Journal of Comparative Economics 51 (2023) 2–14

Contents lists available at ScienceDirect

Journal of Comparative Economics

journal homepage: www.elsevier.com/locate/jce

# Legacies of Vietnam's imperial examinations, 1075–1919: More investment in education and better educational outcomes

Tien Manh Vu [a,*], Hiroyuki Yamada [b]

[a] Miyazaki International College, 1405 Kano-hei, Kiyotake-cho, Miyazaki–shi, Miyazaki, 889–1605, Japan
[b] Faculty of Economics, Keio University, 2–15–45 Mita, Minato–ku, Tokyo, 108–8345, Japan

ARTICLE INFO

*JEL classification:*
I25
N35
O15
Z1

*Keywords:*
Education
Human capital
Imperial examination
Historical legacy
Vietnam

ABSTRACT

This study measures the impact of the number of people who passed the Vietnamese imperial examinations (1075–1919) on the present-day quantity and quality of education in their home districts. Our analyses at the district and individual levels are based mainly on the data of successful imperial test takers, the 2009 population census, and 2009 National Entrance Exams to University test scores. We find a persistent legacy in educational attainment outcomes. Although there may be multiple channels that explain the long-term historical effects, a tradition of human capital investment and cultural elements are among the most important factors.

## 1. Introduction

Investment in human capital is important for economic development, and many models of economic growth take human capital into account (Goldin, 2019). Although the concept of investing in human capital gained prominence following Mincer (1958) and Schultz (1961), such investment has been practiced for centuries. In other words, the study of human capital is inherently historical (Goldin, 2019). This agrees with Nunn (2021)'s view that explaining differences in economic development requires consideration of historical roots.

However, research on investment in human capital over time is scarce and channels for the intergenerational transmission of human capital remain to be explored. Among the few previous studies, Valencia Caicedo (2019) finds that areas with a former Jesuit missionary presence in Guarani (present-day Argentina, Brazil, and Paraguay) from 1609 to 1767 still have higher educational attainment in the present day. In addition, Chen et al. (2020) shows that the population density of Chinese imperial elites (1368–1905) is associated with more years of schooling in the present day in 278 Chinese prefectures, indicating that social capital among clans and views on the value of education and schools would form cultural norms and spread outside the educated elite families via social organizations (e.g., charity organizations, farmer associations, labor unions, and religious associations).

Our study on Vietnam attempts to show that the transmission of human capital investment is crucial. This study seeks empirical evidence for whether the intensity of investment in human capital passed from generation to generation in Vietnam. Our analyses are possible because Vietnamese imperial exam records, particularly those recorded on stele stones (*Bia Văn Miếu* in Vietnamese), provide

* Corresponding author.
*E-mail address:* mvu@sky.miyazaki-mic.ac.jp (T.M. Vu).

https://doi.org/10.1016/j.jce.2022.09.004
Received 30 November 2021; Received in revised form 10 September 2022; Accepted 22 September 2022
Available online 7 October 2022

a long and uninterrupted historical dataset on the nation's top learners. The contents of the stele stones were copied in early 1900s by the French School of the Far East (*École Française d'Extrême-Orient* in French), which would have saved the records from being lost during catastrophic modern wars and conflicts in Vietnam.[1]

Here, we specifically investigate whether an association exists among test takers who passed the 1075–1919 imperial examinations at the national level[2] (hereinafter, "imperial elites") and between imperial elites and the educational achievements of people living in the present-day district that was home to these imperial elites. Our main analysis examines educational outcomes at two levels, namely, the district aggregated level (covering 688 districts) and the individual level (covering about 0.68 million university entrance examination test takers born in 1991). In each district, we consider both the number of imperial elites and the density of imperial elites per square kilometer as an alternative. We apply an instrumental variable (IV) approach for our analysis of both levels.

Our findings indicate a persistent relationship between the success of imperial elites and the quantity and quality of education of subsequent generations. More specifically, we find that the number and density of imperial elites are well connected across dynasties. Both the number and density of imperial elites during 1075–1919 in a given district are associated with a higher school attendance rate in compulsory primary school and twelfth grade, and more average years of schooling in the present day. In contrast, the correlations are negative when primary school dropouts and school non-enrollment are used as the outcomes. From the analysis at the individual level, we find that the number and density of imperial elites in the district led to higher age-standardized test scores (z-scores) in the 2009 National Entrance Exams to University (NEEU). Higher z-scores might also be a reason for a higher school enrollment rate among the district population aged 18–24 years. However, such a legacy might not necessarily be realized in certain clans in these districts. We find that the number of elites in a district is associated with a higher share of the district's average household living expenses reserved for educational purposes. We propose several catalysts for this persistent legacy, including returns on test scores, the strong dependence of wages on academic degrees (especially for university degrees), and the independence of villages from central government control in the present day.

Our paper contributes to the literature in several ways. First, to our knowledge, this paper is the first to examine the transmission of human capital investment and its impacts on present-day education in terms of the education quality, proxied by test scores. Second, we provide empirical evidence for persistence. We find an association across dynasties among the number of imperial elites from the same home district. We also find associations between the number of imperial elites (1075–1919) and present-day investment in education.

The remainder of the paper is organized as follows. We describe the data sources and the integration of the data for analysis in Section 2. Section 3 presents our identification strategy, the construction of the IV, and the methods. We report the results and discuss possible mechanisms behind the results in Section 4. Finally, Section 5 presents the conclusion and notes our research limitations.

## 2. Data

Our analyses are based on several important historical and contemporary data sources. Combining these with geographical data, we overlay the historical data on the contemporary data by using identical concordance district names to construct a dataset of 688 identical districts.

### 2.1. Historical data

We use information extracted from Ngo (2006), who collated the longest known list of imperial elites who passed the 1075–1919 imperial examinations at the national level (see Online appendix A.1 for a description). The list is based on previous research conducted by the French School of the Far East and the Shino-Vietnamese Institute (*Viện Hán Nôm* in Vietnamese) on ancient sources, including the stele stones and imperial examination records (*Đăng Khoa Lục* in Vietnamese). Ngo (2006) provides annotations of the home district of each imperial elite and the corresponding present-day district.

We use a list of 2888 names[3] and home districts of imperial elites obtained from the more than 180 recorded imperial examinations in Vietnam. However, up-to-date concordance geographical locations are available at only the district level for 2844 imperial elites (covering 198 districts).[4] We count the number of imperial elites in the list for each district to create our main variable, *elite number,* as well as the density of the imperial elites per square kilometer, *elite density*. The number of imperial elites per time period is shown in Online appendix A.2 and Fig. 2, and their distribution by geography is shown in Fig. 1.

[1] The French School of the Far East was established in Hanoi in 1900 (and still operates in Paris). It began the work of copying the contents of all available stele stones. Later, the Shino-Vietnamese Institute continued the work for the last examinations. The digitized data and related annotations are available at http://www.hannom.org.vn/default.asp?CatID=564.

[2] We refer to *Kỳ thi Đại khoa (thi Hội)* in Vietnamese as "imperial examinations." The Vietnamese imperial examinations were heavily influenced by the Chinese imperial examinations and were held 15 years longer. The examinations were used to select public officials to work in the imperial government, either at the central or local level. See Online appendix A.1 for further details.

[3] Ngo (2006) originally lists 2,894 names of imperial elites. However, there are a few duplications because some test takers took the test several times to improve their score and obtain the highest rank.

