調査報告書 19－05

Weather and Human Capital

An Empirical Analysis in Vietnam

令和 2（2020）年 3 月

公益財団法人　アジア成長研究所

Preface

This report presents the research results implemented during April 1, 2019 to March 31, 2020 of the Research Project on "Extreme Weather and Human capital Development in Vietnam". The project is to investigate the impact of weather and extreme weather on human capital.

I conduct analysis on a Mathematic score census of approximately 1 million Vietnamese test takers who participated in the National Entrance Exams in 2009 in association with the corresponding weather conditions. The Mathematic test score best measures human capital because the subject was taught continuously from Grade 1 to Grade 12. I examine the effects of two important weather conditions to the Mathematic test scores. One is the accumulated harsh weather (extreme rainfalls and drought) happened during high school time of each student. The other is the weather condition on the test date at the test site.

Using individual first difference, I find that the maximum temperature of the day, 30 to 32°C (86-89.6°F), which is slightly below the usual average in all July between the years 1950-2009, benefitted examinees most. My analysis demonstrates that female testers were more vulnerable to harsh temperature and extreme weather but also more physically adaptive to temperature than males. Extreme weather occurring at the high school, especially during the school calendar, has a negative effect on the test scores.

The findings suggest several important policy implications for other countries and local governments as well. This is because the econometric method that controlled individual fixed effects would guarantee the validity of its inferences to human race. More specifically, my findings imply that at the earliest, every school, especially at lower grades, should be installed with weather adaptation equipment to reduce the impact of harsh weather such as air-conditioners when the cost is bearable to available budget. The installation may be best applied to compulsory education because beneficiaries are the mass population and because this does not raise inequality in the society.

This work is supported by the project grant from the Asian Growth Research Institute; the Japan Society for the Promotion of Science (JSPS) KAKENHI, grant number 18K12784; and research grant from Kyushu University.

I thank Higher Education Department, the Ministry of Education and Training of Vietnam for publishing the data of test scores (2009 National Entrance Exams to University and College) for analysis and all related information in their website in 2009. I acknowledge insightful comments and suggestions from Shinsuke Uchida, and participants from the 2019 Spring Meetings of the Japanese Economic Association, and the 14th Applied Econometrics Conference in Osaka University. I also thank the US National Centers for Environment Information and the National Oceanic and Atmospheric Administration for allowing me to use the Daily (Weather) Summaries.

March 2020

Vu Manh Tien
(Research Assistant Professor, AGI)

Abstract

This study examines the effects of concurrent weather, corresponding to test sites as well as three-year consolidated weather conditions at high school time, on the math test scores of census examinees participating in the Vietnamese national entrance examinations to universities and colleges on July 4 and 15, 2009. Using individual first difference, I find that the maximum temperature of the day, 30 to 32°C (86-89.6°F), which is slightly below the usual average in all July between the years 1950-2009, benefitted examinees most. My analysis demonstrates that female testers were more vulnerable to harsh temperature and extreme weather but also more physically adaptive to temperature than males. Extreme weather occurring at the high school, especially during the school calendar, has a negative effect on the test scores.

1. Introduction

Climate and human capital are two important factors in modern life that the human race cannot alter in the short term. Both are objects of great concern and important literature regarding their interactions, especially in terms of how climate change influences human capital (Garg, Jagnani, and Taraz, 2017; Graff Zivin, Hsiang, and Neidell, 2018; Park, 2017). However, measurement errors and data limitations on the two subjects in past literature, particularly data on human capital, would invite further study in order to improve insight into this important relationship.

During the years 2002-2014, every July witnessed the greatest Vietnamese competition of the year as students vied for placement in universities and colleges. Each year, nearly one million test takers sat in two separate, competitive National entrance examinations to universities and colleges. Test takers needed to earn top scores in order to earn admittance to their desired university or college. Every test taker faced the same test problems at exactly the same time. Students had only two chances to take the exams during the year. If they failed, they had to wait another year for the next exams, which are costly to test takers and their families. Therefore, their test scores, especially the math[1] test scores, would be the best for measuring students' human capital prior to higher education[2], given the selection was specifically among those students who wanted to either proceed to four-year university or three-year college in Vietnam.

In this study, I investigated the effects of weather on human capital development in Vietnam. More specifically, I first examined the concurrent weather conditions onsite for the test day in relation to the math test z-scores of a census of 460,424 test takers in the country in crossed sectional data with district fixed effect and robust district clustered variance. This method allows for the investigation of the influence of weather conditions at a high school's time and location in relation to the math test z-score. Secondly, I undertook the exercise again in panel data consisting of 103,592 examinees (from 460,424) taking the mathematics test twice, one for entrance to university on July 4, 2009, and another for college entrance on July 15, 2009, using individual fixed effect and individual robust clustered variance. I found that where the maximum temperature on the test day was just equal to or two degrees below the average maximum temperature from 1950 through 2009, the test takers were able to gain about

[1] Mathematics is the main subject in general school and taught continuously from grade 1 to 12 in Vietnam.
[2] Students may not try their best if the aim is just to pass. Besides, Park (2017) documents some adaptive grading at score thresholds in the Regents exams in New York city.

0.035 standard deviations in their test scores (compared with the temperature bin from 27 to 29.9°C). Moreover, results suggest that test scores of Vietnamese females tend to be more vulnerable to harsh temperature and extreme weather than those of males. However, females were found to be quicker than males to physically adapt to the harsh weather if facing it twice.

My study contributes several important points to the existing research. To the best of my knowledge, this would be the first analysis using (national level) census test scores and individual fixed effect among test takers above 17 years of age in a developing country, and with temperature bins ranging above 30°C. Unlike Garg, Jagnani, and Taraz (2017), I use the exact weather onsite on test day for each examinee who participated into two separate exams for mathematics. The study is also different from Park (2017) and Cho (2017) because it allowed for individual fixed effect.

My study uniquely considers the effects on the math test scores in the national entrance exams from both the concurrent weather at the test site on test day as well as the harsh weather at the high school during the school year. This is in line with Graff Zivin, Hsiang, and Neidell (2018), but it would contribute new information because air-conditioning application at test sites in Vietnam in 2009 was probably nil.

I take advantage of the uniformity of national tests and the variation of weather throughout the country within two test days in the same month in order to analyze the same set of test takers. Previous studies may have faced the differences in test problem designs in different test periods and the challenges of absenteeism to account for some important changes that could influence the outcomes during longer intervals[3], especially during childhood.

