Environment for Development

Discussion Paper Series

April 2019 ■ EfD DP 19-08

The Origins of Cultural Divergence

Evidence from a Developing Country

Hoang-Anh Ho, Peter Martinsson, and Ola Olsson

The Environment for Development (EfD) initiative is an environmental economics program focused on international research collaboration, policy advice, and academic training. Financial support is provided by the Swedish International Development Cooperation Agency (Sida). Learn more at www.efdinitiative.org or contact info@efdinitiative.org.

Central America
Research Program in Economics and Environment for Development in Central America Tropical Agricultural Research and Higher Education Center (CATIE)

Chile
Research Nucleus on Environmental and Natural Resource Economics (NENRE)
Universidad de Concepción

China
Environmental Economics Program in China (EEPC)
Peking University

Colombia
The Research Group on Environmental, Natural Resource and Applied Economics Studies (REES-CEDE), Universidad de los Andes, Colombia

Ethiopia
Environment and Climate Research Center (ECRC)
Ethiopian Development Research Institute (EDRI)

India
Centre for Research on the Economics of Climate, Food, Energy, and Environment, (CECFEE), at Indian Statistical Institute, New Delhi, India

Kenya
School of Economics
University of Nairobi

South Africa
Environmental Economics Policy Research Unit (EPRU)
University of Cape Town

Sweden
Environmental Economics Unit
University of Gothenburg

Tanzania
Environment for Development Tanzania
University of Dar es Salaam

USA (Washington, DC)
Resources for the Future (RFF)

Vietnam
University of Economics
Ho Chi Minh City, Vietnam

The Environment for Development (EfD) initiative is an environmental economics program focused on international research collaboration, policy advice, and academic training. Financial support is provided by the Swedish International Development Cooperation Agency (Sida). Learn more at www.efdinitiative.org or contact info@efdinitiative.org.

Contents

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

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

University of Gothenburg

April 2019

Abstract

Cultural norms diverge substantially across societies, often within the same country. In the present paper, we propose and investigate the selective migration hypothesis, proposing that cultural differences along the individualism-collectivism dimension are driven by the out-migration of individualistic people from collectivist societies to frontier areas, and that such patterns of historical migration are reflected even in the current distribution of cultural norms. Gaining independence in 939 after 1000 years of Chinese colonization, historical Vietnam occupied the region that is now north Vietnam with a collectivist social organization. From the 11^{th} to the 18^{th} centuries, historical Vietnam gradually expanded its territory southward to the Mekong River Delta through various waves of conquest and migration. Combining findings from a household survey and a 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 robustness checks, we show that the southward out-migration of individualistic people during the territorial expansion is an important driver behind this finding.

Keywords: Culture; Selective Migration; 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, Dang-Thuy Truong, Fabrizio Zilibotti, and seminar participants at the University of Gothenburg and 16^{th} Nordic Conference in Development Economics. We would also like to thank EfD-Vietnam for kindly hosting and supporting our field research and data collection. All remaining errors are our own. A research grant from the Swedish Research Council (ref 348 2013-6348) is gratefully acknowledged. Contact: Hoang-Anh Ho (hoang-anh.ho@economics.gu.se), Peter Martinsson (peter.martinsson@economics.gu.se), and Ola Olsson (ola.olsson@economics.gu.se).

1 Introduction

Economic research has uncovered strong associations between many cultural traits and various indicators of individual behavior and institutional and economic development (e.g., 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 economic and democratic development in a large sample of 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 repeated processes of territorial expansion and migration that happened in historical times. In particular, we put forward a *selective migration hypothesis*, consisting of three building blocks. First, in regions where settled agriculture and states arose early, collectivist societies emerged through a process of *self-domestication* as communities made the transition from hunter-gatherer strategies of food procurement, which were characterized by individualism, to agricultural activities, resulting in a gradual strengthening of collectivism. Second, these collectivist societies triggered the out-migration of individualistic members toward peripheral areas. This pattern then repeated itself as the individualistic migrants inhabited and developed these peripheral areas into less collectivistic 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, owing to the slow-moving nature of culture, these differences have persisted over time and constitute an important feature of the cultural landscapes exhibited in modern times. As a result, the time elapsed since collectivist transformation can predict the strength of collectivism across modern societies.

Testing the selective migration hypothesis requires a historical setting where there was

[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).

a large out-migration of people from a collectivist society to settle down in new regions and that this migration repeated over time once collectivism was gradually strengthened in the new regions. We find such an ideal setting in the process of territorial expansion and migration in historical Vietnam. Gaining independence from the colonization of imperial China during the first millennium, historical Vietnam initially governed the region of what is now north Vietnam (figure 1) with a centralized government and a collectivist social organization. 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 968 to 1757, historical Vietnam gradually expanded its territory southward to the Mekong River Delta to establish the country as it is today (figure 1). This process happened through successive waves of state conquest followed by civil migration, resulting in the displacement of most of the population of local ethnic tribes. Applying the logic of the selective migration hypothesis, we argue that the time elapsed since annexation to historical Vietnam is an important predictor of the strength of collectivist norms across regions within contemporary Vietnam.

[Figure 1 is about here.]

To test the selective migration hypothesis, an ideal empirical strategy would consists of three integral parts. The first part should demonstrate that some early agricultural states were characterized by collectivism, and that people who migrated to the new territories following state expansions were less collectivistic or more individualistic than those who stayed. The second requires historical data to prove that these selective migrations gave rise to early cultural differences along the individualism-collectivism dimension between the initial regions and the new territories. Finally, the third part involves using present-day data to conduct an empirical analysis on the relationship between the time elapsed since collectivist transformation and the strength of collectivism.

To match this ideal empirical strategy in the context of Vietnam, we first present qualitative accounts to demonstrate that the initial society of historical Vietnam was characterized by strong collectivist norms. Second, we examine available primary records on the territorial expansion of historical Vietnam to identify the categories of people who migrated to the new annexed territories. In addition, we also provide qualitative accounts and descriptive statistics to show that cultural differences along the individualism-collectivism dimension across regions were already present in Vietnam in the 17^{th} century.

Most importantly, we differentiate the selective migration hypothesis from other potential theories that can also account for these cultural differences. Third, we provide empirical evidence for a positive relationship between the time elapsed since an area was annexed to historical Vietnam and various indicators of collectivism in the present day. Using different robustness checks, we further show that these empirical findings are consistent with the selective migration hypothesis.

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: (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 potential confounding factors. Using various robustness checks, we further test the selective migration hypothesis against other potential explanations of our empirical results, including different living conditions in the annexed region, omitted-variable bias, the Western influence in south Vietnam in the 19^{th} and 20^{th} centuries, and important population movements in the modern day as a result of state planning and urbanization.

To examine the selective migration 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 not yet significantly exposed to other external influences. By comparing the cooperation levels between the two studied districts, we test the selective migration hypothesis. To ensure that the selective migration hypothesis is examined without confounding effects, we take great care in selecting experimental sites in order to distinguish

the selective migration hypothesis from other potential theories mentioned above. We find that subjects from the earlier-annexed district contribute substantially more in the public goods experiments 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. Thus, 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 today among teenagers of the country.

Our research relates to a growing multidisciplinary literature examining the origins of cultural differences along the individualism-collectivism dimension.[2] 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 (e.g., Vandello and Cohen 1999; Talhelm et al. 2014). Litina (2016) argues further that lower level of natural land productivity increased the return to public agricultural infrastructure, which generated higher incentives for cooperation to solve the problem of collective action. 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) propose the voluntary settlement hypothesis, asserting that settlers in frontier areas are likely to have highly autonomous, independent, and goal-oriented mindsets; see also Bazzi, Fiszbein, and Gebresilasse (2017). Within the Western world, Olsson and Paik (2016) show that collectivism is stronger in regions that adopted agriculture earlier.

The present paper 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 the problem of collective action, especially using a combination of survey and experimental data. Furthermore, most studies so far have either employed cross-country comparisons or concentrated on developed societies. Because these societies have gone through the modernization process

[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).

to a great extent, the reduction of the traditional cultural landscapes makes it harder to study the historical origins of cultural differences. By comparing different regions within a developing country, our research is able to overcome these limitations.