[4] Some home districts are missing or unidentified. In the 2009 population census, Vietnamese administrative divisions comprised 63 provinces (Level 1) and 689 districts (Level 2). The average district has a population of 126,227 people and an area of 434 $km^2$.

### 2.2. Contemporary data

#### 2.2.1. Quantity of education

We rely mainly on the 2009 Vietnamese Population and Housing Census (100% or 86.9 million observations—hereinafter, "population census") conducted by the General Statistics Office of Vietnam (GSO). The census included a question on school attendance status (as of April 1, 2009, which is during the school year), which measured the variable *attendance rate*. The census asked if the individual had ever attended school and whether, at the time of the census, they were still attending school. This information is used to create the variable *dropout rate*. If the individual had completed their education, the information on the highest grade or educational level attained by the individual was recorded. We convert this measure to *years of schooling* for each individual, and we consider the age cohorts that are most likely to have completed their education (e.g., $\geq$18, $\geq$22, and $\geq$25 years).

We calculate several outcomes by age cohort at the district level. We calculate the never enrollment rate (*never enroll*) by dividing the total number of individuals who had never attended school in a given age cohort by the district population of the same age cohort. Similarly, we consider *attendance rate*, *dropout rate*, and the *never enroll* rate at the district level for compulsory education (primary school) in the 7–10-year age cohort (see Online appendix A.1 for a detailed description of the education system). In addition, we consider school *attendance rate* in 2009 for individuals born in 1991 in connection with the NEEU in 2009.

We also merge identical households from the biannual Vietnam Household Living Standard Survey from 2002 to 2008 (conducted by the GSO) to obtain an additional outcome. We calculate the share of household expenditures on education and weight the share with the average age of the household and then aggregate the figure at the district level to obtain the mean household share of education expenditures over living expenditures.

#### 2.2.2. Quality of education

To assess the quality of education, we examine the data of university placement test scores in July 2009 (the NEEU). The NEEU is administered by the Vietnamese Ministry of Education and Training at the same time and date, with the same test problems, using a centralized anonymous scoring system (see Online appendix A.1 for a detailed description). About 980,000 test takers are recorded in the NEEU data (Vu, 2022) according to 11 pre-determined testing groups. We analyze the four main groups, which represent about 95% of all 2009 NEEU test takers: A (Mathematics, Physics, and Chemistry), B (Biology, Mathematics, and Chemistry), C (Literature, History, and Geography), and D (Literature, Mathematics, and Foreign Language). The sum of the three scores (hereinafter, the "total test score") is used to determine university placement.

The NEEU test scores are the best indications of human capital and investment in quality of education. The entrance examinations are held only once a year. Not surprisingly, they are extremely competitive (Vu, 2022). In 2009, the ratio of available seats to test takers ranged from 1:2 to 1:23. Test takers are expected to devote their best effort to all three subjects, for which the pass/fail threshold is never known prior to the tests.

Using the home districts declared by the NEEU test takers on their test application forms, we link the test scores for groups A, B, C, and D with the home districts of imperial elites. We consider only the test takers born in 1991 because they account for about 73% of all NEEU test takers in 2009. Also, most of these test takers were likely taking the NEEU for the first time in their lives.[5] We use the three subject scores and the total scores. More precisely, we use the standardized ($z-score$) test score for each group as an outcome (hereinafter, "z-score") at the individual level. The descriptive statistics of each group are provided in Online appendix A.4.

The act of taking the NEEU presents a selection issue. We consider this selection in the probability of taking the NEEU after the twelfth grade (*taking NEEU after* 12*th grade*) at the district level by dividing the sum of (identical) test takers by the aggregated number of individuals who were born in 1991 and were recorded as attending school as of April 1, 2009 in the population census.

Based on Spolaore and Wacziarg (2013), we use geographical information system (GIS) data from several sources to create control variables at the district level. We prefer the oldest geographical data because they best represent the past characteristics of the districts. The 1992 Global Land Cover Characterization from the United States Geological Survey Earth Resources Observation and Science Center (USGS EROS) is used to generate the urbanization land ratio and cropland ratio for each district. From the USGS EROS Center's Landsat Imagery in 1996, we generate the average elevation for each district. We also calculate the distance from each present-day district center to the coastline (*distance to coastline*) using the shape file provided by the Database of Global Administrative Areas.[6]

Finally, for our analysis we use a dataset on 688 districts as well as a dataset on individuals and their test scores for each group (Online appendix A.3 and Online appendix A.4).

## 3. Methods

First, using the database of imperial elites, we conduct a persistence analysis of the number of elites in the past. We investigate whether the number of elites from home district $i$ who sat for imperial exams during the Nguyen Dynasty (1822–1919)—the last dynasty to use imperial exams per home district ($nelite_{Nguyen,\ i}$)—has any correlation with the number of elites in the same home district who sat for imperial exams during previous $t$ dynasties ($Elite\ number_{t,\ i}$) (Online appendix A.2 for the specifics). Specifically, we estimate the following ordinary least squares (OLS) model:

[5] We choose individuals who graduated from high school in May 2009.

[6] Available at: www.gadm.org

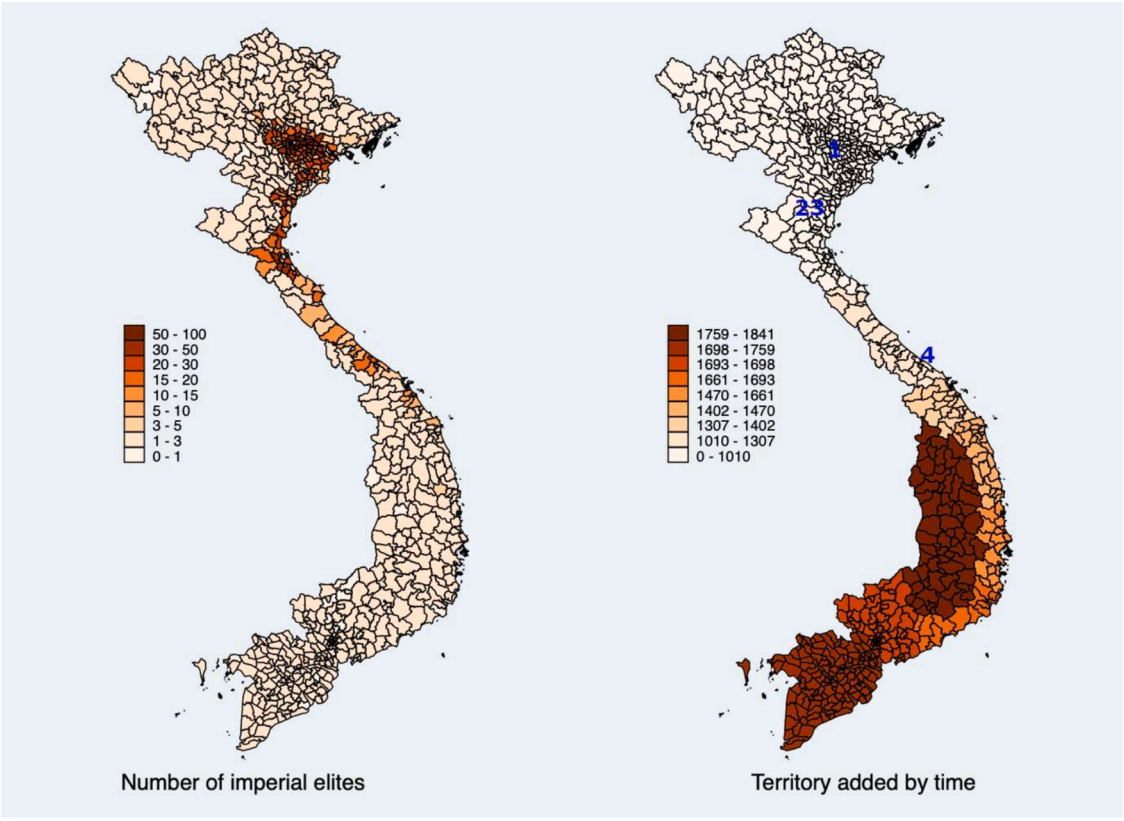


**Fig. 1.** Number of imperial elites and approximate in-land territory expansion timeline.
Notes: 1: Ba Dinh, Hanoi; 2: Tho Xuan, Thanh Hoa; 3: Vinh Loc, Thanh Hoa; 4: Hue city, Thua Thien Hue. The map shape corresponds with the 2015 division. Timelines are approximations based on Tran (1920), Tran et al. (2017), and Truong et al. (2017) and for the calculation of the IV only.

$$Elite\ number_{Nguyen,\ i} = \beta_0 + \beta_1.\ Elite\ number_{t,i} + \beta_2.C_i + \varphi_i, \tag{1}$$

where $C_i$ is a set of control variables, namely, the distance to the coastline and the mean elevation in 1996.

Next, to examine the persistence from the past to the present day, we estimate whether the number of imperial elites from a given home district ($i$) influences present-day outcomes in that district (district people), using an OLS model based on a reduced form equation:

$$Outcome_i = \alpha_0 + \alpha_1.\ Elite\ number_i + \alpha_2.control_i + \varepsilon_i. \tag{2}$$

We add Kinh ethnic ratio, population density in 2009, urban land ratio, and cropland ratio in 1992 to the district control variables mentioned in (1) to form the $control_i$.[7] We also use the density of imperial elites per square kilometer in the present-day district area as an alternative.

Unobservable factors might result in endogeneity issues. For example, the number of elites might be correlated with the abilities of both past and present-day people. Therefore, we use an IV approach. Our *IV* is the product of distance and a weight for those who did not take the imperial examinations

$$IV_i = distance_i \times weight_i, \tag{3}$$

The distance is the average distance from each district $i$ (=1, …, 688) to each imperial examination test venue's district $j$ during 1075–1919

$$distance_i = \frac{1}{n}\sum_{k=1}^{n} Arc\ distance\ \left[district_i; test\ venue_{jk}\right], \tag{4}$$

For each imperial exam year, we calculate the arc distance between the present-day center of each district $i$ to the district $j$ where

[7] We also had a specification without any GIS control variables, as shown in Online appendix A.9, Online appendix A.11, Online appendix A.14, and Online appendix A.16.

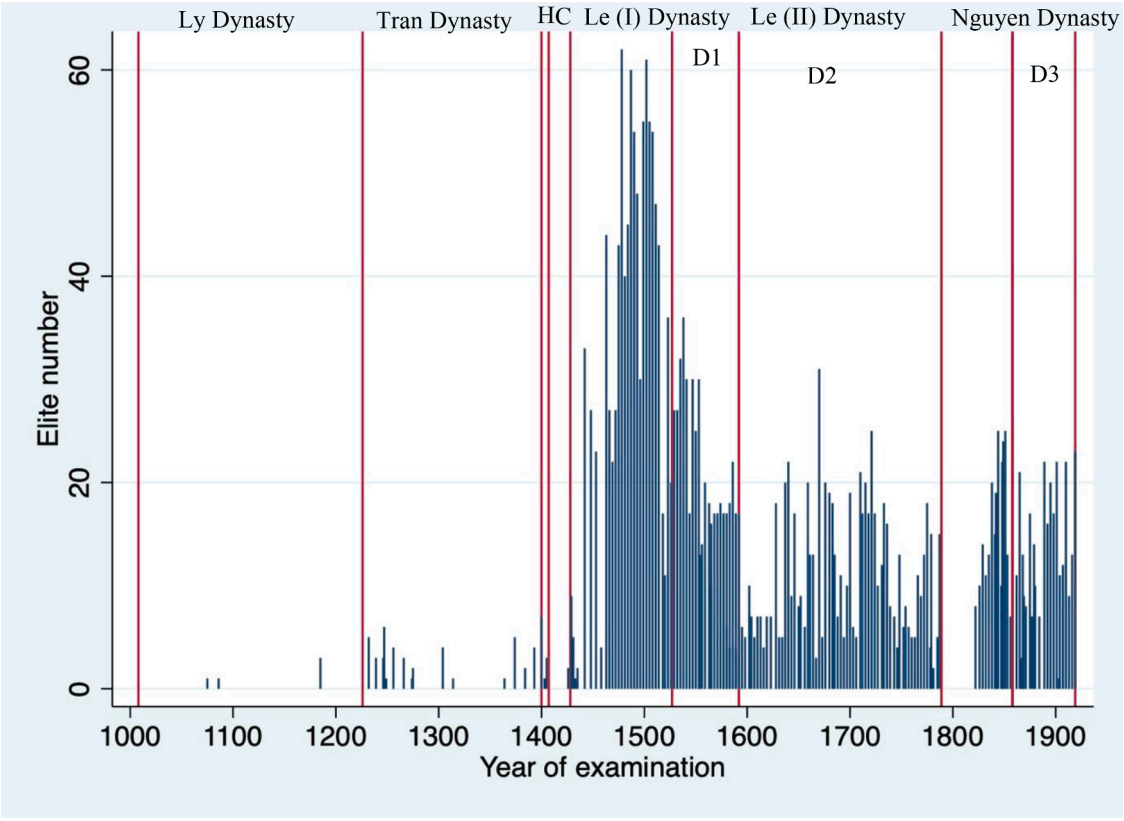


**Fig. 2.** Number of national winners in the imperial examinations by time.
Notes: *Ly* Dynasty (1009–1225). *Tran* Dynasty (1226–1399). H: *Ho* Dynasty (1400–1407). C: Chinese domination (1407–1427). *Le* (I) Dynasty (1428–1526). D1: Division 1: *Mac* Dynasty (1527–1677) with the capital in Ba Dinh, Hanoi until 1592. *Le* Dynasty during division 1 (1533–1789) with the capital in Tho Xuan, Thanh Hoa from 1533 to 1597. *Le* (II) Dynasty and D2: *Le* (II) Dynasty (*Le Trung Hung*) (1533–1789) with division 2. *Le* Dynasty was under the control of *Trinh* lords (1570–1786), with the capital in Ba Dinh, Hanoi. *Nguyen* Lords (later *Nguyen* Dynasty) (1600–1945) with the capital mainly in Hue. *Mac* Dynasty (1527–1677) with the capital in Cao Bang City today from 1593 to 1677. *Tay Son* Dynasty (*Nguyen* brothers) (1778–1802) had capitals in Hue and Binh Dinh today, but no imperial examinations at the national level. *Nguyen* Dynasty (1600–1945) with the capital in Hue. D3: Division 3. French colony (1858–1945). *Nguyen* Dynasty (1600–1945) with the capital in Hue. The last examinations were conducted in 1919.

the imperial examinations were held.[8] In the data for the estimations, the number of the imperial examinations $n$ is 179. Based on Tran (1920), Tran et al. (2017), and Truong et al. (2017), we set the timeline of each dynasty and the corresponding capital, combining the information from Ngo (2006) to determine the imperial examination test venues, which were always where the concurrent emperor was living. This was because the emperor made the final decision on who passed the imperial examination and interviewed the top-ranked candidates (Ngo, 2006). Eight different test locations existed in the period. Vietnam has experienced multiple national divisions and has had two or even three "capitals" concurrently. The emperor of each region held their own respective imperial examinations (Ngo, 2006). Ordinary citizens could move freely to take the tests under different emperors, which were even given in the same year (Ngo, 2006). We take the average of these distances for each district to form the first component of the IV.

The weight is the inverse share of the number of imperial examinations that people in the district could take in practice

$$weight_i = \frac{n}{Number\ of\ eligible\ imperial\ exams_i}, \tag{5}$$

where $n = 179$. Since 1075, the territory of Vietnam has almost doubled in size (see Fig. 1). However, people living in newly acquired territory did not have any incentive to take the imperial examinations when the territory did not belong to Vietnam. Therefore, we assume that the weight accounts for the time during which the area was not part of Vietnam's territory. The shorter the time, the fewer opportunities to take an imperial exams in the district. The weight becomes bigger and adds to the typical cost of moving to the test

[8] We acknowledge the difference between contemporary district boundaries and the corresponding historical ones in each dynasty. However, we assume contemporary district boundaries as proximate common areas to the village from where the imperial elite originated. Each district has an area of 434 $km^2$ on average, or a circle with a 12-km radius.

venue for each imperial examination—the average distance. Thus, we expect the IV to have a negative correlation with the number of elites in the first stage.

The product of the average distance and the weight could be a valid IV for several reasons. The product denotes the cost of education for imperial examination takers. The average distance was important. The narrow S-shaped territory of Vietnam is separated by high mountains and rivers between the north and the south. The weight is related to historical events and is exogenous to people in the past. In addition, the product has a strong correlation with the number or density of elites, as shown in Online appendix A.5, regardless of the specifications.

An exclusion restriction generally refers to a condition in which an IV does not have a direct influence on the outcome variables. It is well known that in the case of one endogenous variable and one instrument, the exclusion restriction cannot be tested. Therefore, we argue the plausibility of the exclusion restriction. First, we control for various geographical variables in our main specification. Second, most of the imperial exam locations, including Hue, Vinh Loc, and Tho Xuan, play an insignificant role in present-day economic activities. One proxy for economic activity and development is the logarithm of nighttime light intensity. Using the logarithm of nighttime light intensity in 2009, we find that most of the imperial test districts were not among the highest ranked areas in terms of economic activity in Vietnam (see Online appendix A.2). While regressing the logarithm of nighttime light intensity on the product shows statistically significant correlations (see column 2 of Online appendix A.6), the raw statistical correlation between the logarithm of nighttime light intensity and the product is merely 0.036, which is negligible. Although we do not include nighttime light in our main specification, omitting this variable is unlikely to influence the main outcomes.[9] School availability is discussed below in the Results section.

We modify (2) to get (7) below using the IV approach. Specifically, in the first stage, we estimated

$$Elite\ number_i = \omega_0 + \omega_1 . IV_i + \omega_2 . C_i + \in_i, \tag{6}$$

In the second stage, we estimate a model similar to (2) as

$$Outcome_i = \gamma_0 + \gamma_1 . \widehat{Elite\ number}_i + \gamma_2 . C_i + \tau_i, \tag{7}$$

where $C_i$ is the set of control variables. $C_i$ in Eqs. (6) and (7) uses the same controls as in Eq. (2). We also use Eqs. (6) and (7) for estimating the other outcomes, such as the probability of attending school and the probability of taking the 2009 NEEU among those born in 1991 who also completed high school.

Using the same procedure as in (6) and (7) but at the individual level, we regress the z-scores of the 2009 NEEU test takers on the number of imperial elites in the corresponding districts. In addition, $C_i$ includes a gender dummy and a dummy indicating whether the individual was ineligible for some contemporary educational policy privileges (1 if ineligible and 0 otherwise). Those eligible for educational policy privileges were children of soldiers wounded or killed in combat and other national heroes (according to Vietnamese government regulations).

We report estimations in which the number of imperial elites are used as an IV for the main results. We use the density of imperial elites as an alternative (see Online appendix A.10, Online appendix A.11, Online appendix A.15, and Online appendix A.16). We also use the number of imperial elites recorded in the stele stones as another alternative (see Online appendix A.7, and Online appendix A.8).

## 4. Results

### 4.1. Persistent correlations across imperial elites by dynasty

We find a strong correlation between numbers of elites from the same home district across dynasties, as shown in Table 1. The number of elites sitting for imperial exams during the Nguyen Dynasty (1822–1919) can be well explained by the corresponding number of elites in the same home district in previous dynasties. The results are robust with or without controls. The only exception is for elites sitting for exams during the Ly Dynasty (1075–1225) (only 11 elite members were recorded during this period).

### 4.2. Persistent effects on present-day quantity of educational attainment

We find that the number of imperial elites in a given district is associated with a higher present-day quantity of educational attainment, as shown in Table 2. The results are robust even when we use density of imperial elites, as shown in Online appendix A.10, and Online appendix A.11. An additional imperial elite is associated with an additional 0.15 years of schooling among those aged $\geq$22 and $\geq$25 years, as shown in columns (4) and (6) of Table 2.

The results from the first stage of the IV approach (column 7, 8, and 9 in Table 2) support our arguments on the validity of the IV. The correlation between the IV and the number of elites is statistically significant at the 1% level. The F-statistic of the first stage is sufficiently large (128.72). The coefficient of the IV (column 7 in Table 2) indicates that, for example, a district belonging to Vietnam since 1010 (*weight*=1) and about 450 km on average away from all test venues (*average distance* = 450; *IV*=450) would bear a higher

[9] The results using nighttime light intensity as a control are shown in Online appendix A.17–A.19. The interpretations of these results are similar to those in our main specifications.

**Table 1**
Transmission to Nguyen Dynasty's imperial exams (1822–1919).

| Elite number in | Nguyen Dynasty(1822–1919) | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Le Dynasty II | | | | | | 0.2696*** |
| (1554–1787) | | | | | | (0.0473) |
| Mac Dynasty | | | | | 0.1700*** | |
| (1529–1592) | | | | | (0.0350) | |
| Le Dynasty I | | | | 0.1317*** | | |
| (1426–1526) | | | | (0.0265) | | |
| Ho Dynasty | | | 1.4207*** | | | |
| (1400–1405) | | | (0.3820) | | | |
| Tran Dynasty | | 1.4620*** | | | | |
| (1232–1393) | | (0.5033) | | | | |
| Ly Dynasty | 0.5377 | | | | | |
| (1075–1225) | (0.6424) | | | | | |
| Controls | No | No | No | No | No | No |
| R-squared | 0.001 | 0.021 | 0.035 | 0.062 | 0.034 | 0.166 |
| | (7) | (8) | (9) | (10) | (11) | (12) |
| Le Dynasty II | | | | | | 0.2625*** |
| (1554–1787) | | | | | | (0.0471) |
| Mac Dynasty | | | | | 0.1574*** | |
| (1529–1592) | | | | | (0.0348) | |
| Le Dynasty I | | | | 0.1247*** | | |
| (1426–1526) | | | | (0.0265) | | |
| Ho Dynasty | | | 1.3454*** | | | |
| (1400–1405) | | | (0.3816) | | | |
| Tran Dynasty | | 1.3081*** | | | | |
| (1232–1393) | | (0.5003) | | | | |
| Ly Dynasty | 0.3529 | | | | | |
| (1075–1225) | (0.6101) | | | | | |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| R-squared | 0.039 | 0.056 | 0.070 | 0.092 | 0.067 | 0.190 |
| N districts | 688 | 688 | 688 | 688 | 688 | 688 |

Notes: Controls are mean elevation in 1996, and distance to the coastline. Robust standard errors are in parentheses (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$).

cost of education compared with a counterfactual district that is always the venue for the imperial exams (national level). The cost is equivalent to having two fewer imperial elites during the whole period 1075–1919. However, the interpretation becomes complicated when the district is farther south and has spent less time as Vietnamese territory. Persons living in such a district had both fewer opportunities to take imperial exams and a longer distance to travel to the imperial test venues, which would also lead to a negative coefficient for the IV.[10]

### 4.3. *Persistent effects on present-day effort to prolong years of schooling*

We find that the number of imperial elites in a given district has a significant positive relationship with school attendance rate in 2009 but a significant negative relationship with school dropouts and non-enrollment, as shown in columns (8) and (10) of Table 3. The results suggest that the presence of imperial elites is associated with greater focus and effort in prolonging education at the very early stages of life in the present day. In addition, primary school attendance is compulsory and free in Vietnam; therefore, the dropout and non-enrollment rates are considered pure because they are unlikely to be distorted by the cost of learning.

### 4.4. *Persistent effects on present-day quality of education*

We find that the sum of z-scores of the 2009 NEEU test takers is associated with the number of imperial elites who shared the same home district. An additional imperial elite in the district leads to an increase of up to 0.02 standard deviation on test scores, as shown in Table 4. The results are consistent regardless of group, test length, and subject (except for Biology, History, and Foreign Language).

In addition, results using Mathematics best reflect the persistent legacy on quality of education and human capital. Mathematics was a compulsory subject for all grades. Most test takers (about 92%) had to take the Mathematics portion of the NEEU.

[10] The kernel density of the IV is shown in Online appendix A.20.

**Table 2**
Effects on present-day quantity of education.

| Age cohort | ≥18 | | ≥22 | | ≥25 | |
|---|---|---|---|---|---|---|
| Variables | Years of schooling | | | | | |
| | OLS | IV 2nd stage | OLS | IV 2nd stage | OLS | IV 2nd stage |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Elite number | 0.0333*** | 0.1668*** | 0.0298*** | 0.1518*** | 0.0297*** | 0.1522*** |
| | (0.0053) | (0.0164) | (0.0043) | (0.0145) | (0.0043) | (0.0146) |
| R-squared | 0.314 | | 0.350 | | 0.347 | |
| Variables | | Elite number IV 1st stage (7) | | Elite number IV 1st stage (8) | | Elite number IV 1st stage (9) |
| IV | | −0.0046*** | | −0.0046*** | | −0.0046*** |
| | | (0.0004) | | (0.0004) | | (0.0004) |
| 2009-Kinh ethnic ratio | | 9.2805*** | | 9.2805*** | | 9.2805*** |
| | | (1.0693) | | (1.0693) | | (1.0693) |
| 2009-Population density | | 0.2357** | | 0.2357** | | 0.2357** |
| | | (0.0924) | | (0.0924) | | (0.0924) |
| 1992-Urban land ratio | | −4.6031 | | −4.6031 | | −4.6031 |
| | | (4.5709) | | (4.5709) | | (4.5709) |
| 1992-Cropland ratio | | 12.3863*** | | 12.3863*** | | 12.3863*** |
| | | (1.6230) | | (1.6230) | | (1.6230) |
| 1996-Mean Elevation | | 0.0053*** | | 0.0053*** | | 0.0053*** |
| | | (0.0011) | | (0.0011) | | (0.0011) |
| Distance to coastline | | −0.0002 | | −0.0002 | | −0.0002 |
| | | (0.0035) | | (0.0035) | | (0.0035) |
| Constant | | −5.0532*** | | −5.0532*** | | −5.0532*** |
| | | (1.0951) | | (1.0951) | | (1.0951) |
| F-statistics† | | 128.72 | | 128.72 | | 128.72 |
| N districts | 688 | 688 | 688 | 688 | 688 | 688 |

Notes:
† Kleibergen-Paap Wald rk F statistic for testing H0: Weak identification test. Robust standard errors are in parentheses (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). All estimations include the Kinh ethnic ratio, population density in 2009, urban land ratio, cropland ratio in 1992, mean elevation in 1996, and distance to the coastline as controls.

### 4.5. Robustness checks

We try several specifications to perform robustness checks. First, we conduct all estimations using the density of imperial elites per square kilometer as the main explanatory variable and find that the main results are consistent (see Online appendix A.10, Online appendix A.11, Online appendix A.15, and Online appendix A.16).

Second, we re-run the estimations without using district GIS control variables. The corresponding results shown in Online appendix A.9, Online appendix A.11, Online appendix A.14, and Online appendix A.16 are similar to those of the original estimations. Of the outcomes, the only exceptional case is the rate of taking the NEEU in 2009.

Third, we consider spatial autocorrelation, as suggested by Conley (1999). Districts might be clustered according to geography rather than administrative boundaries. We apply the procedures of Colella et al. (2019) and try different assumptions for the distance beyond the observations belonging to the same cluster, including 25, 50, and 100 km. We repeat this with the specifications outlined in Tables 2, 3, and 4, and, once again, using the density of imperial elites as in Online appendix A.12, and A.13. The standard errors change slightly; however, the results of the new estimations are consistent with the main results except when the outcome is based on the NEEU taking rate.

Fourth, we account for potential measurement errors due to historical record keeping.[11] For a robustness check, we use the list of imperial elites recorded on stele stones as an alternative variable. Since 1442, the emperors had the names of imperial elites who passed the national exams in each year carved in a stele stone and placed in the Temple of Literature in Hanoi and (later) in Hue.[12] The list of elites carved in the stele stones would not be influenced by regional factors such as (local) wars, anti-feudalism, and record-keeping ability. If a stone was lost, it is likely that the list for all regions in a given year would be lost rather than for a specific region. We re-estimate Tables 2, 3, and 4 using only the number of elites recorded in the stele stones from Hanoi and Hue, rather than the dataset from Ngo (2006). This new variable would be free from regional variation in record keeping. We find that the main interpretations of the main specification are still valid for these exercises (see Online appendix A.7 and Online appendix A.8).

[11] For example, the number of elites in each area in the dataset compiled by Ngo (2006) might be positively correlated with better record keeping or a region with better administration.

[12] We crosschecked the digitized data of the stele stones (available at http://www.hannom.org.vn/default.asp?CatID=564) with Ngo (2006). The number of elites from the stele stones accounted for about 55% of Ngo's (2006) collection (Ngo was a researcher at the Shino-Vietnamese Institute). In addition, in Ngo's (2006) list, the number of elites prior to 1442 (the first stele stone) was 80.

**Table 3**
Effects on prolonging years of schooling.

| Variables | Never-enroll rate | Dropout rate | Attendance rate | | | Taking NEEU after 12th grade | Mean share of education expenditure over living expenditure |
|---|---|---|---|---|---|---|---|
| Age cohort | 7–10 | 7–10 | 7–10 | 18–24 | 1991-born | 1991-born | All |
| **OLS** | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Elite number | −0.0002*** | −0.0004*** | 0.0006*** | 0.0008* | 0.0051*** | 0.0011*** | 0.0003*** |
| | (0.0000) | (0.0000) | (0.0001) | (0.0005) | (0.0007) | (0.0003) | (0.0001) |
| R-squared | 0.354 | 0.209 | 0.324 | 0.313 | 0.310 | 0.523 | 0.193 |
| **IV 2nd stage** | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| Elite number | −0.0017*** | −0.0018*** | 0.0034*** | 0.0073*** | 0.0241*** | 0.0038*** | 0.0013*** |
| | (0.0002) | (0.0002) | (0.0004) | (0.0011) | (0.0022) | (0.0009) | (0.0002) |
| F-statistics[†] | 128.84 | 128.84 | 128.84 | 128.72 | 128.84 | 128.88 | 130.40 |
| N districts | 687 | 687 | 687 | 688 | 687 | 682 | 590 |

Note: Same as Table 2.

**Table 4**
Effects on quality of education using the 2009 NEEU z-scores.

| **Test group** | **A** | | | | **B** | | | |
|---|---|---|---|---|---|---|---|---|
| Variables | Sum | Physics | Mathematics | Chemistry | Sum | Biology | Mathematics | Chemistry |
| **OLS** | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Elite number | 0.0043*** | 0.0049*** | 0.0023*** | 0.0042*** | 0.0051*** | 0.0000 | 0.0066*** | 0.0045*** |
| | (0.0006) | (0.0008) | (0.0005) | (0.0006) | (0.0009) | (0.0005) | (0.0011) | (0.0008) |
| R-squared | 0.036 | 0.035 | 0.029 | 0.021 | 0.058 | 0.020 | 0.063 | 0.041 |
| **IV 2nd stage** | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) |
| Elite number | 0.0205*** | 0.0225*** | 0.0114*** | 0.0201*** | 0.0216*** | −0.0035*** | 0.0292*** | 0.0202*** |
| | (0.0022) | (0.0022) | (0.0017) | (0.0021) | (0.0023) | (0.0013) | (0.0028) | (0.0023) |
| F-statistics† | 151.71 | 151.71 | 151.71 | 151.71 | 147.78 | 147.78 | 147.78 | 147.78 |
| N test takers | 341,085 | 341,085 | 341,085 | 341,085 | 178,840 | 178,840 | 178,840 | 178,840 |
| **Test group** | **C** | | | | **D** | | | |
| Variables | Sum | Literature | History | Geography | Sum | Literature | Mathematics | Foreign language |
| **OLS** | (17) | (18) | (19) | (20) | (21) | (22) | (23) | (24) |
| Elite number | 0.0026*** | 0.0057*** | 0.0001 | 0.0014 | 0.0038*** | 0.0082*** | 0.0028*** | −0.0017** |
| | (0.0009) | (0.0009) | (0.0009) | (0.0009) | (0.0010) | (0.0013) | (0.0009) | (0.0008) |
| R-squared | 0.038 | 0.089 | 0.019 | 0.020 | 0.062 | 0.095 | 0.047 | 0.063 |
| **IV 2nd stage** | (25) | (26) | (27) | (28) | (29) | (30) | (31) | (32) |
| Elite number | 0.0186*** | 0.0351*** | 0.0022 | 0.0119*** | 0.0270*** | 0.0401*** | 0.0212*** | 0.0021 |
| | (0.0023) | (0.0033) | (0.0019) | (0.0022) | (0.0035) | (0.0042) | (0.0028) | (0.0024) |
| F-statistics† | 155.98 | 155.98 | 155.98 | 155.98 | 124.41 | 124.41 | 124.41 | 124.41 |
| N test takers | 50,778 | 50,778 | 50,778 | 50,778 | 111,104 | 111,104 | 111,104 | 111,104 |

Notes:.
† Kleibergen-Paap Wald rk F statistic for testing H0: Weak identification test. Standard errors, clustered at the district level, are in parentheses (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). All estimations include *gender*, *no privilege*, the Kinh ethnic ratio, population density in 2009, urban land ratio, and cropland ratio in 1992, mean elevation in 1996, and distance to the coastline as controls.

### 4.6. Possible channels for persistence

Understanding the mechanisms of persistence is important but challenging (Voth, 2021). This is because there may be multiple channels, including formal institutions (Acemoglu et al., 2002), geography, and historical events, as well as informal institutions, such as trust in institutions (Becker et al., 2016).

In our study, it is highly unlikely that there is only one channel. We consider that a tradition in human capital investment and cultural factors are among the most important. The nature of the legacy is not the contents of Confucianism. The Vietnamese of today have little incentive to study past knowledge on Confucianism. In contrast with China, Chinese characters are not part of Vietnamese life today. To our knowledge, none of the content from Confucian textbooks has been transferred to today's textbooks or influenced today's university entrance examinations or technologies. Therefore, even the forced transfer of Confucian content and knowledge via social interaction would confer no educational advantages on those living in imperial elites' districts today.

Instead, one of the channels likely comes from the long tradition of investing in local human capital. Nguyen (2005) examines the village institution via stele records kept in the villages. Nguyen (2005) notes that villagers donated land for the building of schools in 1767. They even had public rice fields dedicated for educational purposes. The money earned from renting out the rice fields would be spent on private schools for everyone in the village. The money was also used to hire good teachers and pay them bonuses. Villagers also designated a person to manage the village library. Villagers bought books (for learning and exam preparation) and shared them with learners. Nguyen (2005) finds village rules stating that the person in charge of education was required to place Confucian books in direct sunlight on sunny days to help prolong the lives of the books. Nguyen (2005) also finds records of a village program encouraging villagers to learn, giving learners an exemption from forced labor such as public service and offering a bonus when learners passed their examinations. The program also encouraged learners to attain the highest rank in the national level of the imperial examinations. The names of those who succeeded would be carved in the stele stones or documented in records kept in the village common house. This is still considered a source of family pride today, maintained from generation to generation.[13]

We find some statistical evidence to support our arguments on persistence. First, the estimation results shown in Table 1 indicate a persistent correlation among elites from the same home districts across dynasties. Second, focusing on the transmission from the imperial exams during the second Le Dynasty (1554–1787) to the Nguyen Dynasty (1822–1919), we apply the IV approach as in Eq. (6) and (7) to Eq. (1). The IV *distanceBaDinh* is the distance from each district to Ba Dinh, Hanoi, the capital in the second Le Dynasty. During the Nguyen Dynasty, the capital was Hue, about 666 km south of Hanoi. Therefore, *distanceBaDinh* can explain the number of

[13] We note that this tradition exists not just within the clans of elites. Each village in Vietnam generally has several family names.

elites during the second Le Dynasty as a proxy for the cost of education. However, *distanceBaDinh* cannot directly explain the number of elites during the Nguyen Dynasty. Therefore, the corresponding results in columns 2 and 4 of Table 5 indicate a causal effect of the transmission from imperial elites from the second Le Dynasty to subsequent generations during the Nguyen Dynasty.

Third, our estimations in Table 3 (column 14) show the connection between the number of successful imperial test takers and present-day investment in education. An additional successful imperial test taker is associated with a roughly 0.13% increase in the share of the household budget reserved for education expenditures over total household expenditures. This is equivalent to a 1.7% increase in education spending.

Fourth, human capital investment might be associated with a higher density of local schools. We use Eq. (7) for new outcomes as a logarithm of the number of schools and universities per 100,000 people (for each corresponding school-age cohort). The results shown in columns 2 and 3 of Table 6 suggest that the number of imperial elites is associated with a greater density of junior and high schools. Unfortunately, we cannot disentangle whether these results derive from the supply or demand side of local educational services.

However, we consider a few possible explanations for the statistical insignificance of *elite number* for the outcomes of primary schools and universities (see columns 1, 4, 5, and 8 in Table 6). Most primary schools in Vietnam in 2009 were public. To ensure universal primary education, the Vietnamese government would have intervened to address the failure of market forces to provide this service. In addition, it is not necessary for every district to have a university because of economies of scale. A district is small (usually an area with a radius of 12 km and a population of 126,227 people on average). A university would not serve only local residents.

Several economic incentives might explain the persistent legacies of investment in human capital. First, the returns on test scores from the national examinations still reverberate today. A one standard deviation increase in the 2009 NEEU z-score is associated with a 9% gain in the wage rate 9–11 years later (Vu and Yamada, In press). Second, academic degrees play an important role in the wage scale (Vu and Yamada, 2018). Therefore, pursuing more years of schooling should lead to a higher income and consequently lift an individual out of poverty. People living in areas with high levels of investment in human capital over time are more likely to perceive the returns on this investment. Meanwhile, entrance examinations to university are competitive. Demand far exceeds the number of available seats at universities and only 4%–50% of entrance examination test takers are admitted to universities.

Therefore, investment in human capital would have persisted as a part of the culture, being transferred via informal institutions, which are enabled by formal institutions. Present-day formal institutions are based on communist ideology. Vietnamese imperial institutions were abolished in 1945. However, the formal institutions allow village culture and its regulations to prosper, especially because administration at the village level has always been free from central government control (Dell et al., 2018). Present-day government rules (see Government Decree No 92/2009/ND-CP dated November 2009) allow only a few administrative positions at the commune level to receive a salary from the government budget. Moreover, the mandated Government Decree No 34/2019/ND-CP dated April 2019, for the first time, recognized fewer than three positions at the village level, listed them as non-specialized positions, and let them be eligible for a modest subsidy from the government budget. This suggests that village officials are elected mainly based on the will of local voters and the available funding.

Family ties, including the one regarding extended families, are also important (Wantchekon et al., 2015) and provide a straightforward channel for the legacy. Unfortunately, we cannot further investigate this channel in the manner of Chen et al. (2020) with respect to the Chinese imperial examination system. Further complicating the issue is how common some Vietnamese family names are. For example, the family name Nguyen accounted for nearly a third of both imperial elites and 2009 NEEU test takers. In addition, a given family name written in Vietnamese characters in the present day might be associated with several different names in the past written with Chinese characters that had the same pronunciation. Finally, Vietnamese middle names and the lack of consistency in naming conventions made all our attempts at identifying family relationships futile.

## 5. Conclusions

In this paper, we have examined the relationship between the number (or density) of imperial elites and the quantity and quality of present-day education in Vietnam. We found that the association is positively statistically significant for educational variables such as average years of schooling (among those aged $\geq$18, $\geq$22, and $\geq$25 years) and share of household expenditures on education. NEEU test scores were positively proportional to the number and density of imperial elites. The results were obtained using an IV that was calculated from the arc distance to imperial test venues and a weight adjusted by territorial expansion.

However, this study has some limitations. First, the data we collected on imperial elites do not include every imperial elite that ever existed in Vietnamese history. It is likely that some names or entire lists of imperial elites are missing because records were destroyed, ruined, or lost. Prior to the first stele stone (1442), this was particularly true for the Ly (1009–1225) and Tran (1226–1399) dynasties.[14] Among the few surviving records, work by the French School of the Far East and later by the Shino-Vietnamese Institute to document the stele stones remains an important reference source because most of the stele content was copied before the devastating Vietnam War. In addition, the names and homes of imperial elites in the dataset were taken from available and up-to-date documents. Thus, in general, the number of imperial elites in our estimations is systematically smaller than the value it should be in some districts. This may yield nonclassical measurement errors and lead to bias in the estimations. For example, compared with the corresponding coefficients from IV 2nd stage estimations reported in Tables 2 and 3 (using the number of imperial elites from Ngo, 2006), the coefficients in

[14] Perhaps, the short Chinese domination (1407–1427) may be among the reasons. As suggested by Ta (2017: page 76), invaders would have had incentive to erase the evidence of the pre-existing state's culture.

**Table 5**
Causal transmission from the Second Le dynasty (1554–1787) to the Nguyen Dynasty (1822–1919).

| Variables | **Nguyen Dynasty** | | | |
|---|---|---|---|---|
| | Elite number | Elite number | Elite density | Elite density |
| | OLS | IV 2nd stage | OLS | IV 2nd stage |
| | (1) | (2) | (3) | (4) |
| **Second Le Dynasty** | | | | |
| Elite number | 0.2625*** | 0.5258*** | | |
| | (0.0471) | (0.0665) | | |
| Elite density | | | 0.5407*** | 0.4436*** |
| | | | (0.1142) | (0.0711) |
| F-statistics† | | 78.49 | | 17.70 |
| R-squared | 0.190 | | 0.787 | |
| N districts | 688 | 688 | 688 | 688 |

Notes:.
Robust standard errors are in parentheses (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). Mean elevation in 1996, and distance to the coastline are controls in all estimations. † Kleibergen-Paap Wald rk F statistic for testing H0: Weak identification test.

**Table 6**
Effects on the density of educational institutes.

| Variables | Logarithm density of | | | |
|---|---|---|---|---|
| | Primary school | Junior high school | High school | University |
| **OLS** | (1) | (2) | (3) | (4) |
| Elite number | 0.0047 | 0.0176*** | 0.0091*** | 0.0017 |
| | (0.0032) | (0.0035) | (0.0031) | (0.0080) |
| R-squared | 0.151 | 0.086 | 0.016 | 0.261 |
| **IV 2nd stage** | (5) | (6) | (7) | (8) |
| Elite number | 0.0179 | 0.0641*** | 0.0565*** | 0.0196 |
| | (0.0140) | (0.0140) | (0.0152) | (0.0152) |
| F-statistics† | 128.84 | 128.72 | 128.84 | 128.84 |
| N districts | 687 | 688 | 687 | 687 |

Notes:.
We used 2007 Establishment Census from the GSO to count the number of primary, junior high, and high schools and universities located in the districts. The density of schools/university is per 100,000 children in the corresponding-school-age cohort in 2009. The logarithm density is adjusted using the formula $In\ (x)=In\ (x+(x\hat{}2 + 0.001)\hat{}(1/2))$ from the raw value $x$.
Robust standard errors are in parentheses (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$).
† Kleibergen-Paap Wald rk F statistic for testing H0: Weak identification test.
Kinh ethnic ratio, population density in 2009, urban land ratio, and cropland ratio in 1992, mean elevation in 1996, and distance to the coastline are controls in all estimations.
An island district (Con Co district, Quang Tri) does not have any people within the school-age cohort.

Online appendix A.7 (using the smaller number of imperial elites from the stele stones) tend to be larger.[15] When econometrics methods dealing with such bias problems become available, this will be an important part of our future research agenda.

Second, the culture of learning would be rooted at the village level rather than the district level. Therefore, if information were to be identified at the village level, the effects may be even higher than those estimated. Third, we did not have the information necessary to account for internal migration among imperial elites, although we acknowledged the movement among present-day Vietnamese, which was accounted for 6.6% of the 15% population sample in 2009. Fourth, we did not have sufficiently detailed data about several historical events that occurred between 1919 and 2009 to be included in the controls, including the impacts of French colonization, the land reforms implemented between 1953 and 1956 in North Vietnam, the re-unification event in 1975, and reforms specific to South Vietnam.

## Acknowledgements

This work was supported by JSPS (Japan Society for the Promotion of Science) KAKENHI Grant Numbers 18K12784, 18K01580, 19H00619, 20H01506, 21K01455, and 22K18534 and project grants from the Asian Growth Research Institute, Kyushu University, and Keio Gijuku Academic Development Funds.

We thank the editor, Timur Kuran, and two anonymous referees for their invaluable suggestions. We are grateful to the participants

[15] We also observe a similar pattern when comparing the corresponding coefficients (IV approach) of Table 4 with those in Online appendix A.8.

of the 2021 Asian Meeting of the Econometric Society, Asian and Australasian Society of Labour Economics 2021 Conference for their useful comments and suggestions.

## Supplementary materials

Supplementary data associated with this article can be found, in the online version, at 10.1016/j.jce.2022.09.004.

## References

Acemoglu, D., Johnson, S., Robinson, J.A, 2002. Reversal of fortune: geography and institutions in the making of the modern world income distribution. Q. J. Econ. 117 (4), 1231–1294. https://doi.org/10.1162/003355302320935025.

Becker, S.O., Boeckh, K., Hainz, C., Woessmann, L., 2016. The empire is dead, Long live the empire! Long-run persistence of trust and corruption in the bureaucracy. Econ. J. 126 (590), 40–74. https://doi.org/10.1111/ecoj.12220.

Chen, T., Kung, J.K., Ma, C., 2020. Long live Keju! The persistent effects of China's civil examination system. Econ. J. 130 (631), 2030–2064. https://doi.org/10.1093/ej/ueaa043.

Colella, F., Lalive, R., Sakalli, S.O., Thoenig, M., 2019. Inference with arbitrary clustering. IZA Discussion Paper Series 12584, 1–15 (accessed July 8, 2021).

Conley, T.G., 1999. GMM estimation with cross sectional dependence. J. Econ. 92 (1), 1–45. https://doi.org/10.1016/S0304-4076(98)00084-0.

Dell, M., Lane, N., Querubin, P., 2018. The historical state, local collective action, and economic development in Vietnam. Econometrica 86 (6), 2083–2121. https://doi.org/10.3982/ECTA15122.

Goldin, C., 2019. Human capital. In: Diebolt, C., Haupert, M. (Eds.), Handbook of Cliometrics. Springer, Cham, pp. 147–177. https://doi.org/10.1007/978-3-030-00181-0_23.

Mincer, J., 1958. Investment in human capital and personal income distribution. J. Political Econ 66 (4), 281–302. https://www.jstor.org/stable/1827422.

Ngo, D.T. 2006. List of Vietnamese imperial Elites 1075-1919 (Các Nhà Khoa Bảng Việt Nam in Vietnamese). Edited by Ngo Duc Tho, Literature Publishing House: Hanoi.

Nguyen, H.M., 2005. Several documented imperial education at village and commune level from recorded stele (Vài nét về tình hình giáo dục Nho học ở cấp làng xã qua tư liệu văn bia in Vietnamese). Shino-Nom proceedings Tạp chí Hán Nôm in Vietnamese 4, 71. Link. http://www.hannom.org.vn/web/tchn/DATA1/0504.HTM (accessed July 8, 2021).

Nunn, N., 2020. The historical roots of economic development. Science 367, 6485. https://doi.org/10.1126/science.aaz9986.

Schultz, T.W., 1961. Investment in human capital. Am. Econ. Rev. 51 (1), 1–17. https://www.jstor.org/stable/1818907.

Spolaore, E., Wacziarg, R., 2013. How deep are the roots of economic development? J. Econ. Lit. 51 (2), 325–369. https://doi.org/10.1257/jel.51.2.325.

Ta, N.L., 2017. The History of Vietnam: Part III from 15th to 16th Century (Lịch Sử Việt Nam, Tập 3 Từ Thế Kỷ XV đến thế Kỷ XVI in Vietnamese). Edited by Ta Ngoc Lien. Social Sciences Publishing House, Hanoi, Vietnam, p. 2017.

Tran, T.K. 1920. A Brief History of Vietnam (*Việt-Nam Sử-Lược* in Vietnamese). Edited by Tran Trong Kim. Literature Publish, 2008: Hanoi, Vietnam.

Tran, T.V., Do, D.H., Truong, T.Y., and Nguyen, T.P.C. 2017. The History of Vietnam: Part IV from 17th to 18th Century (Lịch Sử Việt Nam, Tập 4 Từ Thế Kỷ XVII đến thế Kỷ XVIII in Vietnamese) Edited by Tran Thi Vinh. Social Science Publishing House, Hanoi, Vietnam, p. 2017.

Truong, T.Y., Vu, D.M., Nguyen, D.N., Nguyen, H.T., and Pham, A.P. 2017. The History of Vietnam: Part V from 1802 to 1858 (*Lịch Sử Việt Nam, Tập 5 Từ Năm 1802 đến năm 1858* in Vietnamese). Edited by Truong Thi Yen. Social Sciences Publishing House, 2017: Hanoi, Vietnam.

Valencia Caicedo, F., 2019. The mission: human capital transmission, economic persistence, and culture in South America. Q. J. Econ. 134 (1), 507–556. https://doi.org/10.1093/qje/qjy024.

Voth, H.-.J, 2021. Persistence – myth and mystery. In: Bisin, A., Federico, G.B.T.-T.H.of H.E. (Eds.), The Handbook of Historical Economics, pp. 243–267. https://doi.org/10.1016/B978-0-12-815874-6.00015-0.

Vu, T.M., 2022. Effects of heat on mathematics test performance in Vietnam. Asian Econ. J. 36 (1), 72–94. https://doi.org/10.1111/asej.12259.

Vu, T.M., Yamada, H., 2018. Decomposing Vietnamese gender equality in terms of wage distribution. Pacific Econ. Rev. 23 (5), 705–731. https://doi.org/10.1111/1468-0106.12269.

Vu, T.M., and Yamada, H. In press. Returns in wage and employment from test scores: evidence from Vietnamese national university entrance exams 2009 and household data 2018–2020. Appl. Econ. Lett.. https://doi.org/10.1080/13504851.2022.2075538.

Wantchekon, L., Klasnja, M., Novta, N., 2015. Education and human capital externalities: evidence from Colonial Benin. Q. J. Econ. 130 (2), 703–757. https://doi.org/10.1093/qje/qjv004.