Querubin (2017) show that living standards are currently higher in areas governed for a longer period of time by the centralized states of historical Vietnam. Therefore, our research is com-

² The modernization hypothesis, arguing that societies become more individualistic as they reach higher levels of economic development, essentially focuses on the convergent tendency toward individualism, rather than pre-existing cultural differences across modern societies (Inglehart and Baker 2000).

plementary by investigating the role of individualism-collectivism as another important cultural dimension for understanding long-run economic development.

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 accompanying 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, sample selection, and results. Section 5 closes the paper with some concluding remarks.

2 Historical Background

In this section, we present detailed information to justify the three building blocks of our theory: (i) the initial region of historical Vietnam was home to a collectivist society; (ii) individualistic people were induced to migrate southward as the country expanded its territory, eventually giving rise to cultural differences along the individualism-collectivism dimension; and (iii) these cultural differences have persisted to the present day.

2.1 Initial Region of Historical Vietnam Was a Collectivist Society

Archaeological evidence indicates that ancient populations had settled down in the Red River Delta with rice agriculture around 2000 BCE during the Neolithic Revolution (Nguyen, Pham, and Tong 2004). These populations inhabited together, without a centralized state, in the region that is now north Vietnam (Figure 1). From 111 BCE to 939 CE, the whole region was brought under the colonization of the centralized imperial bureaucracy of historical China. During this period, “the Vietnamese evolved from a preliterate society within a “south-sea civilization” into a distinctive member of the East Asian cultural world” (Taylor 1983, p. xvii).

After the victory over historical China in 939 CE, the first unified state of historical Vietnam was founded in 968 CE and inherited with a centralized bureaucratic system from the colonizer (Taylor 2013, p. 51-77). Subsequent dynasties governing historical Vietnam continued to build stronger structures and orders into the society, which emphasized the values of social groups above the needs and desires of its constituent members (Whitmore 1984, 1997). The collectivist nature of historical Vietnam was best exemplified by its village-based bureaucracy. Village was the lowest bureaucratic level, which was responsi-

ble for regulating almost all aspects of daily living for its members (Nguyen 2003). Two important functions of the village were to allocate public land under its management to its needed members (Dao 1993), and to organize unpaid labor for public goods production, such as irrigation facilities, roads, and communal buildings (Adams and Hancock 1970). The collectivist strength of historical Vietnam was also demonstrated in its warfare history, e.g., being the only land-based polity in Eurasia to defeat the Mongols in the 13th century (Taylor 2013, p. 123-137).

The area bordering historical Vietnam in the south, that is now central Vietnam (Figure 1), was inhabited by various ethnic groups that formed the Champa Kingdom. Next to the Champa Kingdom in the south, that is now south Vietnam (Figure 1), was the land belonging to the Khmer Empire. In contrast to the centralized state of historical Vietnam, both the Champa Kingdom and Khmer Empire were basically networks of small political entities (Hall 2011, p. 67-102, 159-210).

2.2 Individualistic People Migrated toward the South

From the 11th to the 18th centuries, historical Vietnam expanded its territory southward along the coast to the Mekong River Delta. This so-called Vietnamese *Southern Advance* (*Nam Tien*) took place gradually through various steps and completed in 1757 CE, by which the border of Vietnam was established as it is today.³ In each step, the state conquered and established its bureaucracy in the annexed territory, while Vietnamese migrants settled in. These migrants consisted of landless peasants, traders, adventurers, pioneers, fugitives, refugees, slaves, and criminals (Taylor 2013, p. 221, 252, 303, 322). In the society of historical Vietnam, Taylor (2013, p. 211) supposes that these people were “marginal and undesirable”. In the same line, Tana (1998, p. 111) argues that these people “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”.⁴ In other words, these groups of people were unlikely to be imprinted with the cultural characteristics of the

³ In their study, Dell, Lane, and Querubin (2017) assume that the southern region in modern Vietnam was divided by a fixed border between the Khmer Empire and historical Vietnam from 1698 CE until 1833 CE, i.e., just prior to the French Colonization in 1859 CE. For a discussion about the support for such a fixed border assumption, see details in Appendix B based on historical sources.

⁴ One may recall the analogous image of European immigrants to the United States during the age of mass migration (Abramitzky, Boustan, and Eriksson 2012).

societies they left behind.⁵ Regarding the local ethnic groups, most of their populations ran away, while others stayed and acculturated to the Vietnamese culture.⁶

Studying historical Vietnam in the 17th and 18th centuries, Tana (1998, p. 99-116) provides many historical accounts to demonstrate that the social environment in the annexed region was characterized by greater openness, mobility and autonomy compared to the initial region; “a new way of being Vietnamese”. In other words, the annexed region was less collectivistic or more individualistic than the initial region. Available statistics of land allocation in the early 19th century also illustrate this cultural divergence. In the initial region, public land was allocated to only village members; this was intended to tie people to their villages (Nguyen 2003, p. 107). In the annexed region, however, the in-village/out-village distinction was loosened and allocating public land to people from other villages became a common phenomenon. For example, studies on the land registries (cadastres) of four provinces in the annexed region in the early 19th century show that the proportions of public land allocated to people from other villages were 1.4% in Quang Nam (Nguyen 2010a), 3% in Quang Ngai (Nguyen 2010b), 14.7% in Binh Dinh (Nguyen 1996), and 1.2% in Khanh Hoa (Nguyen 1995).

Following the above discussion, we argue that Vietnamese migrants in the annexed region were more inclined toward an individualistic mindset (typical of pioneers in a frontier), which was the dominant driver of cultural differences found between the annexed region and the initial region. It is not unreasonable to think that the state of historical Vietnam had a longer time to instill norms into the people in the initial region than into those in the annexed region. Nevertheless, because most of the population that inhabited the annexed region was Vietnamese migrants from the initial region, people in both regions should have a similar length of time in terms of state experience. Thus, it is unlikely that the state influence was the dominant force driving the cultural differences across regions. Another potential driver was the frontier environment in the annexed areas. For example, Vietnamese migrants inhabiting the annexed region had to rely on their own resources and initiative to a degree not possible in the initial region (Taylor 2013, p. 252).

⁵ Alesina and Giuliano (2010) find that people with weak family ties are more likely to migrate.

⁶ In the end of the 18th century, these ethnic groups accounted for about one third of the total population of the Mekong River Delta (Wook 2003). In modern times, their descendants only constituted minor fractions in the total population of Vietnam; for example in 1999, the Cham were 0.17% and the Khmer were 1.38% (General Statistics Office of Vietnam 2001, p. 167).

In addition, they were certainly in contact with and to some extent influenced by the indigenous cultural characteristics in these areas (Tana 1998, p. 99-116). We believe that the frontier environment did play an important role in explaining the cultural differences across regions. Nevertheless, if it is the dominant force, then the voluntary settlement hypothesis should not hold when restricting the analysis to the subsample of the annexed areas. We conduct this test in the following section.

2.3 Cultural Differences Have Persisted to the Present Day

The last block of the voluntary settlement hypothesis argues that the cultural differences across regions of historical Vietnam found around the 18th century have persisted to modern day. Thus, we expect to find that the areas annexed earlier to historical Vietnam are currently more prone to collectivist norms. Our hypothesis is built upon a well-established literature showing that cultural norms are transmitted through both parental investment (vertically) and social influence (horizontally) (Bisin and Verdier 2011). In traditional Vietnam, parents had strong authority over their children on almost all aspects of life (e.g., education, marriage, and housing), while children had to serve and obey their parents with the utmost respect (Haines 1984). Meanwhile, the village also had a non-negligible influence on its members through defining and enforcing the rules governing behaviors and relationships within the village (Nguyen 2003).

Ending his Vietnamese history, Taylor (2013, p. 624) argues that cultural differences across regions in Vietnam have persisted well to modern day, and 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 Vietnam, although regarding collectivism as the main cultural theme, Ashwill and Diep (2005, p. 71-72) note that “northerners are considered to be more intelligent, conservative, austere, serious, and frugal, ..., [and] are more apt to save for a rainy day”, while “southerners are perceived as fun-loving, easy-going, open people who rarely think of saving for a rainy day”. 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).

Besides the persistent effect of the territorial expansion and selective migration in the remote past, there were two recent events (both took place after 1757 CE) in the history of Vietnam that might account for the current cultural differences described above. The

first was the French colonization (starting in 1858 and ending with the Vietnamese victory in the First Indochina War, 1946-1954), during which the French colonizers concentrated most of their activities in south Vietnam (Taylor 2013, p. 484-523). Following the French defeat was the American intervention in south Vietnam (i.e., the Second Indochina War, commonly known as the Vietnam War), which ended with reunification of the country in 1975 (Taylor 2013, p. 561-614). Because these two recent experiences with the Western world are much shorter than eight centuries of the southern advance, they are unlikely to be the dominant drivers of contemporary cultural differences. Nevertheless, if it is the case, then the voluntary settlement hypothesis should not hold when restricting the analysis to the subsample of the areas that were heavily influenced by the Western world. We conduct this robustness test in the following section.