My study would be categorized within rare research that examines physical adaptations to weather in large-scale observations, which contribute to other research on weather adaptation with technology (i.e. Park, 2017; Barreca et al., 2016). The unique feature of analyzing two consecutive tests within 10 days allows me to examine the physical adaptations to weather conditions (by gender) under a reasonable assumption that air-conditioners were not used at test sites in Vietnam in 2009 and that given students were allowed to bring only few items into the test rooms (see Appendix 1).

Moreover, using the data of individuals above 17 years of age, the study contributes and prolongs the time line research to a significant portion of literature, such as Garg, Jagnani,

[3] Dell, Jones and Olken (2014) argue weather adaptation can be one of several important issues. Another example is additional investment to education after each test.

and Taraz (2018), Hoddinott and Kinsey (2001), and Thai and Falaris (2014), using test scores/educational outcomes at primary and secondary schools.

The paper is organized as follows. Section 2 describes how the Vietnamese National Entrance Exams for university and college are uniquely organized. Section 3 reviews related literature on weather, human development, and test scores. Section 4 follows with data descriptions. Section 5 describes econometric methods and specifications. Section 6 presents the results and leads to the conclusions in Section 7.

2. Vietnamese National Entrance Exams to university and college

Being a university/college student guarantees significantly higher lifetime income and improved future opportunities. The economic returns from college and post-college degrees are greatest, at 2.46 times higher than those of high school graduates in 2006 in Vietnam (Oostendorp and Doan, 2013).

Despite the increase in the number of universities during the early and mid 2000s, the competition for acceptance is fierce. A highly competitive university would require a candidate's score to be ranked in the top 4.59% among applicants (430 slots for 9,366 applicants) (Nguyen, 2009). The least competitive schools require successful applicants to be in at least the 45.45% of top scorers (Nguyen, 2009). Moreover, the total score must be at least 13 (a perfect score is 30) to be eligible for consideration among short listed candidates. Thus, all test takers have no other choice than to try their best to get the highest possible score. The entrance examination score thus would essentially represent the best measure of all efforts and investments in Vietnamese human capital of 12 years of education.

Vietnamese students must prepare and plan for the entrance exams early. Their choice of three universities for the entrance exams cannot be changed after students finish the test registration (MOET, 2008). The choice of three schools for the college entrance exams to colleges follows exactly the same procedures. In general, the test locations are decided by the university/college for which each student registered in the preference list with the National Entrance Exams Council (NEEC). The universities and colleges often make use of their own (or, if necessary, nearby schools') facilities and staff for testing and monitoring in order to reduce cost, which is particularly important because the test fee was set as low as VND 40,000 (approximately USD 2) for all three test subjects by an inter-ministry guideline (between

Ministry of Education and Training (MOET), and Ministry of Finance, numbered 71/2004/TTLT-BTC-BGDĐT).

For these tests, there were about 11 test major classifications for arts and sciences. Students generally chose 3 test subjects, and the sum of test scores was used to decide placement in university or college, each of which has a given number of available slots. For example, a majority of test takers (461,861of 984,477 in 2009) sat in A classification, including mathematics, physics and chemistry test subjects. Test problems were based on content of the corresponding subjects taught in three years of national high school curriculum (MOET, 2008). Therefore, students might have had to prepare for the tests as early as the 10th grade: deciding test majors, focusing on problem solving techniques, for example. Therefore, the weather on the test date at a test site was plausibly exogenous to test takers' choice, which would help to claim the causal effect of weather.

All test takers of the same test major classification encountered the same test subject problems developed by MOET at exactly the same time of the same day across the country in different locations, all of which factors serve to secure the fairness[4]. The test structures in the exams for university and college were similar. For example, the math test duration was 180 minutes and based on the National High School Curriculum (NHSC) (MOET, 2008). The NHSC is uniform throughout the country, monitored, and controlled by MOET. However, compared with the entrance exams to college, math test problems for university would be more difficult in order to help determine placement. Another difference is that Vietnamese universities are concentrated in major cities while colleges tend to be located throughout provinces.

3. Weather, human development, task performance, and test score

Rich literature has investigated the influence of weather on human development. Correlations have been found between warmer temperatures or extreme rainfall and major human conflicts occurring since 10,000 BCE (Hsiang, Burke, and Miguel, 2013). However, the mechanisms behind such statistical results remain unclear. Hsiang, Burke, and Miguel (2013) have found that, in the US, violent personal crime, including rape and violent inter-

[4] See Appendix 1 for detailed information.

group retaliation, occurred more often in higher ambient temperatures, but they acknowledged that the physiological mechanism was unknown. Similarly, Dell, Jones, and Olken (2012) found that higher temperatures deferred economic growth, but only among poor countries in the last half century. In particular, the non-linear effect of temperature on economic production has been applied globally (Burke, Hsiang, and Miguel, 2015).

Weather would not always have negative, crop-specific, effects. Deschênes and Greenstone (2007) found that climate change could raise annual US agriculture profits by 4 percent. Schlenker and Roberts (2009) showed that temperature has a non-linear effect on crop yields. For example, while high temperature increases yield up to 30°C for soybeans and 32°C for cottons in the US, temperature above these cutting points harms these crops. Welch et al. (2010) found that a higher maximum temperature increases those yields in tropical and sub-tropical Asian countries, while a lower minimum temperature reduces rice yields. Similarly, using district level data, Jayachandran (2006) suggested crop yields increase with rainfall increases in India.

Nevertheless, the channel via agricultural production is just one among many possible explanation channels. Dell, Jones, and Olken (2012) found there was a decrease in industrial output in poor countries as temperature rose. Hsiang (2010) showed that a rise in temperature is associated with losses in wholesale and retail; restaurants and hotel industries; other service areas; transport and communications sectors; and construction businesses among 28 countries in the Caribbean and Central America during the years 1970-2006.

Furthermore, extreme weather can interfere with individuals' work or test performance. Extreme ambient temperature can decrease the labor productivity of adults and the task performance of students (Dell, Jones, & Olken, 2014). Similarly, Garg, Jagnani, and Taraz (2017) indicated that warmer days during growing seasons have adverse effects on both test scores and yields compared with non-growing seasons.

The mechanism at play in determining individual performance in terms of climate would be via health and heat tolerance channels. For example, mathematical problem solving depends on the functioning of the brain located in an area of the prefrontal cortex and neural circuits, an area which is sensitive to heat (Hocking et al., 2001; Graff Zivin, Hsiang and Neidell, 2018). In addition, Cho's (2017) reviews implied that high temperature would decrease cerebral blood flow. Moreover, since extreme weather can lower agriculture yields, it can cause shortages of food and therefore lower caloric intakes for families with budget constraints. Straus (1986) suggested lower caloric intakes would in turn lower farm productivity. Besides, the extreme weather might at times work like a shock and cause death. Deschênes and Greenstone (2011) found a nonlinear association between daily temperature and annual death records, which suggests that death records accelerate at the high extremes of temperature. In addition, high temperatures would also spread diseases via spoilt food, especially in low-income countries.