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 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 (Lowes et al. 2017). In a related study, Dell, Lane, and Querubin (2018) 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 complementary in that it investigates 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 discusses in detail the conceptual framework behind the selective migration hypothesis. Section 3 provides the historical background of the southward territorial expansion of historical Vietnam and the accompanying migration process, with a focus on the three building blocks of the selective migration hypothesis. Section 4 presents the empirical analysis with survey data. Section 5 describes the sample selection, experimental design, and corresponding results. Section 6 closes the paper with some concluding remarks.

2 Conceptual Framework

In this section, we first define the individualism-collectivism dimension in the cultural repertoire of a population. We then outline a theory of *selective migration* and cultural divergence along the individualism-collectivism dimension. This theory is the backbone of the selective migration hypothesis.

2.1 Individualism versus Collectivism

Research on the individualism-collectivism dimension of culture was first initiated within social psychology. Many of the key insights were summarized by Triandis (1995) and in subsequent cross-country empirical research by Hofstede (2001). In this voluminous

literature, *collectivism* is considered to be characterized by a strong focus on the goals of a collective that forms the in-group boundary. In other words, the goals of the individual are subordinate to the goals of the collective, and the individual willingly makes costly sacrifices for the group. The individual typically has low self-expression and self-esteem, and an interdependent sense of agency. Family, duty, honor, and respect for elders are central for collectivists. On the macro level, collectivist societies are typically characterized by a highly stratified or autocratic leadership (sometimes referred to as *vertical collectivism*) and hostility towards out-groups.

Individualism is orthogonal to collectivism on all the features mentioned above. There is a strong focus on the goals of the individual, and the in-group identity is weak. The goals of the individual are superior to the goals of the collective, and the individual is typically unwilling to make costly contributions to the group at the expense of himself or herself. The individual has a strong sense of personal agency, high self-expression and self-esteem. The extended family does not play a central role, and individual preferences and fulfillment are more important than duty and honor. Individualists tend to live in egalitarian societies, are not very loyal to their fellow in-group members, and are willing to cooperate with out-group members (Triandis 1995).

How do these cultural norms translate into economic behavior? This issue has recently been studied in a series of papers by Gorodnichenko and Roland (2011, 2017). The authors outline a hypothesis and demonstrate empirically that societies characterized by individualism are less bound by rules and authority, reward personal achievement, and hence tend to be associated with fewer constraints and stronger incentives for *innovation*. Analogously, the strong norms towards in-group cooperation, combined with subordination of the self to the goals of the collective, give collectivists a comparative advantage for *collective action* and *public goods production* that under certain circumstances might be necessary for the in-group to survive. Individualistic societies are thus more loosely held together but are on the other hand more dynamic, whereas collectivist societies have tight social ties and effective cooperation but limited growth potential in the longer run.

2.2 Selective Migration and Cultural Divergence

To develop the selective migration hypothesis, we build on Triandis (1995) and Olsson and Paik (2016), and assign a crucial role to the rise and consolidation of the early agricultural states. Before the first transition to agriculture about 12000 year ago, all societies relied on

hunting-gathering-fishing where the household of core family members was often the main unit of social organization. Households only stayed in larger camps during shorter periods but then splintered in order to avoid crowding and social tensions. In some environments, hunting required greater coordination, which sometimes led to larger and semi-sedentary social groups, but whenever possible, the basic tendency in pre-agricultural societies was autonomous households without stronger bonds or obligations to other in-group members (Johnson and Earle 2000).[3]

The first agricultural societies emerged in regions such as Mesopotamia and China. In these regions, a highly productive irrigation agriculture gave rise to a dense and sedentary population, living in crowded villages and depending on the cultivation of a few domesticated crops and animals. Compared to hunter-gatherer households in pre-agricultural societies, these early farming villages were characterized by a much greater degree of collectivism, where the goals of the collective were far more important than individual aspirations. The survival of such villages often required sophisticated public goods such as irrigation canals, protective walls, military defense, public granaries, and deep wells. Such projects were initiated, coordinated and supervised by a social elite that managed to control the great majority of the population. By the 4th millennium BCE, the first states arose from such complex farming communities in Mesopotamia (Borcan, Olsson, and Putterman 2018; Scott 2017).

It has been argued by Johnson and Earle (2000), Scott (2017) and others that the new type of social organization in dense agricultural communities and states was only possible through a gradual *self-domestication of humans* whereby the natural inclinations towards family-level social units was overcome through a strong selective pressure favoring individuals and groups who could more successfully adapt to the new lifestyle in the farming villages, with a higher pathogen load, more toil and work hours in the fields, a new diet with more carbohydrates and less protein, and more children per woman. In addition to these changes, we argue that there must have been a very strong pressure towards the adoption of collectivist norms. It is well known that social stratification expanded with agriculture and the goals of the individual were suppressed for the benefit of the common good, involving larger and larger collective action projects such as irrigation,

[3] See, for instance, Johnson and Earle (2000)'s description of the contemporary Kung people of the Kalahari, a society which has been considered a modern remnant of prehistorical social organization.

city walls, the construction of cult centers, and even massive burial complexes for divine rulers (Diamond 1997). This kind of social organization would not have been possible without a great increase in the proliferation of collectivist norms.

We argue that this self-domestication process first emerged in regions with favorable conditions for agriculture, but it then repeated itself all around the world when farming replaced hunting-gathering and states arose from the dense farming communities. The self-domestication process probably had several mechanisms. *Evolution* provides a selective advantage for individuals with genes that helped them cope with the physiological, psychological and cultural challenges of intensive farming. In addition, there were presumably also push factors such as a conscious "weeding out" of individualists who did not adapt to the new collectivist norms. Social exclusion or ostracism might be one mechanism whereby individualists were pushed from the collectivist core to peripheral areas. There were surely also *pull factors*, inducing individualists to leave the collectivist core voluntarily in order to have a freer life at the periphery. Such a "pioneer spirit" was emphasized in Turner (1920)'s work on the westward expansion of settlers in North America, and more recently studied by Bazzi, Fiszbein, and Gebresilasse (2017). It is also similar to Kitayama et al. (2006, 2009)'s notion of *voluntary settlement* of peripheral areas by individualistic people. As discussed by Bazzi, Fiszbein, and Gebresilasse (2017), the adaptation to the living conditions in the "rugged frontier", where a strong sense of individual agency most likely was necessary for survival, surely also contributed to a greater degree of individualism even among people with collectivistic inclinations.

Typically, the peripheries to the original agricultural core region were soon colonized by a collectivist farmer-state through territorial expansion. Evolutionary adaptation, push and pull forces then played out in a similar manner, making the peripheral population more collectivistic as well. But as described by Olsson and Paik (2016) in their application of the selective migration logic to the expansion of Neolithic agriculture throughout the Western hemisphere, the most individualistic people in the periphery would soon once again take off towards more peripheral areas in repeated frontier colonizations. Since self-domestication, just like evolution, is a function of time, the penetration of collectivist norms was typically also an increasing function of the time elapsed since the collectivist transformation. In this manner, a gradient arose with the greatest degree of collectivism in the oldest regions and the highest degree of individualism in the youngest territories

of the farmer-state. The slow-moving nature of culture implied that, centuries or even millennia after the first settlement of individualistic farmers, signals from these early migration processes are still visible in contemporary cultural record.[4]

Nevertheless, as already argued by Triandis (1995), the Industrial Revolution, with innovation as a key driving factor, once again turned the tables and gave individualism an economic advantage in north European countries such as Britain and the Netherlands. Thus, we might expect that the collectivist legacy of the transition from hunting-gathering to farming should be weaker in countries where an industrial economy has existed for a longer period. In addition, Western colonization of regions outside Europe might also change the indigenous cultural landscapes to a large extent. In some developing countries that only experienced industrialization recently and had strong indigenous states, the cultural imprint from the historical expansion of the collectivist farmer-state is more likely to be observable in the present day.

3 Historical Background

In previous section, we have outlined a theory of selective migration and cultural divergence along the individualism-collectivism dimension. In this section, we survey historical materials to examine three building blocks of our theory in the context of Vietnam: (i) the initial region of historical Vietnam was home to a collectivist society; (ii) individualistic people migrated 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.

3.1 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 lived together, without a centralized state, in the region that is now north Vietnam (figure 1). From 111 BCE to 939, the whole region was brought under the colonization of the centralized bureaucracy of imperial China. During this period, "the Vietnamese evolved from a preliterate society within

[4] For instance, Olsson and Paik (2016) show that peripheral Scandinavia and other North European countries still have much stronger individualistic norms than people in the old agricultural regions in the Middle East.

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, the first unified state of historical Vietnam was founded in 968 and inherited a centralized bureaucratic system from the Chinese 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 administrative system and family organization. The village was the lowest administrative level, which was responsible for regulating almost all aspects of the daily living of its members (Nguyen 2003). Two important responsibilities of the village were to allocate public land under its management to its members (Dao 1993), and to organize unpaid labor for public goods production such as irrigation facilities, roads, and communal buildings (Adams and Hancock 1970). With respect to the family, parents had absolute authority over their children in almost all aspects of life (e.g., education, marriage, and housing), while children had to serve and obey their parents with the utmost respect throughout their lives (Haines 1984).

The area bordering historical Vietnam in the south, which 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, which 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). Nevertheless, available historical materials do not allow us to draw any comparison between these societies and historical Vietnam along the individualism-collectivism dimension.

3.2 Selective Migration and Cultural Divergence

From 968 to 1757, 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 annexations and was completed in 1757, by which time the border of Vietnam was established as it is today.[5] Historical Vietnam first annexed

[5] See appendix B for a chronological description of the Vietnamese Southern Advance.

the land from modern Quang Binh to modern Binh Dinh from 968 to 1471. This land was effectively governed by the Nguyen Lords since the early 16$^{\text{th}}$ century, when the fight to control the throne erupted between them and the Trinh Lords in the initial region. From 1611 to 1757, the Nguyen Lords continued to expand the country southwards to the Mekong River Delta to establish the border as it is today. Compared to the initial region, the annexed region under the government of the Nguyen Lords was more open towards foreign trade (Tana 1998, p. 59-98).

After historical Vietnam conquered an area, Vietnamese migrants started to settle in. A few records from two official chronicles of historical Vietnam, *Dai Viet Su Ky Toan Thu* (from 204 BCE to 1675) and *Dai Nam Thuc Luc* (from 1558 to 1888), indicated that Vietnamese migrants to the annexed region ranged from landless farmers to rich adventurers, who took advantage of the opportunities in the new land, and from exiled criminals to recruited soldiers, who were sent to the new land by the government. There are no records available to identify who were the dominant settlers, let alone their cultural characteristics along the individualism-collectivism dimension.[6] Regarding the local ethnic groups, most of their populations migrated away, while those who decided to stay acculturated to the Vietnamese culture (Wook 2004).[7] The logic of the selective migration hypothesis implies that areas annexed earlier to historical Vietnam were more collectivistic. The historical evidence presented below is in support of this prediction.

Within historical Vietnam, cultural differences along the individualism-collectivism dimension between the annexed region and the initial region were already remarkable as early as the 17$^{\text{th}}$ century. For example, Tana (1998, p. 99-116) provides many historical accounts to demonstrate that the social environment in the annexed region was

[6] Historians seem to be in agreement that these migrants were not recognized as true members of the collectivistic society in the initial region. Such a migrant "lacked standing in a social group, like the family or the village, was less than a full person and could hope for no better future in traditional village society" (Tana 1998, p. 111), or even was "marginal and undesirable" (Taylor 2013, p. 211). One may recall the analogous image of European immigrants to the United States during the age of mass migration (Abramitzky, Boustan, and Eriksson 2012). In a related study, Alesina and Giuliano (2010) also find that people with weak family ties are more likely to migrate.

[7] 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).

characterized by greater openness, mobility and autonomy compared to the initial region. Available statistics of land allocation in the early 19th century also illustrate this cultural divergence. In the initial region, land was only allocated to or owned by village members (Nguyen 2003). In the annexed region, however, the in-village/out-village distinction was loosened and land was commonly allocated to or owned by people from other villages. For example, studies on the land registries (cadastres) in the annexed region in the early 19th century show that the proportions of land allocated to or owned by people from other villages were 20% to 30% in southernmost provinces (Nguyen 1994) compared to 8% to 15% in more northern provinces (Nguyen 1997, 2010, 1996a, 1996b).

Besides the selective migration of individualistic people as proposed by our theory, there are certainly other potential explanations for the cultural differences along the individualism-collectivism dimension between the annexed region and the initial region of historical Vietnam as described above. First, the frontier environment in the annexed region (e.g., sparsely populated) might induce Vietnamese migrants to be more individualistic. Second, Vietnamese migrants might be influenced by the cultural characteristics of the Champa Kingdom and Khmer Empire, which in turn might be more individualistic than historical Vietnam. Finally, Vietnamese migrants to the annexed region might pick up individualistic traits from foreigners because of the open trade policy of the Nguyen Lords. The main difference between our theory and these explanations is that our theory predicts cultural differences even within the annexed region, i.e., areas annexed earlier are predicted to be more collectivistic.

3.3 Cultural Differences Have Persisted to the Present Day

The last block of the selective migration hypothesis argues that the cultural differences across regions of historical Vietnam found around the 17th century have persisted and made up a key characteristic of the cultural landscape of modern Vietnam. In other words, the time elapsed since annexation to historical Vietnam is an important predictor of the strength of collectivism within contemporary Vietnam. The north-south cultural differences along the individualism-collectivism dimension in modern Vietnam have been documented in details in many anthropological studies, e.g., Hickey (1964), Rambo (1973), and Luong (1992). This north-south cultural divergence is also a typical characteristic

that is normally mentioned in descriptions about modern Vietnam.[8]

Besides the persistence of cultural differences in the past as proposed by our theory, divergent developments between the north and the south in modern times might also account for the north-south cultural differences along the individualism-collectivism dimension in modern Vietnam. The French colonization started in 1858 and ended with the Vietnamese victory in the First Indochina War (1946-1954), during which the French colonizers concentrated most of their activities in south Vietnam. 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 the reunification of the country in 1975. Meanwhile, Communism started to develop in north Vietnam in the early 20^{th} century and gained control of this part of the country in 1954. The main difference between our theory and an explanation relying on divergent developments in the modern times is that our theory predicts cultural differences even within south Vietnam, i.e., areas annexed earlier are predicted to be more collectivistic.

To sum up, the north-south cultural differences along the individualism-collectivism dimension were already in place as early as the 17^{th} century and are currently a central theme of Vietnam. Although there are certainly many potential explanations, our theory of selective migration uniquely predicts that areas annexed earlier to historical Vietnam are currently more prone to collectivist norms, and this relationship holds even within the annexed region and within south Vietnam. We now turn to investigate these predictions empirically using survey data in section 4 and experimental data in section 5.

4 Survey Data Analysis

4.1 Empirical Model

In this section, we use survey data to investigate the selective migration hypothesis in the context of Vietnam. The key argument of the hypothesis is that collectivist societies trig-

[8] Ending his Vietnamese history, Taylor (2013, p. 624) notes 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". Although regarding collectivism as the main cultural theme in their practical guide to Vietnam, 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".

gered the out-migration of individualistic members toward peripheral areas, and, owing to the slow-moving nature of culture, these differences have persisted over time. Our empirical strategy revolves around regressing a measure capturing the strength of collectivism on the time elapsed since an area was annexed to historical Vietnam, while controlling for potential confounding factors. We use the Vietnam Household Living Standard Survey (VHLSS) as our main dataset. We measure cultural norms at the district level (Vietnam consists of 630 districts in 61 provinces) by aggregation of individual data from the VHLSS.[9] The core regression model 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 potential confounding factors, and ϵ_i is a random error term. Our hypothesis postulates that β is positive with respect to the strength of the collectivist measure, i.e., the longer the time since annexation, the stronger the collectivist norms. For most of the regressions, we use the ordinary least squares (OLS) estimator. Because the sample contains nearly all districts of all provinces in Vietnam and the treatment is at the district level, we do not adjust standard errors for clustering, following the guideline from Abadie et al. (2017).[10] Instead, we use robust standard errors in all regressions.

4.2 Variables

The Individualism-Collectivism Trait

In the present paper, we follow the conventional definition of culture in economic research as "decision making heuristics or 'rules of thumb' that have evolved given our need to make decisions in complex and uncertain environments" (Nunn 2012, p. S109).[11]

[9] 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).

[10] In particular, Abadie et al. (2017) show that cluster adjustments for standard errors should only be performed if the data were collected by cluster sampling (e.g., first taking a subset of provinces, and then drawing a sample of districts from sampled provinces) or treatment occurs at a higher level of aggregation than the unit of observation.