3 Survey Data Analysis

3.1 Empirical Model

In this section, we use survey data to investigate the voluntary settlement hypothesis. Our empirical strategy revolves around regressing a measure capturing the individualism-collectivism dimension on an explanatory variable measuring the time elapsed since an area was annexed to historical Vietnam, while controlling for other potential confounding factors. We use the Vietnam Household Living Standard Survey (VHLSS) as our main dataset. We measure cultural norms at the district level by aggregation of individual data from the VHLSS.⁷ We use the ordinary least squares (OLS) estimation method as our baseline empirical model, which takes the following form:

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

In this equation, IC_i is a measure of the average expression of cultural norms in district i , $TimeSinceAnnexation_i$ is the time since annexation to 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 with respect to the collectivist measure, i.e., the longer time since annexation the stronger the collectivist norms.

⁷ There is certainly individual heterogeneity within a society, but in aggregation one can observe what social psychologists call the “cultural syndrome” of each society (Triandis 1995). This is also a standard exercise in economics (Alesina and Giuliano 2015).

3.2 Variables

The Individualism-Collectivism Trait

In the present paper, we follow the conventional definition of culture in economic research 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).⁸ Many observable outcomes have been used in the literature to capture the individualism-collectivism trait such as family structure, marriage stability, inventiveness, or infrequent children’s names (Vandello and Cohen 1999; Talhelm et al. 2014; Bazzi, Fiszbein, and Gebresilasse 2017). To be a good measure of the individualism-collectivism trait, an outcome must be both theoretically and practically relevant. In other words, the outcome must capture an important aspect of the individualism-collectivism trait and feature as a traditional practice in the society under study.

For our main analysis, we employ labor contribution to public goods production (i.e., building or repairing public infrastructure such as roads, wells, irrigation, schools, or health clinics) to capture the strength of collectivism, relying on two underpinnings. First, the ability to solve collective action problems is the main feature of collectivism in related economic models (Gorodnichenko and Roland 2015, 2017). Because collectivist societies are considered to be better in this aspect, one should observe their members to contribute more resources to public goods production. Second, from a practical perspective, labor contribution to public goods production was also a typical activity among people in historical Vietnam (Adams and Hancock 1970). This tradition is still prevalent in modern day because labor is probably the most abundant resource in a developing country as Vietnam. Using the VHLSS, we construct three related variables at the district level based on labor contribution to public goods production. First, we calculate the percentage of households contributing labor in the district to measure the prevalence of labor contributions. Second, we calculate the average number of persons making labor contributions per household. Finally, we calculate the average number of labor days contributed per household. These last two variables capture the intensity of labor contributions.

The VHLSS covers all provinces in Vietnam with information on geographical location

⁸ 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.”

at the district level, which enables us to conduct a detailed empirical analysis. We employ the 2002 survey, the only survey containing detailed information on labor contribution to public goods production. The 2002 survey covers 607 districts (out of 630) across 61 provinces. Table A1 (Appendix A) shows that, in 2001, around 26% percent of households contributing labor to public goods production, whereas the average number of persons making labor contributions per household is 0.44 and the average number of labor days contributed per household is 3.24. To avoid blank areas for non-surveyed districts on the map, we calculate the average values of all three variables at the province level and depict them in Figure 2. A visual comparison with Figure 1 gives the impression that provinces annexed earlier to historical Vietnam currently have higher percentages of households contributing labor, more members per household making labor contributions, and more labor days contributed per household.

[Figure 2 is about here.]

The Time since Annexation to Historical Vietnam

As previously mentioned, our main explanatory variable is the time elapsed since an area was annexed to historical Vietnam. Following the historical background discussed earlier, we choose the first unified state of historical Vietnam in 968 CE as the beginning year, while the terminal year is set to 1990. In our analyses, we measure the time since annexation in centuries (100 years) to make the estimated coefficients easy to read in the reported tables. The descriptive statistics in Table A1 (Appendix A) show that the annexation took place between 2.33 to 10.22 centuries from the terminal year of 1990.

To construct the time since annexation to historical Vietnam for each modern district, two dimensions are needed: (i) its corresponding area in historical Vietnam and (ii) the year that this area was annexed. For the year that a historical area was annexed, we rely on two official chronicles of historical Vietnam, *Dai Viet Su Ky Toan Thu* and *Dai Nam Thuc Luc*, recording events from the beginning to 1675 CE and in the 1558-1888 CE period, respectively. These chronicles were written by state officials of historical Vietnam to keep track of events and, to the best of our knowledge, constituted the primary sources for Vietnamese histories. We code an area to be annexed when there is a record in the chronicles demonstrating that this area was under the control of historical Vietnam. To link historical areas to their modern counterparts, we rely on two seminal works of

Vietnamese historians: Dao (2005) and Phan et al. (2011). All details on the coding procedure are presented in Appendix B.

Control Variables

We select a set of control variables to be included in the regression analysis, including many potential confounding factors normally found in previous studies. First, agricultural suitability might have both attracted historical Vietnam to conquer a region and promoted the development of collectivism. We use the percentage of cropland area in total land area to control for the prevalence of agriculture, which is argued to have given rise to collectivism by the ecological context theory (Vandello and Cohen 1999). We also control for natural land productivity, as it was shown by Litina (2016) that high productivity reduced the incentive to cooperate in the traditional agricultural economy. Second, geographical conditions such as elevation and ruggedness might affect the difficulty in conquering a region. Isolated areas are also conducive to the development of a collectivist culture (Triandis 1995). Thus, we control for elevation and ruggedness in our analysis. Descriptive statistics of all variables can be found in Table A1 (Appendix A).

The percentage of cropland area is taken from the Global Land Cover Characteristics Database, which is constructed at 30 arc-second resolution (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 project of the Food and Agriculture Organization. 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 based on agro-climatic conditions, which are arguably unaffected by human intervention. Elevation is taken from the Global 30 Arc-Second Elevation Dataset (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 prey and lookout posts. This index is calculated by Nunn and Puga (2012) based on the GTOPO30 dataset.

To examine the robustness of the results to other potential omitted variables, we further conduct an instrumental variable (IV) estimation. Although many factors might

influence the annexation of an area, we argue that the north-south geographical order is given by nature, and hence exogenous to the annexation. In other words, from the Red River Delta in the north, one could not conquer the Mekong River Delta in the south without annexing all areas located in between.⁹ Thus, in the subsample of the annexed areas, the north-south geographical order can serve as a valid instrument for the time since an area was annexed to historical Vietnam. We propose to use the distance from an annexed area to a reference point in the north as a measure of the north-south geographical order. Nghe An (Figure 1) is arbitrarily chosen as the reference point in the north (the result is robust to any other choice), and the walking distance along the coast (instead of the geodesic, “bird-fly” distance) is calculated to capture the military route in historical times. Distance is measured in 100 kilometers using the district centroids, where borders are taken from the Global Administrative Unit Layers. From Nghe An to the furthest district in the south is roughly 1400 kilometers.

3.3 Baseline Results

To begin with, we regress the percentage of households contributing labor on the time since a district was annexed to historical Vietnam and various control variables discussed above. Table 1 shows that the estimated coefficients of the time since annexation are positive and significant, even when all control variables are included. Thus, districts annexed earlier to historical Vietnam today have a higher percentage of households contributing labor on average. Relative to the mean value of the dependent variable, the marginal effect is economically significant and stable through various regressions. When all control variables are included, for example, one century increase in the time since annexation is associated with an additional 3% of households contributing labor, which is roughly 10% of the mean value of the variable. The time since annexation also accounts for approximately 10% of the total variation in the percentage of households contributing labor.

The estimated coefficient of the percentage of cropland area is negative and significant, suggesting that a higher percentage of cropland area is associated with a lower percentage of households contributing labor, which is contrary to the prediction of the

⁹ Theoretically, one could conquer an area in the Mekong River Delta by traveling either along the coastline in the east or over the mountainous band in the west separating Vietnam from Laos and Cambodia. Both strategies were infeasible given the logistic and transportation technologies in historical Vietnam. Indeed, we do not find any attempt to do so in the historical accounts.

ecological context theory (Vandello and Cohen 1999). The estimated coefficient of the caloric suitability index is significant when included alone, but not so when other control variables are added. Its negative sign suggests that a higher natural land productivity is associated with a lower percentage of households contributing labor, which concurs with Litina (2016). In line with Triandis (1995), the estimated coefficients of elevation and ruggedness are significant and positive, indicating that more elevated and rugged areas have higher percentages of households contributing labor. The estimated coefficient of ruggedness, however, turns insignificant when other control variables are added.

[Table 1 is about here.]