Particularly, weather shocks in early life can influence children's later schooling and overall health. The primary mechanism for weather effects would be through the family investment in children in early life (Cunha and Heckman, 2007). Hoddinott and Kinsey (2001) documented effects of drought during 12 to 24-months of age lowers Zimbabwean children's annual growth rates to about 1.5-2 cm of height, and the effects are more pronounced among poor households. Following a similar procedure, Thai and Falaris (2014) found an adverse effect on later schooling and health outcomes from rainfall shocks near the time of children's births in Vietnam. In contrast, research suggests that early-life rainfall benefits health, school, and socioeconomic status for Indonesian females (Maccini and Yang, 2009). In later life, students, especially in developing countries and in areas relying on agricultural production, might (temporarily) defer their educations in order to work to provide earnings to help their families during times of cultivation loss due to extreme weather (Garg, Jagnani, and Taraz, 2017).

Furthermore, adaptation over the long term would reduce many of the negative effects of extreme weather. Adaptation can be a product of applied technology; for example, the widespread use of home air conditioning explained a significant reduction in temperature-related causes of death after 1960 in the US (Barreca et al., 2016). Adaptation can include some compensation from policy makers. For example, educators can change the testing calendar to avoid hot days (Garg, Jagnani, & Taraz, 2017) or use adaptive grading after exams occurring during heat waves (Park, 2017). Besides, individuals can simply alter their schedules. Graff, Zivin, and Neidell (2014) showed that people change their time use (labor supply) depending on the temperature, for example, reducing time on work exposing people to weather when temperatures are extremely high, or shifting from indoor to outdoor leisure when temperatures warmer after severe cold. However, few studies were able to examine the physical adaptation to weather using large-scale individuals.

4. Data

For this study, I employed test score data from the 2009 Vietnamese national entrance examinations to universities and colleges and merge this data with the Daily Climate Summary.

The test score data of are from a census for all Vietnamese examinees in 2009 in two entrance exams: to university on July 4-5 (984,477 individuals) and to college (277,988 individuals) on July 15-16, 2009. The data were available for analysis from the website of Higher Education Department, Ministry of Education and Training of Vietnam in 2009.

As shown in Appendix 1, there were several subject classifications into which students had to self-select during high school. I focus on only the A type/classification, which has the largest concentration of test takers. Self-selected to A classification, 461,861 students took university entrance exams and 183,388 took college entrance exams. Among those, 105,396 examinees took both of the exams. For the entrance exams to university, the test takers at A classification were required to take mathematics, physics, and chemistry subjects consecutively on the morning and afternoon of July 4, 2009, and the morning of July 5, 2009. For the entrance exams to college, participants took tests in a different order: physics, mathematics, and chemistry on the morning and afternoon of July 15, 2009, and the morning of July 16, 2009.

I selected only math test scores because students study mathematics continuously from 1st grade until 12th grade and because math problems were not presented on the test in multiple-choice format[5]. Therefore, math test scores could be a better proxy for human capital than other test scores.

After omitting individuals with missing information on gender, high school, and unidentified colleges' codes[6], I obtained 460,424 university entrance test takers and 103,592 test takers for both university and college. The descriptive statistics and score distributions can be seen in Appendices 7-10. I acknowledge that the score distribution is not normal distribution, perhaps due to self-selection of test takers. Also, some students would self-select to take both of the two exams. Those students more likely had lower scores than their peers, who participated only the entrance exams to university.

---

[5] Meanwhile, physics and chemical were taught from 7th grade and above, and questions were written in multiple-choice problem format.

[6] Five colleges were included in the data, but their information was not in the national entrance exams guidebook published by MOET. I was unable to identify the names.

The data contain information including full names, date of birth, gender, raw test scores (maximum point=10), [7] codes of high schools in each grade from 10 to 12, year of high school graduation, and relative resident location (province, district, and area types by definition of Vietnamese government such as large cities, townships, rural and mountainous areas, and those areas with most economic difficulty). I used a combination of relevant information to identify those examinees who took both entrance exams.

Available from the US National Oceanic and Atmospheric Administration (NOAA), the Daily Climate Summary contains information on the Vietnamese daily weather since 1950. The data specify climate indicators including highest (or lowest) daily temperatures, including daily temperature and daily rainfall for 353,257 days from January 1, 1950, to December 31, 2009, from every meteorological station in service (44 existing stations during the years 1950-2009: from 5 stations in 1950 to 30 stations in 2009) in Vietnam.[8]

I preferred the weather (ground) station data to gridded weather data for reasons that were pointed out by Auffhammer et al. (2013). First, gridded weather data produce similar average temperatures in each cell. Second, the gridded data often have significant spatial correlation depending on the data inference method. Third, the gridded data are often based on weather station data and missing information in the weather station data leads to artificial variation in the gridded data. Besides, the highest resolution available for gridded data is about 0.25 degrees $\times$ 0.25 degrees (about 27-28 km squared block). Meanwhile, using the weather station data, the distance between location and weather station, as seen in Appendix 2, is about 37 km for high school location and about 15-24 km for test sites. Besides, the "missing" weather station on the two test days would not be systematically serious (see Appendix 3), compared with examples of other transition economies listed in Auffhammer et al. (2013). The reasons that the weather station was "missing" would be because no test sites on the specific date were close to the weather station and/or the nearest weather station did not measure on that specific date.

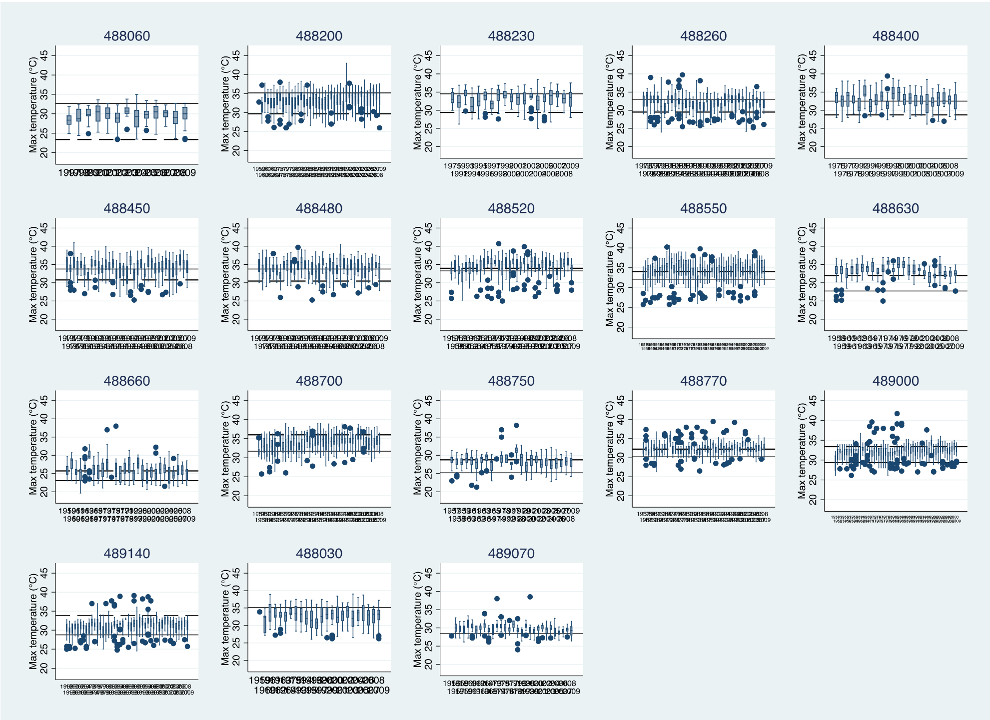
From the test score data and the MOET exams guidebook, I have been able to identify the corresponding latitude and longitude of school locations from their full addresses. I merge the corresponding information from the Daily Climate Summary to the data of test scores by choosing the nearest weather (ground) station[9] to the testing sites (high school locations). I use

[7] In the data, I use the scale of 1000 for the maximum point.
[8] On average, daily weather was recorded in about 20 stations all over the country from 1950 to 2009.
[9] Some weather stations were located on islands and did not correspond to any test or school locations.

the exact weather on each test day at the test locations and the aggregated weather of the year corresponding to the year and location of the individuals' high schools.

The detailed information of each weather station in the two test days corresponding to all test locations is provided in Graph 1. The main body of each station-graph is a box graph of all maximum temperatures ever measured in July during the years 1950-2009. The reference lines corresponding to the weather on the two test days tend to be in either (or both) of two extremes of each station.

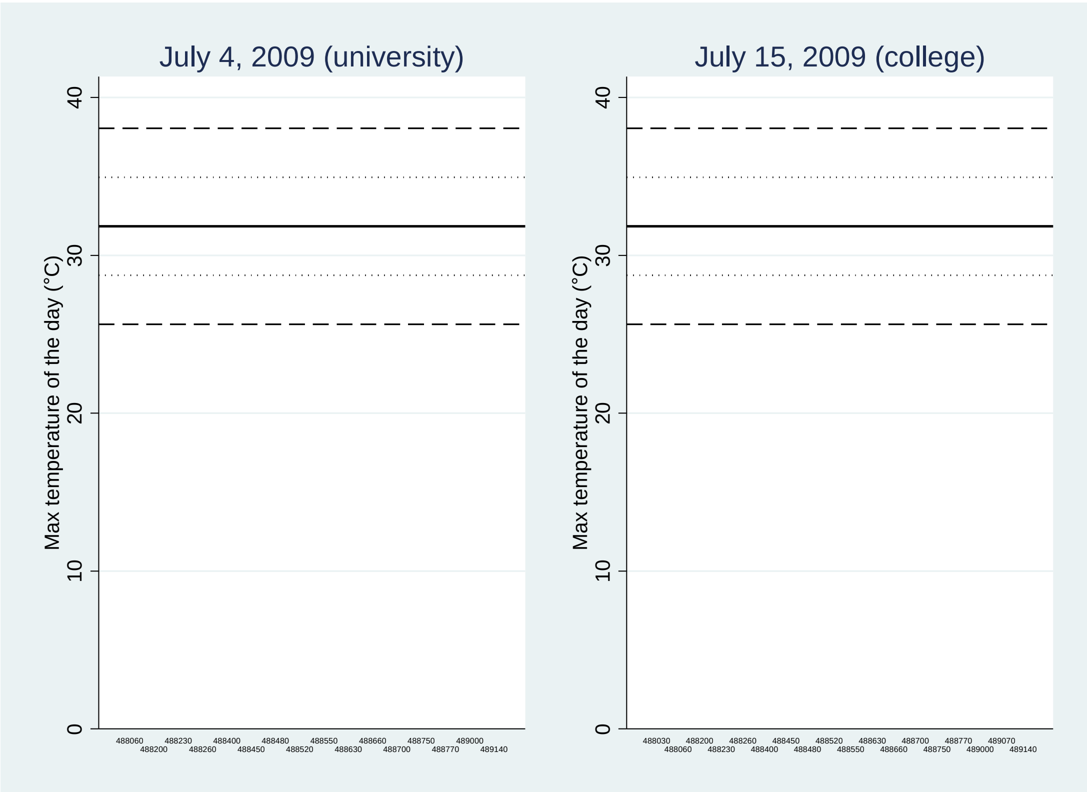
Graph 1 Max temperature on the test days in comparison with 1950–2009 July averages



Notes:
Station names/locations:
488060 Son La; 488030 Lao Cai; 488200 Noibai Intl (Hanoi); 488230 Nam Dinh; 488260 Phu Lien (Hai Phong); 488400 Thanh Hoa; 488450 Vinh; 488480 Dong Hoi (Quang Binh); 488520 Phu Bai (Hue); 488550 Danang intl (Da Nang); 488630 Quang Ngai; 488660 Pleiku; 488700 Quy Nhon; 488750 Ban Me Thuot; 488770 Nha Trang; 489070 Rach Gia (Phan Thiet); 489000 Tansonnhat intl (Ho Chi Minh City); 489140 Ca Mau.

Reference lines:
- Dash lines: Max temperature of the test day, university entrance exams July 4, 2009, at weather stations corresponding with test locations.
- Solid lines: Max temperature of the test day, college entrance exams July 15, 2009, at weather stations corresponding with test locations.

Graph 2 Max temperature at weather stations corresponding to test venues on two test days



Notes:
1950 – 2009 reference lines for July-average maximum temperature of the day
Mean of the period (solid bold line): 31.83737°C (89.307°F)
2-SD upper bound (dash line): 38.045514°C (100.482°F)
2-SD lower bound (dash line): 25.629226°C (78.133°F)
1-SD upper bound (dot line): 34.941442°C (94.895°F)
1-SD lower bound (dot line): 28.733298°C (83.72°F)
Station names/locations:
488060 Son La; 488030 Lao Cai; 488200 Noibai Intl (Hanoi); 488230 Nam Dinh; 488260 Phu Lien (Hai Phong); 488400 Thanh Hoa; 488450 Vinh; 488480 Dong Hoi (Quang Binh); 488520 Phu Bai (Hue); 488550 Danang intl (Da Nang); 488630 Quang Ngai; 488660 Pleiku; 488700 Quy Nhon; 488750 Ban Me Thuot; 488770 Nha Trang; 489070 Rach Gia (Phan Thiet); 489000 Tansonnhat intl (Ho Chi Minh City); 489140 Ca Mau.