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

Many observable outcomes have been used in the literature to capture the individualism-collectivism trait, such as family structure, marriage stability, inventiveness, or unusual 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 dimension, 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 respect, one should observe their members contributing 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 the modern day because labor is probably the most abundant resource in a developing country such 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.

[Table 1 is about here.]

The VHLSS covers all provinces in Vietnam with information about geographical location at the district level, which enables us to conduct a detailed empirical analysis. We employ the VHLSS 2002, covers almost 30000 households in 607 districts (roughly 50 households per district) and is the only round that contains detailed information about labor contribution to public goods production. Table 1 shows that, in 2001, around 26% percent of households contributed labor to public goods production, whereas the average

unchanged from generation to generation."

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 missing data for non-surveyed districts on the map, we calculate the average values of each of the 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 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 1 show that the annexations took place between 2.33 to 10.22 centuries before the terminal year of 1990.

To construct the time since annexation to historical Vietnam for each modern district, two dimensions are needed: (i) the district's 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* (2012) and *Dai Nam Thuc Luc* (2002), recording events from the beginning to 1675 and in the 1558-1888 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 as having been 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

To cope with the problem of endogeneity, we identify a set of potential confounding factors, i.e., factors that influence both the time since annexation to historical Vietnam

and the strength of collectivism. In this context, a necessary condition for a variable to be a confounding factor is that it must have existed before the annexation to historical Vietnam. Variables realized after that might be the results of the annexation, and hence are bad controls (Angrist and Pischke 2009). First, agricultural suitability might have both attracted historical Vietnam to conquer a region and promoted the development of collectivism. Thus, we control for natural land productivity, which is shown to influence the incentive to cooperate in the traditional agricultural economy (Litina 2016). Second, geographical conditions 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 distance to the coast, elevation, and ruggedness in our analysis. Descriptive statistics of all variables can be found in table 1.

Natural land productivity is measured by caloric suitability 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 land productivity, the production conditions are set at a low level of inputs and rain-fed agriculture based on agro-climatic conditions, which are unaffected by human intervention. Distance to the coast is measured by the shortest (bird-fly) distance to the coastal line. Elevation is taken from the Global 30 Arc-Second Elevation Dataset (GTOPO30) provided by the Earth Resources Observation and Science Center. The terrain ruggedness index was 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.[12] Thus, in the subsample of the annexed

[12]Theoretically, one could conquer an area in the Mekong River Delta by traveling either along the coast-

areas, the north-south geographical order can serve as a valid instrument for the time since an area was annexed to historical Vietnam.[13] We propose to use the distance from an annexed area to a northern reference point as a measure of the north-south geographical order. Quang Binh (figure 1), the first area that was annexed to historical Vietnam, is arbitrarily chosen as the northern reference point (the result is robust to other choices), 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 district borders are taken from the Global Administrative Unit Layers. The walking distance from Quang Binh to the farthest district in the south is roughly 1350 kilometers.

4.3 Baseline Results

To begin with, we regress the percentage of households contributing labor in a district on the time since a district was annexed to historical Vietnam, controlling for potential confounding factors as discussed above. Table 2 shows that the estimated coefficients of the time since annexation are positive and significant, whether or not 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. When all control variables are included, for example, a one century increase in the time since annexation is associated with an additional 1.7% of households contributing labor, which is roughly 5.5% of the mean value of the variable. The reduction in the magnitude of the estimated coefficient of the time since annexation when control variables are added indicates that these variables do confound the impact of the time since annexation on the prevalence of collectivism to some extent. The time since annexation accounts for almost 10% of the variation in the percentage of households contributing labor.

[Table 2 is about here.]

line in the east or over the mountainous band in the west separating Vietnam from Laos and Cambodia. Both strategies were infeasible given the logistical and transportation technologies in historical Vietnam. Indeed, we do not find any attempt to do so in the historical accounts.

[13] Two assumptions are involved: (i) the geographical location of a district is independent of the time since annexation and the strength of collectivism, and (ii) the north-south geographical order only affects the strength of collectivism through its effect on the time since annexation (Angrist and Pischke 2009).

The estimated coefficient of caloric suitability is negative and significant (column 2), suggesting 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 distance to the coast, elevation, and ruggedness are all significant and positive (columns 3 to 5), indicating that areas farther from the coastal line, more highly elevated and rugged have higher percentages of households contributing labor. When all control variables are added together, their estimated coefficients decrease substantially in magnitude, which is expected given that these variables are correlated (table A1 in the appendix). The estimated coefficient of caloric suitability is not different from zero. Together, these control variables account for almost 20% of the total variation in the percentage of households contributing labor.

Table 3 reports the results of similar regressions for the two other dependent variables measured at district level, i.e., the average number of persons per household making labor contributions (panel A) and the 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 or not all control variables are included. 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, i.e., around 6% of the respective mean values. The time since annexation accounts for approximately 10% of the total variations in both dependent variables, while control variables altogether account for another 20%. The estimated coefficients of the control variables are qualitatively similar to the results with respect to the percentage of households contributing labor. The only exception is the estimated coefficient of elevation with respect to the average number of labor days contributed per household, which is not different from zero when other control variables are added (column 6 of panel B).

[Table 3 is about here.]

4.4 Robustness Analysis

In this subsection, we tackle various challenges in taking the above findings as evidence for the selective migration hypothesis. First, as discussed earlier, there might be various characteristics in the annexed region that can explain its cultural differences along the

individualism-collectivism dimension compared to the initial region. The selective migration hypothesis differentiates itself by predicting a positive relationship between the time since annexation to historical Vietnam and the strength of collectivism, even within the subsample of the annexed areas. As can be seen in panel A of table 4, 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 an important driver behind the contemporary cultural differences across Vietnam.

[Table 4 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 Quang Binh as an instrumental variable for the time since annexation to historical Vietnam. Table 4 (panel B) reports the results of the instrumental variable estimation using the two-stage least squares (TSLS) estimator. The estimated coefficients of the time since annexation remain economically and statistically significant whether or not all control variables are added (all columns). In addition, the first-stage results ensure that the walking distance to Quang Binh is a strong predictor of the time since annexation, i.e., it has significant and negative estimated coefficients and large F statistics (full results are available upon request). 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, which is now south Vietnam, and hence influenced this region more than anywhere else in Vietnam (Taylor 2013, p. 484-523). Following the French colonization and the First Indochina War (1946-1954) was the American intervention in south Vietnam during the Second Indochina War (1954-1975), during which Vietnam was divided into two regions along the 17^{th} parallel, i.e., the communist north and the capitalist south, according to the Geneva Conference in 1954. In reality, however,

the communist north occupied and controlled a non-negligible area of the part of the country belonging to the capitalist south (Taylor 2013, p. 561-614). Nevetherless, the former Cochinchina was the homeland of the capitalist South, and hence was under its strongest control. Although this Western influence in south Vietnam might account for its cultural differences along the individualism-collectivism dimension compared to north Vietnam, the selective migration hypothesis differentiates itself by predicting a positive relationship between the time since annexation to historical Vietnam and the strength of collectivism, even within the subsample of areas belonging to the former Cochinchina. Panel A of table 5 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 across regions left by eight centuries of territorial expansion and selective migration in the remote past.

[Table 5 is about here.]

Finally, historical Vietnamese immigrants (the Kinh ethnicity) often inhabited 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 on a 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 those that happened in historical times. To examine this issue, we exclude from the estimation all districts in the highland areas, i.e., where the average elevations are above 500 meters (the results are robust to other values such as 400 and 600 meters).[14] 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 5 (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

[14]This definition of highland is taken from Wikipedia: https://en.wikipedia.org/wiki/Highland

autocorrelation. The corresponding results are respectively presented in tables A2 and A3 (appendix A) and show that the estimated coefficients of the time since annexation to historical Vietnam remain qualitatively intact.

4.5 Discussion

The above analysis has demonstrated that labor contribution to public goods production is both more prevalent and more 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 19^{th} and 20^{th} centuries, or important population movements in the modern day. These findings altogether lend support to the out-migration of individualistic people during eight centuries of territorial expansion of historical Vietnam as an important 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 in 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.

5 Experimental Data Analysis

5.1 Sample Selection

We apply a public goods game to test the selective migration hypothesis experimentally. For a general discussion on public goods experiments, see Zelmer (2003) and Chaudhuri (2011). A crucial aspect is the selection of experimental sites in such a way that we control for other potential explanations. First, we focus on the annexed areas to rule out differences in the frontier environment. Second, we restrict the selection 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 historically was the typical route of migration and settlement. 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 randomly choose one rural district in Thua Thien Hue province and one rural district in Ben Tre province; the the former is located more to the north and thus has a 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 not yet affected by living outside their communities, which potentially could make it harder to capture the local cultural norms.[15] The experimental design thus allows a direct test of the selective migration hypothesis by comparing the contribution levels to 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 stronger norms of cooperation, and hence on average to contribute at a higher level compared to subjects from Ben Tre (henceforth the "southern site"). Each rural district in Vietnam has three to five 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 its own norms, when we are aiming at measuring norms in the society in which they lived.

5.2 The Public Goods Experiment

We build our experimental design on the one-shot linear public goods experiment developed by Fischbacher, Gächter, and Fehr (2001).[16] We begin by describing the general

[15] This strategy of focusing on high school students has also been used earlier in the literature on public goods experiments when investigating cultural differences, e.g., Kocher, Martinsson, and Visser (2012).

[16] For 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-

features of a public goods experiment before discussing the specific features of the design in Fischbacher, Gächter, and Fehr (2001).

The basic idea of a public goods experiment is to create a social dilemma situation where there is a conflict between the social and private optima. In our setting, 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 of measuring norms of cooperation in the places where they reside. All subjects receive an endowment of 20 tokens and must decide simultaneously how much of their endowments to invest in a public good, and the residual is kept for themselves, which is labeled as a private good. The marginal per capita return ($MPCR$) from the public good is 0.5, which means that each token contributed to the public good by a group member 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. Thus, 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 good (i.e., what is kept as a private good), and (ii) the return from the public good. The payoff function for subject i is given by:

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

Each token earned in the experiment is exchanged for money at the exchange rate of 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 public goods experiment developed by Fischbacher, Gächter, and Fehr (2001) is that it is based on the strategy method. In their design, each subject

Palacio, and Wollbrant (2015) among others.

makes two types of contribution decisions to the public good: (i) unconditional contribution and (ii) conditional contribution. In the unconditional contribution decision, which is the standard public goods experiment described above, each subject states how much he or she would like to contribute to the public good from his or her endowment of 20 tokens. The additional feature of the design of Fischbacher, Gächter, and Fehr (2001) is the introduction of the contribution table in which subjects make contribution decisions conditional on the other group members' average contributions. 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. The final feature of the design of Fischbacher, Gächter, and Fehr (2001) is to ensure that all decisions, i.e., both unconditional and conditional contributions, are 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 all the decisions will be pay-off relevant, and hence has no incentive to choose anything other than the 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.[17]

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, i.e., how much they decided to contribute to the public good conditional on the average contribution of the other two group members for all integers in the range 0 to 20. 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

[17]For a discussion on using incentivized guesses in a public goods experiment to increase guess accuracy, see, e.g., Gächter and Renner (2010).

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 the 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 the 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".

At each school, we conducted the public goods experiment, including two sessions: the first one is to elicit unconditional contribution and the second one is to elicit conditional contribution. Subjects received written instructions for the experiment and the instructions were also read aloud.[18] Before the experiment began, 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 socioeconomic questions. Finally, subjects were called one at a time for payment done in private. 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 the socioeconomic questionnaire, we ask a number of basic questions including where their parents grew up. In accordance with our expectation, 98% of the subjects and 95% of their parents were born in the sampled districts, while the others were born in other districts in the sampled provinces.[19]

[18]Details on the experimental instructions 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 the first one. To eliminate potential spillover between experiments, the subjects were informed about the second experiment only after completing the first experiment.

[19]The following results are robust to the omission of subjects who were not born in the sample district, or whose parents were not born in the sample district. These results are available upon request.

5.3 Results

Table 6 shows that subjects from the northern site and southern site on average unconditionally contributed 7.50 tokens and 6.58 tokens respectively out of the endowment of 20, and the difference is statistically significant (p-value = 0.024, Mann-Whitney U test).[20] Thus, subjects from the northern site on average contribute higher than subjects from the southern site by 0.92 tokens, which is 13% of the mean level of contribution in the sample. 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 the 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). At the aggregate level, the results indicate conditional cooperative behavior. It is common that guesses about the average contribution by others are higher than own contribution levels because, on average, people are imperfect conditional cooperators, and there is also a fraction of free-riders.

[Table 6 is about here.]

The innovative part of the design developed by Fischbacher, Gächter, and Fehr (2001) is that it allows us to classify subjects into different contributor types. The lower panel of table 6 shows the distribution of types. By far, conditional cooperators are the most frequent type both in the northern site (52.17%) and in the southern site (54.04%), while the fractions of free riders are low (3.62% and 5.53% in the northern and the southern site, respectively). 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), and the distribution of types is similar with previous findings in the literature (Chaudhuri 2011). Furthermore, table 6 shows that, except for free riders, other types in the northern site on average have higher levels of unconditional contribution and guesses about the average contribution of other group members compared to their counterparts in the southern site. The largest contributor type is the conditional cooperators, and this type is considered the key group for contri-

[20] The contribution levels are similar to what has been found in the literature of public goods experiments (Zelmer 2003; Chaudhuri 2011).

butions to public goods. The average unconditional contributions are 7.29 and 6.39, in the northern and the southern site, respectively, with guesses of 8.19 and 7.88, respectively. This finding suggests that the north-south difference in contribution behaviors is driven by belief rather than the composition of contributor types.

We use regression models to examine the unconditional contribution behaviors further and the results are reported in table 7. 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 the regression alone (column 1). In the literature of public goods experiments, belief about the contribution of others has been shown to be an important determinant of contribution, i.e., a subject matches his or her own contribution level with the expected levels of others. In the next regression model, we also add belief about the average contribution of the other two group members and find that its estimated coefficient is positive and significant (column 2). In the same regression, the estimated coefficient of the northern site dummy is reduced substantially in magnitude, indicating that higher levels of contribution in the northern site are mainly driven by higher beliefs about the average contribution of the other two group members. These estimated coefficients remain positive and significant when the socioeconomic characteristics (gender, household size, and a wealth index) are also added to the regression (columns 3 to 6).[21] 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 an important driver behind these cultural differences.

[Table 7 is about here.]

6 Conclusion

The individualism-collectivism contrast has been found to be a powerful predictor of economic and democratic development in a large sample of countries (Gorodnichenko and Roland 2011, 2015, 2017). Thus, why some societies have become more collectivistic or

[21] The 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.

individualistic than others is a crucial question in understanding long-run comparative development. In the present paper, we propose and investigate the selective migration hypothesis, stating that cultural differences along the individualism-collectivism dimension are driven by the out-migration of individualistic people from collectivist societies to settle down in frontier areas, and that such patterns of historical migration are reflected even in the current distribution of cultural norms. We use the territorial expansion of historical Vietnam from the 11^{th} to the 18^{th} centuries as an ideal 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 of other subjects. Relying on various Vietnamese historical accounts, together with many robustness checks, we show that the southward out-migration of individualistic people during eight centuries of territorial expansion of historical Vietnam is an important driver 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

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey Wooldridge. 2017. "When Should You Adjust Standard Errors for Clustering". *NBER Working Paper* No. 24003.

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.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

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 Gebresilasse. 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* (Forthcoming).

Borcan, Oana, Ola Olsson, and Louis Putterman. 2018. "Transition to Agriculture and First State Presence: A Global Analysis". *Working Paper, University of Gothenburg.*

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. 2018. "The Historical State, Local Collective Action, and Economic Development in Vietnam". *Econometrica* (Forthcoming).

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

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.

Gächter, Simon, and Elke Renner. 2010. "The Effects of (Incentivized) Belief Elicitation in Public Goods Experiments". *Experimental Economics* 13 (3): 364–377.

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 Transfroming 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.

Hickey, Gerald Cannon. 1964. *Village in Vietnam*. New Haven: Yale University Press.

Hofstede, Geert. 2001. *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*. 2nd. London: Sage.

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

Johnson, Allen W., and Timothy Earle. 2000. *The Evolution of Human Societies: From Foraging Group to Agrarian State*. 2nd. Stanford: Stanford University Press.

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.

Lowes, Sara, Nathan Nunn, James A. Robinson, and Jonathan Weigel. 2017. "The Evolution of Culture and Institutions: Evidence from the Kuba Kingdom". *Econometrica* (forthcoming).

Luong, Hy Van. 1992. *Revolution in the Village: Tradition and Transformation in North Vietnam, 1925-1988*. Honolulu: University of Hawaii Press.

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. 1996a. *Nghien Cuu Dia Ba Trieu Nguyen: Binh Dinh*. Vol. I. Ho Chi Minh City: TP. Ho Chi Minh.

— . 1996b. *Nghien Cuu Dia Ba Trieu Nguyen: Binh Thuan*. Ho Chi Minh City: TP. Ho Chi Minh.

— . 2010. *Nghien Cuu Dia Ba Trieu Nguyen: Quang Nam*. Ho Chi Minh City: Dai Hoc Quoc Gia TP. Ho Chi Minh.

— . 1997. *Nghien Cuu Dia Ba Trieu Nguyen: Thua Thien*. Ho Chi Minh City: TP. Ho Chi Minh.

— . 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.

Rambo, A. Terry. 1973. *A Comparison of Peasant Social Systems of Northern and Southern Vietnam*. Carbondale: Center for Vietnamese Studies, Southern Illinois University.

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.

Scott, James C. 2017. *Against The Grain: A Deep History of the Earliest States*. New Haven: Yale University Press.

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.

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.

Turner, Frederick Jackson. 1920. *The Frontier in American History*. New York: Holt.

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

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

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

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

Wook, Choi Byung. 2004. *Southern Vietnam Under the Reign of Minh Mang (1820–1841)*. New York: Cornell Southeast Asia Program Publications.

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

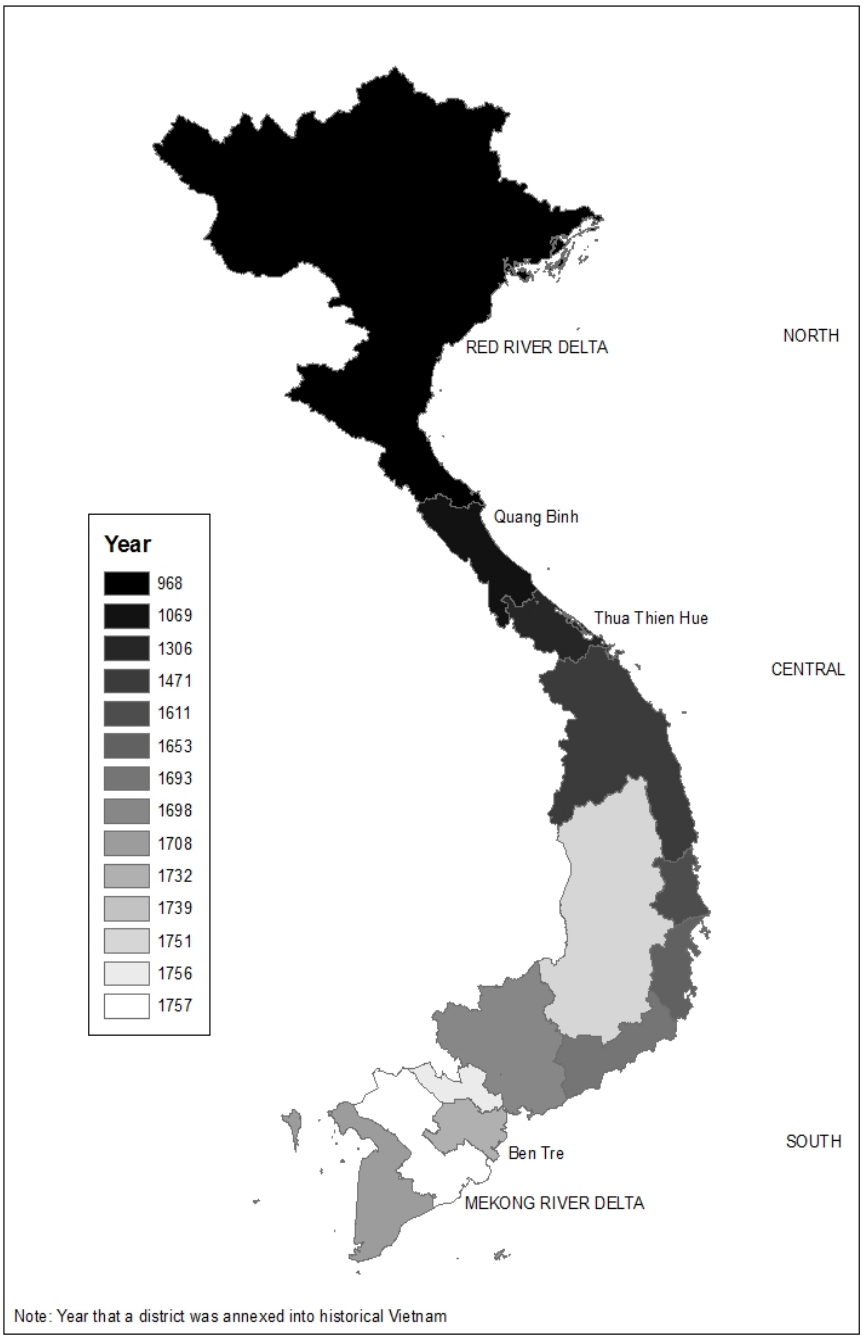
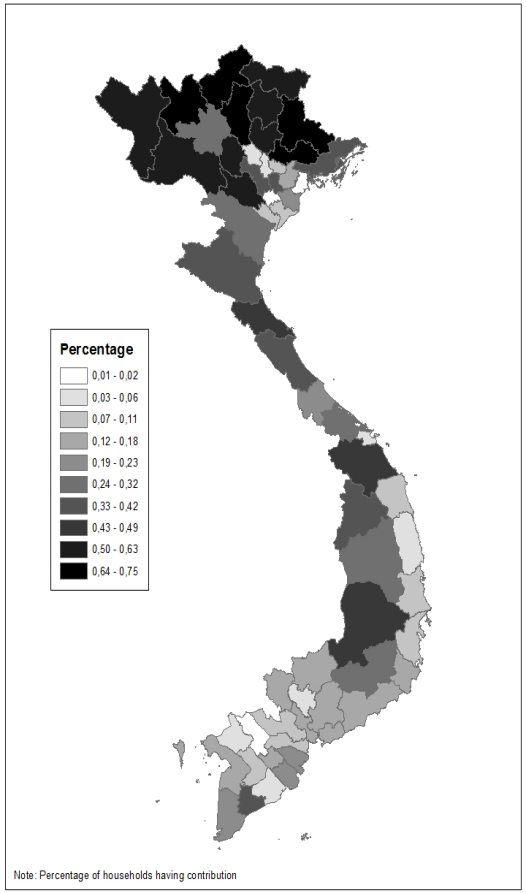
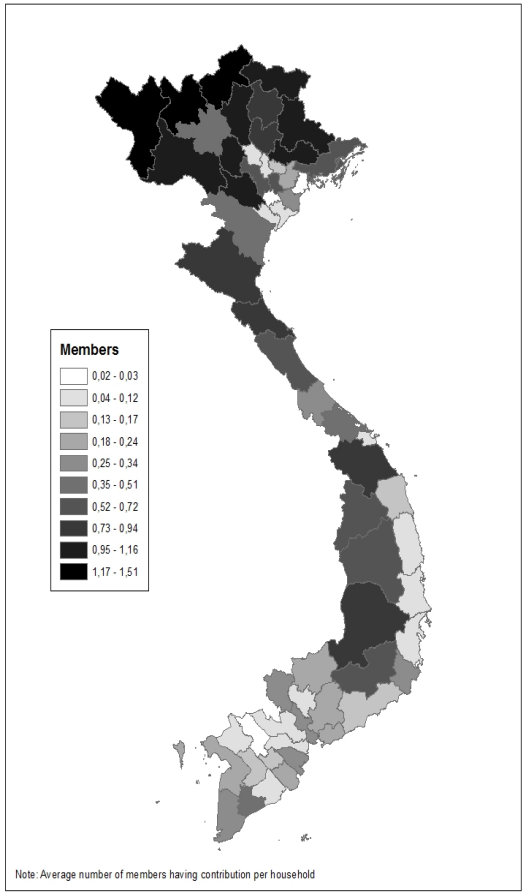
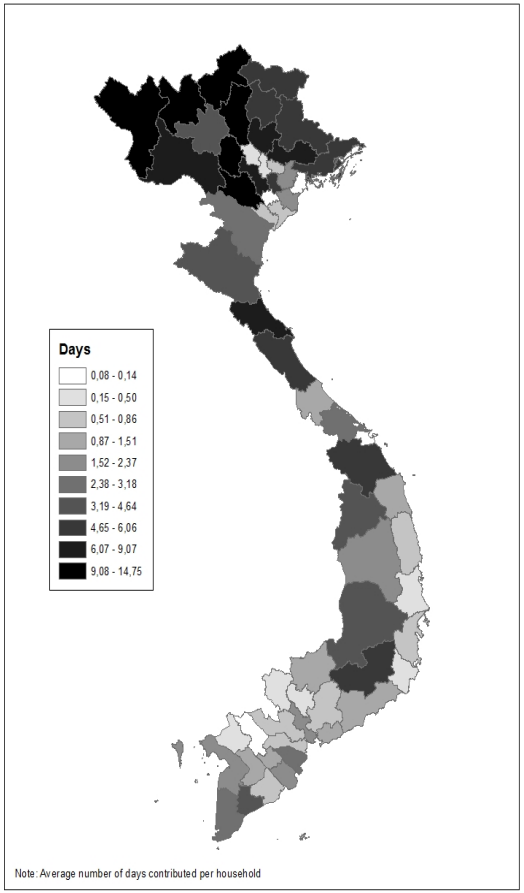


Figure 1. The Vietnamese Southern Advance

Source: Authors' construction.







A. Percentage of Households B. Number of Persons per Household C. Number of Days per Household

Figure 2. Labor Contribution to Public Goods Production

Source: VHLSS 2002.

Table 1. Variable Description

Variable	Description	Mean	SD	Min	Max	N	Source
Main Variables							
Having contributed	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
Control Variables							
Caloric suitability	Average yield (million kilo calories per km^2 per year)	25.08	2.50	15.88	31.74	607	Galor & Özak (2016)
Distance to coast	Shortest distance to the coastal line (100 km)	0.73	0.81	0	4.30	607	GAUL (FAO)
Elevation	Average height above sea level (km)	0.20	0.30	0.001	1.56	607	GTOPO30 (EROS)
Ruggedness	Average topographic heterogeneity (10000 km)	1.14	1.43	0	6.27	607	Nunn & Puga (2012)
Distance to Quang Binh	Walking distance to Quang Binh (100 km)	6.12	8.43	0	13.49	311	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. GAUL: Global Administrative Unit Layers. FAO: Food and Agriculture Organization, United Nations. GTOPO30: Global 30 Arc-Second Elevation Dataset. EROS: Center for Earth Resources Observation and Science, United States Geological Survey.

Table 2. The Prevalence of Collectivism

	Percentage of households contributing labor					
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.028***					0.017***
	(0.004)					(0.004)
Caloric suitability		-0.036***				0.003
		(0.005)				(0.007)
Distance to coast			0.184***			0.070**
			(0.016)			(0.030)
Elevation				0.483***		0.182*
				(0.045)		(0.097)
Ruggedness					0.108***	0.040**
					(0.009)	(0.019)
Constant	0.122***	1.213***	0.179***	0.215***	0.190***	-0.016
	(0.024)	(0.135)	(0.015)	(0.013)	(0.013)	(0.201)
Mean value of dep. var.	0.31	0.31	0.31	0.31	0.31	0.31
R^2	0.094	0.075	0.203	0.195	0.219	0.282
Observations	607	607	607	607	607	607

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

Table 3. The Strength of Collectivism

A.	Average number of persons making contributions per household					
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.056***					0.033***
	(0.007)					(0.007)
Caloric suitability		-0.071***				0.007
		(0.011)				(0.015)
Distance to coast			0.379***			0.130**
			(0.038)			(0.066)
Elevation				1.037***		0.418**
				(0.102)		(0.192)
Ruggedness					0.230***	0.089**
					(0.020)	(0.039)
Constant	0.173***	2.339***	0.274***	0.339***	0.288***	-0.138
	(0.044)	(0.280)	(0.030)	(0.024)	(0.023)	(0.415)
Mean value of dep. var.	0.55	0.55	0.55	0.55	0.55	0.55
R^2	0.094	0.076	0.222	0.230	0.255	0.316
Observations	607	607	607	607	607	607
B.	Average number of days contributed per household					
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.533***					0.255***
	(0.058)					(0.057)
Caloric suitability		-0.828***				-0.189
		(0.104)				(0.143)
Distance to coast			3.362***			1.302**
			(0.382)			(0.591)
Elevation				7.695***		1.640
				(1.034)		(1.681)
Ruggedness					1.840***	0.765**
					(0.195)	(0.323)
Constant	0.436	24.797***	1.595***	2.483***	1.946***	4.894
	(0.317)	(2.676)	(0.259)	(0.223)	(0.218)	(3.988)
Mean value of dep. var.	4.05	4.05	4.05	4.05	4.05	4.05
R^2	0.114	0.136	0.232	0.168	0.217	0.302
Observations	607	607	607	607	607	607

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

Table 4. 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***	0.028**	0.066***	0.059**	0.392***	0.330**
	(0.011)	(0.012)	(0.023)	(0.024)	(0.150)	(0.149)
Constant	0.121***	0.514	0.143*	1.291	0.872*	9.102
	(0.039)	(0.425)	(0.079)	(0.920)	(0.506)	(6.769)
R^2	0.030	0.191	0.043	0.227	0.032	0.121
B. IV (TSLS)	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.064***	0.035***	0.140***	0.076***	0.677***	0.314*
	(0.013)	(0.013)	(0.026)	(0.024)	(0.171)	(0.176)
Constant	-0.005	0.414	-0.117	1.068	-0.122	9.313
	(0.039)	(0.435)	(0.074)	(0.944)	(0.529)	(7.092)
Exogeneity (p-value)	0.000	0.347	0.000	0.272	0.005	0.884
First-stage F statistic	284	334	284	334	284	334
Mean value of dep. var.	0.22	0.22	0.37	0.37	2.24	2.24
Control variables	NO	YES	NO	YES	NO	YES
Observations	311	311	311	311	311	311

Note: Robust standard errors are in parentheses. Control variables include caloric suitability, distance to the coast, elevation, and ruggedness. All regressions only include districts in the annexed region. Walking distance to Quang Binh 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 5. 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***	0.199***	0.309***	0.345***	1.372**	1.657**
	(0.049)	(0.059)	(0.079)	(0.095)	(0.681)	(0.812)
Constant	-0.337***	0.247	-0.619***	0.487	-2.204	2.457
	(0.124)	(0.828)	(0.197)	(1.238)	(1.730)	(11.472)
Mean value of dep. var.	0.14	0.14	0.20	0.20	1.44	1.44
Control variables	NO	YES	NO	YES	NO	YES
R^2	0.058	0.069	0.071	0.083	0.019	0.040
Observations	169	169	169	169	169	169
B. Lowland areas	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.027***	0.020***	0.052***	0.037***	0.468***	0.260***
	(0.004)	(0.004)	(0.007)	(0.007)	(0.054)	(0.063)
Constant	0.094***	0.116	0.116***	0.327	0.221	8.429*
	(0.024)	(0.272)	(0.042)	(0.516)	(0.288)	(4.701)
Mean value of dep. var.	0.19	0.19	0.31	0.31	1.87	1.87
Control variables	NO	YES	NO	YES	NO	YES
R^2	0.097	0.256	0.104	0.265	0.123	0.235
Observations	478	478	478	478	478	478

Note: Robust standard errors are in parentheses. Control variables include caloric suitability, distance to the coast, 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 6. 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	8.25	100	6.58	7.60
		(3.92)	(3.46)		(4.07)	(4.11)
Conditional cooperators	57.12	7.29	8.19	54.04	6.39	7.88
		(3.20)	(3.72)		(3.35)	(4.07)
Free riders	3.62	0.60	6.20	5.53	2.69	6.54
		(0.55)	(4.55)		(3.88)	(3.55)
Hump-shaped cooperators	5.80	10.50	8.38	4.26	6.90	6.70
		(5.45)	(3.42)		(5.92)	(5.19)
Others	38.41	7.98	8.49	36.17	7.42	7.44
		(4.03)	(2.98)		(4.51)	(4.15)

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

Table 7. Unconditional Contribution: Regression Analysis

	Unconditional contribution					
	(1)	(2)	(3)	(4)	(5)	(6)
North	0.921**	0.728*	0.808**	0.686*	0.735*	0.789*
	(0.426)	(0.403)	(0.405)	(0.405)	(0.407)	(0.410)
Belief		0.297***	0.299***	0.290***	0.297***	0.294***
		(0.060)	(0.060)	(0.059)	(0.060)	(0.059)
Male			0.376			0.378
			(0.437)			(0.440)
Household size				0.118		0.131
				(0.130)		(0.137)
Wealth					-0.011	-0.046
					(0.125)	(0.133)
Constant	6.579***	4.324***	4.138***	3.877***	4.318***	3.618***
	(0.266)	(0.494)	(0.517)	(0.745)	(0.491)	(0.805)
R^2	0.012	0.093	0.095	0.095	0.093	0.098
Observations	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. Correlation Matrix

	(1)	(2)	(3)	(4)	(5)	(6)
(1) Time since annexation	1.00					
(2) Caloric suitability	-0.51	1.00				
(3) Distant to coast	0.36	-0.60	1.00			
(4) Elevation	0.11	-0.21	0.68	1.00		
(5) Ruggedness	0.29	-0.36	0.65	0.82	1.00	
(6) Distance to Quang Binh	-0.91	0.37	-0.24	-0.17	-0.31	1.00

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***	0.018***	0.056***	0.034***	0.526***	0.258***
	(0.004)	(0.004)	(0.007)	(0.007)	(0.059)	(0.059)
Constant	0.120***	-0.052	0.172***	-0.145	0.501	5.268
	(0.024)	(0.202)	(0.044)	(0.418)	(0.330)	(4.210)
Control variables	NO	YES	NO	YES	NO	YES
R^2	0.094	0.285	0.092	0.318	0.107	0.292
Observations	607	607	607	607	607	607
B. Annexed areas	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.028**	0.029**	0.062***	0.059**	0.376**	0.327**
	(0.011)	(0.012)	(0.023)	(0.023)	(0.157)	(0.156)
Constant	0.122***	0.516	0.153*	1.246	0.962*	9.115
	(0.039)	(0.421)	(0.079)	(0.911)	(0.533)	(6.746)
Control variables	NO	YES	NO	YES	NO	YES
R^2	0.028	0.207	0.038	0.239	0.028	0.135
Observations	311	311	311	311	311	311
C. Cochinchina areas	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.173***	0.189***	0.297***	0.325***	1.297*	1.510*
	(0.049)	(0.059)	(0.079)	(0.094)	(0.667)	(0.779)
Constant	-0.320**	-0.567	-0.585***	-0.972	-1.955	-5.797
	(0.125)	(0.812)	(0.199)	(1.174)	(1.720)	(11.072)
Control variables	NO	YES	NO	YES	NO	YES
R^2	0.055	0.064	0.065	0.074	0.016	0.035
Observations	169	169	169	169	169	169
D. Lowland areas	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.027***	0.021***	0.052***	0.039***	0.463***	0.264***
	(0.004)	(0.004)	(0.007)	(0.007)	(0.055)	(0.065)
Constant	0.091***	0.033	0.116***	0.167	0.246	8.393*
	(0.024)	(0.273)	(0.043)	(0.523)	(0.301)	(5.013)
Control variables	NO	YES	NO	YES	NO	YES
R^2	0.098	0.250	0.101	0.248	0.118	0.216
Observations	478	478	478	478	478	478

Note: OLS estimator, robust standard errors are in parentheses. Control variables include caloric suitability, distance to the coast, 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***	0.017***	0.056***	0.033***	0.533***	0.255***
	(0.004)	(0.004)	(0.007)	(0.007)	(0.060)	(0.065)
Constant	0.122***	-0.016	0.173***	-0.138	0.436	4.894
	(0.027)	(0.184)	(0.054)	(0.354)	(0.462)	(3.102)
Control variables	NO	YES	NO	YES	NO	YES
R^2	0.525	0.623	0.473	0.602	0.417	0.540
Observations	607	607	607	607	607	607
B. Annexed areas	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.028***	0.028***	0.066***	0.059***	0.392***	0.330**
	(0.009)	(0.010)	(0.018)	(0.020)	(0.122)	(0.144)
Constant	0.121***	0.514	0.143**	1.291	0.872*	9.102
	(0.036)	(0.444)	(0.068)	(0.834)	(0.472)	(6.105)
Control variables	NO	YES	NO	YES	NO	YES
R^2	0.410	0.508	0.361	0.484	0.301	0.365
Observations	311	311	311	311	311	311
C. Cochinchina areas	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.179***	0.199***	0.309***	0.345***	1.372*	1.657*
	(0.056)	(0.062)	(0.086)	(0.096)	(0.770)	(0.851)
Constant	-0.337**	-0.438	-0.619***	-0.499	-2.204	-3.188
	(0.149)	(0.806)	(0.230)	(1.247)	(2.055)	(11.078)
Control variables	NO	YES	NO	YES	NO	YES
R^2	0.398	0.405	0.377	0.384	0.264	0.279
Observations	169	169	169	169	169	169
D. Lowland areas	Percentage		No. of persons		No. of days	
	(1)	(2)	(3)	(4)	(5)	(6)
Time since annexation	0.027***	0.020***	0.052***	0.037***	0.468***	0.260***
	(0.004)	(0.004)	(0.007)	(0.008)	(0.057)	(0.064)
Constant	0.094***	0.116	0.116**	0.327	0.221	8.429**
	(0.029)	(0.260)	(0.053)	(0.480)	(0.435)	(4.049)
Control variables	NO	YES	NO	YES	NO	YES
R^2	0.498	0.587	0.459	0.556	0.418	0.492
Observations	478	478	478	478	478	478

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 caloric suitability, distance to the coast, 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* (from 204 BCE to 1675) and *Dai Nam Thuc Luc* (from 1558 to 1888). 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, by which time the border of Vietnam was established as it is today.

The Dinh Dynasty

In 938, 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 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, as having been annexed in 968.

The Ly Dynasty

In 1069, Ly Nhat Ton attacked the former Champa Kingdom to retaliate against a territorial intrusion (*Toan Thu* p. 197). After being defeated and captured, the Champa King ceded 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 notes 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) as having been annexed in 1069.

The Tran Dynasty

In 1306, 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 as having been annexed in 1306.

The Le Dynasty

In 1471, 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.[22] 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 comprised the 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, in which modern Kon Tum province can be said to be included (Dao 2005, p. 203). We code all districts in the modern provinces of Quang Nam, Quang Ngai, Binh Dinh, and Kon Tum as having been annexed in 1471.

The Nguyen Lords

In 1611, 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 as having been annexed in 1611.

In 1653, 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) as having been annexed in 1653.

[22] Under the Ho Dynasty, historical Vietnam already controlled these regions from 1402 to 1407 (*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, King Ba Tranh of the former Champa Kingdom raided the border of historical Vietnam (*Thuc Luc* p. 106). In 1693, 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 as having been annexed in 1693.

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, 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 allow Vietnamese migrants to move in and exploit the Khmer land in the far south of Vietnam (*Thuc Luc* p. 72). In 1698, 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 described as corresponding to Bien Hoa and Gia Dinh provinces under the Nguyen Dynasty, i.e., the time when 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, this new territory basically corresponds to the south-east region of modern Vietnam (Nguyen 1994, p. 64). We code all districts in the modern provinces of Binh Phuoc, Dong Nai, Ba Ria-Vung Tau, Ho Chi Minh City, Binh Duong, and Tay Ninh as having been annexed in 1698.

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, Mo Jiu ceded 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 the modern provinces of Kien Giang, Ca Mau, and Bac Lieu (Phan et al. 2011, p. 448). We code all districts in these provinces as having been annexed in 1708.

In 1731, some inhabitants in the former Khmer Empire raided the south of historical Vietnam. In 1732, 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 the modern provinces of Vinh Long, Ben Tre, and Tien Giang (Phan et al. 2011, p. 448). We code all districts in these provinces as having been annexed in 1732.

In 1755, the former Khmer Empire attacked historical Vietnam. Following successful retaliation by historical Vietnam, King Ang Sngoun of the former Khmer Empire ceded the former Tam Bon and Loi Lap in 1756 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 as having been annexed in 1756.

In 1756, King Ang Sngoun of the former Khmer Empire died, which ignited a fight for the throne between different royal forces. After one side asked for help in 1757, 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 the 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 as having been annexed in 1757.

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