Table 2 reports the results of similar regressions for the two other dependent variables, i.e., average number of persons per household making labor contributions (Panel A) and average number of labor days contributed per household (Panel B). With respect to both variables, the estimated coefficients of the time since annexation are significant and positive whether control variables are included or not. Districts annexed earlier to historical Vietnam currently have more members per household making labor contributions and more labor days contributed per household on average. For both dependent variables, the marginal effects are economically significant, roughly 10 percent of the respective mean values, and stable through various regressions. The time since annexation also accounts for approximately 10% of the total variations in both dependent variables. The results with respect to the control variables (not shown) are similar to those in Table 1 in terms of sign and significance.

[Table 2 is about here.]

3.4 Robustness Analysis

In this subsection, we tackle various challenges in taking the above findings as evidence for the voluntary settlement hypothesis. First, as discussed earlier, the frontier environment might be the dominant driver of the differences in cultural norms in the annexed region. If this is the case, then we should not observe significant relationships between the time since a district was annexed to historical Vietnam and our three dependent variables when restricting the estimation to the subsample of the annexed areas. As can be seen in Panel A of Table 3, the estimated coefficients of the time since annexation to historical Vietnam

remain positive and significant with respect to all three dependent variables. We take this finding as a strong indication that the selective migration of individualistic people in the past is the dominant driver behind the contemporary cultural differences across Vietnam.

[Table 3 is about here.]

Although many control variables have been introduced, it is not enough to claim that there is no omitted-variable bias left in the regression model. As mentioned above, a further test is to restrict the analysis to the subsample of the annexed areas and employ the walking distance to Nghe An as an instrumental variable for the time since annexation to historical Vietnam. Table 3 (Panel B) reports the results of the instrumental variable estimation using the two stages least squares (TSLS) estimator. The estimated coefficients of the time since annexation remain economically and statistically significant with roughly equal marginal effects as regards all three dependent variables. In addition, the first-stage results ensure that the walking distance to Nghe An is a strong predictor of the time since annexation, i.e., it has significant and negative estimated coefficients (not shown) and large F statistics. Finally, we reject the exogeneity assumption of the time since annexation when the variable enters the regression alone, but cannot do so when all control variables are added. Conditioning on the control variables, the time since annexation is exogenous, and hence the results from the OLS estimator are valid.

The next challenge was the recent experiences with the Western world in south Vietnam. During the French colonization (1858-1954), the colonizers concentrated most of their activities in the former Cochinchina, that is now south Vietnam (Figure 1), and hence influenced this region more than anywhere else in Vietnam (Taylor 2013, p. 484-523). Following the French colonization was the American intervention in south Vietnam during the Second Indochina War (1954-1975), a time when Vietnam was divided into two regions along the 17th parallel, i.e., the communist north and the capitalist south. Despite this division, a non-negligible part of the capitalist south was in fact under the control of the communist north (Taylor 2013, p. 561-614). Although the exact border is not known, the former Cochinchina was the homeland of the capitalist South, and hence was in its strongest control. Therefore, if these experiences with the Western world were the dominant drivers of the contemporary cultural differences, then we should not observe significant relationships between the time since a district was annexed to historical Vietnam and our three dependent variables when restricting the estimation to the districts in the

former Cochinchina. Panel A of Table 4 shows that the estimated coefficients of the time since annexation remain economically and statistically significant with respect to all three dependent variables. This finding indicates that the recent experiences with the Western world, well-known as they are, did not significantly change the cultural differences left by eight centuries of territorial expansion and selective migration in the remote past.

[Table 4 is about here.]

Finally, historical Vietnamese immigrants (the Kinh ethnicity) often inhabited along the coastal plain with their traditional rice agriculture. At the same time, the highland areas were mainly inhabited by various ethnic groups. After the Reunification in 1975, the Kinh started to migrate to the highland areas in large scale through state-sponsored programs under the central planning economy to establish new production zones (Hardy 2003). These later migrations, therefore, might be different from the one that happened in historical times. To examine this issue, we exclude from the estimation all districts in the highland areas, i.e., where average elevation are above 500 meters (the results are robust to other values such as 400 and 600 meters). Furthermore, we also exclude two provinces, Ha Noi (in the north) and Ho Chi Minh City (in the south), which are the two biggest venues for immigrants in modern times. Table 4 (Panel B) shows that the estimated coefficients of the time since annexation remain economically and statistically significant with respect to all three dependent variables.

We also conduct two other robustness checks. In the first one, we incorporate sampling weights in constructing our three dependent variables. In the second check, we employ the estimation method developed by Conley (1999) to adjust the standard errors for spatial autocorrelation. The results are present in Table A2 and Table A3 (Appendix A) and show that the qualitative results with respect to the time since annexation to historical Vietnam remain intact.

3.5 Discussion

The above analysis has demonstrated that labor contribution to public goods production is both more prevalent and intensive in districts annexed earlier to historical Vietnam. More importantly, the results are not driven by the frontier environment, potential omitted-variable bias, the Western influence in south Vietnam in the 19th and 20th centuries, or important population movements in modern day. These findings altogether lend support

to the out-migration of individualistic people during eight centuries of territorial expansion of historical Vietnam as the dominant driver behind the cultural differences found across regions in modern Vietnam.

So far, we have used naturally occurring data on labor contribution to public goods production. As argued by Falk and Heckman (2009), different types of data sources have their own pros and cons, but complement each other overall. The disadvantage of the large and rich dataset obtained from surveys such as the VHLSS is that they are naturally occurring data, i.e., obtained from un-controlled environments. To complement the survey data analysis, we collected experimental data by conducting a lab-in-the-field experiment to further examine contributions to public goods in Vietnam.

4 Experimental Data Analysis

4.1 Sample Selection

To test the voluntary settlement hypothesis experimentally, it is crucial to select experimental sites in such a way to control for other potential channels. First, we focus on the annexed areas to rule out differences in the frontier environment. Second, we restrict to the areas under control of the former capitalist South to eliminate regime differences in the north and the south during the Second Indochina War (1954-1975). Third, we select provinces, and rural districts in them, located along the coast, which was the route that historical Vietnamese typically migrated and settled down. Finally, 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, i.e., neither any significant immigration nor emigration from these places. Thus, this procedure leaves us with coastal, rural, and Kinh-dominated districts in the annexed areas under control of the former capitalist South during the Second Indochina War. From this subsample, we randomly select one of the districts with the longest time since annexation to historical Vietnam and one of the districts with the shortest time. This process narrows our selection to one rural district in Thua Thien Hue and one rural district in Ben Tre, where the former has the longer time since annexation (Figure 1).

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 have not yet been affected by living or working outside their communities, which potentially

could make it harder to capture the local cultural norms.¹⁰ 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, which are similar in many background factors except the time since annexation to historical Vietnam. In other words, we expect subjects in Thua Thien Hue (henceforth the “northern site”) to share a stronger cultural norm of cooperation, and hence on average contribute at a higher level compared to subjects from Ben Tre (henceforth the “southern site”). Each district in Vietnam has 3 to 5 high schools. To keep similarities between the selected districts, we randomly selected one school located in the center of the district among the schools that had at least six classes for the oldest age cohort, which means that students come from a larger catchment area where they have attended different secondary schools. The latter requirement was imposed to avoid measuring cooperation norms within a specific class, which might have developed their own norms, but rather that we aimed at measuring norms in the society wherein they lived.

4.2 The Public Goods Experiment

We conduct a public goods experiment with high school students in Vietnam to analyze attitudes toward cooperation.¹¹ We build our experimental design on the one-shot linear public goods experiment developed by Fischbacher, Gächter, and Fehr (2001).¹² We begin by describing the general features of the public goods experiment before discussing the specific features of the design in Fischbacher, Gächter, and Fehr (2001).

The subjects are randomly assigned to groups of three, where each member comes from a different class at the high school, and this was clearly stated in the instructions of the experiment. This feature of the design was chosen to avoid having subjects allocated to groups consisting of classmates with whom subjects might have developed a specific norm of behavior, reducing the possibility to measure norms of cooperation in the places where they reside. All subjects receive an endowment of 20 tokens and must decide

¹⁰This strategy to focus on high school students has also been used earlier in the literature on public goods experiment when investigating cultural differences, e.g., Kocher, Martinsson, and Visser (2012).

¹¹For a general discussion on public goods experiment, see Zelmer (2003) and Chaudhuri (2011).

¹²For other experiments using this specific design, see Kocher et al. (2008), Herrmann and Thöni (2009), Fischbacher and Gächter (2010), Fischbacher, Gächter, and Quercia (2012), and Martinsson, Villegas-Palacio, and Wollbrant (2015) among others.

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 contributed to the public goods results in 0.5 token to all group members, including the member who contributes the token. If a subject is rational and selfish, then a *MPCR* below 1 leads to a dominant strategy to free ride (i.e., to contribute zero to the public good), because the return from the public good is lower than the return from the private 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 consists of two components: (i) the amount of the endowment that is not invested in the public goods (i.e., what is kept as private good), and (ii) the return from the public good. It is given by:

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

Each token earned in the experiment is exchanged to money, i.e., one token equals 3000 Vietnamese Dong. This experiment is calibrated, partially based on pilot studies, such that each student on average receives a monetary payoff worth roughly three meals at the local restaurants. They receive no show-up fee.

The specific feature of the Fischbacher, Gächter, and Fehr (2001)'s design is the strategy method. Each subject makes two types of contribution decisions to the public good: (i) unconditional contribution and (ii) conditional contribution. In the unconditional contribution decision, which is similar to a standard public goods experiment, each subject states how much he or she would like to contribute to the public goods from his or her endowment of 20 tokens. In a contribution table, which includes all possible average contributions of the two other players in the group, rounded to integers and ranging from 0 to 20 points, a subject indicates how much he or she would contribute to the public good if these were the average contributions to the public good by the other two group members. The contributions reported in the table are referred to as conditional contributions. All decisions are made incentive compatible by using the following approach. 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 a subject makes his or her decisions, he or she does not know which of the decisions will be pay-off relevant, and hence has the incentive to choose nothing else than his or her preferred option. After the experiment, we also elicited beliefs by asking a subject what he or she thought that the other two group members had contributed unconditionally on average. We pay subjects for the accuracy of their guesses to create stronger incentives for truthful revelation.

The strength of the strategy method is that subjects can be categorized into different contribution types based on their 21 conditional contribution decisions to the public good. We use the same classification as proposed in the original paper by Fischbacher, Gächter, and Fehr (2001). A subject is classified as a “conditional cooperator” if his or her conditional contribution increases weakly monotonically with the average contribution of the other group members or if the relationship between his or her conditional contribution and the average contribution of the others is positive and significant at 1% 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 average contribution by others. A “hump-shaped” contributor is characterized by a subject who shows weakly monotonically increasing contributions or a positive Spearman rank correlation coefficient at 1% significance level, which is the same classification strategy as applied to a conditional contributor, but it only holds up to an inflection point. For average contribution levels by others above this level, the subject’s own conditional contributions decrease weakly monotonically or show a significant and negative Spearman rank correlation coefficient at the 1% significance level. Those who cannot be categorized based on any of the above criteria are referred to as “others”.

Subjects were recruited by teachers, and the participation rates of students are similar across schools: 70% in the northern site (140 out of 200) and 73% (235 out of 320) in the southern site. In accordance with our expectation, 98% of the subjects and 95% of their parents were born in the chosen districts, while the others were born in other districts in the chosen provinces. At each school, we conducted the pencil and paper experiment, including two sessions, in a large lecture hall. Subjects received written instructions for the experiment and the instructions were also read aloud.¹³ Before the experiment began,

¹³Details on the experimental instruction can be found in the Online Appendix. The public goods experiment is the first experiment out of two, and the second experiment is pay-off independent from

various examples were given to facilitate understanding of the experiment and the subjects also completed some exercises. When the experiments were finished, subjects answered a survey about socio-economic questions. Finally, subjects were called one at a time for payment done in private.

4.3 Results

The average contribution to the public good is 6.92 out of the endowment of 20, which is similar to what has been found in the literature (Zelmer 2003; Chaudhuri 2011). Table 5 shows that subjects from the northern site and southern site on average unconditionally contributed 7.50 tokens and 6.58 tokens respectively, and the difference is statistically significant (p -value = 0.024, Mann-Whitney U test). Previous studies have indicated that a large fraction of subjects are conditional cooperators, i.e., their contributions are positively correlated with contribution levels by others. We also elicited guesses about the average contributions by other two group members, in which subjects from the northern site and southern site on average guessed 8.25 tokens and 7.60 tokens respectively (p -value = 0.053, Mann-Whitney U test). It is common that guesses about the average contribution by others are higher than own contribution levels since, on average, people are imperfect conditional cooperators.

[Table 5 is about here.]

We use regression models to examine the contribution behaviors further and the results are shown in Table 6. In all models, we include a dummy variable if a subject comes from the northern site. In line with the descriptive statistics, the estimated coefficient of the northern site dummy is positive and significant when entering alone (Column 1) and also when socioeconomic characteristics (gender, household size, and a wealth index) are included (Columns 2 to 5).¹⁴ In previous studies, conditional cooperation has been shown to be an important determinant of contributions, i.e., a subject matches his or her own contribution level with the levels of others. In the last two regression models, we include

the first one. To eliminate potential spillover between experiments, the subjects were informed about the second experiment only after completing the first experiment.

¹⁴The wealth index is constructed by extracting the first principal component of six variables measuring the numbers of mobile phones, computers, motorbikes, refrigerators, gas cookers, and air conditioners that the households possess.

belief about the average contribution of the other two group members. As expected, its estimated coefficient is significant and positive. In Column (7), where we interact beliefs and northern site as one additional variable, we find that the estimated coefficient of the northern site 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 own contribution behavior is found to be stronger in the northern site, which is reflected through the significant and positive estimated coefficient of the interaction term between belief and the northern site dummy. In other words, the positive effect in the northern site is explained by a strong norm for conditional cooperation.

[Table 6 is about here.]

The innovative part of the Fischbacher, Gächter, and Fehr (2001)'s design is that it allows us to classify subjects into different contributor types. Table 5 shows that, except for free riders, other types in the northern site on average have higher levels of unconditional contribution and guess about the average contribution of other group members compared to their counterparts in the southern site. The distributions of contributor types are similar across the two experimental sites, i.e., the fractions of free riders are 3.62% in the northern site and 5.53% in the southern site, and the corresponding numbers for conditional cooperators are 52.17% and 54.04%. We indeed cannot reject the null hypothesis that the compositions of contributor types in the northern site and the southern site are drawn from the same distribution (p -value = 0.930, Pearson's χ^2 test). This finding suggests that the north-south difference in contribution behaviors is driven by belief rather than the composition of contributor types. To summarize, the experimental findings corroborate the survey data analysis that districts annexed earlier to historical Vietnam currently have stronger norms for cooperation, and that the out-migration of individualistic people during eight centuries of territorial expansion of historical Vietnam is the dominant driver behind these cultural differences.

5 Conclusion

The individualism-collectivism contrast has been found to be a powerful predictor of democratic and economic development in a large sample of countries. Thus, why some societies have become more collectivistic or individualistic than others is a crucial question in understanding comparative development in the long run. In the present paper, we

study the voluntary settlement hypothesis, postulating that cultural differences along the individualism-collectivism dimension are driven by the out-migration of individualistic people from collectivist societies to settle down in less collectivistic or more individualistic regions. We recognize the grand territorial expansion from the 11th to the 18th centuries of historical Vietnam as an appropriate setting to empirically examine this hypothesis. 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 the country as it is today.

We examine the ability to solve collective action problems, which is both the main feature of collectivism in related economic models and the most typical collective action in daily life in Vietnam, by using both survey and experimental data on contributions to public goods. Using household survey, we find that areas annexed earlier to historical Vietnam currently have higher levels of labor contribution to public goods production in terms of not only intensity, but also prevalence. Conducting a public goods experiment with high school students, we find that subjects from areas annexed earlier to historical Vietnam contribute substantially more to the public good compared to subjects from areas annexed later, and that the result is mainly driven by the belief about the contributions by other subjects. Relying on various historical accounts, together with many checks and tests, we show that the southward out-migration of individualistic people during eight centuries of territorial expansion of historical Vietnam is the dominant mechanism behind these cultural differences.

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

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 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.
- . 2010. “The Power of the Family”. *Journal of Economic Growth* 15 (2): 93–125.
- 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.
- Bazzi, Samuel, Martin Fiszbein, and Mesay Gebresilassee. 2017. “Frontier Culture: The Roots and Persistence of "Rugged Individualism" in the United States”. *NBER Working Paper* No. 23997.
- Becker, Sascha O., Katrin Boeckh, Christa Hainz, and Ludger Woessmann. 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, Stefania Bortolotti, Marco Casari, Diego Gambetta, and Francesca Pancotto. 2016. “Amoral Familism, Social Capital, or Trust? The Behavioural Foundations of the Italian North-South Divide”. *Economic Journal* 126 (594): 1318–1341.
- Bigoni, Maria, Stefania Bortolotti, Marco Casari, and Diego Gambetta. 2018. “At the root of the North-South cooperation gap in Italy Preferences or beliefs?” *Economic Journal*, no. Forthcoming.

- Bisin, Alberto, and Thierry Verdier. 2011. “The Economics of Cultural Transmission and Socialization”. In *Handbook of Social Economics*, ed. by Jess Benhabib, Alberto Bisin, and Matthew Jackson, vol. 1A, 339–416. North-Holland: Elsevier.
- Briggs, Lawrence Palmer. 1947. “A Sketch of Cambodian History”. *The Far Eastern Quarterly* 6 (4): 345–363.
- Chaudhuri, Ananish. 2011. “Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature”. *Experimental Economics* 14 (1): 47–83.
- Conley, Timothy G. 1999. “GMM Estimation with Cross Sectional Dependence”. *Journal of Econometrics* 92 (1): 1–45.
- Nguyen, Ngoc Tinh, trans. *Dai Nam Thuc Luc*. 2002. Vol. 1. Ha Noi: Giao Duc.
- Cao, Huy Giu, trans. *Dai Viet Su Ky Toan Thu*. 2012. Ha Noi: Thoi Dai.
- Dao, Duy Anh. 2005. *Dat Nuoc Viet Nam Qua Cac Doi*. Ha Noi: Van Hoa Thong Tin.
- Dao, Minh-Quang. 1993. “History of Land Tenure in Pre-1954 Vietnam”. *Journal of Contemporary Asia* 23 (1): 84–92.
- Dell, Melissa, Nathan Lane, and Pablo Querubin. 2017. “The Historical State, Local Collective Action, and Economic Development in Vietnam”. *NBER Working Paper No. 23208*.
- 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.
- 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.
- Falk, Armin, and James J. Heckman. 2009. “Lab Experiments Are a Major Source of Knowledge in the Social Sciences”. *Science* 326 (5952): 535–538.
- 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., Randy Thornhill, Damian R. Murray, and Mark Schaller. 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.
- General Statistics Office of Vietnam. 2001. *Vietnam Population and Housing Census 1999*. Ha Noi: General Statistics Office of Vietnam.
- 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.
- 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.
- Haines, David W. 1984. “Reflections of Kinship and Society under Vietnam’s Le Dynasty”. *Journal of Southeast Asian Studies* 15 (2): 307–314.
- Hall, Kenneth R. 2011. *A History of Early Southeast Asia*. Plymouth: Rowman / Littlefield Publishers.
- 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.

- Hardy, Andrew. 2003. "State Visions, Migrant Decisions: Population Movements since the End of the Vietnam War". In *Postwar Vietnam: Dynamics of a Transforming Society*, ed. by Hy V. Luong, 107–137. Singapore: Institute of Southeast Asian Studies.
- 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, Hyekyung Park, A. T. Sevincer, Mayumi Karasawa, and Ayse K. Uskul. 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, Keiko Ishii, Toshie Imada, Kosuke Takemura, and Jenny Ramaswamy. 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., Todd Cherry, Stephan Kroll, Robert J. Netzer, and Matthias Sutter. 2008. "Conditional Cooperation on Three Continents". *Economics Letters* 101 (3): 175–178.
- Litina, Anastasia. 2016. "Natural Land Productivity, Cooperation and Comparative Development". *Journal of Economic Growth* 21 (4): 351–408.
- Loveland, T. R., B. C. Reed, J. F. Brown, D. O. Ohlen, Z. Zhu, L. Yang, and J. W. Merchant. 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, Nathan Nunn, James A. Robinson, and Jonathan Weigel. 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.
- Nguyen, Dinh-Dau. 1996. *Nghien Cuu Dia Ba Trieu Nguyen: Binh Dinh*. Ho Chi Minh City: Ho Chi Minh City Publisher.
- . 1995. *Nghien Cuu Dia Ba Trieu Nguyen: Khanh Hoa*. Ho Chi Minh City: Ho Chi Minh City Publisher.

- . 2010a. *Nghien Cuu Dia Ba Trieu Nguyen: Quang Nam*. Ho Chi Minh City: National University, Ho Chi Minh City.
- . 2010b. *Nghien Cuu Dia Ba Trieu Nguyen: Quang Ngai*. Ho Chi Minh City: National University, Ho Chi Minh City.
- . 1994. *Tong Ket Nghien Cuu Dia Ba Nam Ky Luc Tinh*. Ho Chi Minh City: Ho Chi Minh City.
- 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, The-Anh. 2003. “Village versus State: The Evolution of State-Local Relations in Vietnam until 1945”. *Southeast Asian Studies* 41 (1): 101–123.
- Nguyen, Thi-Dieu. 1999. *The Mekong River and the Struggle for Indochina: Water, War, and Peace*. Westport: Praeger.
- 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. 2016. “Long-Run Cultural Divergence: Evidence From the Neolithic Revolution”. *Journal of Development Economics* 122:197–213.
- Phan, Huy Le, Xuan Bien Phan, Thi Kim Chuyen Truong, Duc Cuong Tran, Minh Giang Vu, Minh Huan Doan, Van Kim Nguyen, Van Le Ngo, Quang Ngoc Nguyen, Cong Nguyen Vo, Van Quan Vu, and Van Sen Vo. 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, X. Zhang, S. Oishi, C. Shimin, D. Duan, X. Lan, and S. Kitayama. 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.
- . 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. 2003. “Vietnamisation of Southern Vietnam during the First Half of the Nineteenth Century”. *Asian Ethnicity* 4 (1): 47–65.
- Zelmer, Jennifer. 2003. “Linear Public Goods Experiments: A Meta-Analysis”. *Experimental Economics* 6 (3): 299–310.

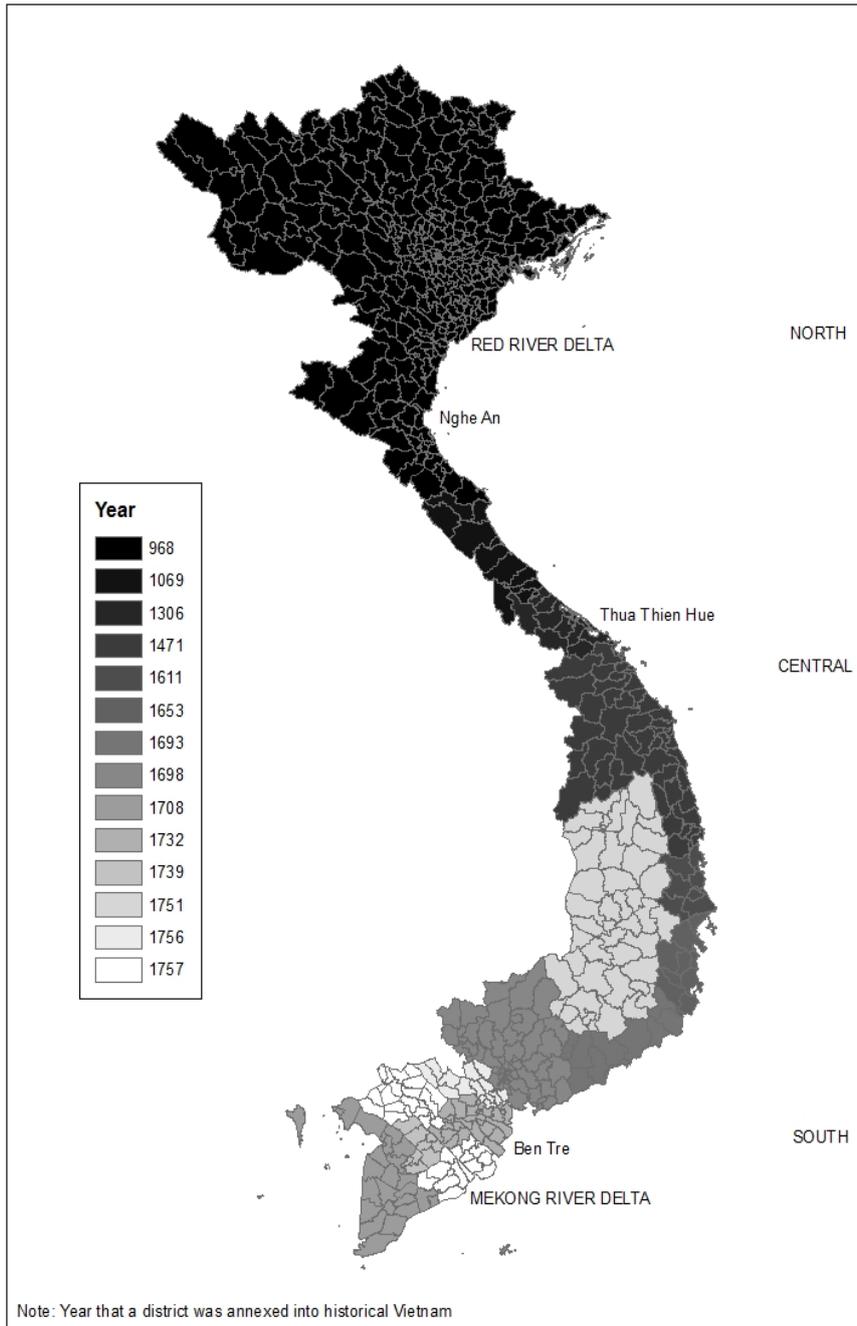
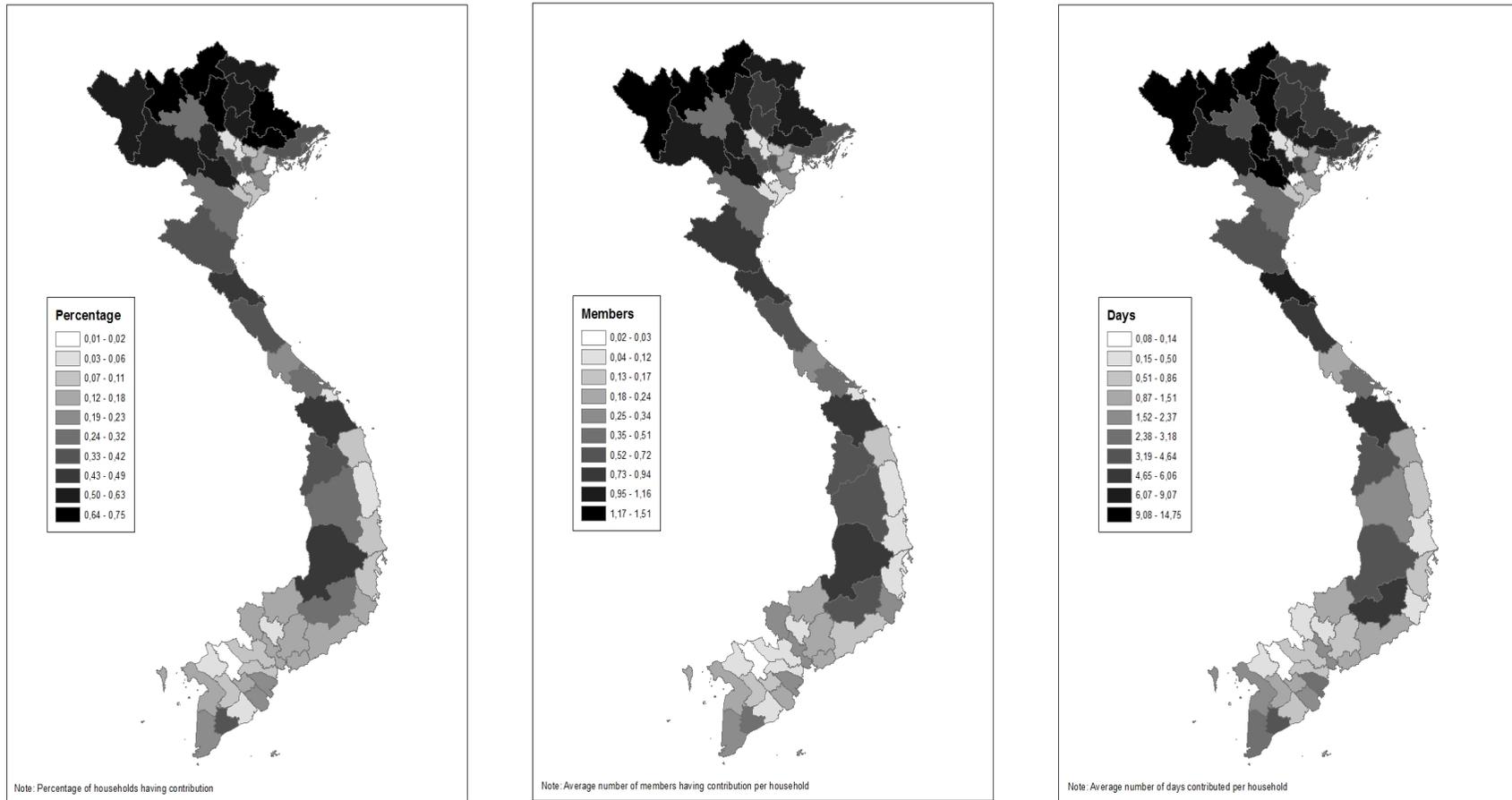


Figure 1. The Vietnamese Southern Advance

Source: Authors' construction, district border is from the Global Administrative Unit Layers



A. Percentage of Households

B. Number of Persons per Household

C. Number of Days per Household

Figure 2. Labor Contribution to Public Goods Production

Source: VHLSS 2002, province border is from the Global Administrative Unit Layers

Table 1. The Prevalence of Collectivism

	Percentage of households contributing labor					
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.028*** (0.004)	0.026*** (0.003)	0.024*** (0.005)	0.024*** (0.003)	0.017*** (0.003)	0.024*** (0.004)
Percentage of cropland area		-0.320*** (0.034)				-0.180*** (0.047)
Caloric suitability index			-0.019*** (0.007)			-0.005 (0.006)
Elevation				0.452*** (0.043)		0.286*** (0.081)
Ruggedness					0.095*** (0.009)	0.005 (0.021)
Constant	0.122*** (0.024)	0.283*** (0.031)	0.655*** (0.189)	0.058*** (0.021)	0.089*** (0.022)	0.318* (0.178)
Adjusted R^2	0.092	0.213	0.122	0.259	0.247	0.291
Observations	607	607	555	607	607	555

Note: OLS estimator, robust standard errors are in parentheses.

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

Table 2. The Strength of Collectivism

A.	Average number of persons making contributions per household					
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.056*** (0.007)	0.050*** (0.006)	0.047*** (0.009)	0.047*** (0.006)	0.031*** (0.006)	0.045*** (0.007)
Constant	0.173*** (0.044)	0.503*** (0.059)	1.254*** (0.372)	0.036 (0.038)	0.102** (0.040)	0.448 (0.353)
Mean value of dep. var.	0.55	0.55	0.58	0.55	0.55	0.58
Adjusted R^2	0.092	0.222	0.122	0.292	0.279	0.323
B.	Average number of days contributed per household					
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.533*** (0.058)	0.499*** (0.054)	0.374*** (0.069)	0.467*** (0.053)	0.347*** (0.053)	0.337*** (0.062)
Constant	0.436 (0.317)	2.613*** (0.436)	16.681*** (3.458)	-0.556* (0.331)	-0.111 (0.318)	10.197*** (3.628)
Mean value of dep. var.	4.05	4.05	4.27	4.05	4.05	4.27
Adjusted R^2	0.113	0.188	0.179	0.253	0.259	0.296
Percentage of cropland area	NO	YES	NO	NO	NO	YES
Caloric suitability	NO	NO	YES	NO	NO	YES
Elevation	NO	NO	NO	YES	NO	YES
Ruggedness	NO	NO	NO	NO	YES	YES
Observations	607	607	555	607	607	555

Note: OLS estimator, robust standard errors are in parentheses.

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

Table 3. Robustness: Frontier Environment and Omitted-Variable Bias

A. OLS	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.028*** (0.011)	0.024* (0.013)	0.066*** (0.023)	0.050* (0.026)	0.392*** (0.150)	0.352** (0.160)
Constant	0.121*** (0.039)	0.676 (0.485)	0.143* (0.079)	1.628 (1.052)	0.872* (0.506)	9.198 (7.761)
Adjusted R^2	0.026	0.156	0.040	0.193	0.029	0.116
B. IV (TSLS)	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.064*** (0.013)	0.032** (0.014)	0.140*** (0.026)	0.067** (0.028)	0.677*** (0.171)	0.315* (0.193)
Constant	-0.005 (0.039)	0.532 (0.498)	-0.117 (0.074)	1.332 (1.091)	-0.122 (0.529)	9.823 (8.411)
Exogeneity (p -value)	0.000	0.287	0.000	0.249	0.005	0.753
First-stage F statistic	284	312	284	313	284	313
Mean value of dep. var.	0.22	0.23	0.37	0.39	2.24	2.31
Control variables	NO	YES	NO	YES	NO	YES
Observations	311	289	311	289	311	289

Note: Robust standard errors are in parentheses. Control variables include percentage of cropland area, caloric suitability, elevation, and ruggedness. All regressions only include districts in the annexed region. Walking distance to Nghe An is employed as an instrumental variable for the time since annexation into historical Vietnam, see the main text for more detail.

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

Table 4. Robustness: French Colonization and Modern Population Movements

A. Cochinchina areas	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.179*** (0.049)	0.237*** (0.065)	0.309*** (0.079)	0.366*** (0.097)	1.372** (0.681)	1.547** (0.777)
Constant	-0.337*** (0.124)	0.247 (0.739)	-0.619*** (0.197)	0.123 (1.091)	-2.204 (1.730)	3.012 (9.264)
Mean value of dep. var.	0.14	0.14	0.20	0.20	1.44	1.40
Control variables	NO	YES	NO	YES	NO	YES
Adjusted R^2	0.053	0.062	0.066	0.069	0.013	0.006
Observations	169	153	169	153	169	153
B. Lowland areas	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.027*** (0.004)	0.023*** (0.004)	0.052*** (0.007)	0.044*** (0.008)	0.468*** (0.054)	0.297*** (0.066)
Constant	0.094*** (0.024)	0.472** (0.209)	0.116*** (0.042)	0.724* (0.396)	0.221 (0.288)	14.045*** (4.240)
Mean value of dep. var.	0.19	0.20	0.31	0.32	1.87	1.94
Control variables	NO	YES	NO	YES	NO	YES
Adjusted R^2	0.095	0.242	0.102	0.254	0.122	0.216
Observations	478	446	478	446	478	446

Note: Robust standard errors are in parentheses. Control variables include percentage of cropland area, caloric suitability, elevation, and ruggedness. Panel A only includes districts in the former Cochinchina. Panel B excludes districts in Ha Noi and Ho Chi Minh City and districts whose elevations are above 500 meters.

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

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

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

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

Table 6. Unconditional Contribution: Regression Analysis

	Unconditional contribution						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Northern site	0.921** (0.426)	0.973** (0.428)	0.835* (0.425)	0.891** (0.436)	0.892** (0.435)	0.789* (0.410)	-1.409 (0.966)
Belief						0.294*** (0.059)	0.215*** (0.072)
Northern site×Belief							0.276** (0.123)
Male		0.235 (0.455)			0.241 (0.457)	0.378 (0.440)	0.474 (0.439)
Household size			0.219* (0.132)		0.222 (0.141)	0.131 (0.137)	0.117 (0.132)
Wealth Index				0.050 (0.140)	-0.010 (0.150)	-0.046 (0.133)	-0.046 (0.135)
Constant	6.579*** (0.266)	6.475*** (0.302)	5.657*** (0.593)	6.590*** (0.268)	5.533*** (0.680)	3.618*** (0.805)	4.237*** (0.822)
Adjusted R^2	0.010	0.008	0.014	0.007	0.009	0.085	0.097
Observations	373	373	373	373	373	373	373

Note: OLS estimator with robust standard errors in parentheses. The sample includes 138 subjects from the northern site and 235 subjects from the southern site.

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

Appendix A. Additional Tables

Table A1. Variable Description

Variable	Description	Mean	SD	Min	Max	N	Source
<i>Main Variables</i>							
Having contribution	Percentage of households contributing labor	0.31	0.33	0	1	607	VHLSS 2002
Members contributed	Average no. of persons per household making contributions	0.55	0.65	0	4.10	607	VHLSS 2002
Days contributed	Average no. of days contributed per household	4.05	5.63	0	41	607	VHLSS 2002
Time since annexation	Number of centuries since annexation into historical Vietnam	6.77	3.57	2.33	10.22	607	DVSKTT and DNTL
<i>Control Variables</i>							
Percentage of cropland area	Percentage of cropland in total area	0.45	0.36	0	1	607	GLCCv2 (EROS)
Caloric suitability index	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)
Ruggedness	Average topographic heterogeneity (100 m)	1.14	1.43	0	6.27	607	Nunn & Puga (2012)
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 measured at district level. Variables from the VHLSS are aggregated without sampling weights. VHLSS: Vietnam Household Living Standard Survey. DVSKTT and DNTL: Dai Viet Su Ky Toan Thu and Dai Nam Thuc Luc. GLCCv2: Global Land Cover Characteristics Database Version 2.0. EROS: Center for Earth Resources Observation and Science, United States Geological Survey. GTOPO30: Global 30 Arc-Second Elevation Dataset. GAUL: Global Administrative Unit Layers. FAO: Food and Agriculture Organization, United Nations.

Table A2. Sampling Weights

A. Full sample	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.029*** (0.004)	0.025*** (0.004)	0.056*** (0.007)	0.046*** (0.008)	0.526*** (0.059)	0.335*** (0.064)
Constant	0.120*** (0.024)	0.302* (0.180)	0.172*** (0.044)	0.437 (0.356)	0.501 (0.330)	10.172*** (3.889)
Control variables	NO	YES	NO	YES	NO	YES
Adjusted R^2	0.093	0.294	0.091	0.322	0.106	0.287
Observations	607	555	607	555	607	555
B. Annexed areas	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.028** (0.011)	0.024* (0.013)	0.062*** (0.023)	0.049* (0.026)	0.376** (0.157)	0.351** (0.168)
Constant	0.122*** (0.039)	0.643 (0.484)	0.153* (0.079)	1.515 (1.043)	0.962* (0.533)	8.397 (7.745)
Control variables	NO	YES	NO	YES	NO	YES
Adjusted R^2	0.025	0.162	0.035	0.194	0.025	0.125
Observations	311	289	311	289	311	289
C. Cochinchina areas	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.173*** (0.049)	0.226*** (0.065)	0.297*** (0.079)	0.347*** (0.096)	1.297* (0.667)	1.496* (0.785)
Constant	-0.320** (0.125)	0.142 (0.773)	-0.585*** (0.199)	-0.219 (1.131)	-1.955 (1.720)	2.769 (10.155)
Control variables	NO	YES	NO	YES	NO	YES
Adjusted R^2	0.053	0.062	0.066	0.069	0.013	0.006
Observations	169	153	169	153	169	153
D. Lowland areas	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.027*** (0.004)	0.024*** (0.004)	0.052*** (0.007)	0.045*** (0.008)	0.463*** (0.055)	0.305*** (0.069)
Constant	0.091*** (0.024)	0.482** (0.214)	0.116*** (0.043)	0.705* (0.403)	0.246 (0.301)	14.205*** (4.668)
Control variables	NO	YES	NO	YES	NO	YES
Adjusted R^2	0.097	0.236	0.100	0.237	0.117	0.200
Observations	478	446	478	446	478	446

Note: OLS estimator, robust standard errors are in parentheses. Control variables include percentage of cropland area, caloric suitability, elevation, and ruggedness. Panel B only includes districts in the annexed region. Panel C only includes districts in the former Cochinchina. Panel D excludes districts in Ha Noi and Ho Chi Minh City and districts whose elevations are above 500 meters.

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

Table A3. Spatial Autocorrelation

A. Full sample	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.028*** (0.004)	0.024*** (0.004)	0.056*** (0.007)	0.045*** (0.008)	0.533*** (0.060)	0.337*** (0.070)
Constant	0.122*** (0.027)	0.318* (0.166)	0.173*** (0.054)	0.448 (0.321)	0.436 (0.462)	10.197*** (2.850)
Control variables	NO	YES	NO	YES	NO	YES
Adjusted R^2	0.523	0.642	0.471	0.618	0.415	0.545
Observations	607	555	607	555	607	555
B. Annexed areas	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.028*** (0.009)	0.024** (0.011)	0.066*** (0.018)	0.050** (0.021)	0.392*** (0.122)	0.352** (0.149)
Constant	0.121*** (0.036)	0.676 (0.499)	0.143** (0.068)	1.628* (0.940)	0.872* (0.472)	9.198 (6.682)
Control variables	NO	YES	NO	YES	NO	YES
Adjusted R^2	0.406	0.498	0.357	0.469	0.296	0.367
Observations	311	289	311	289	311	289
C. Cochinchina areas	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.179*** (0.056)	0.237*** (0.066)	0.309*** (0.086)	0.366*** (0.099)	1.372* (0.770)	1.547* (0.865)
Constant	-0.337** (0.149)	0.247 (0.817)	-0.619*** (0.230)	0.123 (1.235)	-2.204 (2.055)	3.012 (10.756)
Control variables	NO	YES	NO	YES	NO	YES
Adjusted R^2	0.391	0.410	0.369	0.387	0.255	0.262
Observations	169	153	169	153	169	153
D. Lowland areas	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.027*** (0.004)	0.023*** (0.004)	0.052*** (0.007)	0.044*** (0.008)	0.468*** (0.057)	0.297*** (0.070)
Constant	0.094*** (0.029)	0.472** (0.196)	0.116** (0.053)	0.724** (0.362)	0.221 (0.435)	14.045*** (3.073)
Control variables	NO	YES	NO	YES	NO	YES
Adjusted R^2	0.496	0.584	0.457	0.555	0.416	0.484
Observations	478	446	478	446	478	446

Note: OLS estimator, standard errors are in parenthesis and calculated following Conley (1999) with the assumption that autocorrelation decreases in the distance between district centroids and equals zero for districts that are more than 0.5 degree apart. Control variables include percentage of cropland area, caloric suitability, elevation, and ruggedness. Panel B only includes districts in the annexed region. Panel C only includes districts in the former Cochinchina. Panel D excludes districts in Ha Noi and Ho Chi Minh City and districts whose elevations are above 500 meters.

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

Appendix B. Constructing the Time since Annexation to Historical Vietnam

In this appendix, we present the construction of the time since annexation to historical Vietnam in chronological order, relying on 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). We also provide many secondary sources in English that are in agreement with these primary sources that the Southern Advance (Nam Tien) of historical Vietnam ended in 1757 CE, by which the border of Vietnam was established as it is today.

The Dinh Dynasty

In 938 CE, Ngo Quyen defeated the Southern Han Kingdom and ended a millennium of being colonized by historical China (*Toan Thu* p. 118). Nevertheless, his dynasty was short-lived and followed by civil wars among 12 independent feudal lords. Only in 968 CE was Dinh Bo Linh able to pacify these feudal warlords and establish the first unified state of historical Vietnam (*Toan Thu* p. 127). The territory of this state included the whole area that is now north Vietnam (Dao 2005, p. 114-118). We code all districts in the modern north, i.e., from the border with China down to Ha Tinh Province, to be annexed in 968 CE.

The Ly Dynasty

In 1069 CE, Ly Nhat Ton attacked the former Champa Kingdom to retaliate a territorial intrusion (*Toan Thu* p. 197). After being defeated and captured, the Champa King submitted the former Bo Chinh, Dia Ly, and Ma linh to compensate for his release. *Toan Thu* notes that Bo Chinh and Dia Ly were the north and south of modern Quang Binh Province, while Ma Linh was the north of modern Quang Tri Province. Dao (2005, p. 162) further identifies that Ma Linh included the northern area of modern Quang Tri Province down to Thach Han River. We code all districts in modern Quang Binh Province and in the north of Thach Han River in modern Quang Tri Province (Vinh Linh, Gio Linh, Cam Lo, Dong Ha, and Huong Hoa) to be annexed in 1069 CE.

The Tran Dynasty

In 1306 CE, Tran Anh Tong fulfilled a promise made by his father to arrange a marriage between Princess Huyen Tran of historical Vietnam and King Che Man of the former

Champa Kingdom (*Toan Thu* p. 340). In return, King Che Man submitted the former O and Ly as wedding presents. The former O was in the south of Thach Han River in modern Quang Tri Province, while the former Ly was modern Thua Thien Hue Province (Dao 2005, p. 163-164). We code all districts in the south of Thach Han River in modern Quang Tri Province (Trieu Phong, Quang Tri, Da Krong, and Hai Lang) and in modern Thua Thien Hue Province to be annexed in 1306 CE.

The Le Dynasty

In 1471 CE, Le Tu Thanh mobilized a military campaign against the former Champa Kingdom in response to its attack on historical Vietnam one year before (*Toan Thu* p. 659-662). The campaign was a victory in which historical Vietnam annexed the former Dai Chiem and Co Luy.¹⁵ These areas correspond to a territory from modern Hai Van Pass to modern Cu Mong Pass, where former Dai Chiem was modern Quang Nam Province while former Co Luy was modern provinces of Quang Ngai and Binh Dinh (Dao 2005, p. 201-203). Historical Vietnam also had control over the highlanders down to modern Binh Dinh Province, which modern Kon Tum Province can be said to be included (Dao 2005, p. 203). We code all districts in modern provinces of Quang Nam, Quang Ngai, Binh Dinh, and Kon Tum to be annexed in 1471 CE.

The Nguyen Lords

In 1611 CE, responding to a territorial intrusion by the former Champa Kingdom, Nguyen Hoang attacked and annexed what is now modern Phu Yen Province (*Thuc Luc* p. 36). We code all districts in modern Phu Yen Province to be annexed in 1611 CE.

In 1653 CE, King Ba Tam of the former Champa Kingdom invaded modern Phu Yen Province, Nguyen Phuc Tan retaliated and annexed a new area down to Phan Rang River (*Thuc Luc* p. 62). This area corresponds to modern Khanh Hoa Province and the north of modern Ninh Thuan Province. We code all districts in modern Khanh Hoa Province and in the north of Phan Rang River in modern Ninh Thuan Province (Bac Ai, Ninh Hai, and Phan Rang-Thap Cham) to be annexed in 1653 CE.

¹⁵Under the Ho Dynasty, historical Vietnam already controlled these regions from 1402 CE to 1407 CE (*Toan Thu* p. 436-437). When the Ming Dynasty of historical China conquered historical Vietnam, the former Champa Kingdom took them back. Thus, we do not count this event as an annexation.

In 1692 CE, King Ba Tranh of the former Champa Kingdom raided the border of historical Vietnam (*Thuc Luc* p. 106). In 1693 CE, Nguyen Phuc Chu conquered and annexed the final territory of the former Champa Kingdom (*Thuc Luc* p. 107). This area included the south of Phan Rang River in modern Ninh Thuan Province and modern Binh Thuan Province. We code all districts in the south of Phan Rang River in modern Ninh Thuan Province (Ninh Phuoc and Ninh Son) and in modern Binh Thuan Province to be annexed in 1693 CE.

After annexing the whole former Champa Kingdom, historical Vietnam gradually expanded its territory into the land of the former Khmer Empire in the far south. Before that, in 1658 CE, King Ponhea Chan of the former Khmer Empire invaded the border of historical Vietnam. Following his defeat, King Ponhea Chan had to adopt a tributary position towards historical Vietnam and accept Vietnamese migrants to move in and exploit the Khmer land in the far south of Vietnam (*Thuc Luc* p. 72). In 1698 CE, historical Vietnam officially established two new provinces in this land, Tran Bien and Phien Tran, to register land and collect tax, as well as to mobilize more Vietnamese migrants to settle down (*Thuc Luc* p. 111). In *Thuc Luc* (p. 111), Tran Bien and Phien Tran provinces are noted to correspond to Bien Hoa and Gia Dinh provinces under the Nguyen Dynasty, i.e., the time where this chronicle was written. Based on *Dai Nam Nhat Thong Toan Do*, the official map of historical Vietnam produced by the Nguyen Dynasty in 1838 CE, this new territory basically corresponds to the south east region of modern Vietnam (Nguyen 1994, p. 64). We code all districts in modern provinces of Binh Phuoc, Dong Nai, Ba Ria-Vung Tau, Ho Chi Minh City, Binh Duong, and Tay Ninh to be annexed in 1698 CE.

During the Qing conquest of historical China, a Chinese refugee named Mo Jiu ran away and successfully maneuvered himself into the court of the former Khmer Empire and was appointed to govern a coastal area in the Mekong River Delta. In 1708 CE, Mo Jiu submitted his land to historical Vietnam and received an appointment to govern this territory, i.e. the former Ha Tien (*Thuc Luc* p. 122). This territory roughly corresponds to modern provinces of Kien Giang, Ca Mau, and Bac Lieu (Phan et al. 2011, p. 448). We code all districts in these provinces to be annexed in 1708 CE.

In 1731 CE, some inhabitants in the former Khmer Empire raided the south of historical Vietnam. In 1732 CE, historical Vietnam invaded the former Khmer Empire to track

down these raiders. King Ang Chey of the former Khmer Empire proposed a cease-fire and promised to deliver the raiders. Historical Vietnam then continued to establish its bureaucracy in an area in the far south, i.e. the former Long Ho (*Thuc Luc* p. 141-143). This area roughly corresponds to modern provinces of Vinh Long, Ben Tre, and Tien Giang (Phan et al. 2011, p. 448). We code all districts in these provinces to be annexed in 1732 CE.

In 1755 CE, the former Khmer Empire attacked historical Vietnam. Being retaliated, King Ang Sngoun of the former Khmer Empire ceded the former Tam Bon and Loi Lap in 1756 CE as compensation (*Thuc Luc* p. 164-165). This area roughly corresponds to modern Long An Province (Phan et al. 2011, p. 448). We code all districts in this province to be annexed in 1756 CE.

In 1756 CE, King Ang Sngoun of the former Khmer Empire died, which ignited a fight for throne between royal forces. Being called for help from one side of the fight in 1757 CE, historical Vietnam intervened in return for more land, i.e. the former Tra Vinh, Ba Thac, and Tam Phong Long (*Thuc Luc* p. 166-167). These areas roughly correspond to modern provinces of Dong Thap, An Giang, Can Tho, Soc Trang, and Tra Vinh (Phan et al. 2011, p. 448). We code all districts in these provinces to be annexed in 1757 CE.

For modern provinces of Gia Lai, Dak Lak, Lam Dong, and Dak Nong in the Central Highland Region, we code 1751 CE as the year they were annexed as it was the first year recording that highlanders came to pay tributes (*Thuc Luc* p. 157). Historical studies have also argued that this is the time when historical Vietnam started collecting taxes from highlanders (Tana 1998).

Secondary Sources on the Vietnamese Southern Advance

Below, we provide extracts from a standard Vietnamese history and some related historical studies in English that are in agreement that the Southern Advance (Nam Tien) of historical Vietnam ended in 1757 CE, by which historical Vietnam had already gained control over a territory as today.

From Taylor (2013, p. 336):

“One year later, in 1757, Chei Chéttha IV died. Two men who endeavored to succeed him were successively murdered. Vietnamese armies were again mobilized to intervene. . . . As part of the final settlement,

Outeireachea III ceded more territories along the borders of Ha Tien to Mac Thien Tu, who then passed them along to Phu Xuan. With this, the Khmer-Viet border was drawn more or less as it exists today.”

From Briggs (1947, p. 358):

“Thus, by infiltration and by occupation, the Annamites acquired most of the lower delta, early in the seventeenth century. From that time, Cambodia was forced to pay Annam a regular tribute. By the end of the century, the Annamites had absorbed all the lower delta east of the Mekong and had organized it into Annamite administrative units. In 1758, they completed the occupation of the entire delta and fortified it.”

From Nguyen (1999, p. 18)

“This *Nam Tien*, which was spearheaded by individuals, families, and now by the authorities, reached its end toward the middle of the eighteenth century in the Transbassac region, dangerously close to the Khmer Kingdom.”

From Engelbert (1994, p. 173):

“By the end of this conquest the Vietnamese occupied the two areas situated between the mouths of the Mekong and the Bassac River: Bassac (Ba Thac, i.e., Soc Trang) and Tra Vinh, which the Nguyen got as late as in 1757 (or 1758).”