The average of maximum daily temperatures is 31.83°C (89.307°F) for any month in the whole period 1950 – 2009 (see Appendix 4). Applying this average to the test days, I compared variations around this average in Graph 2. The northern area of Vietnam (stations on the left-hand side in each sub-graph) experienced "cooler" weather than usual for the entrance exams for university (July 4, 2009), while it faced a "hotter" wave on the days of the other exams for college (July 15, 2009). Meanwhile, the southern area of Vietnam experienced the opposite weather pattern. Although this variation occurs within 20 distinguished ground stations, the variation in each test day covered all ranges, from 24.9% number of days with the lowest temperature (maximum temperature of the day is used to compare), 25-49.9%, 50-

74.9%, 75-89.9%, and 90-94.9% number of the days, as shown in Appendix 6. This helps to overcome the challenge of lacking spatial variation in weather condition raised by Dell, Jones, and Olken (2014). However, since July is summer in Vietnam, the maximum temperature is normally above 25.6°C (78.1°F). Thus, I acknowledge that my figures do not cover the lower boundary of temperature.

5. Methods

First, following Dell, Jones, and Olken (2014), I used a classic reduced equation form to estimate the influence of weather on the math test z-score of individual *i* (with district fixed effect and robust district clustered variance).

$$Math\ z-score_{ij} = \beta Weather_{ij} + \alpha Weather_{ikt} + \gamma_l + \varepsilon_{ij}, \quad (1)$$

I standardized the raw math score of each test taker to form a comparable math z-score based on the mean and standard deviations from the origin sample of (461,861) test takers in the A classification. $Weather_{ij}$ is a vector of conditions such as maximum temperature, temperature changes in the day, rainfall >199mm, at the weather (ground) station nearest to the test site *j*. Meanwhile, $Weather_{ikt}$ is another vector of the weather conditions (drought, probability of having heavy rainfalls >199mm) at the weather station during one year or 3-years (*t*) corresponding to and nearest to their high school location[10]. I use γ_l to represent the district where the test takers registered for family residency ("*ho khau*" in Vietnamese). Districts have over 700 distinguished units and are the second level of administrative area in Vietnam. A test taker could have a family registration location different from where he or she studied high school. This is because Vietnamese high schools are often located in higher population density area. Students also could move to a "better" high school located far away from their family residence. The test location is different from the high school location and the location of family residence because Vietnamese universities are highly concentrated in big cities such as Ho Chi Minh City and Hanoi. Thus, γ_l can be used to control for any time-invariant in the origin of test takers, such as past climate and past socio-economic situation (economic growth, educational services, and health care) in the district.

[10] I used the first principal component product of weather variables if the correlation among weather variables of high school time-location is high. Otherwise, I used the average value for the whole period. I also checked the correlations among weather variables before including them in the estimations.

The specific weather dummies are defined in Appendix 6 and 9. The rationale for each range of temperature was based on the weather statistics for the whole period 1950-2009 in Vietnam. This is similar to the temperature bins approach (as suggested by Deschênes and Greenstone, 2011), but at slightly wider spacing because I used only two different test days' weather in July 2009. For example, a dummy 27_29.9°C was set to 1 if the maximum temperature of the test day at the test site was below 30°C but at least equal to or above 27°C. This corresponds with 25% to 49.9% of lowest temperature days during 1950-2009 (the maximum temperature of the day is used to compare). Employing similar logic, other dummies are set based on the cutting point of 25%, 75%, 90%, 95%, and 99% lowest temperature days (see Appendix 6).

Following Garg, Jagnani, and Taraz (2018), I also identified months with drought (heavy rainfalls), i.e. *Drought_School* (*Heavy Rain_School*) and *Drought_No School* (*Heavy Rain_No School*), overlapping with the high school calendar and summer vacation months (June to August) into two different variables. However, unlike India, rice cultivation months in Vietnam are primarily in two seasons up to three times in a year and highly correlated with the school calendar. The school year coincides with rice cultivation months; thus a design to test mechanisms as suggested by Garg, Jagnani and Taraz (2018) is difficult. Another difficulty is that Vietnamese types of rice prefer more rainfall than none (Institute of Agriculture, 2018).

In equation (1), reverse causality is less likely[11]. Regarding the testing groups, the stable endowment of students or student ability locates in ε_{ij}; however, ε_{ij} cannot interfere with $Weather_{ikt}$ or $Weather_{ij}$. This is because the test dates were determined several months prior to testing. Besides, students may have decided to study in a high school with favorable weather, which is a selection. However, this decision could not change the weather at the location of their high school. Therefore, both $Weather_{ikt}$ and $Weather_{ij}$ would be exogenous. If any correlation between $Weather_{ikt}$ ($Weather_{ij}$) and ε_{ij} occurs, the causal would be from the weather to individual characteristics remaining in ε_{ij} rather than the opposite direction. Therefore, I would argue $Weather_{ikt}$ and $Weather_{ij}$ would have causal effects on the math test z-scores. However, I acknowledge that the issue of omitted variables would remain in equation (1). For example, the long-term effect of climate in the areas where the test takers were living might stay in the error terms and therefore be correlated with the variables of high

[11] Another issue may that occur in (1) if I put some economic control variables is that the weather variables might correlate with them and result in a biased estimation of β and α. I acknowledge inherent bias since weather conditions at high school are lagged weather conditions for short (3 year), continuous periods.

school time/weather. Therefore, individual fixed effect (individual first differencing) is necessary to identify the instant effect of weather to the math test z-score.

I use individual first differencing to remove time-invariant factors (μ_i) belonging to individuals as follows:

$$Math\ zscore_{i,j,time} = \rho Weather_{i,j,time} + \mu_i + \epsilon_{i,j,time}, \tag{2}$$

$$Math\ zscore_{i,j,time+1} = \rho Weather_{i,j,time+1} + \mu_i + \epsilon_{i,j,time+1}, \tag{3}$$

where time=0 for the entrance exams to university and time=1 for the entrance exams to college.

Taking the difference between (3) and (2), I could reduce the biasness in the estimation ρ for the causal effect of weather on test score. ρ measures the influence of time-variant weather conditions on test score and would be independent from the time-invariant ability of the test takers and any time-invariant factors of the larger area associated with the test takers.

$$\Delta Math\ zscore_{ij} = \rho \Delta Weather_{ij} + \epsilon_{ij}, \tag{4}$$

In addition, individuals might have different physical weather endurance/adaptations against weather conditions. The characteristics might vary by gender as well. Therefore, I included an interaction among time, gender, and weather conditions for further analysis:

$$\Delta Math\ zscore_{ij} = \theta \Delta Weather_{ij} + \lambda_1 \Delta Weather_{ij} \times time + \lambda_2 \Delta Weather_{ij} \times Gender_i + \lambda_3 \Delta Weather_{ij} \times time \times Gender_i + \lambda_4 t + \omega_{ij}, \tag{5}$$

Here, λ_2 suggests the degree of difference between females and males when facing difficult weather conditions. Meanwhile, λ_3 shows the difference between the two genders at when facing it twice.

6. Results

6.1 Instant effect of weather at test sites

I found that the maximum temperature of the day at test sites would influence the math z-scores. When time-invariant factors were not controlled, the range between 30 and 31.9°C would result in the best score for test takers, equal to an increase of 0.094 to 0.144 standard deviation compared with the baseline temperature (27 to 29.9°C). Otherwise, temperatures either colder or hotter than the baseline temperatures would inhibit individuals from earning the best possible test scores. This result suggests that temperature effect would be people-specific and an add-on to Graff Zivin, Hsiang, and Neidell (2018)[12]. The result agrees with the findings of Cho (2017) on the adverse effect of heat (above 34°C, compared with temperatures between 28°C and 30°C).

Table 1 Effects of Test Day Temperatures on Entrance Examinations' Math z-scores.

VARIABLES	To university		To both university and college
	(1)	(2)	(3)
	Math z-score	Math z-score	Math z-score
Temperature on test day (Baseline: 27_29.9°C)			
Under_27°C	–0.529***	–0.462***	–0.106***
	(0.020)	(0.030)	(0.024)
30_31.9°C	0.095**	0.144***	0.035***
	(0.037)	(0.041)	(0.012)
32_33.9°C	–0.141***	–0.268***	–0.127***
	(0.034)	(0.076)	(0.004)
34_37.37°C	–0.020	–0.112*	0.106***
	(0.039)	(0.063)	(0.004)
Heavy Rain_Test Day		Yes	
Temperature Shock_Test Day		Yes	
Drought		Yes	
Heavy Rain_High School		Yes	
District fixed effect	Yes	Yes	
Individual fixed effect			Yes
Observations	460,424	460,424	207,184
R-squared	0.081	0.082	0.023

When first differencing controls for time-invariant factors, including the stable ability of individuals, the effect of temperature was less pronounced. However, it was still an increase of 0.035 standard deviation for the range between 30 and 31.9°C. The average of maximum daily temperatures during 1950 and 2009 in Vietnam in July is 31.84°C. Therefore, slightly below this average might work best.

6.2 Gender and physical weather adaptation

My study found that ability for physical adaptation to weather is different by gender. Females in my study were more vulnerable to temperature on the test-day site as well as extreme weather, such as drought and heavy rainfall, during high school time-location.

Table 2 Effects of weather upon gender and adaptation

[12] Graff Zivin, Hsiang, and Neidell (2018) find higher temperature always decreases math test score of US children beyond 26°C.

VARIABLES	To university	To both university and college		
	Math z-score	Math z-score	Math z-score	Math z-score
	(1)	(2)	(3)	(4)
Gender × Under_27°C	–0.031		–0.321***	–0.227***
	(0.025)		(0.059)	(0.063)
Gender × 30_31.9°C	–0.036*		0.198***	0.119***
	(0.020)		(0.023)	(0.039)
Gender × 32_33.9°C	–0.105***		–0.000	–0.556***
	(0.022)		(0.013)	(0.030)
Gender × 34_37.37°C	–0.063***		–0.020**	–0.388***
	(0.019)		(0.010)	(0.019)
Time × Heavy Rain_Test Day		0.724***	0.771***	0.809***
		(0.080)	(0.082)	(0.080)
Time × Under_27°C		–0.631***	–0.647***	–0.627***
		(0.062)	(0.062)	(0.087)
Time × 30_31.9°C		–0.197***	–0.184***	–0.189***
		(0.026)	(0.026)	(0.034)
Time × 32_33.9°C		–0.156***	–0.153***	–0.281***
		(0.016)	(0.016)	(0.022)
Time × Temperature Shock_Test Day		0.046***	0.047***	0.048***
		(0.001)	(0.001)	(0.001)
Gender × Time × Under_27°C				0.110
				(0.094)
Gender × Time × 30_31.9°C				0.105**
				(0.050)
Gender × Time × 32_33.9°C				0.742***
				(0.037)
Gender × Time × 34_37.37°C				0.612***
				(0.030)
Gender × Time ×				–0.012***
Temperature Shock_Test Day				(0.001)
Gender × Drought	–0.008**			
	(0.004)			
Gender × Heavy Rain_High School	–0.011***			
	(0.003)			
Gender	Yes			
Time		Yes	Yes	Yes
Dummies of max temperature on test day	Yes	Yes	Yes	Yes
Heavy Rain_Test Day	Yes	Yes	Yes	Yes
Temperature Shock_Test day	Yes	Yes	Yes	Yes
Drought & Heavey Rain_High	Yes			
Gender × Temperature	Yes		Yes	Yes
District fixed effect	Yes			
Individual fixed effect		Yes	Yes	Yes
Observations	460,424	207,184	207,184	207,184
R-squared	0.083	0.206	0.208	0.213
Number of individuals		103,592	103,592	103,592

As seen in column (1) of Table 2, after general gender difference was controlled, females still had a non-linear lower test-score in any temperature range, and the decrease ranges from 0.036 to 0.105 of a test score standard deviation. Perhaps, this is because the female students were not as strong as male students in terms of endurance to temperature. When individual fixed effect was applied, the results were similar to cross sectional data estimation with one exception. Females tended to do better on the examinations at favorable temperatures, ranging from 30 to 31.9°C. The negative effect of extreme weather at high school time-location for females may be the result of lower physical endurance to weather changes. However, the results might suggest that females receive lower priority in household investment to their education as the result of a society of son preference (Vu, 2014b; Vu and Matsushige, 2016). The gender role of housework division might also cause the difference (Vu, 2014a). For example, females might be in charge of collecting water for the household or taking care illness people. And thus, during drought (harsh weather) time, female students might have to alter their time-use for learning.

Despite the fact females did not perform well on testing in most of the unfavorable temperature ranges, females did tend to have quicker weather adaptations compared with males, as shown in column (4) of Table 2. Females increased their test scores, from 0.105 to 0.742 standard deviation compared with males if they all met the same unfavorable temperature range twice.

6.3 Effects of extreme weather conditions during high school time

My study found that extreme weather occurring at high school time/location had a negative effect on the math z-score. More specifically, drought at high school time/location was associated with a decrease of 0.028 standard deviation in math test z-score. The average heavy rain might be associated with higher precipitation in the area, which results in a positive coefficient. However, when using the first principal component product of heavy rainfalls in the three-year period, I found that extreme, heavy rain also had an adverse effect on the test score, but this tended to be lower than the drought effect. Besides, the effect of drought during the school calendar year seemed to be stronger during school year than during school summer vacation. However, the adverse effect of heavy rain was difficult to distinguish between school calendar year and vacation time because of the high correlation between two variables.

Table 3 Effects of weather during high school on math z-scores for entrance exams to university

VARIABLES	Math z-score	Math z-score	Math z-score	Math z-score	Math z-score
	(1)	(2)	(3)	(4)	(5)
Drought_Grade 10	–0.315***				
	(0.081)				
Drought_Grade 11	–0.528***				
	(0.086)				
Drought_Grade 12	0.572***				
	(0.194)				
Average Heavy Rain	0.009*				
	(0.005)				
Drought_School				–0.028***	–0.038***
				(0.003)	(0.004)
Drought_No School				–0.009***	–0.010***
				(0.003)	(0.003)
Heavy Rain_No School		–0.013*		–0.003	–0.129***
		(0.006)		(0.007)	(0.026)
Heavy Rain_School			–0.015**		0.130***
			(0.006)		(0.028)
Heavy Rain_Test Day	Yes	Yes	Yes	Yes	Yes
Temperature Shock_Test Day	Yes	Yes	Yes	Yes	Yes
Dummies of max temperature on test day	Yes	Yes	Yes	Yes	Yes
District fixed effect	Yes	Yes	Yes	Yes	Yes
Observations	460,424	460,424	460,424	460,424	460,424
R-squared	0.082	0.082	0.082	0.082	0.083

The different results by months both outside of and within the school calendar year provide insight into the mechanism of how weather can interfere with the test scores. The results by school calendar contribute to Nguyen and Pham (2018) and Bui et al. (2014) about the influence of weather shocks.

Moreover, the results among rural and mountainous and extremely poverty-stricken areas were more significant than those in large cities, which may reflect the channel of influence via agricultural production. The interaction term between drought (heavy rainfalls) and school calendar year is higher and more statistically significant in rural, extremely poor, mountainous, remote area and townships compared with large cities (as seen in Table 4).

Compared with Garg, Jagnani and Taraz (2018), my results suggest that all kinds of students, whether in large cities or rural areas, are affected by drought, regardless of whether students' families are involved in agricultural activities. This may be similar to findings by Dell, Jones, and Olken (2012), which found negative effects of "hotter" years on industrial output among poor countries. However, I would add another plausible reason: drought causes temporary shortage of food supplies and increases the price of food. Since the proportion of food expense among the total household living expenses is large (about 46.12% in 2006, 47.87% in 2008[13]), this could shrink the basket of food, resulting in lower nutritional alternatives for the students' families in any location.

Table 4 Effects of weather on test scores by test taker's origin

	Township	Rural	Extremely poor, mountainous, remote area	Large city
VARIABLES	Math z-score	Math z-score	Math z-score	Math z-score
	(1)	(2)	(3)	(4)
Under_27°C	–0.417***	–0.507***	–0.383***	–1.731***
	(0.135)	(0.033)	(0.034)	(0.218)
30_31.9°C	0.108	0.032	0.304***	0.035
	(0.101)	(0.041)	(0.053)	(0.358)
32_33.9°C	–0.839***	–0.537***	0.014	0.081
	(0.197)	(0.060)	(0.068)	(0.622)
34_37.37°C	–0.761***	–0.232***	0.166***	–0.295
	(0.168)	(0.052)	(0.048)	(0.440)
Drought_School	–0.051***	–0.029***	–0.041***	0.023*
	(0.009)	(0.005)	(0.007)	(0.014)
Drought_No School	–0.016**	–0.004	0.002	–0.017**
	(0.007)	(0.004)	(0.006)	(0.007)
Heavy Rain_No School	–0.181***	–0.129***	–0.257***	0.065
	(0.049)	(0.030)	(0.039)	(0.068)
Heavy Rain_School	0.212***	0.144***	0.260***	–0.119
	(0.051)	(0.030)	(0.043)	(0.081)
Heavy Rain_Test Day	Yes	Yes	Yes	Yes
Temperature Shock_Test Day	Yes	Yes	Yes	Yes
District fixed effect	Yes	Yes	Yes	Yes
Observations	84,291	179,850	145,126	51,135
R-squared	0.056	0.058	0.051	0.053

7. Conclusions

I examined the immediate effect of weather at the test site/day and weather at high school time/location on the math test z-score of census test takers in A classification on the 2009 Vietnamese national entrance exams for university and college. I found the maximum temperature of the test day ranging between 30-32°C worked best. I also found a difference

[13] I calculated from the Vietnam household living standard survey 2006 and 2008.

between females and males in terms of physical adaptation to weather. To the best of my knowledge, few studies about adaptation consider the physical adaptations to weather in individuals using large-scale observations. In the longer term, extreme weather in high school time/location, especially during school calendar had adverse effects on the test scores.

However, I acknowledge some drawbacks of this study that should be considered in future work. First, I provided only statistical results, but the underlying mechanism is just my extrapolations. Second, there are selection issues in both crossed-sectional data and "panel" data because some students did not take the entrance exams since they did not have plans to continue in higher education. Third, ambient temperature was proxied by maximum temperature while maximum temperature would occur during midday in Vietnam. Moreover, ambient temperature can be influenced by the heat prevention measures at the test sites, for example, with ceiling fans, for which I do not have information. Fourth, two tests occurred at different times of the day, one in the morning and the other in the afternoon. Therefore, the second time would be more sensitive to the max temperature of the day (evidence is that the interaction between time and temperature is negative and significant).

References

Auffhammer, M., Hsiang, S. M., Schlenker, W., & Sobelz, A. (2013). Using Weather Data and Climate Model Output in Economic Analyses of Climate Change. *Review of Environmental Economics and Policy*, *7*(2), 181–198. http://doi.org/10.1093/reep/ret016

Barreca, A. I., Clay, K. B., Deschenes, O., Greenstone, M., & Shapiro, J. S. (2016). Adapting to Climate Change: The Remarkable Decline in the U.S. Temperature-Mortality Relationship Over the 20th Century. *Journal of Political Economy*, *124*(1), 105–159.

Bui, A. T., Dungey, M., Nguyen, C. V., & Pham, T. P. (2014). The Impact of Natural Disasters on Household Income, Expenditure, Poverty and Inequality: Evidence from Vietnam. *Applied Economics*, *46*(15), 1751–1766. http://doi.org/10.1080/00036846.2014.884706

Burke, M., Hsiang, S. M., & Miguel, E. (2015). Global Non-Linear Effect of Temperature on Economic Production. *Nature*, *527*(7577), 235–239. http://doi.org/10.1038/nature15725

Cho, H. (2017). The Effects of Summer Heat on Academic Achievement: A Cohort Analysis. *Journal of Environmental Economics and Management*, *83*, 185–196. http://doi.org/10.1016/j.jeem.2017.03.005

Cunha, F., & Heckman, J. (2007). The Technology of Skill Formation. *American Economic Review*, *97*(2), 31–47. http://doi.org/10.1257/aer.97.2.31

Deschênes, O., & Greenstone, M. (2007). The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather. *The American Economic Review*, *97*(1), 354–385. http://doi.org/10.1257/000282807780323604

Deschênes, O., & Greenstone, M. (2011). Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US. *American Economic Journal: Applied Economics*, *3*(4), 152–185. http://doi.org/10.1257/app.3.4.152

Dell, M., Jones, B. F., & Olken, B. A. (2012). Temperature Shocks and Economic Growth: Evidence from the Last Half Century. *American Economic Journal: Macroeconomics*, *4*(3), 66–95. http://doi.org/10.1257/mac.4.3.66

Dell, M., Jones, B., & Olken, B. (2014). What Do We Learn from the Weather? The New Climate-Economy Literature. *Journal of Economic Literature*, *52*(3), 740–798. http://dx.doi.org/10.1257/jel.52.3.740

Garg, T., Jagnani, M., & Taraz, V. (2017). Temperature and Human Capital in India. *SSRN*. http://doi.org/10.2139/ssrn.2941049

Graff Zivin, J., & Neidell, M. (2014). Temperature and the Allocation of Time: Implications for Climate Change. *Journal of Labor Economics*, *32*(1), 1–26. http://doi.org/10.1086/671766

Graff Zivin, J., Hsiang, S. M., & Neidell, M. (2018). Temperature and Human Capital in the Short and Long Run. *Journal of the Association of Environmental and Resource Economists*, *5*(1), 77–105. http://doi.org/10.1086/694177

Hocking, C., Silberstein, R. B., Lau, W. M., Stough, C., & Roberts, W. (2001). Evaluation of Cognitive Performance in the Heat by Functional Brain Imaging and Psychometric Testing. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, *128*(4), 719–734. http://doi.org/10.1016/S1095-6433(01)00278-1

Hoddinott, J., & Kinsey, B. (2001). Child Growth in the Time of Drought. *Oxford Bulletin of Economics and Statistics*, *63*(4), 409–436. http://doi.org/10.1111/1468-0084.t01-1-00227

Hsiang, S. M. (2010). Temperatures and Cyclones Strongly Associated with Economic Production in the Caribbean and Central America. *Proceedings of the National Academy of Sciences*, *107*(35), 15367–15372. http://doi.org/10.1073/pnas.1009510107

Hsiang, S. M., Burke, M., & Miguel, E. (2013). Quantifying the Influence of Climate on Human Conflict. *Science*, *341*(6151), 1235367–1235367. http://doi.org/10.1126/science.1235367

Institute of Agriculture. (2018). Agriculture Encyclopedia: Rice Seasons in Vietnam (In Vietnamese: "Từ điển Bách khoa Nông nghiệp Việt Nam: Vụ lúa ở Việt Nam"). Link: http://vitc.edu.vn/tudiennn/home/view/7220/Vu-lua-o-Viet-Nam. Last accessed December 18, 2018.

Jayachandran, S. (2006). Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries. *Journal of Political Economy*, *114*(3), 538–575. http://doi.org/10.1086/503579

Maccini, S., & Yang, D. (2009). Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall. *American Economic Review*, *99*(3), 1006–1026. http://doi.org/10.1257/aer.99.3.1006

Ministry of Education and Training of Vietnam. (MOET). (2008). Decision of the Minister of MOET on regulations on the national entrance exams to full-time regular university and college programs (In Vietnamese: Quyết định của Bộ trưởng Bộ Giáo dục và Đào tạo về việc ban hành "Quy chế tuyển sinh đại học, cao đẳng hệ chính quy"). Link: https://thuvienphapluat.vn/van-ban/Giao-duc/Quyet-dinh-05-2008-QD-BGDDT-Quy-che-tuyen-sinh-dai-hoc-cao-dang-he-chinh-quy-62356.aspx. Last accessed December 18, 2018.

Nguyen, H. (2009). Overview of competition among test takers in the 2009 national examinations to university (In Vietnamese: "Toàn cảnh tỷ lệ "chọi" các trường ĐH năm 2009"). URL: https://dantri.com.vn/giao-duc-khuyen-hoc/toan-canh-ty-le-choi-cac-truong-dh-nam-2009-1245106720.htm. Last accessed: December 18, 2018.

Nguyen, C. V., & Pham, M. N. (2018). The Impact of Natural Disasters on Children's Education: Comparative Evidence from Ethiopia, India, Peru, and Vietnam. *Review of Development Economics*, *22*(4), 1561–1589. http://doi.org/10.1111/rode.12406

Oostendorp, R. H., & Doan, Q. H. (2013). Have the Returns to Education Really Increased in Vietnam? Wage versus Employment Effect. *Journal of Comparative Economics*, *41*(3), 923–938. http://doi.org/10.1016/j.jce.2012.12.002

Park, J. (2017). Temperature, Test Scores, and Human Capital Production. Working Paper. Link: http://scholar.harvard.edu/files/jisungpark/files/temperature_test_scores_and_human_capital_production_-_j_park_-_1-18-17.pdf. Last accessed: December 18, 2018.

Schlenker, W., & Roberts, M. J. (2009). Nonlinear Temperature Effects Indicate Severe Damages to U.S. crop Yields under Climate Change. *Proceedings of the National Academy of Sciences*, *106*(37), 15594–15598. http://doi.org/10.1073/pnas.0906865106

Strauss, J. (1986). Does Better Nutrition Raise Farm Productivity? *Journal of Political Economy*, *94*(2), 297–320. http://doi.org/10.1086/261375

Thai, T. Q., & Falaris, E. M. (2014). Child Schooling, Child Health, and Rainfall Shocks: Evidence from Rural Vietnam. *The Journal of Development Studies*, *50*(7), 1025–1037. http://doi.org/10.1080/00220388.2014.903247

Vu, TM. (2014a). Are Daughters Always the Losers in the Chore War? