

Confucian Literati and Long-run Development in Northern Vietnam

Meng Liu* and Tomoki Fujii†

Abstract

Following Max Weber's thesis, studies have suggested that Confucianism could impede economic growth despite its positive effects. In this paper, we revisit the impact of Confucianism on long-run development in a less explored context – northern Vietnam. Using the variation in Confucian literati across 217 historical districts between the Primitive Le and Nguyen dynasties (1426-1919), we find that districts with a greater exposure to Confucianism have experienced better economic outcomes over the past century. The result is robust to controlling for a battery of confounders and using the distance to exogenously located hermit scholars as an instrument. We show that the impact of Confucianism can be attributed to the persistence of a culture of respect for education and norms of collective action, which facilitated human and social capital accumulation, public goods provision, and economic transition.

JEL classification: N0, O12, Z10.

Keywords: Confucianism; long-run development; Vietnam.

*School of Economics, Singapore Management University, 90 Stamford Rd, Singapore 178903. E-mail: meng.liu.2018@phdecons.smu.edu.sg, Phone: +65 98857987.

†School of Economics, Singapore Management University, 90 Stamford Rd, Singapore 178903. E-mail: tfujii@smu.edu.sg, Phone: +65 68280279, Fax: +65 68280833.

1. Introduction

Over the past few decades, East and Southeast Asian economies such as China, Singapore, South Korea, and Vietnam have achieved astounding rates of economic growth.¹ One explanation for this Asian miracle is the Confucian root of these Asian societies, as most prominently supported by Lee Kuan Yew, the former Prime Minister of Singapore (Gong and Jang, 1998; Hofstede and Bond, 1988; Kahn, 1979; Lee, 2013; Pezzutto, 2019; Redding, 1990; Tu, 1996, 2000). This starkly contrasts with Max Weber’s contention that the traditionalism and conservativeness of Confucian values could impede the economic growth (Weber, 1951). As with the two opposing arguments, recent empirical studies have also suggested various channels through which Confucian values and institutions could either facilitate or hinder development (Chen et al., 2020; Chen et al., 2021; Kung and Ma, 2014; Zhang, 2020). However, few studies have systematically investigated the effects of Confucianism on economic development over the long run.

In this paper, we use Northern Vietnam as a laboratory to examine the impact of Confucianism on long-run economic development. Northern Vietnam held Confucianism as the state ideology and adopted the Confucian civil exam as the primary bureaucrat selection institution for nearly five hundred years until the early twentieth century (see section 2 for a discussion). This long history left behind detailed official records of Confucian literati for empirical analyses.

Combining the list of Tiensi (individuals who passed the national level Confucian civil exam) awarded during 1426-1919 for 217 historical districts with the proxies for modern economic development from various sources, we find that areas with greater exposure to Confucianism have experienced better economic outcomes in the modern period. Specifically, the baseline ordinary least squares (OLS) estimates show that a one percentage point increase in the density of Tiensi is associated with 0.097 percentage point increment in colonial road density in 1929, 0.260 in degree of urbanization in 1975 and 0.258 in 1990-2015, 0.194 in average nighttime light intensity in 1992-2013, and 0.029 in average annual household consumption per capita in 2010-2018. This positive association is robust to controlling for geographic correlates, initial endowments, other ideologies, and French colonization, all with province-fixed effects.

This finding, however, is likely to be biased by measurement errors and omitted variables. For instance, Tiensi density might fail to capture the influence of other Confucian scholars who did not take the civil exam, resulting in underestimates. Or the areas with better village institutions might have produced more Tiensi and enjoyed better economic outcomes, leading to an upward bias. To address this concern, we instrument Tiensi density with the shortest distance from the border of the district to its nearest hermit Confucian scholar. We identify hermit Confucian scholars as individuals who were renowned for their expertise in Confucianism, nurtured many students, but refused to serve as government officials. This instrument is relevant for Tiensi density since historical

¹World Bank data on the annual growth rate of real GDP per capita show that during 1980-2010, the four countries have an average annual growth rate of 8.911%, 4.482%, 6.068%, and 4.943%, respectively. This is in contrast to a growth rate of 1.735% for the European Union, and 1.722% for North America, and other Southeast Asian countries who were less influenced by Confucianism such as the Philippines (0.945%), Cambodia (3.283%), and Indonesia (3.359%).

accounts state that their expertise attracted a lot of followers from other areas and that some of the followers actually obtained the Tiensi degree. Considering the lack of private Confucian academies and accesses to Confucian book, these hermit scholars could have been one of the primary forces in driving the regional variation in Tiensi density (Chen, 2015; Nguyen, 2020; Woodside, 2002).

The instrumental exogeneity premises on the fact that the emergence, activities, and locations of hermit scholars did not depend on local economic conditions. We justify the presumption in the following ways. First, we provide anecdotal evidence that the decisions of the scholars to be hermits mainly depended on their passion for Confucian learning and indifference to fame and power and that they did not intentionally choose where to retreat. Second, regression results show that our IV is not correlated with other confounders, such as state capacity, the presence of communication infrastructure, and access to Confucian books, after controlling for baseline correlates and province-fixed effects. Furthermore, we conduct a zero-first-stage test to show that the effects of our IV on economic development are unlikely to be mediated by variables other than Tiensi density. The instrumented results are consistent with the baseline results that Confucianism has positive effects on economic development, while the magnitude of the estimates is on average 3.26 times bigger. The instrumented results are also robust to the imperfect satisfaction of exclusion restrictions, as evidenced by the estimates generated by applying Conley et al. (2012)'s method.

To explain the positive association between Confucian literati and economic development, we explore multiple channels suggested in the current literature and historical accounts. First, the long pursuit of academic success in the civil exam and the high social status held by Tiensi degree holders led to a culture of respect for education (Chen et al., 2020; Tran, 1971). This could encourage investment in education and result in better human capital outcomes, which are vital for modern development (Gennaioli et al., 2012). To establish this channel, we rely on the description of local customs in the historical chronicle of the Nguyen dynasty, the Vietnam House Living Standards Survey (VHLSS) in 2010-2018, and the Population Census in 2009. Both OLS and IV estimates indicate that districts with a higher density of Tiensi are more likely to have an expression for respect for literature or literati in the description of local customs, and have a larger average share of household expenditure on education, a lower rate of illiterate population, a longer average years of schooling, and a higher proportion of individuals with the secondary and above degree. This finding suggests that human capital accumulation is a possible channel.

Secondly, Confucianism in Vietnam is characterized by strong family ties and collaborations among village members (Haines, 1984; Nguyen, 1998). The network, shared norms, trust and cooperation among family and community members constitute social capital that enable public goods provision (Putnam, 2000). Using the VHLSS data on surnames and Economic Census data on the locations of communal (village) houses and non-government organizations (NGOs), we find districts with a higher Tiensi density tend to have more social capital, as proxied by a higher level of surname similarity within districts, communal house density, and NGO density. In additions, these districts also tend to have a higher density of public facilities for education and training, health treatment, and cultural and sports activities.

Lastly, the clan and community network, and code of conduct governed by Confucian ethics could form supportive business environments that are conducive to the development of the non-state-owned and non-agricultural sectors (Peng, 2004; Zhang, 2020). This is especially helpful in Vietnam, as the development of formal legal and financial institutions is not commensurate with economic growth. To test this channel, we calculate the density of registered individual business and the share of individuals employed in different sectors and industries using the economic and population census data. We find that districts with more Tiensi have a higher density of registered individual business, and a larger share of individuals employed in private and foreign sectors, and secondary and tertiary industries. By contrast, having more Tiensi has no effects on public sector employment and a negative effect on primary industry employment. All these suggest that Confucian values could contribute to development by creating better environments for non-state-owned and non-agricultural sector development.

This paper contributes to several strands of literature. First, it adds to the work which identifies culture and historical institutions as important determinants of economic growth and development (e.g. Acemoglu et al., 2001, 2002; Banerjee and Iyer, 2005; Becker and Woessmann, 2009; Michalopoulos and Papaioannou, 2013). In particular, it provides a new perspective to the debate about whether Confucianism could facilitate or hinder development by establishing a causal link and identifying potential channels of influence. Second, it serves as a logical extension of the emerging studies on Confucian values and institutions in China (Bai, 2019; Bai and Jia, 2016; Chen et al., 2020; Chen et al., 2021; Kung and Ma, 2014; Zhang, 2020). Given the similar civil exam system and development trajectory, the results generated in northern Vietnam can be comparable to that of China, which could help gauge how Confucianism affects a country as a whole. Third, it provides an alternative explanation for the cross-district differences in the level of development within Vietnam. This complements the previous findings on the role of the historical institution of village governance, foreign intervention, the Vietnam war, and land property rights played in the development of Vietnam (Dell et al., 2018; Dell and Querubin, 2018; Ho, 2021).

The remainder of the paper proceeds as follows. In section 2, we introduce the history and features of Confucian values and institutions in Vietnam. In section 3, we explain the sample selection, data sources and variable construction. The baseline results are discussed in section 4, while in section 5 we address the endogeneity issues using an instrumental variable approach. Section 6 tests the potential channels. Finally, section 7 concludes the study.

2. Research Background

In this section, we first describe the how Confucianism emerged as the state ideology in Vietnam and the policies employed by the central state to spread Confucian ethics. Then we move to introduce the Confucian civil exam system. Finally, we elaborate on two cultural traits associated with Confucianism.

2.1 *Confucianism as State Ideology*

Ever since conquered by the Han China (206 BC - 211 CE) in 111 BC, northern Vietnam (called Nam Viet back then) was under the Chinese rule until 938 CE. During this long rule, Confucian moralities, especially those governing the familial social relation, were introduced and practiced since 43 CE, which essentially transformed the Nam Viet from a “south-sea civilization” society to a member of Eastern Asian culture (Taylor, 1983). However, during the whole Chinese rule and the following Ly and Tran dynasties (1009 CE - 1400 CE), Buddhism, instead of Confucianism, was the most dominant ideology among the Viet people.

The watershed event happened in the fifteenth century when Northern Vietnam was occupied by Ming dynasty of China, who tried to “civilize” Viet people with Confucian norms. Even though the Ming Occupation only lasted less than twenty years (1407 CE - 1426 CE), it convinced the Vietnamese ruler that Confucianism was a more suitable ideology for maintaining a centralized state and containing the power of military oligarchies (Whitmore, 1977). Starting from the emperor Le Thanh Tong (1460 CE - 1497 CE) in the Primitive Le dynasty (1427 CE - 1527 CE), Confucianism was promoted as the state ideology through the stipulation in legal code and national policy, school education, erecting Confucian Temples and dismantling Buddhist temples, and most importantly the adopting Confucian civil exam to select bureaucrats (see the discussion on the civil exam in section 2.2). For instance, the penal Le Code formalized the Confucian social relationships, in which the subject was subordinate to the ruler, wife was subordinate to husband, and child was subordinate to parent (i.e. sangang, the three principles). Articles on code of good conduct governed by Confucian moralities were published, in which harmony, filial piety and loyalty were greatly emphasized. The whole population were ordered to learn and practice them. Examples of such code include the Twenty-Four Articles in the Le dynasty, and the Ten Articles of Good Conduct in the Nguyen dynasty (Woodside, 1971).

Although rivaled by Buddhism and thoughts of military oligarchies, Confucianism remained as the state ideology for the imperial periods between the Le and Nguyen dynasties (1427 CE - 1945 CE) until the French colonisation motivated the modernization of ideologies in the early twentieth century.

2.2 *Civil Exam*

First held in 1075 CE during the Ly dynasty, the civil exam in historical Vietnam was organized in a less regulated manner and did not exclusively test Confucianism between the eleventh and fourteenth centuries.² It was until the occupation of the Ming dynasty of China (1407-1426) that historical Vietnam came to direct contact with the fully institutionalized Confucian civil exam system.³ After overthrowing the Ming, the Viet people established the Primitive Le dynasty and

²Buddhism and Taoism were also tested back then.

³Northern Vietnam (called Jiaozhi) was treated as a province of Ming China during the occupation. According to *Veritable Records of the Ming Dynasty*, it was allocated a quota of ten for the provincial-level civil exam, the same as the Yunnan province. During the twenty years of occupation, 161 students were selected from Jiaozhi to study in the National Academy in the Ming capital, Nanjing (Whitmore, 1977). These students usually returned Jiaozhi to serve

institutionalized the Confucian civil exam by catering the Chinese model to its own needs. Ever since then, the civil exam remained the primary institution for recruiting government officials until it was abolished in 1919 in the Nguyen dynasty.

The civil exam intended to select individuals based on their abilities to memorize, understand, and apply Confucian knowledge. As summarized by Chen (2015), a typical exam contained four tests. The first one required candidates to dictate the content of *Four Books and Five Classics*. The second and third tests were to compose different types of official documents, poems, and verses. The last one asked candidates to write a long essay to discuss how they would address economic, social, political, military, or demographic issues by combining Confucian thoughts with the history and realities of Vietnam.

Like the Chinese model, the civil exam in Vietnam also adopted a bottom-up selection process, though it was only administered at the provincial and national levels. The provincial level exam was open to qualified students of government schools, and any individuals preselected from the commune or district exam.⁴ Passing the first three tests of the provincial exam led to the award of the *tu tai* degree (equivalent to *Xiucai* in the Chinese system) and the qualification for the fourth test. Among the *tu tai*, about 10% could pass the fourth test, which granted them the *cu nhan* degree (equivalent to *juren* in the Chinese system) and the eligibility for the national level exam (Chen, 2015).⁵ *Cu nhan* would compete with the students selected from the National Academy in the national level exam. Successful candidates in the national exam would be awarded the *tiensi* degree, which was equivalent to *jinshi* in Ming-Qing China. The passing rate at this stage was 0.1%-7% (Chen, 2015).

Success in the civil exam could lead to substantial improvement in economic and social status. Once became a *tu tai*, a commoner could enter the gentry class and be qualified to teach at local schools and advise on village affairs (Whitmore, 1997). A *cu nhan* degree often guaranteed a position at the district government which was responsible for the population registration, land management, and agricultural production of 30-90 villages. *Tiensis* degree holders were assigned to higher positions in the imperial courts and received private land as compensations. In addition to the material rewards, the imperial court would also establish *tiensis* as role models by carving their names into the stelae erected in the Temple of Literature. Family and community members regarded the *tiensis* degree as a great honor that should be recorded in genealogies and local gazetteers. All these rewards motivate aspiring individuals, especially those with a commoner background to pursue success in the civil exam. Such pursuits could have not just made Confucian values well-ingrained in the Vietnamese society, but also led to the formation of cultural traits that continue to influence economic development today. We discuss the cultural traits in the following subsections.

as local officials, which allowed them to implement and spread the knowledge of Confucian thoughts and institutions acquired in China.

⁴Individuals could also enter the provincial level exam by making pecuniary donations, but they could neither gain any title nor proceed to the national level exam.

⁵The estimates are provided by Chen (2015). This 10% passing rate was slightly higher than the 6% given by Chen et al. (2020).

2.3 *Culture of Respect for Education*

The Confucian civil exam gave rise to a new social class, literati (or *nho* in Vietnamese), who identified themselves, particularly by their perspectives acquired from texts (Taylor, 2002; Whitmore, 1997). The literati were seen as the highest social class for a commoner, which was better respected than peasants, artisans, and merchants. For most commoners, the only way to become a literatus was to succeed in the civil exam, which could not be achieved without a great commitment to learning and an emphasis on education. Over time, this led to a culture of respect for education.

Indeed, Cooke (1994) states that even the poorest families in the dynasty of Vietnam would make every effort to provide for Confucian education. Folktales, family rules, and books that stress the importance of education were written and transmitted through generations of the Viet people. For example, the *Doc Thu Cach Ngon (Mottoes for Study)* included a saying that “One could not establish themselves as a person without learning”. Even today, the Vietnamese still attach great importance to education, as evidenced by their high ranking in worrying about not being able to give good education to their children in the World Values Survey and the preserving of the culture among the Viet people who moved to other countries (Caplan et al., 1991).

This respect for education did not just contribute to the learning of Confucianism and success in the civil exam, but also to the accumulation of general knowledge and human capital. Confucian scholars in historical Vietnam were also well-established in other fields like literature, history, cartography, and printing technology (Chen, 2015). During the French Colonization, it was those Confucianists who promoted the modernization of education. In more recent periods, Vietnamese ranked high in the Programme of International Student Assessment, in spite of being a developing country.

2.4 *Collective Actions among Family and Village Members*

Another cultural trait associated with the rise of Confucian literati is the collective actions of family and community members to support each other. The most prominent example of such collective actions is probably the setup and management of agricultural land for education in the villages. The land was donated by villagers to generate profits to hire teachers, reward students, or build schools (Adams and Hancock, 1970; Li, 2011; Nguyen, 1998; Nguyen, 2020). The farming of the land also relied on the voluntary labour contribution of students’ parents and family members.

When the Confucian students passed the exam and became officials, they often returned the favour by helping other family and community members to pass the exam. Moreover, Confucian literati were still willing to transfer resources to their hometowns, even when their success did not depend on their community (Do et al., 2017). For instance, Luong Nhu Hoc, a mandarin of the Le dynasty, passed the woodblock printing technology which he acquired in China to two villages in his hometown. These two villages later became the centres of commercial woodblock printing. Nearly every woodblock printing house in the country was run by people from the two villages. Gleaning from the examples, we can imply that the collective action and reciprocity between family and community members constitute social capital that could facilitate public goods provision and

commercial activities.

3. Data

To examine the effects of Confucianism on long-term economic development, we compile a cross-sectional dataset from various historical and contemporary sources, including official chronicles that have not been used in other economic studies. In this section, we describe in detail the sample selection, data source, and variable construction of our analyses.

3.1 *The Sample*

The sample for our baseline and IV analyses consists of 217 historical districts from 19 provinces in the Nguyen dynasty. It covers the area from the most northern provinces bordering China to the Phu Loc district of Thua Thien Hue province of Vietnam today (see Figure A1 in Appendix).⁶ The choice of this sample is primarily determined by data availability. Our analysis requires historical data at the district level, especially the data on local customs. To our best knowledge, *Dong Khanh Dia Du Chi* (*the Descriptive Geography of the Emperor Dong Khanh, hereafter DKDDC*) is the only official chronicle that provides such information. Unfortunately, it does not cover districts further south.

Nevertheless, this sample still provides a good setting for our study. The northern part of our sample region was the homeland of Dai Viet, in which individuals started observing Confucian traditions as early as 43 CE.⁷ Although annexed to Dai Viet only much later, the remaining districts in our sample had been exposed to Confucianism for at least a century longer than those excluded. Consequently, more than 95% of Tiensi holders listed in Ngo (2006) originated in our sample region, despite the region representing only around 50% of Vietnam’s land area. As such, this sample allows us to observe the persistence of Confucianism in Vietnam and provides considerable cross-district variations for examining its effects.

In addition, our sample districts are relatively similar in terms of their experience of the French colonization, the Vietnam War, and communism, which could help mitigate the confounding impact of these experiences. The region belonged to the French Tonkin and Annam Protectorates (later unified into Indochina) in the colonial period, which was subject to a weaker penetration of French culture, values, and political and legal practice than the societies further south (Brocheux and Hémery, 2011). 95% of our sample districts were located north of the 17 Parallel. They were supported by the Soviet Union and China during the Vietnam War and led by the Vietnamese Communist Party since the August Revolution in 1945.⁸

⁶The current locations of these historical districts are retrieved by following the work of Vietnamese historian Dao (1964). The details on retrieving procedure are described in section A1 in Appendix.

⁷The territory of the initial Dai Viet covers the area ranging from the contemporary provinces bordering China to the north border of Quang Binh province. From 111 BC to 936 CE, this area was under the rule of the imperial dynasties of China. Initially, the area was administered by Viet feudal lords, who practiced matriarchal traditions. Starting from 43 CE, the Han dynasty of China sent civil servants to govern the area and instituted an assimilation policy, which required the population to observe Confucian morality (see Volume 3 of *Dai Viet Su Ky Thoan Tu*).

⁸The remaining districts were within the U.S. Military Corps Region I. This region was commanded by the U.S. Marine Corps, which was less aggressive with airstrikes than the U.S. Army (Dell and Querubin, 2018). In a robustness

3.2 Key Independent Variable: *Tiensi* Density

We use the density of *Tiensi* degree holders in a historical district to proxy the strength of Confucianism. The selection of this proxy is based on the presumption that places with stronger Confucian traditions were more likely to produce successful candidates in civil exams, who would, in return, reinforce local Confucianism-learning and -spreading after they gained the prestigious *Tiensi* title (Chen, 2015).

To calculate the variable, we first count the total number of *Tiensi* degree holders in a district in 1426 CE – 1919 CE (spanning the Primitive Le, Mac, Revival Le, and Nguyen dynasties), and then normalize it with the average population (in thousand) in 1400 CE – 1900 CE.⁹ The period of analysis is limited to 1426 CE – 1919 CE, since civil exam was held in a different manner prior to 1426 CE and was abolished in 1919 CE (see section for a discussion). To correct for the skewness and preserve zero entries, we add 0.01 to the normalized figure and then raise it to the natural log.

The data on the list of *Tiensi* degree holders come from Ngo (2006). They are compiled from the inscriptions on the stelae in the Temple of Literature in Hanoi and Hue, and documents stored in local commune houses. Each scholar on the list is given a short biography that provides information on their names, place of birth, year of taking civil exam, and career path. In total, Ngo (2006) records 2,889 *Tiensi* who sat for 183 exams in 1017 CE – 1919 CE. To our knowledge, this is the most comprehensive record for *Tiensi* in Vietnam. Narrowing down to our sample region and period of analysis, we are able to georeference the locations of 2,739 *Tiensi*. An average district has 1.343 *Tiensi* per thousand population. The spatial distribution of *Tiensi* density is shown in Figure 1a.

3.3 Outcome Variables: Proxies for Economic Development

We use several measures to proxy the level of economic development over different time periods in Vietnam. The construction and data source of these measures are discussed in this subsection.

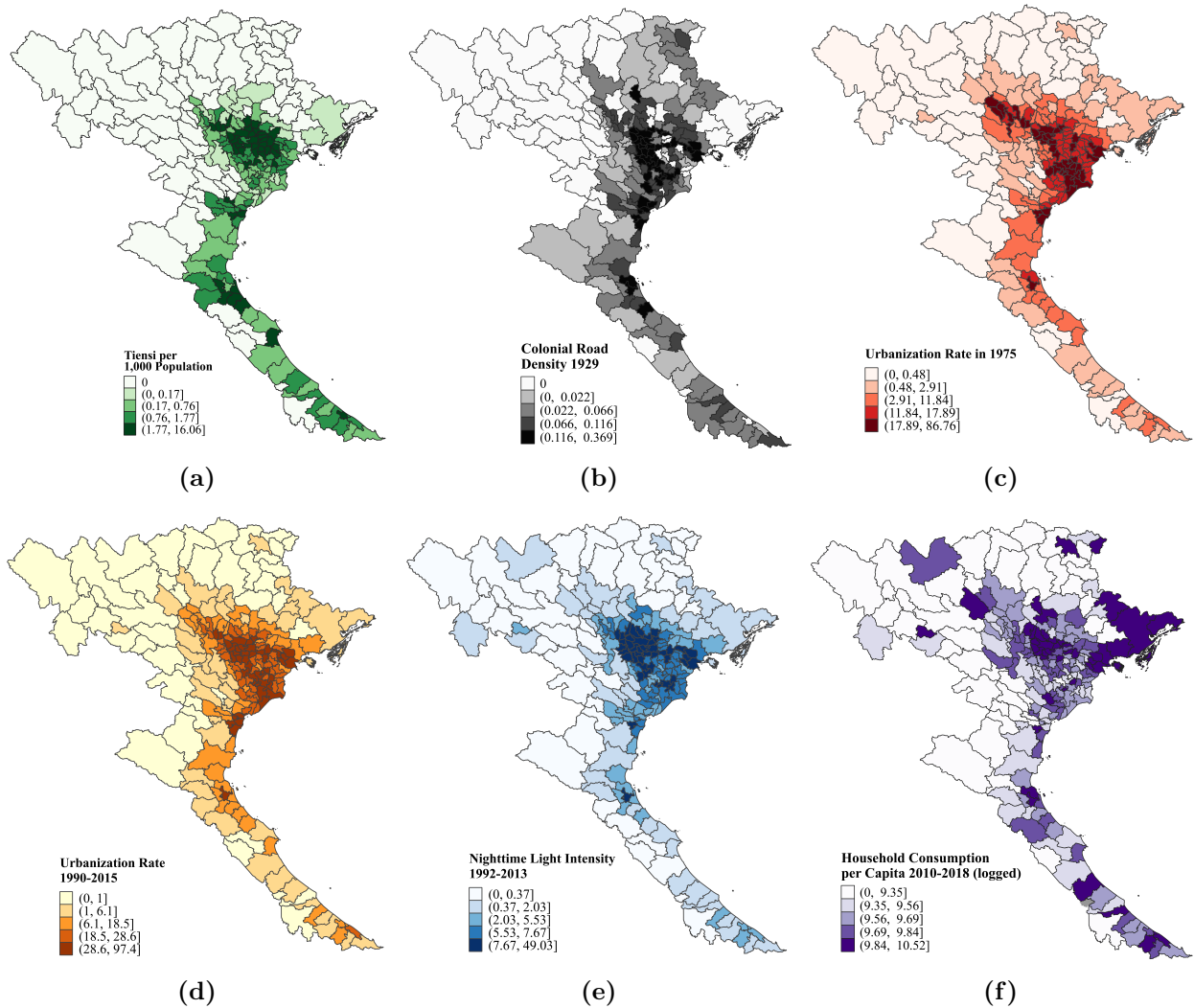
Colonial Roads Density in 1929 Due to the lack of fine-grained data during the colonial period in Vietnam, we follow Dell et al. (2018) to use the density of stone roads paved by the French colonizers as a proxy for economic development. Infrastructure construction was an important domain of colonial development, which was intended to better fit the exploitation activities and trade (Brocheux and Hémery, 2011). Thus, it is plausible to expect that the colonizers would only invest in the road construction in areas with the greatest potential to generate profits at that time.

The data on the density of paved road is extracted from the map of *Réseau Routier de l'Indochine* (*Road Network of Indochina*) produced by Service Géographique de l'Indochine (1929). Specifically, we georeference the map, create a polyline shapefile of the route empierrée (stone road), and

check, we drop these districts to mitigate the influence of the U.S. military (see Table A5 in Appendix).

⁹The population data is obtained from Goldewijk et al. (2017), which provides historical population estimates at the 5×5 arc-minutes grid cell level around the globe for every century between 1400 CE – 1900 CE. To obtain the population for a historical district in a certain time, we sum the population of all grid cells that fall within the district boundary. The average population in 1400 CE – 1900 CE is then calculated by averaging the district-level population in 1400 CE, 1500 CE, 1600 CE, 1700 CE, 1800 CE and 1900 CE. As robustness checks, we also try normalizing by taxable population in Nguyen dynasty and land area. The results hold with these alternative measures (see Table A5 in Appendix).

Figure 1: Spatial Distribution of Main Variables



Notes: This figure shows the spatial distribution of main variables used in our analyses, namely (a) total number of Tiensi per 1,000 population, (b) density of colonial roads in 1929, (c) degree of urbanization in 1975, (d) degree of urbanization in 1990-2015, (e) average nighttime light intensity in 1992-2013, (f) average household consumption per capita in 2010-2018. The variables are categorized into quintiles. Darker shades represent higher quintiles

intersect the shapefile with district boundaries to get the roads for each district (see Figure A2 in Appendix). The density of paved road for a district is then calculated as the total length of the paved road (in kilometres) divided by the total land area of the district (in squared kilometres).

Degree of Urbanization The second proxy for economic development is the average degree of urbanization in a district. The degree of urbanization is a measure developed by six international organizations to capture the urban-rural continuum.¹⁰ It has been shown to be highly correlated with GDP by Dijkstra et al. (2021).

The data on the degree of urbanization is obtained from the Global Human Settlement Layer (GHSL). It provides 30×30 arc-seconds raster files of the globe for each epoch 1975, 1990, 2000, and 2015. In each raster file, grid cells are categorized into urban centres, dense urban clusters, semi-dense clusters, suburban or peri-urban cells, or one of the four classes of rural grid cells, according to their population density, total population, and built-up areas.¹¹

To calculate the average degree of urbanization of a district, we first assign a score of 100 to urban centre grids, 75 to dense urban cluster grids, 50 to semi-dense cluster grids, 25 to suburban or peri-urban grids, and 0 to rural grids. Then we take the area-weighted average pixel value of all grids within the district boundary. To mitigate skewness and preserve zero entries, we add 0.01 to the weighted average and take a natural log. The degree of urbanization in 1975 is used to proxy the level of development before the reunification of north and south Vietnam, while the average of the measure in 1990, 2000, and 2015 is used to capture the level of development after the Doi Moi.

Nighttime Light Intensity Another proxy for economic prosperity used in our study is the average nighttime light intensity in 1992-2013. The luminosity data have been widely used to proxy economic activities since the seminal work by Henderson et al. (2012). In the case of Vietnam, Min and Gaba (2014) show that the luminosity data provided by satellite images are strongly correlated with actual lights at night on the ground; and Ho (2021) uses it to test the effects of the privatization of agricultural land on development at the commune level.

The data come from the National Oceanic and Atmospheric Administration (NOAA), which provides a composite measure of annual light intensity at night based on time-stable satellite images captured every night at 20:00 - 21:30 local time.¹² The intensity is scaled between 0-63, with 0 representing the darkest grids and 63 the brightest. The data are available for the years in 1992-2013 from several satellites, and are provided in gridded format with a resolution of 30×30 arc-seconds.

To compute the average intensity for each district, we first take the average cell value of all

¹⁰The six organizations are the Food and Agricultural Organization of the United Nations, the United Nations Human Settlement Programme, the International Labour Organization, the Organization for Economic Cooperation and Development, the World Bank, and the European Commission.

¹¹The four rural classes are rural clusters, low density rural grid cells, very low density rural grid cells, and water grid cells. Please refer to the documentation provided by GHSL for the details on the classification. The documentation is available on <https://ghsl.jrc.ec.europa.eu/datasets.php>

¹²The raw satellite images are provided by the Defense Meteorological Satellite Program's Operational Linescan System (DMS-P-OLS) of the U.S. Air Force Weather Agency. The raw satellite images include the lights from human settlements, lightning, aurora, fires, and gas flares. In constructing the composite measure, NOAA removed ephemeral lights and dropped images with clouds or solar and aurora glares before overlaying daily images to generate the annual measure. As a result, the composite measure captures only the nighttime lights from persistent human settlements.

grid cells lying within the district boundary using the raster produced by each satellite each year. Then we average across satellites to generate a yearly district-level measure. The final district-level intensity is obtained by averaging the yearly district-level measure over the period 1992-2013. In the analysis, we also raise the measure to natural log to reduce the skewness.

Annual Household Consumption per Capita The last measure of economic development employed in our analysis is the average annual household consumption per capita of a district in 2010-2018. We use the household expenditure survey data collected by the Vietnam Household Living Standards Surveys (VHLSS) to compute the household consumption. While VHLSS was conducted every two years since 2002, we focus on using the rounds during 2010-2018 because the location information of the households is incomplete prior to 2010. From 2012 onwards, around 50% of the households from the previous round were repeatedly surveyed. We use the most recent observation of the same households to avoid repeated observations.

We calculate the annual household consumption as the total living expenditure for the past 12 months, including the expenditure on education, health, food and non-food stuff, durable goods, and other regular expenses such as housing, electricity and sanitation. To ensure comparability across survey rounds, we adjust the nominal values of household consumption using the Consumer Price Index (CPI) data from the World Bank. To obtain the district-level annual household consumption per capita, we take the average household-level measure of all households residing in the district from all survey rounds. This measure is also log-transformed to correct for the skewness.

Some Remarks The proxies discussed above are significantly and positively correlated with each other, with a correlation coefficient of at least 0.5 and significance at the 1% level by the Pearson’s test (see Table A3 in Appendix). The summary statistics are presented in Table A2 in Appendix, while the spatial distribution is illustrated in Figure 1.

3.4 *Baseline Controls*

We control for a set of variables that might be correlated with the strength of Confucianism and economic development in our baseline analysis. These include absolute latitude, distance to coast, average elevation, terrain ruggedness, agricultural suitability, and population density in 1400 CE.

The current territory of Vietnam was formed through a series of expansions from the north to the south, with the north being the core region of culture and centralized bureaucracy (Dell et al., 2018; Ho et al., 2022). The north-south stretch of the territory also leads to varying climatic conditions, which matter for development. To account for these north-south differences, we control for the absolute value of the centroid latitude of each district.

With long coastlines, the access to trade and ideology for places in Vietnam is highly correlated with the distance to the coast. For instance, the initial merchants and settlers from mainland China were mainly located in coastal areas (Whitmore, 2006). The maritime trade in the seventeenth and eighteenth centuries involved the exchange of raw silk materials and mineral resources, as well as books on frontier Confucian knowledge (Li, 2011). We thus include distance to coast as a baseline control, which is constructed using the data from NOAA (2009).

The high mountainous regions in Vietnam have long been less developed and more influenced by the culture of ethnic minorities (van de Walle and Gunewardena, 2001), whereas places with more abundant land resources attracted more inhabitants and enjoyed a headstart in development. To control for the influences of these factors, we include average elevation and terrain ruggedness to capture the influence of terrain, and agricultural suitability and the share of urban population in 1400 CE for the impact of early development.

The data on elevation come from GOTOP30 (1996). The average terrain ruggedness is constructed using the data and method provided by Nunn and Puga (2012). Land productivity is proxied by the post-1500 maximum caloric suitability index by Galor and Özak (2016). The data for population in 1400 CE is obtained from Goldewijk et al. (2017). Table A1 in the Appendix provides a detailed description of the definition of the variables, while the summary statistics can be seen in Table A2.

4. Baseline Analysis

4.1 Baseline Regression Model

We examine the effect of Confucianism on economic development using a regression model of the following specification:

$$y_{ip} = \alpha + \beta \text{Tiensi density}_i + X_i \lambda + f e_p + \epsilon_{ip} \quad (1)$$

where y_{ip} is the proxy for the economic development of district i in province p ; Tiensi density_i is the key independent variable defined in section 3.2; the vector X_i includes the set of baseline controls, namely, absolute latitude, distance to coast, elevation, terrain ruggedness, caloric suitability, and urban population share in 1400. We also include province fixed effects $f e_p$ to account for the unobserved province-level characteristics that affect both Confucianism and development. To deal with the correlations of error terms within the province, we cluster the standard errors at the province level in all specifications. β is the coefficient of our interest, which captures the effects of Confucianism on economic development.

4.2 Baseline Results

We report the OLS estimates of equation (1) in Table 1. In the first five columns, we regress each measure of economic development on Tiensi density while only controlling for the province fixed effects. The coefficient estimates of Tiensi density suggest that the strength of Confucian literati is positively and significantly correlated with development. This positive and significant association holds when we add the baseline controls in columns (6) to (10). Specifically, 1% increase in Tiensi density is associated with 0.097% increase in colonial road density in 1929, 0.26 % in degree of urbanization in 1975, 0.258% in degree of urbanization in 1990-2015, 0.194% in nighttime light intensity in 1992-2013, and 0.029 % in annual household consumption per capita in 2010-2018.

Table 1: Effects of Confucian Literati on Economic Development - OLS Estimates

Dep var =	(1) Colonial road density 1929 (logged)	(2) Degree of urbaniza- tion 1975 (logged)	(3) Degree of urbaniza- tion 1990- 2015 (logged)	(4) Nighttime light intensity 1992- 2013 (logged)	(5) Household consump- tion per capita 2010- 2018 (logged)	(6) Colonial road density 1929 (logged)	(7) Degree of urbaniza- tion 1975 (logged)	(8) Degree of urbaniza- tion 1990- 2015 (logged)	(9) Nighttime light intensity 1992- 2013 (logged)	(10) Household consump- tion per capita 2010- 2018 (logged)
Tiensi density (logged)	0.171*** (3.777)	0.520*** (5.232)	0.504*** (7.148)	0.456*** (6.928)	0.067*** (6.018)	0.097** (2.303)	0.260*** (4.158)	0.258*** (5.980)	0.194*** (4.558)	0.029* (1.967)
Absolute latitude						0.104 (0.379)	-0.270 (-0.997)	-0.015 (-0.055)	-0.009 (-0.025)	-0.059 (-0.600)
Distance to coast (logged)						0.020 (0.091)	-0.369** (-2.309)	-0.255* (-2.039)	-0.165 (-1.415)	-0.050 (-1.221)
Elevation						-0.523 (-0.631)	-1.671 (-1.698)	-0.519 (-1.261)	-0.684 (-0.484)	-0.190 (-0.914)
Terrain ruggedness						-0.153 (-1.077)	-0.538*** (-3.236)	-0.647*** (-6.027)	-0.706*** (-4.022)	-0.078* (-1.785)
Caloric suitability						0.150 (1.631)	-0.324 (-1.697)	0.035 (0.566)	-0.011 (-0.076)	-0.018 (-0.598)
% Urban pop 1400						0.027 (1.520)	0.046* (1.846)	0.039** (2.359)	0.075*** (3.530)	0.016*** (3.017)
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
adj. R^2	0.410	0.681	0.777	0.743	0.334	0.420	0.808	0.877	0.847	0.419
Observations	217	217	217	217	216	217	217	217	217	216

Notes: Robust standard errors clustered at the province level are used and the corresponding t-statistics are reported in the parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

The sign of the coefficients of baseline controls is also plausible. In particular, distance to coast and terrain ruggedness are negatively and significantly correlated with development, while urban population share in 1400 have a positive and significant impact. The results support our presumptions that inland and mountainous districts are less developed, and that districts with a headstart in the initial period are more likely to be prosperous in the later periods.

4.3 *Additional Controls*

The strength of Confucianism for a district could be influenced by many other factors that are not controlled in our baseline. In particular, there were different ideologies and foreign interventions that could complement or substitute the role of Confucianism in the development. Therefore, we include a set of proxies for other ideologies and foreign interventions as additional controls for robustness checks. The variable definition and data source can be found in Table A1 in Appendix.

Buddhist Temple Density Buddhism was the most influential ideology when the Vietnamese first gained its independence in the tenth century. At that time, Buddhist monks were the most learned men in the societies. They did not just serve as the consultants of the imperial courts to give policy advice, but also spread Confucian philosophies to individuals who wanted to take administrative positions (Nguyen, 1998). Contemporary Vietnamese societies also see the juxtaposition of Buddhism and Confucianism, as exemplified by the coexistence of Buddhist temples and communal houses in villages as venues for religious and Confucian rituals, respectively. To disentangle the effects of Confucianism, we control for the density of Buddhist temples in the Nguyen dynasty.

Catholic Mission Density Catholic missionaries brought cultural and intellectual changes to Vietnam ever since its presence in sixteenth century. Most prominently, *Quoco Ngu*, a Latin-script-based writing system of Vietnamese developed by two Catholic missionaries, replaced the Chinese-character-based *Han Nom* as the national standard writing system (Haudricourt, 2010). This could gravely undermine the influence of Chinese culture and literature. Hence, we include the density of Catholic missions in 1887 as an additional control.

French Colonial Concession Density Driven by the incentive to exploit mineral deposits and agricultural products, the decades long French colonial rule led to economic crisis of Indochinese capitalism and underdevelopment (Brocheux and Hémerly, 2011; Murray, 1980). At the same time, the colonial intrusion provoked anti-French movements that aimed at modernizing Vietnam by abandoning Confucian values (Tran, 2013). To rule out these influences, we control for the density of agricultural and mineral concessions in 1926.

Communism Communism was both anti- and pro-Confucianism on its transition into the core societal value of contemporary Vietnam. On the one hand, revolutionary communists in 1930s opposed Confucianism as feudalistic and detrimental to modernization (Huynh, 1986). On the other hand, the most influential communist leader Ho Chi Minh saw Confucianism compatible with Marxism, and incorporated the Confucian values of *Datong* (great harmony) and *Minben* (people-oriented thought) into the communist party, government, and institutions (Brocheux, 2007; Nam,

2018). Thus we control for the effects of communism using the share of land within the war zones of the August Revolution in 1945.

Table 2: Robustness to Additional Controls

Dep var =	(1) Colonial road density 1929 (logged)	(2) Degree of urbanization 1975 (logged)	(3) Degree of urbanization 1990-2015 (logged)	(4) Nighttime light intensity 1992-2013 (logged)	(5) Household consumption per capita 2010-2018 (logged)
Tiensi density (logged)	0.053 (1.104)	0.235*** (4.114)	0.243*** (5.996)	0.194*** (4.012)	0.025* (1.836)
Buddhist temple density (logged)	0.106*** (3.988)	0.011 (0.324)	0.030 (1.225)	0.072** (2.575)	0.021** (2.108)
Catholic mission density (logged)	0.046 (1.170)	0.109*** (3.699)	0.063*** (3.709)	0.030 (1.658)	0.004 (0.515)
Agricultural concession density (logged)	-0.004 (-0.150)	-0.002 (-0.106)	0.021 (0.973)	0.078** (2.300)	0.009 (0.967)
Mineral concession density (logged)	0.174*** (4.530)	0.163 (1.422)	0.109 (1.182)	0.012 (0.108)	0.019 (0.705)
% Communist war zone (logged)		0.010 (0.554)	0.014 (0.988)	0.036*** (2.956)	0.006 (1.104)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes
adj. R^2	0.464	0.815	0.882	0.855	0.425
Observations	217	217	217	217	216

Notes: Notes: Robust standard errors clustered at the province level are used and the corresponding t-statistics are reported in the parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively. Baseline controls include absolute latitude, distance to coast (logged), elevation, terrain ruggedness, caloric suitability and the share of urban population in 1400 (logged).

Table 2 reports the results with additional controls. In column (1), we omit the proxy for communism, since it was developed after the French colonization. In columns (4)-(5), we control for all proxies discussed above. As shown by the estimates, the positive effects of Tiensi density on development remain significant for all development outcomes except for the colonial road density. One possible explanation for the insignificant results is that colonial road density is a rather indirect measure of economic prosperity. It is also possible that the distribution of colonial concession is highly correlated with that of colonial roads so that it shuts down the effects of Tiensi density. Nevertheless, as we will show in section 5.2, the significance is restored when we apply an instrumental variable approach.

5. Instrumental Variable Approach

Although we have accounted for a plethora of correlates, Tiensi density is still likely to be subject to endogeneity issues. For example, Tiensi density might fail to capture the strength of influential scholars who did not take civil exam but strengthened local Confucian norms. This measurement

error could lead to downward bias. In addition, there can be unobserved variables that contribute to both Confucianism and economic development. Omitting those variables can result in overestimating the effects of Tiensi density. In this section, we address these endogeneity issues by adopting an instrumental variable approach.

5.1 *Distance to the Nearest Hermit Scholar as An Instrumental Variable*

We use the geodesic distance from the border of a district to the nearest hermit scholar as the instrument for Tiensi density. The regression model for the two stage least estimation is specified in equation (2).

$$\begin{aligned} y_{ip} &= \alpha_2 + \beta_2 \widehat{\text{Tiensi density}}_i + X_i \lambda_2 + C_i \gamma_2 + \pi_p + \epsilon_{ip} \\ \widehat{\text{Tiensi density}}_i &= \alpha_1 + \beta_1 \text{Distance to hermit}_i + X_i \delta + C_i \gamma_1 + \pi_p + \mu_{ip} \end{aligned} \quad (2)$$

where y_{ip} and $\widehat{\text{Tiensi density}}_i$ are the outcome of interest and key independent variable, respectively; $\text{Distance to hermit}_i$ measures the shortest distance from the border of district i to its nearest hermit scholar; X_i and C_i are the vectors of the baseline and additional correlates, respectively; π_p account for province fixed effects; ϵ_{ip} is the error term.

We defined hermit scholars as individuals who *a.* were renowned for their expertise in Confucianism to the degree that their names were recorded in official chronicles; *b.* taught many Confucian students; *c.* refused to serve as a government official. These characteristics, as discussed in sections 5.1.1 and 5.1.2, make the distance to the nearest hermit scholar a plausible IV.

The list of hermit scholars come from two official chronicles, *Lich Trieu Hien Chuong Loai Chi* (*Categorized Chronicles of Successive Dynasties, hereafter CCSD*) and *Dai Nam Thuc Luc* (*Chronicle of Nguyen Dynasty, hereafter CND*), which together cover the periods from the Tran to Nguyen dynasties (1225 CE - 1945 CE) in Vietnam. In CCSD, the famous historian, Phan Huy Chu enumerated 29 virtuous scholars who were active during the Tran to Revival Le dynasties (1225 CE - 1789 CE) and provided short biographies about their main contribution and career path. In CND, there are two volumes named *Biographies of Hermits* that focus on introducing the 11 hermits that were discovered and recommended by local officials to the Nguyen government. After reading the biography of each scholar recorded in the above-mentioned two sources, we identify 9 scholars who meet all three conditions in the definition of hermit scholars. Table A4 in the appendix summarizes the main characteristics of the scholars, while Figure A3 gives an example of how we identify the characteristics of the hermit scholars from historical records. The geographic locations of the scholars are shown in Figure 2a.

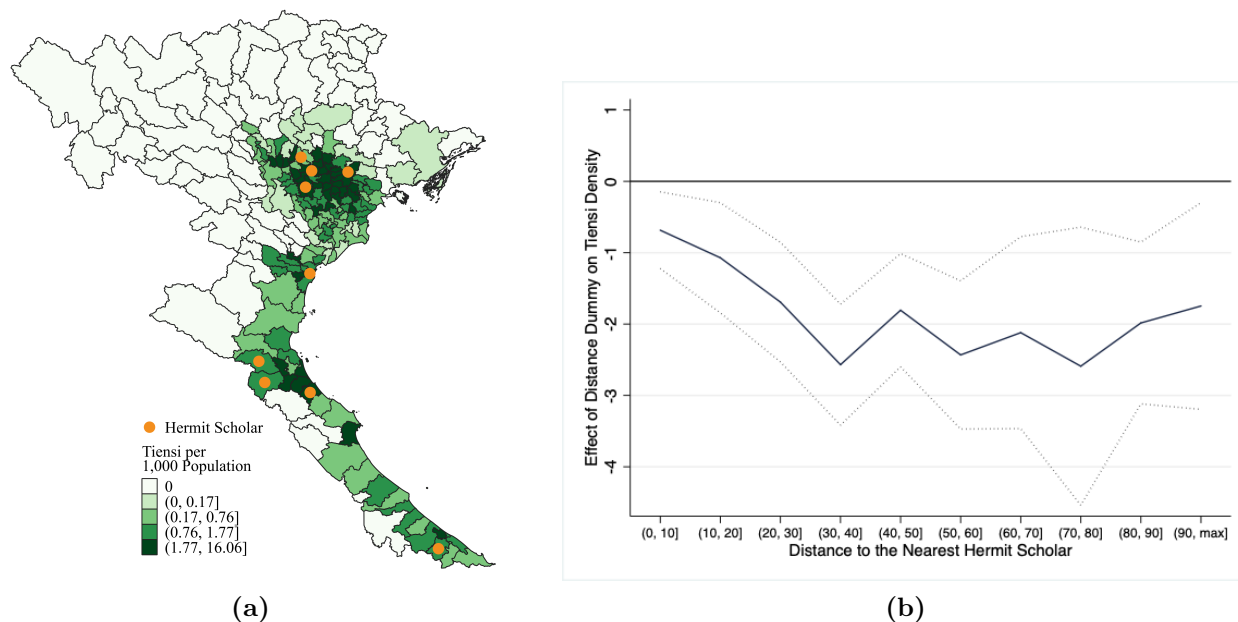
5.1.1 *Instrumental Relevance*

The nurturing of Confucian students and spreading of Confucianism among commoners relied heavily on individual efforts in historical Vietnam (Chen, 2015; McHale, 2002; Nguyen, 2020; Woodside,

2002).¹³ Being dedicated to Confucianism studying and teaching, hermit scholars were a prominent example of such efforts. The first hermit scholar on our list, Chu Van An, is the first Confucian sage venerated in the Temple of Literature in Hanoi and recognized as the master of all Vietnamese Confucianists (see Volume 7 of *Dai Viet Su Ky Thoan Tu*). Although not as influential as Chu, other hermit scholars also fostered many Confucian students and Tiensi. For example, the biography of Nguyen Dinh Tru documented that thousand of students had studied with him, of whom 70 obtained the Tiensi title. As with Nguyen, the remaining hermit scholars on our list were described as renowned for their accomplishment, which attracted hundreds or even thousands of aspiring students to study with them (see a summary of the descriptions in Table A2 in Appendix).

Given the difficulties in travelling and communication back then, the influence of the hermit scholar should vary with the geodesic distance to them. Individuals in nearby regions were more likely to get to know and visit the hermit scholars, whereas those in distant regions might be more oblivious to such influence. Overtime, places closer to hermit scholars could have produced more Tiensi and developed a stronger inclination to learn and observe Confucian values and norms.

Figure 2: Hermit Scholars



Notes: (a) the locations of hermit scholars, as represented by the orange dots. (b) illustrate the effects of distance to hermit scholars on Tiensi density based on a non-parametric regression. The solid line depicts the coefficient estimates of the effects of distance dummy on Tiensi density, while the grey dash lines delineate the 95% confidence intervals.

¹³This is because historical Vietnam lacked influential academies, and the Confucianism teaching at the lower level were largely provided by private schools organized by local scholars (Nguyen, 2020; Woodside, 2002). In addition, the relatively underdeveloped print culture impeded the formation of institutions, network and media of spreading Confucianism teachings (McHale, 2002).

Table 3: Balance Checks of IV

Dep var =	(1) State history until 1500 (logged)	(2) Province or prefecture seat	(3) Distance to Hanoi (logged)	(4) Tax in cash per capita (logged)	(5) Courier sites	(6) Guanbao road	(7) Tributary route	(8) Local market	(9) Buddhist temple density	(10) Distance to nearest harbour (logged)
Distance to hermit	0.000 (0.599)	-0.023 (-1.014)	0.095 (1.707)	-0.017 (-1.660)	0.004 (0.274)	-0.030 (-1.213)	0.001 (0.406)	-0.032*** (-3.665)	-0.068 (-1.045)	0.109 (0.835)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
adj. R^2	0.888	0.048	0.684	0.580	0.177	0.340	0.529	0.532	0.127	0.360
Observations	217	217	217	217	217	212	217	217	217	217

Notes: Robust standard errors clustered at the province level are used and the corresponding t-statistics are reported in the parentheses. * * *, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

To formally test the relationship between Tiensi density and the distance from the border of a district to the nearest hermit scholar, we regress Tiensi density on the binned distance to hermit scholar while controlling for the baseline correlated and province fixed effects. We plot the coefficient and 95 confidence interval estimates of the dummies for distance bins in Figure 2b.¹⁴ As shown in the Figure 2b, Tiensi density diminishes almost linearly with the distance to hermit scholar in the first 40 kilometres. Beyond the point of 40 kilometres, there is no further reduction.

5.1.2 Exclusion Restrictions

The validity of our IV also relies on the assumption that distance to the nearest hermit scholar can only impact economic development through its effects on Tiensi density. We provide several justifications for why this assumption is likely to hold.

Anecdotal Evidence Historical accounts suggest that the emergence, activities and locations of hermit scholars are plausibly exogenous. According to their biographies, the scholars on our list chose to live as hermits primarily due to their passion for Confucianism and indifference towards fame or power. This implies that the emergence of hermit scholars has nothing to do with local economic conditions. In addition, the fact that the hermit scholars refused to be important government officials suggests that they had little direct impact on local economy. Furthermore, hermit scholars did not seem to have chosen the location of retreats intentionally. All of the hermit scholars established Confucianism studying and teaching venues at their places of birth. Since one cannot choose their birth places, the locations of hermit scholars are arguably exogenous. Indeed, as shown in Figure 2a, the hermit scholars were scattered in different part of our sample region.

Balance Checks However, one might still raise concerns with the violation of exclusion restrictions. For example, it might be argued that hermits from places with better state capacity or communication infrastructure were more likely to be discovered and reported to the government. Or individuals born in places with a better access to Confucian books were more likely to be drawn to Confucianism studies. To alleviate the concerns, we examine whether distance to the nearest hermit is correlated with aforementioned factors in Table 3. Specifically, we regress each one of proxies for these confounders on our IV while controlling for baseline correlates and province fixed effects. The details on the data source and variable definition of the proxies can be found in Table A1 in appendix.

The results in columns (1) to (4) of Table 3 indicate that distance to the nearest hermit scholar is not correlated with state capacity, which is proxied by state history in 1500, tax paid in cash per taxable population in Nguyen dynasty, a dummy for being the province or prefecture seat in Nguyen dynasty, and the distance to Hanoi. The presence of communication infrastructure does not appear to be correlated our IV either, as suggested by the results in columns (5) and (6), in which dummies for the presence of courier site or Guanbao road (i.e. the road for official document delivery) in the

¹⁴The regression model is $Tiensi_i = \sum_j \beta_{ij} \times D_{ij} + X_i \lambda + fe_p + \epsilon_i$, where D_{ij} is a dummy variable that equals 1 if the distance of district i to its nearest hermit scholar falls into the j^{th} bin, X_i include the full set of baseline controls, fe_p accounts for province fixed effects. The base group in this regression is those districts with a zero distance, i.e. the 9 districts where the hermit scholars were located.

Table 4: Zero-First Stage Test

Dep var =	(1) pre-1722 Tiensi density (logged)	(2) % Urban pop 1500-1600 (logged)	(3) post-1722 Tiensi density (logged)	(4) Degree of Urbanization 1975 (logged)
Distance to hermit scholar (logged)	-0.140 (-1.564)	0.003 (1.220)	-0.146** (-2.877)	-0.122*** (-3.160)
Absolute latitude	-0.560 (-0.693)	0.020 (0.505)	0.109 (0.316)	-0.309 (-0.964)
Distance to coast (logged)	-0.525 (-1.442)	-0.009 (-0.399)	-0.556* (-1.867)	-0.735*** (-3.092)
Elevation	0.622 (0.405)	-0.011 (-0.220)	0.645 (0.758)	-1.278 (-1.500)
Terrain ruggedness	-0.949** (-2.865)	-0.002 (-0.087)	-0.445** (-2.447)	-0.700*** (-3.975)
Caloric suitability	0.547* (1.905)	0.002 (0.243)	0.345 (1.009)	-0.081 (-0.376)
% Urban pop 1400	0.025 (0.605)	1.025*** (412.324)	0.049* (2.124)	0.028 (1.000)
Province FE	Yes	Yes	Yes	Yes
adj. R^2	0.626	0.999	0.556	0.781
Observations	132	132	132	132

Notes: Robust standard errors clustered at the province level are used and the corresponding t-statistics are reported in the parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Nguyen dynasty are used as the dependent variables, respectively. Results in columns (7) to (9) show little association between the access to Confucian books and our IV as well, where the major sources of book imports (tributary route to China and maritime trade), and printing (woodblock printing by local Buddhist temples) are considered.¹⁵ However, results in column (10) indicate that our IV is correlated with the presence of local market in the Nguyen dynasty.¹⁶ To shut down the effect, we include the dummy for the access to local market as an additional control in our IV estimations.

Zero-First-Stage Test To provide additional evidence that exclusion restrictions are not violated, we conduct the zero-first-stage test by following the practice of Angrist et al. (2010) and Nunn and Wantchekon (2011). The intuition of this test is as follows. There can be a subsample of observations, for whom the effects of IV on the instrumented variable is arguably zero, that is, the coefficient estimates of the IV at the first stage is not statistically different from zero. Since an IV satisfying exclusion restrictions can only affect the outcome variable through its effects on the instrumented variable, the reduced-form effects of IV on the outcome for this subsample should also be insignificant. Any significant effects of the IV on the outcome using the subsample would

¹⁵See Li (2011) for the historical account on the role of diplomats and maritime trade in bringing in Confucian books, and Nguyen et al. (2018) for the role of Buddhist temples in woodblock printing.

¹⁶The interpretation of association can be a bit tricky, since we do not have the information on the time of formation of local markets. If the local markets were formed before Nguyen dynasty, the results should be interpreted as the distance to hermit scholar is correlated with market access. However, if the local markets were formed in late Nguyen dynasty, it can be interpreted as the reduced-form results for the effects of our IV on development, which is proxied by the presence of local market.

indicate the violation of exclusion restrictions.

To apply the zero-first-stage test to our analysis, we use a subsample of districts whose nearest hermit scholar was born after the 1722. For this subsample of districts, we would not expect to see any significant effects of the distance to the nearest hermit scholar on pre-1722 Tiensi density, since their nearest hermit scholars were born after 1722. Accordingly, the reduced-form effects of distance to the nearest hermit scholar on pre-1722 development outcomes are also expected to be insignificant. Any significant effects of the IV on Tiensi density and pre-1722 development outcome would suggest the violation of exclusion restrictions.

The zero-first-stage test results in Table 4 lend further support to our IV. As shown by columns (1) and (2), our IV does not have any significant effects on pre-1722 Tiensi density or urban population share (as a proxy for degree of urbanization) when the subsample of districts whose nearest hermit was born after the 1722 is used. By contrast, the post-1722 Tiensi density and degree of urbanization are significantly correlated with our IV using the same subsample.

5.2 *Instrumented Results*

Having established the validity of our IV, we report the IV estimates of the effects of Tiensi density on each of the five proxies for economic development in Table 5. In odd columns, we controls for baseline correlates and province fixed effects, while in even columns additional controls are included. Panel A reports the the reduced-form effects of distance to hermit scholars on development outcomes. All columns report a negative and significant association, suggesting that the further away a district was from the hermit scholar, the less developed it was in different periods.

Panel B reports the 2SLS results. The first-stage estimates show a negative and significant effect of distance to hermit scholars on Tiensi density, as discussed in section 5.1.1. And the Keleibergen-Paap rank Wald F statistics indicates that our IV is not weak. The second-stage results report a positive and significant effects of the predicted Tiensi density on economic development, which are consistent with the OLS estimates. The magnitude of IV estimates is larger than the OLS estimates for all specifications, suggesting that OLS estimates are biased towards zero. One possible explanation for the downward bias is that Tiensi density measure can fail to capture the strength of Confucian literati in the districts that did produce any Tiensi in the history. Given the fact there are a lot of such districts, the increase in magnitude is not surprising. Another possibility is that in the OLS analysis we omit some variables that might be negatively correlated with Confucianism but can contribute to development. For example, the inter-ethnic conflicts might inhibit the spreading of Confucianism but could contribute to state capacity building and development (Dincecco et al., 2021).

Table 5: Effects of Tiensi Density on Economic Development - 2SLS Estimates

Dep var =	Colonial road density 1929 (logged)		Degree of urbanization 1975 (logged)		Degree of urbanization 1990-2015 (logged)		Nighttime light intensity 1992-2013 (logged)		Household consumption per capita 2010-2018 (logged)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Reduced-form results										
Distance to hermit scholar (logged)	-0.089** (-2.571)	-0.085** (-2.499)	-0.120*** (-3.748)	-0.120*** (-4.056)	-0.135*** (-4.305)	-0.129*** (-4.289)	-0.117*** (-3.156)	-0.098** (-2.556)	-0.030** (-2.436)	-0.029** (-2.180)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
adj. R^2	0.423	0.466	0.788	0.800	0.856	0.864	0.838	0.842	0.427	0.429
Observations	217	217	217	217	217	217	217	217	216	216
Panel B: 2SLS results										
<i>Second-stage</i>										
Tiensi density (logged)	0.418** (2.304)	0.401** (2.443)	0.564*** (4.338)	0.568*** (4.595)	0.632*** (3.877)	0.614*** (4.152)	0.549*** (2.639)	0.467** (2.496)	0.143** (2.187)	0.140** (2.120)
<i>First-stage: Dep var = Tiensi density (logged)</i>										
Distance to hermit scholar (logged)	-0.213*** (-4.911)	-0.211*** (-6.009)	-0.213*** (-4.911)	-0.211*** (-6.009)	-0.213*** (-4.911)	-0.211*** (-6.009)	-0.213*** (-4.911)	-0.211*** (-6.009)	-0.212*** (-4.790)	-0.208*** (-5.772)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap rk Wald F	24.117	36.110	24.117	36.110	24.117	36.110	24.117	36.110	22.947	33.320
Observations	217	217	217	217	217	217	217	217	216	216

Notes: Robust standard errors clustered at the province level are used and the corresponding t-statistics are reported in the parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively. Baseline controls include absolute latitude, distance to coast (logged), elevation, terrain ruggedness, caloric suitability, and the share of urban population in 1400 (logged). Additional controls are dummy for the presence of local market, density of Buddhist temple (logged), density of agricultural concession (logged), density of mine concession (logged), and the share of communist war zone (logged) (not used in columns (1)-(2)).

5.3 Robustness to the Violation of Exclusion Restrictions

Although we have provided multiple justifications for the validity of our IV, it is still possible that one could find a specific scenario under which the exclusion restrictions do not hold perfectly. To address this concern, we adopt the method developed by Conley et al. (2012) to check the sensitivity of our results to the degree of exclusion restrictions violation. To illustrate the method, we use the regression model specified in equation (3)

$$\begin{aligned} Y &= X\beta + Z\gamma + \epsilon \\ X &= Z\Pi + V \end{aligned} \tag{3}$$

where Y is an $N \times 1$ vector of outcome variables, X is an $N \times 1$ vector of instrumented variables, and Z is an $N \times 1$ vector of instruments. γ captures the direct effects of the instrument on the outcome (i.e. the degree of violation of exclusion restrictions when $\gamma \neq 0$). Conley et al. (2012) provide consistent estimates for β , when the assumption $\gamma = 0$ is relaxed, and replaced with a prior on γ . Different inference procedures are employed, when the prior is specified as a support or a distribution of γ . In our analysis, we adopt the local to zero approach and assume that γ follows a normal distribution $N(\mu_\gamma, \Omega_\gamma)$. With the prior, the consistent estimator of β takes the form:

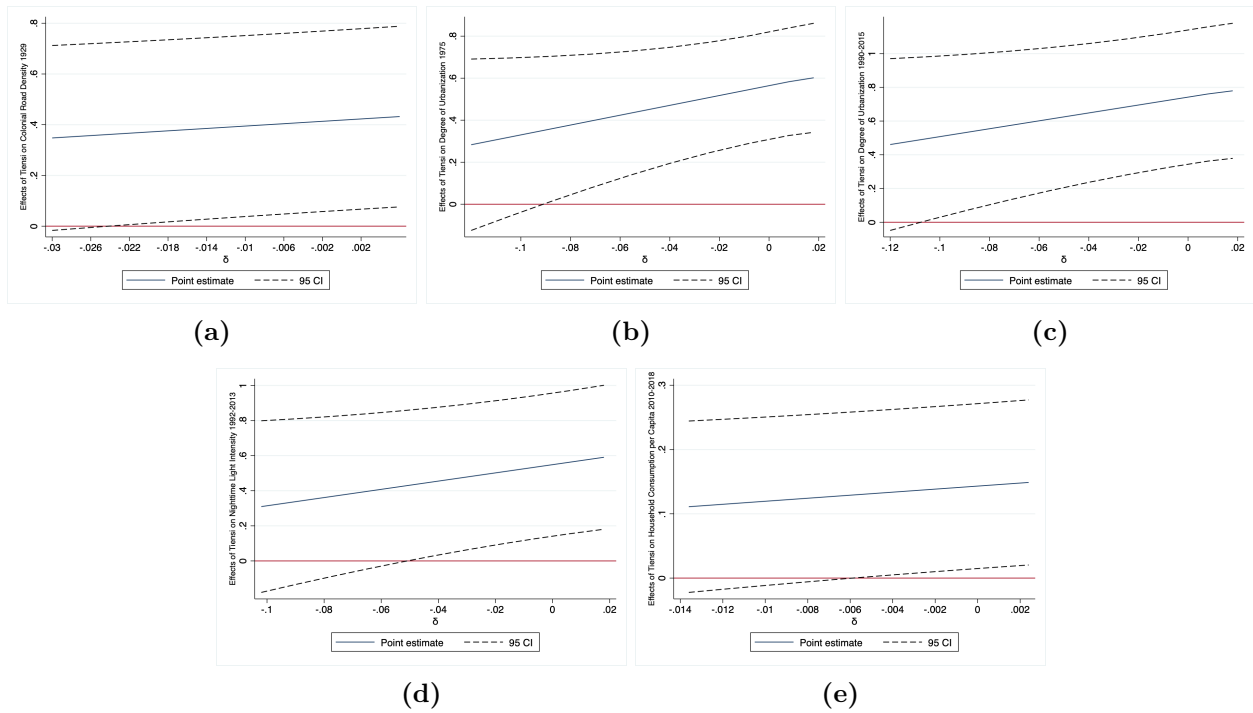
$$\hat{\beta} \sim N(\beta_{2SLS} + A\mu_\gamma, W_{2SLS} + A\Omega_\gamma A') \tag{4}$$

where $A = (X'Z(Z'Z)^{-1}Z'X)^{-1}(X'Z)$, and β_{2SLS} and W_{2SLS} are the conventional 2SLS β coefficient estimator and variance-covariance matrix, respectively.

Conley et al. (2012) do not provide a guidance on how to choose the value of μ_γ and Ω_γ , neither does the literature provides empirical evidence on the direct effects of our IV on Tiensi density. To resolve this issue, we follow the practice of van Kippersluis and Rietveld (2018) to use the reduced-form effects of our IV on the outcome in the zero-first-stage test results (column (2) of Table 4) as the benchmark, and see how the estimates change when different values in the neighborhood of the benchmark are specified as the input for μ_γ and Ω_γ . Specifically, we assume that $\mu_\gamma = \frac{\delta}{2}$ and $\Omega_\gamma = \frac{\delta^2}{12}$, where δ takes on different values depending on the proxy for economic development used. For the degree of urbanization, δ varies around 0.008 since the direct effects of IV on the share of urban population in the zero-first-stage test is estimated to be 0.004. For other proxies, we scale the values using the ratio of the effects of Tiensi density on the proxy and degree of urbanization in the baseline analysis. The point estimates and 95% confidence intervals based on and the priors specified above are shown in Figure 3.

As shown by the figures, the positive effects of Tiensi density on economic development remain significant at the 5% level when the direct effects of our IV on development are positive for all measures of development. The significance disappears when the direct effects of IV became more negative than a certain threshold value. For example, the 95% confidence interval for the degree of urbanization in 1975 starts to include zero when the mean value of the direct effects of our IV is smaller than -0.045 (i.e. when δ is smaller than -0.09). However, this is unlikely to happen, since

Figure 3: Sensitivity to the Imperfect Satisfaction of Exclusion Restrictions



Notes: This figure presents the consistent IV estimates of the effects of Tiensi density on (a). density of colonial roads in 1929, (b). degree of urbanization in 1975, (c). degree of urbanization in 1990-2015, (d). average nighttime light intensity in 1992-2013, (e). average household consumption per capita in 2010-2018, when the exclusion restrictions do not hold perfectly. The estimates are generated based on the Local to Zero approach of Conley et al. (2012). We assume that the direct effects of our IV follows $N(\frac{\delta}{2}, \frac{\delta^2}{12})$ and δ takes on the values specified on the horizontal axis.

the estimated direct effects of our IV is 0.004, which is not only positive but also ten times smaller than -0.045 in magnitude. Similar inferences can be drawn when other proxies are used. Therefore, our results remain consistent even when a considerable degree of violation of exclusion restricts is allowed for.

6. Mechanism

Our results suggest a persistent positive association between the strength of Confucian literati and economic development. This contradicts the prediction of Max Weber that Confucianism impeded economic growth. A natural question then ensues: What mechanism can account for the positive association? In this section, we propose and test several channels emerged from historical accounts and relevant literature.

6.1 *Culture of Respect for education and Human Capital Accumulation*

As well established by Chen et al. (2020), the pursuit of academic success in civil exam led to a culture of respect for education that persists and bears upon contemporary human capital outcomes in China. A similar story can be told for Vietnam. The aspiration of Viet people to become well-educated Confucianists can be traced to the invention of vernacular lexicon, *nho* (ru in Chinese, or Confucianist in English) in the sixth century (Taylor, 2002). This concept was repeatedly exercised by generations of Vietnamese, strengthened by the adoption civil exam, and eventually developed into what the historian calls “literati culture” (Tran, 1971; Whitmore, 1997). The legacy of this literati culture on human capital outcome can be exemplified by the second-generation Indochinese migrants in America. These migrant children came from humble background but achieved exceptional academic records among their local peers largely due to the value of education in the culture of origin (Caplan et al., 1992).¹⁷

Given the importance of human capital in development, we test whether the effects of Tiensi density work through its influence on the literati culture and human capital accumulation. In doing so, we construct a dummy variable for the presence of literati culture in a district using the information in the chronicle DKDDC of the Nguyen dynasty. Specifically, the variable equals 1 if there is an expression of respect for literature or literati in the description of local customs.¹⁸ To measure contemporary educational investment and outcomes, we calculate the district-level average share of household expenditure on education from 2010-2018 VHLSS; and the proportion of illiterate individuals, average years of schooling, and the share of individuals with secondary and above education from 2009 population census.¹⁹ We then estimate the effects of Tiensi density on each one of the outcomes while controlling for the baseline and additional correlates.

¹⁷As much as we want to establish the persistence of literati culture through intergenerational cultural transmission, we did not find any data that allows us to do so at the time of writing this paper.

¹⁸An example of the expression of respect for literature or literati is “Shi Shang Fengya” in Chinese, which means scholars appreciate literary refinement and elegance. See Figure in Appendix for a snapshot of the chronicle for such description.

¹⁹The illiteracy, years of schooling and educational attainment variables are based on individuals aged between 25-65. The details on the variable definitions can be found in Table A1 in Appendix.

Table 6: Channel: Literati Culture and Human Capital Accumulation

Dep var =	Literati culture		% Expenditure on education		illiteracy rate (logged)		Years of schooling (logged)		Secondary and above degree	
	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)	OLS (7)	IV (8)	OLS (9)	IV (10)
Tiensi density (logged)	0.063** (2.179)	0.283*** (3.852)	0.050*** (3.421)	0.143*** (5.153)	-0.045** (-2.207)	-0.227*** (-2.942)	0.010*** (3.130)	0.060*** (3.210)	0.033* (1.933)	0.262** (2.512)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
adj. R^2	0.261	-0.532	0.529	0.032	0.889	0.541	0.636	-0.232	0.358	-0.575
Observations	217	217	216	216	217	217	217	217	217	217

Notes: Robust standard errors clustered at the province level are used and the corresponding t-statistics are reported in the parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively. Baseline controls include absolute latitude, distance to coast (logged), elevation, terrain ruggedness, caloric suitability, and the share of urban population in 1400 (logged). Additional controls are dummy for the presence of local market, density of Buddhist temple (logged), density of agricultural concession (logged), density of mine concession (logged), and the share of communist war zone (logged).

The results are shown in Table 6. The odd columns show the OLS estimates while even columns report IV estimates. As suggested by the results, districts with higher Tiensi density are more likely to be recorded to have literati culture (columns (1)-(2)), spend more on education (columns (3)-(4)), and have better educational outcomes (columns (5)-(10)). It worth noting that our IV estimate of the effect on average years of schooling (0.06) is smaller than that of Chen et al. (2020) (0.085). As a byproduct, it gives a rough comparative statistics on the size of the effect.

6.2 *Social Capital and Public Goods Provision*

Confucian norms in Vietnam are characterised by strong kinship and public contribution of village members. The collaborative endeavours among family and community members could constitute social capital which is essential for public goods provision that promotes economic development. As early as in fifteenth century, the national code of conduct (called Le Code) which is of Confucian nature stipulated that families should contribute 5% of their property to the veneration of ancestors (Haines, 1984). Dating back to seventeenth century, village regulations (called huong uoc in Vietnamese) created by village-based Confucian literati institutionalized the norm of voluntary contribution of labour and land to supporting the construction of irrigation facilities and roads, hiring teacher, and rewarding accomplished students (Adams and Hancock, 1970; Li, 2011; Nguyen, 1998; Nguyen, 2020).

To test whether social capital is a potential channel, we construct three variables. First, we calculate the surname similarity index within a district which measures the probability of two randomly picked individuals sharing the same surname. The surname data come from 2018 VHLSS and the comprehensive list of Vietnamese surnames are obtained from Nguyen (2000). Surname similarity is a proper proxy for kinship capital, because *a*. Vietnam is a patrilineal society in which children share the same surname with the fathers; *b*. collaboration at the village level relies heavily on lineage groups with large population shares (Xu and Yao, 2015).

Second, we capture the village level coordination using the density of communal (or village) houses in a district. Communal houses symbolize collective actions among village members: they are mainly funded by village members' voluntary contributions and serve as the venues for performing Confucian rituals and discussing village fairs (Endres, 2001). Third, the density of non-government organizations (NGOs) within a district is used as a proxy for general social capital. The data on communal house and NGO are taken from Economic Census.

OLS and IV estimates of the effects on social capital are reported in Table 7. We find that Tiensi density is positively and significantly correlated with surname similarity (column (2)), communal house density (columns (3)-(4)), and density of NGOs (columns (5)-(6)). This supports our hypothesis that the social capital based on kinship ties and village-level collaboration could be a possible channel through which the strength of Confucian literati contributes to development.

Table 7: Channel: Social Capital and Public Goods Provision

Dep var =	Surname similarity (logged)		Commune house density (logged)		NGO density (logged)		Education unit density (logged)		Health unit density (logged)		Culture and sports unit density (logged)	
	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)	OLS (7)	IV (8)	OLS (9)	IV (10)	OLS (11)	IV (12)
Tiensi density (logged)	0.033 (1.106)	0.169*** (3.357)	0.364*** (5.156)	0.913*** (4.233)	0.130*** (5.940)	0.381*** (2.896)	0.126*** (5.501)	0.378*** (3.027)	0.128** (2.306)	0.549* (1.807)	0.162** (2.436)	0.294 (0.769)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yes	Yes											
adj. R-sq	0.376	-0.072	0.862	0.390	0.812	0.332	0.860	0.375	0.356	-0.092	0.300	0.041
Observations	216	216	217	217	217	217	217	217	217	217	217	217

Notes: Robust standard errors clustered at the province level are used and the corresponding t-statistics are reported in the parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively. Baseline controls include absolute latitude, distance to coast (logged), elevation, terrain ruggedness, caloric suitability, and the share of urban population in 1400 (logged). Additional controls are dummy for the presence of local market, density of Buddhist temple (logged), density of agricultural concession (logged), density of mine concession (logged), and the share of communist war zone (logged).

Table 8: Channel: Economic Transitions

Dep var =	Individual business density (logged)		% Employed in private sector (logged)		% Employed in foreign sector (logged)		% Employed in public sector (logged)		% Employed in primary industry (logged)		% Employed in secondary industry (logged)		% Employed in tertiary industry (logged)	
	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)	OLS (7)	IV (8)	OLS (9)	IV (10)	OLS (11)	IV (12)	OLS (13)	IV (14)
lpopconf	0.186*** (4.521)	0.539*** (3.136)	0.151*** (3.307)	0.478* (1.947)	0.216* (1.881)	0.441* (1.951)	0.014 (0.514)	0.222 (1.478)	-0.040** (-2.127)	-0.266** (-2.112)	0.095*** (3.234)	0.249** (2.535)	0.042* (2.086)	0.232* (1.896)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
adj. R^2	0.686	0.210	0.438	-0.108	0.486	-0.041	0.207	-0.405	0.235	-0.517	0.682	0.216	0.254	-0.442
Observations	217	217	217	217	217	217	217	217	217	217	217	217	217	217

Notes: Robust standard errors clustered at the province level are used and the corresponding t-statistics are reported in the parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively. Baseline controls include absolute latitude, distance to coast (logged), elevation, terrain ruggedness, caloric suitability, and the share of urban population in 1400 (logged). Additional controls are dummy for the presence of local market, density of Buddhist temple (logged), density of agricultural concession (logged), density of mine concession (logged), and the share of communist war zone (logged).

Having established the effects on social capital, we proceed to test whether the collaborative norms of Confucianism are also translated into better public goods provision. To proxy public goods provision, we calculate the density of different types of public units provided by non-public entities. The data are obtained from 2021 Economic Census. The results in Table 7 lend support to our presumption. Districts with stronger local Confucian literati have a better public goods provision, as proxied by higher density of public units for education and training (columns (1)-(2)), health (columns (3)-(4)), and cultural and sports activities (column (6)).

6.3 *Economic Transition*

Over the past few decades, Vietnam has witnessed rapid development in the non-state and non-agricultural sectors, which have contributed greatly to economic growth (Jaax, 2020).²⁰ The development was, however, achieved without commensurate strengthening of formal institutions.²¹ Instead, informal institutions like personal network and family ties have been shown to have facilitated the development of private sectors (Steer and Sen, 2010; Zhang, 2020). In the Confucian traditions, family and community members acted collectively to support the education of Confucian students. In return, Confucian students gave back to their family and communities once they passed the civil exam and became part of the network of prestigious gentry class. This reciprocity formed strong family and community ties that could facilitate development through resource poolings and transfers, risk sharing and creating a trusting environment for non-state and non-agricultural sector development. As a result, more individuals would choose to work in these sectors.

To test the this channel, we use the density of registered individual business and the proportions of individuals employed in each sector and industry to capture the structure of the economy. The results in Table 8 show that individuals from districts with more tiensi are more likely to establish individual business (columns (1)-(2)), to work in private and foreign sectors (columns (3)-(6)), and to take jobs in secondary and tertiary industries (columns (9)-(12)). By contrast, having more tiensi would not affect the public sector employment (columns (7)-(8)) and appears to reduce the employment in agricultural sector (columns (9)-(10)). This provides support for the argument that Confucian traditions can promote development through facilitating economic transitions.

7. Conclusion

Using Northern Vietnam as a laboratory, we find a strong and positive association between historical exposure to Confucianism and modern economic development over the past century. This finding could not be explained away by other ideologies and foreign influences and appears to be causal when the endogenous strength of Confucianism is instrumented by the distance to the nearest exogenously located hermit scholar. It is consistent with the Neo-Confucianism hypothesis that

²⁰The private firm registrations increased six folds from 1999 to 2006, while the share of GDP contribution of state-owned enterprises decreased about seven percentage points from 2000 to 2010 (Malesky and Taussig, 2008).

²¹Officials are often have hostile attitudes toward private sector (Steer and Sen, 2010). Private property rights are implemented with restrictive regulations. For example, individuals can apply for the private use right of agricultural land but the type of crops that could be planted is stipulated by the district-level government (Resolution 17/2011/QH13, 2011)

argues Confucian values have been conducive to economic development for the past few decades. In addition, we show that the positive effects of Confucianism could be explained by the persistence of a culture of valuing education and cooperation among family and community members, which facilitate the human capital accumulation, public goods provision, and economic transitions.

In a nutshell, these findings suggest that traditional norms could favor contemporary development, particularly in developing countries that lack well-developed formal institutions. Moreover, by providing empirical evidence generated in Vietnam, this paper extends the growing literature on Confucianism which has predominantly focused on the story of China.

References

- Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, 91:1369–1401.
- Acemoglu, D., Johnson, S., and Robinson, J. A. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *The Quarterly Journal of Economics*, 117(4):1231–1294.
- Adams, J. and Hancock, N. (1970). Land and Economy in Traditional Vietnam. *Journal of Southeast Asian Studies*, 1(2):90–98.
- Angrist, J., Lavy, V., and Schlosser, A. (2010). Multiple Experiments for the Causal Link between the Quantity and Quality of Children. *Journal of Labor Economics*, 28:773–824.
- Bai, Y. (2019). Farewell to confucianism: The modernizing effect of dismantling China’s imperial examination system. *Journal of Development Economics*, 141:102382.
- Bai, Y. and Jia, R. (2016). Elite Recruitment and Political Stability: the Impact of the Abolition of China’s Civil Service Exam. *Econometrica*, 84(2):677–733.
- Banerjee, A. and Iyer, L. (2005). History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India. *American Economic Review*, 95(4):1190–1213.
- Becker, S. O. and Woessmann, L. (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *The Quarterly Journal of Economics*, 124(2):531–596.
- Brocheux, P. (2007). *Ho Chi Minh A Biography*. UK: Cambridge University Press.

- Brocheux, P. and Hémery, D. (2011). The Structures of Domination. In *Indochina: An Ambiguous Colonization, 1858-1954*, pages 70–115. California: University of California Press.
- Caplan, N., Choy, M., and Whitmore, J. K. (1992). Indochinese Refugee Families and Academic Achievement. *Scientific American*, 266(2):36–45.
- Caplan, N., Choy, M. H., and Whitmore, J. K. (1991). *Children of the Boat People: A Study of Educational Success*. Michigan: University of Michigan Press.
- Chen, T., Kung, J. K.-S., and Ma, C. (2020). Long Live Keju! The Persistent Effects of China’s Civil Examination System. *Economic Journal*, 130:2030–2064.
- Chen, W. (2015). *Yuenan Keju Zhidu Yanjiu (A Study on the Civil Examination in Vietnam)*. Beijing: the Commerical Press.
- Chen, Z., Ma, C., and Sinclair, A. J. (2021). Banking on the Confucian Clan: Why China Developed Financial Markets So Late. *Economic Journal*, 132:1378–1413.
- Conley, T. G., Hansen, C. B., and Rossi, P. E. (2012). Plausibly Exogenous. *The Review of Economics and Statistics*, 94(1):260–272.
- Cooke, N. (1994). Nineteenth-century vietnamese confucianization in historical perspective: Evidence from the palace examinations (1463-1883). *Journal of Southeast Asian Studies*, 25(2):270–312.
- Dao, D. A. (1964). *Dat Nuoc Viet Nam Qua Cac Doi (Yuenam Lidai Jiangyu, the Territory of Vietnam through the Ages)*. Ha Noi: Van Hoa Thong Tin.
- Dell, M., Lane, N., and Querubin, P. (2018). The Historical State, Local Collective Action, and Economic Development in Vietnam. *Econometrica*, 86:2083–2121.
- Dell, M. and Querubin, P. (2018). Nation Building through Foreign Intervention: Evidence from Discontinuities in Military Strategies. *Quarterly Journal of Economics*, 133:701–764.
- Dijkstra, L., Florczyk, A. J., Freire, S., Kemper, T., Melchiorri, M., Pesaresi, M., and Schiavina, M. (2021). Applying the Degree of Urbanisation to the Globe: A New Harmonised Definition Reveals a Different Picture of Global Urbanisation. *Journal of Urban Economics*, 125:103312.
- Dincecco, M., Fenske, J., Menon, A., and Mukherjee, S. (2021). Pre-Colonial Warfare and Long-Run Development in India. *The Economic Journal*, 132(643):981–1010.
- Do, Q.-A., Nguyen, K.-T., and Tran, A. N. (2017). One mandarin benefits the whole clan: Home-town favoritism in an authoritarian regime. *American Economic Journal: Applied Economics*, 9(4):1–29.
- Don, L. Q., editor (1776). *Phu Bien Tap Luc (Fubian Zalu, Various Records of Phu Bien)*. Hanoi: Vien Khoa Hoc Xa Hoi Viet Nam.

- Endres, K. W. (2001). Local Dynamics of Renegotiating Ritual Space in Northern Vietnam: The Case of the “Dinh”. *Sojourn: Journal of Social Issues in Southeast Asia*, 16(1):70–101.
- Galor, O. and Özak, Ö. (2016). The Agricultural Origins of Time Preference. *American Economic Review*, 106:3064–3103.
- General Statistics Office of Vietnam (2021a). 2009 Population Census.
- General Statistics Office of Vietnam (2021b). 2021 Economic Census.
- Gennaioli, N., La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2012). Human Capital and Regional Development. *The Quarterly Journal of Economics*, 128(1):105–164.
- Goldewijk, K. K., Beusen, A., Doelman, J., and Stehfest, E. (2017). Anthropogenic Land Use Estimates for the Holocene – HYDE 3.2. *Earth System Science Data*, 9:927–953.
- Gong, Y. and Jang, W. (1998). Culture and Development: Reassessing Cultural Explanations on Asian Economic Development. *Development and Society*, 27(1):77–97.
- GOTOP30 (1996). Global Digital Elevation Model (DEM). <https://earthexplorer.usgs.gov/>.
- Haines, D. W. (1984). Reflections of Kinship and Society under Vietnam’s Le Dynasty. *Journal of Southeast Asian Studies*, 15(2):307–314.
- Haudricourt, A.-G. (2010). The Origin of the Peculiarities of the Vietnamese Alphabet. *Mon-Khmer Studies*, 39:89–104.
- Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring Economic Growth from Outer Space. *American Economic Review*, 102:994–1028.
- Ho, H. A. (2021). Land Tenure and Economic Development: Evidence from Vietnam. *World Development*, 140:105275.
- Ho, H. A., Martinsson, P., and Olsson, O. (2022). The Origins of Cultural Divergence: Evidence from Vietnam. *Journal of Economic Growth*, 27:45–89.
- Hofstede, G. and Bond, M. H. (1988). The Confucius Connection: from Cultural Roots to Economic Growth. *Organizational Dynamics*, 16(4):5–21.
- Huynh, K. K. (1986). *Vietnamese Communism, 1925–1945*. Ithaca: Cornell University Press.
- Indochine Française Direction des Affaires Économiques (1888). *Tonkin et Nord-Annam (Tonkin and North Annam)*. Hanoi: Imprimerie d’Extrême-Orient.
- Jaax, A. (2020). Private Sector Development and Provincial Patterns of Poverty: Evidence from Vietnam. *World Development*, 127:104747.

- Joint Research Centre (JRC) and European Commission and Center for International Earth Science and Information Network (CIESIN) and Columbia University (2021). *Global Human Settlement Layer: Degree of Urbanization Settlement Model Grid*. New York: NASA Socioeconomic Data and Applications Center (SEDAC).
- Kahn, H. (1979). *World Economic Development: 1979 and Beyond*. New York: Routledge.
- Kung, J. K.-S. and Ma, C. (2014). Can Cultural Norms Reduce Conflicts? Confucianism and Peasant Rebellions in Qing China. *Journal of Development Economics*, 111:132–149.
- Launay, A. and Hausermann, G. R. (1888). *Les Missions Catholiques dans l'Indo-Chine Française (Catholic Missions in French Indo-China)*. Paris: Imp. A. Lemerancier et C.ie.
- Lee, K. Y. (2013). *The Wit & Wisdom of Lee Kuan Yew*. Singapore: Editions Didier Millet.
- Li, T. (2011). Imported Book Trade and Confucian Learning in Seventeenth and Eighteenth Century Vietnam. In Aung-Thwin, M. A. and Hall, K. R., editors, *New Perspectives on the History and Historiography of Southeast Asia: Continuing Explorations*, pages 167–182. Oxford: Routledge.
- Malesky, E. J. and Taussig, M. (2008). Where Is Credit Due? Legal Institutions, Connections, and the Efficiency of Bank Lending in Vietnam. *The Journal of Law, Economics, and Organization*, 25(2):535–578.
- McHale, S. (2002). Mapping a Vietnamese Confucian Past and Its Transition to Modernity. In Elman, B. A., Ooms, H., and Duncan, J. B., editors, *Rethinking Confucianism: Past and Present in China, Japan, Korea, and Vietnam*, page 397–430. California: UCLA Asian Pacific Monograph Series.
- Michalopoulos, S. and Papaioannou, E. (2013). Pre-Colonial Ethnic Institutions and Contemporary African Development. *Econometrica*, 81:113–152.
- Min, B. and Gaba, K. M. (2014). Tracking Electrification in Vietnam Using Nighttime Lights. *Remote Sensing*, 6:9511–9529.
- Murray, M. J. (1980). *The Development of Capitalism in Colonial Indochina (1870-1940)*. California: the University of California Press.
- Nam, N. (2018). The Noble Person and the Revolutionary: Living with Confucian Values in Contemporary Vietnam. In Ames, R. T. and Hershock, P. D., editors, *Confucianisms for a Changing World Cultural Order*, page 128–162. Hawaii: University of Hawaii Press.
- National Library of Vietnam (1909). *Dai Nam Nhat Thong Chi (Danam Yitong Zhi, the Comprehensive Chronicle of Great Nam)*.
- Ngo, D. T., editor (2006). *Cac Nha Khoa Bang Viet Nam 1075-1919 (Yuenan Lichao Kebang, List of Tiensi through Imperial Dynasties of Vietnam)*. Hanoi: Literature Publishing House.

- Ngo, S. L. (1479). *Dai Viet Su Ky Toan Thu (Dayue Shiji Quanshu, Complete Annals of Dai Viet)*. Cambridge: Harvard-Yenching Institute.
- Nguyen, N. H. (1998). The Confucian Incursion into Vietnam. In Slote, W. H. and de Vos, G. A., editors, *Confucianism and the Family*, pages 91–104. New York: State University of New York Press.
- Nguyen, T.-C. (2020). Private Academies and Confucian Education in 18th-Century Vietnam in East Asian Context: The Case of Phuc Giang Academy. In Vladimír Glomb, Eun-Jeung Lee, M. G., editor, *Confucian Academies in East Asia*, page 89–125. Leiden: Brill.
- Nguyen, T. C., Pham, V. T., and Nguyen, V. T. (2018). Buddhist Print Culture in Nineteenth-Century Northern Vietnam: A Case Study of the Woodblock Collection in Khe Hoi Temple. *Journal of Vietnamese Studies*, 13:51–78.
- Nguyen, T. M., editor (2000). *Ho Van Ten Nguoi Viet Nam (Surnames and Given Names of Vietnam)*. Ho Chi Minh: Nha Xuat Ban Khoa Hoc Xa Hoi Chi Nhanh Tai.
- NOAA (2009). GMT Intermediate Distance from Coast Project. *National Ocean Service and NASA Ocean Biology Processing Group*. <https://oceancolor.gsfc.nasa.gov/docs/%0Ddistfromcoast/>.
- NOAA (2014). DMSP-OLS Nighttime Lights Time Series Version 4. *DMSP Data Collected by US Air Force Weather Agency*.
- Nunn, N. and Puga, D. (2012). Ruggedness: The Blessing of Bad Geography in Africa. *The Review of Economics and Statistics*, 94:20–36.
- Nunn, N. and Wantchekon, L. (2011). The Slave Trade and the Origins of Mistrust in Africa. *American Economic Review*, 101:3221–3252.
- Open Development Vietnam (2016). Harbor and Port of Vietnam.
- Peng, Y. (2004). Kinship Networks and Entrepreneurs in China’s Transitional Economy. *American Journal of Sociology*, 109(5):1045–1074.
- Pezzutto, S. (2019). Confucianism and Capitalist Development: from Max Weber and Orientalism to Lee Kuan Yew and New Confucianism. *Asian Studies Review*, 43(2):224–238.
- Phan, H. C., editor (2005). *Lich Trieu Hien Chuong Loai Chi (Lichao Xianzhang Leizhi, Categorized Chronicles of Successive Dynasties)*. Hanoi: Vien Khoa Hoc Xa Hoi Viet Nam. Translated by Vien Su Hoc.
- Putnam, R. D. (2000). *Bowling Alone: the Collapse and Revival of American Community*. New York: Simon and Schuster.
- Quoc Su Quan Trieu Nguyen (1961). *Dai Nam Thuc Lu (Chronicle of Nguyen Dynasty)*.

- Redding, G. (1990). *The Spirit of Chinese Capitalism*. Berlin, Boston: De Gruyter.
- Resolution 17/2011/QH13 (2011). On the Land Use Matter Plan up to 2020 and National Five-Year (2011-2015) Land Use Plan.
- Service Géographique de l'Indochine (1929). *Réseau Routier de l'Indochine (Road Network of Indochina)*.
- Steer, L. and Sen, K. (2010). Formal and Informal Institutions in a Transition Economy: The Case of Vietnam. *World Development*, 38(11):1603–1615.
- Taylor, K. W. (1983). *The Birth of Vietnam*. California: University of California Press.
- Taylor, K. W. (2002). Vietnamese Confucian Narratives. In Elman, B. A., Ooms, H., and Duncan, J. B., editors, *Rethinking Confucianism: Past and Present in China, Japan, Korea, and Vietnam*, page 337–369. California: UCLA Asian Pacific Monograph Series.
- Tho, N. D., Nguyen, N. V., and Papin, P., editors (2004). *Dong Khanh Dia Du Chi (Tongqing diyuzhi, the Descriptive Geography of the Emperor Dong Khanh)*. Hanoi: Ban Do Press.
- Tran, T. K. (1971). *Nho Giao (Confucianism)*. NXB Van Hoa Thong Tin.
- Tran, T. P. H. (2013). From Liberally-Organized to Centralized Schools Education in Tonkin, 1885–1927. *Journal of Vietnamese Studies*, 8:27–70.
- Tu, W.-m., editor (1996). *Confucian Traditions in East Asian Modernity: Moral Education and Economic culture in Japan and the Four Mini-Dragons*. Cambridge: American Academy of Arts & Sciences.
- Tu, W.-m. (2000). Implications of the Rise of “Confucian” East Asia. *Daedalus*, 129(1):195–218.
- van de Walle, D. and Gunewardena, D. (2001). Sources of Ethnic Inequality in Viet Nam. *Journal of Development Economics*, 65:177–207.
- van Kippersluis, H. and Rietveld, C. A. (2018). Beyond Plausibly Exogenous. *The Econometrics Journal*, 21(3):316–331.
- VHLSS (2020). Vietnam Household Living Standards Survey 2010-2018. *General Statistics Office of Vietnam*.
- Vietnam Ministry of Education (1975). *Cach Mang Thang Tam (August Revolution)*. Hanoi: Second Publishing House of Survey Geodesy and Map.
- Weber, M. (1951). *The Religion of China: Confucianism and Taoism*. Free Press.
- Whitmore, J. K. (1977). Chiao-Chih and Neo-Confucianism: The Ming Attempt to Transform Vietnam. *Ming Studies*, 1977:51–92.

- Whitmore, J. K. (1997). Literati Culture and Integration in Dai Viet, C. 1430-C. 1840. *Modern Asian Studies*, 31:665–687.
- Whitmore, J. K. (2006). The Rise of the Coast: Trade, State and Culture in Early Dai Viet. *Journal of Southeast Asian Studies*, 37:103–122.
- Woodside, A. (1971). *Far East: Vietnam and China 1938-1954*. London: SAGE Publications.
- Woodside, A. (2002). Classical Primordialism and Historical Agendas of Vietnamese Confucianism. In Elman, B. A., Ooms, H., and Duncan, J. B., editors, *Rethinking Confucianism: Past and Present in China, Japan, Korea, and Vietnam*, page 116–136. California: UCLA Asian Pacific Monograph Series.
- Xu, Y. and Yao, Y. (2015). Informal Institutions, Collective action, and Public Investment in Rural China. *American Political Science Review*, 109(2):371–391.
- Zhang, C. (2020). Clans, Entrepreneurship, and Development of the Private Sector in China. *Journal of Comparative Economics*, 48:100–123.

Appendix

Table A1: Variable Definition and Data Source

Variable	Description	Source
Tiensi density (logged)	Total number of Tiensi holders in a district in 1426-1919 normalized by the average population of the district in 1400, 1500, 1600, 1700, 1800, and 1900. 0.01 is added to the normalized figure before taking log to avoid losing zero entries.	Ngo (1479)
Colonial road density 1929 (logged)	Total length of stone roads (in kilometres) paved by French colonizers by the year 1929 divided by the area of the districts (in square kilometres). 0.01 is added to the normalized figure before taking log to avoid losing zero entries.	Service Géographique de l'Indochine (1929)
Degree of urbanization 1975 (logged)	Area-weighted average degree of urbanization for a district in 1975. The underlying data are provided at the grid-cell level with a resolution of at 30×30 arc-seconds. Each grid cell is categorized into urban centre, dense urban cluster, semi-dense cluster, suburban or peri-urban cell, or one of the four classes of rural grid cells. A value of 100, 75, 50, 25, or 0 is assigned to each category, respectively. The district level degree of urbanization is obtained by taking the grid-cell-area-weighted average of all grids within a district. 0.01 is added to the normalized figure before taking log to avoid losing zero entries.	GHSL
Degree of urbanization 1990-2015 (logged)	Area-weighted average degree of urbanization for a district for the period 1990-2015. The underlying data are provided at the grid-cell level with a resolution of at 30×30 arc-seconds for 1990, 2000, and 2015. Each grid cell is categorized into urban centre, dense urban cluster, semi-dense cluster, suburban or peri-urban cell, or one of the four classes of rural grid cells. A value of 100, 75, 50, 25, or 0 is assigned to each category, respectively. The district level degree of urbanization is obtained by taking the grid cell area weighted average cell value of all grids within a district. The district level degree of urbanization is computed for the years 1990, 2000, and 2015, and then averaged across the three years to obtain the final value. 0.01 is added to the normalized figure before taking log to avoid losing zero entries.	GHSL
Nighttime light intensity 1992-2013 (logged)	Average nighttime light intensity of a district. The underlying data are available for 30×30 arc-seconds grids and for the period 1992-2013. The grid cell values are measured on an increasing scale from 0 to 63. The district level average nighttime light intensity for a certain year is computed as the average cell value of all grids within a district. The district level nighttime light intensity is computed for each year between 1992 and 2013, and then averaged across years to obtain the final value. 0.01 is added to the normalized figure before taking log to avoid losing zero entries.	NOAA (2014)
Household consumption per capita 2010-2018 (logged)	Average household consumption per capita for the period 2010-2018. The household consumption is measured as the total living expenditure, including the expenditure on education, health, food and non-food stuff, and durable goods, and other regular expenses such as housing, electricity and sanitation.	VHLSS (2020)
Absolute latitude	Absolute latitude of the centroid of a district.	Calculated using R
Distance to coast (logged)	Average nearest distance to coastline for a district in kilometres. The underlying data are provided at the grid-cell level with a resolution of at 0.01×0.01 degree. Each grid cell carries the value of the nearest distance from the centroid of grid cell to the coastline. The average value for a district is obtained by taking the average value of all grids within the district.	NOAA (2009)

Table A1 Continued: Variable Definition and Data Source

Variable	Description	Source
Elevation	Average elevation of a district in kilometres. The original elevation data are provided at the grid-cell level with a resolution of 30×30 arc-seconds. Each grid cell carries the value of the elevation above the sea. The average value for a district is obtained by taking the average cell value of all grids within the district.	GOTOP30 (1996)
Terrain ruggedness	Average ruggedness of terrain in a district. The original elevation data are provided at the grid-cell level with a resolution of 30×30 arc-seconds. Each grid cell carries the value of the square root of the sum of the squared differences between the elevations of the grid cell and all its surrounding grids. The district level value is obtained by taking the area-weighted average cell value of all grids within the district.	Nunn and Puga (2012)
Caloric suitability	Average caloric suitability for a district. The original data is available for 5×5 arc-minutes grids. Each grid carries a value which indicates the maximum potential caloric yields attainable given the set of crops that are suitable for cultivation in the grid in the post 1500 period. The average value for a district is obtained by taking the average of cell value of all grids within the district.	Galor and Özak (2016)
% Urban population 1400	Share of urban population for a district in 1400. The underlying data are provided for each 5×5 arc-minutes grids. Each grid carries a value of the count of urban population in the grids. We first calculate the total urban population of a district by summing up the values of grids lying within the district boundary and then divide the total urban population by total population calculated in the same manner.	GHSL
Buddhist temple density (logged)	Total number of Buddhist temples per hundred squared kilometres in the Nguyen dynasty.	Dai Nam Nhat Thong Chi
Catholic mission density (logged)	Total number of Catholic missions per hundred squared kilometres in 1888.	Launay and Hausermann (1888)
Agricultural concession density (logged)	Total number of agricultural concessions per hundred squared kilometres in 1927.	Indochine Française Direction des Affaires Économiques (1888)
Mineral concession density (logged)	Total number of mineral concessions per hundred squared kilometres in 1927.	Indochine Française Direction des Affaires Économiques (1888)
% Area within war zone of the August Revolution	Share of land areas within the war zones during the August Revolution in 1945.	Vietnam Ministry of Education (1975)
Distance to hermit scholar (logged)	The shortest distance from the border of a district to its nearest hermit scholar.	Lich Trieu Hien Chuong Loai Chi, and Dai Nam Thuc Luc
State history until 1500 (logged)	The time since the district was recorded in official chronicles in 1500.	Ho et al. (2022)
Province or prefecture seat	Dummy variable equal 1 if the district was prefecture or province seat in the Nguyen dynasty, 0 otherwise.	Dong Khanh Dia Du Chi
Distance to Hanoi (logged)	Geodesic distance from the district to Hanoi.	Calculated using R
Tax in cash per capita (logged)	Tax paid in cash per 1,000 population in the Nguyen dynasty.	Dong Khanh Dia Du Chi
Courier sites	Dummy variable equals 1 if the district had at least one courier site in the Nguyen dynasty, 0 otherwise.	Dai Nam Nhat Thong Chi

Table A1 Continued: Variable Definition and Data Source

Variable	Description	Source
Guanbao road	Dummy variable equals 1 if the district had at least one road for the delivery of official documents in the Nguyen dynasty, 0 otherwise.	Dong Khanh Dia Du Chi
Tributary route	Dummy variable equals 1 if the district lied along the tributary route between Vietnam and China in the Revival Le dynasty, 0 otherwise.	Phu Bien Tap Luc
Local market	Dummy variable equals 1 if the district had at least one local market in the Nguyen dynasty, 0 otherwise.	Dai Nam Nhat Thong Chi
Distance to nearest harbour (logged)	Geodesic distance from the district to the nearest harbour.	Open Development Vietnam (2016)
Literati culture	Dummy variable equals 1 if there was an expression for respect for literature or literati in the description of local customs in official chronicle in the Nguyen dynasty, 0 otherwise.	Dai Nam Nhat Thong Chi
Illiteracy rate (logged)	Share of illiterate individuals aged between 24-64 in a district in 2009 population census.	GSO (2009)
Average years of schooling (logged)	Average years of schooling of individuals aged between 24-64 in a district in 2009 population census.	GSO (2009)
% with secondary and above education (logged)	Share of individuals aged between 24-64 who have a secondary and above educational attainment in a district in 2009 population census.	GSO (2009)
Surname similarity (logged)	The probability of two randomly picked individuals having the same surname in a district in 2018.	VHLSS (2020)
Commune house density (logged)	Total number of commune houses per 100 squared kilometres in a district in 2021.	Economic Census (2021)
NGO density (logged)	Total number of non-government organizations per 100 squared kilometres in a district in 2021.	Economic Census (2021)
Education unit density (logged)	Total number of public facilities for education and training per 100 squared kilometres in a district in 2021.	Economic Census (2021)
Health unit density (logged)	Total number of health facilities per 100 squared kilometres in a district in 2021.	
Culture and sports unit density (logged)	Total number of public facilities for cultural and sports activities per 100 squared kilometres in a district in 2021.	Economic Census (2021)
Individual business density (logged)	Total number of individual business per 100 squared kilometres in a district in 2021.	Economic Census (2021)
% employed in private sector (logged)	Share of individuals aged between 24-64 who are employed in the private sector for a district in 2009 population census.	GSO (2009)
% employed in foreign sector (logged)	Share of individuals aged between 24-64 who are employed in the foreign sector for a district in 2009 population census.	GSO (2009)
% employed in public sector (logged)	Share of individuals aged between 24-64 who are employed in the public sector for a district in 2009 population census.	GSO (2009)
% employed in primary industry (logged)	Share of individuals aged between 24-64 who are employed in the primary industry for a district in 2009 population census.	GSO (2009)
% employed in secondary industry (logged)	Share of individuals aged between 24-64 who are employed in the secondary industry for a district in 2009 population census.	GSO (2009)
% employed in Tertiary industry (logged)	Share of individuals aged between 24-64 who are employed in the tertiary industry for a district in 2009 population census.	GSO (2009)

Table A2: Summary Statistics

Variables	Obs	Mean	SD	Min	Max
Tiensi density (logged)	217	-1.479	2.354	-4.605	2.777
Colonial road density 1929 (logged)	217	-3.141	1.096	-4.605	-0.969
Degree of urbanization 1975 (logged)	217	1.217	2.004	-4.605	4.463
Degree of urbanization 1990-2015 (logged)	217	1.834	1.820	-4.605	4.579
Nighttime light intensity 1992-2013 (logged)	217	0.595	1.910	-6.642	3.893
Household consumption per capita 2010-2018 (logged)	216	9.595	0.327	8.078	10.521
Absolute latitude	217	20.529	1.423	16.274	22.880
Distance to coast (logged)	217	3.796	1.260	-0.117	6.024
Elevation	217	0.174	0.247	0.002	1.073
Terrain ruggedness	217	1.209	1.362	0.006	5.347
Caloric suitability	217	9.677	1.195	6.867	12.389
% Urban population 1400	217	-3.002	2.583	-4.605	3.369
Buddhist temple density (logged)	217	-3.151	1.986	-4.605	2.404
Catholic mission density (logged)	217	-3.260	1.907	-4.605	0.838
Agricultural concession density (logged)	217	-3.110	2.045	-4.605	1.039
Mineral concession density (logged)	217	-4.435	0.714	-4.605	-0.278
% Area within war zone of the August revolution	217	-1.500	4.064	-4.605	4.605
Distance to hermit scholar (logged)	217	3.129	1.893	-4.605	5.565
State history until 1500 (logged)	217	6.066	0.256	4.787	6.127
Province or prefecture seat	217	0.323	0.469	0.000	1.000
Distance to Hanoi (logged)	217	4.335	1.178	-4.605	6.318
Tax in cash per capita (logged)	217	0.893	0.538	-1.190	2.141
Courier sites	217	0.253	0.436	0.000	1.000
Guanbao road	212	0.448	0.498	0.000	1.000
Tributary route	217	0.092	0.290	0.000	1.000
Local market	217	0.641	0.481	0.000	1.000
Distance to nearest harbour (logged)	217	2.534	2.333	-4.605	5.164
Literati culture	217	0.217	0.413	0.000	1.000
Illiteracy rate (logged)	217	0.905	1.099	-1.188	4.099
Average yearsz of schooling (logged)	217	2.151	0.116	1.866	2.519
% with secondary and above education (logged)	217	2.847	0.393	1.415	4.153
Surname similarity (logged)	216	2.747	0.467	1.556	4.368
Commune house density (logged)	217	0.546	3.441	-4.605	4.951
NGO density (logged)	217	3.284	1.032	-0.123	7.014
Education unit density (logged)	217	3.146	1.019	-0.123	6.350
Health unit density (logged)	217	-0.467	1.462	-4.605	4.195
Culture and sports unit density (logged)	217	-1.287	1.878	-4.605	4.013
Individual business density (logged)	217	6.025	1.279	-0.805	10.518
% employed in private sector	217	0.717	1.027	-1.994	2.999
% employed in foreign sector	217	-1.227	1.738	-4.605	1.998
% employed in public sector	217	2.200	0.479	1.129	3.709
% employed in primary industry	217	4.102	0.461	-0.227	4.486
% employed in secondary industry	217	2.516	0.774	0.056	3.646
% employed in Tertiary industry	217	2.893	0.417	1.900	4.323

Notes: This table presents the summary statistics for all variables used in the analyses.

Table A3: Pairwise Correlation Coefficients between Proxies for Economic Development

Variables	(1)	(2)	(3)	(4)	(5)
(1) Colonial road density 1929	1				
(2) Degree of urbanization 1975	0.661***	1			
(3) Degree of urbanization 1990-2015	0.761***	0.892***	1		
(4) Nighttime light intensity 1992-2013	0.740***	0.826***	0.915***	1	
(5) Household consumption per capita 2010-2018	0.514***	0.524***	0.585***	0.703***	1
Observations	216	216	216	216	216

Table A4: Information on Hermit Scholars

Name	Information on Year of Birth	Born after 1722	Teaching Activities and Students	Reason to Be Hermit	Source
Chu Van An	1292	No	Known as the teacher of all Confucian scholars	He was disappointed by the fact that the emperor refused to behead corrupted officials	Lich Trieu Hien Chuong Loai Chi
Nguyen Dinh Tru	1627	No	He taught thousands of students, 70 of whom obtained Tiensi degree	He left his position after being accused by the imperial court	Lich Trieu Hien Chuong Loai Chi
Nguyen Dang Dan	About 70 during 1802-1820	Yes	His students amounted to a hundred. He was known as "the master"	He was not attracted by fame and wealth but want to pursue kindness and truth	Dai Nam Thuc Luc
Chu Doan Tri	1779	Yes	His students followed him to work on the field	He determined not to pursue the success in civil exam and fame	Dai Nam Thuc Luc
Le Bat Trieu	Summoned by the emperor during 1802-1820	Yes	His students amounted to a thousand	He was not interested in taking civil exam	Dai Nam Thuc Luc
Le Man Duc	Admitted to district-level government school in 1821	Yes	His followers were mounting after he returned home to teach	He became bored of civil exam after passing the district level exam	Dai Nam Thuc Luc
Ton Duc Tien	Died at the age of 81 during 1847-1883	Yes	He was known for being knowledgeable. Hundreds of his students became well established	He was content with not being rich and stayed at home to teach after took the district level civil exam for several times	Dai Nam Thuc Luc
Bui Chu	Took the national level civil exam in 1826	Yes	He taught Confucianism and nurtured students	He refused to be admitted to the national academy	Dai Nam Thuc Luc
Le Khac Phoi	1772	Yes	Confucian scholars followed him to study	He refused to take civil exam when he reached his twenties	Dai Nam Thuc Luc

Table A5: Alternative Measurements and Sample

	Colonial road density 1929 (logged)		Degree of urbanization 1975 (logged)		Degree of urbanization 1990-2015 (logged)		Nighttime light intensity 1992-2013 (logged)		Household consumption per capita 2010-2018 (logged)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Panel A: Normalizing by taxable population in the Nguyen dynasty										
Tiensi density (logged)	0.087** (2.462)	0.363*** (2.657)	0.183*** (3.602)	0.564*** (4.338)	0.184*** (4.757)	0.549*** (4.885)	0.162*** (4.958)	0.476*** (2.985)	0.027** (2.516)	0.125** (2.317)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
adj. R^2	0.425	-0.286	0.804	0.483	0.873	0.361	0.850	0.375	0.426	-0.192
Observations	217	217	217	217	217	217	217	217	216	216
Panel B: Normalizing by land area										
Tiensi density (logged)	0.127*** (4.200)	0.312*** (2.957)	0.249*** (4.551)	0.422*** (5.933)	0.250*** (7.119)	0.472*** (7.425)	0.193*** (6.334)	0.410*** (3.691)	0.032* (2.068)	0.106*** (2.752)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
adj. R^2	0.438	-0.059	0.815	0.549	0.887	0.615	0.853	0.518	0.428	0.013
Observations	217	217	217	217	217	217	217	217	216	216
Panel C: Drop districts to the south of 17 Parallel										
Tiensi density (logged)	0.100** (2.295)	0.389** (2.315)	0.276*** (4.477)	0.507*** (4.807)	0.269*** (6.311)	0.592*** (3.889)	0.207*** (4.615)	0.495** (2.570)	0.029* (1.851)	0.134** (2.035)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
adj. R^2	0.443	-0.144	0.816	0.537	0.883	0.554	0.856	0.501	0.419	-0.072
Observations	206	206	206	206	206	206	206	206	205	205

Notes: Robust standard errors clustered at the province level are used and the corresponding t-statistics are reported in the parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively. Baseline controls include absolute latitude, distance to coast (logged), elevation, terrain ruggedness, caloric suitability and the share of urban population in 1400 (logged).

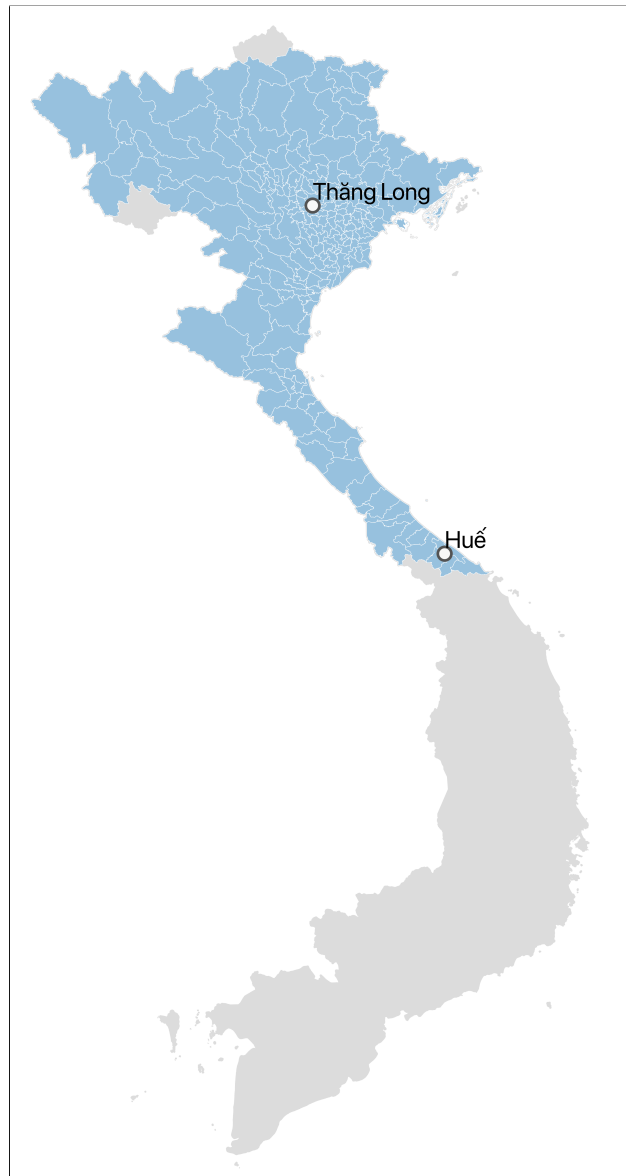
A1. Contemporary Locations of Historical Districts in the Nguyen Dynasty

To link the historical districts to their contemporary locations, we rely on four sources: (a). the list of historical districts of the Nguyen Dynasty in *Dong Khanh Dia Du Chi*, (b). the matching of historical districts to modern districts in 1960s by Dao (1964), (c). the boundary shapefile of communes in Vietnam in 2018 provided by GADM, and (d). the administrative division history provided by Wikipedia.¹

To recreate the boundaries of historical districts, we use the following steps: (a). find the corresponding district(s) in the early 1960s for each historical district using Dao (1964); (b). identify the component communes for the corresponding district(s) in the early 1960s using the administrative division history provided by Wikipedia; (c). link each component commune to its counterpart in the 2018 GADM shapefile; (d). create the boundary for the historical district by dissolving the boundaries of all component communes. We are able to reconstruct the boundaries for all historical districts listed in *Dong Khanh Dia Du Chi*, except for Trinh Co, Man Duy, and Sam Da, since their current locations are not provided by Dao (1964).

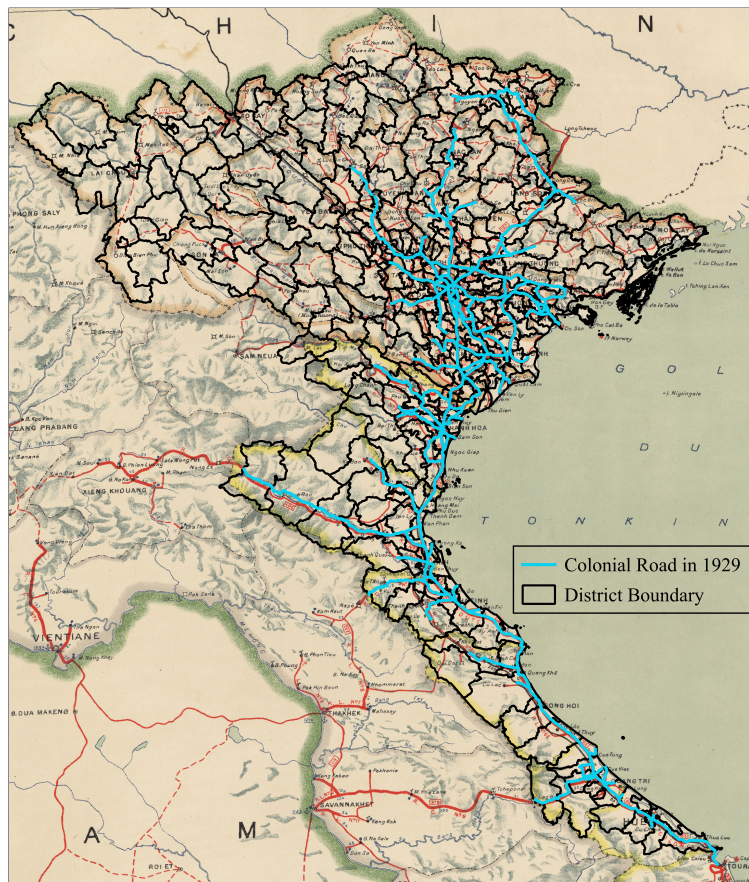
¹We use the Version 3.6 of the shapefile. It can be downloaded from <https://gadm.org/about.html>. Contemporary Vietnam administrative divisions have three levels: provinces are divided into districts, and districts into communes. Wikipedia provides the detailed boundary change history for each district. Every change is validated by relevant official documents.

Figure A1: The Sample



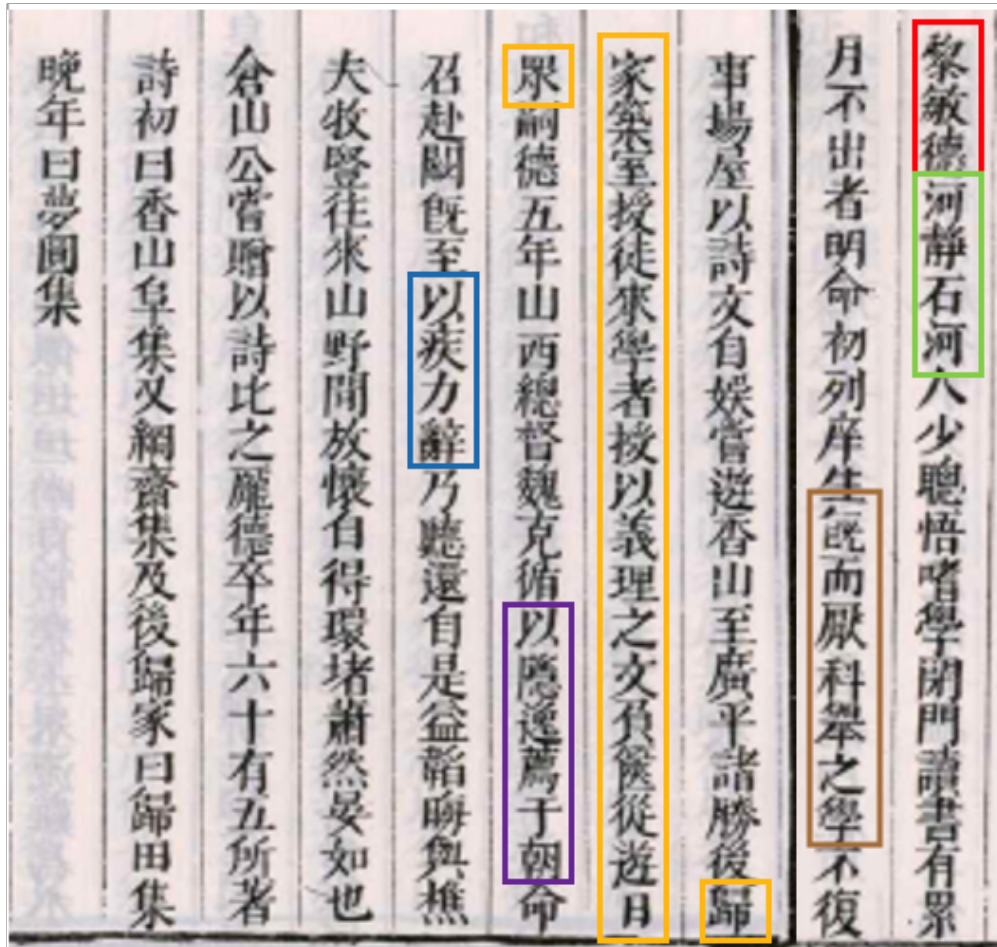
Notes: This figure depicts the boundaries of the districts used in our analyses. The sample districts are shaded in blue.

Figure A2: The Map for Colonial Road Network



Notes: This figure presents the map that is used to generate the density of colonial roads. The stone roads are depicted in blue. District boundaries are in black

Figure A3: An Example of Hermit Scholar's Biography



Notes: This figure presents a snapshot of a hermit scholar's biography, Le Man Duc. The content in the red frame indicates the name of the scholar. The green frame includes the place of birth. The brown frame reads that the scholar was tired of learning for civil exam. The yellow frame describes that "the scholar returned home, established a learning space, and taught followers with righteousness and principles. The number of followers were mounting as time goes by." The purple frame indicates that the scholar was recommended to the imperial court as a hermit scholar and asked to serve the court, whereas the blue frame suggests that the scholar turned down the offer of the imperial court due to (with an excuse of) unsound health conditions.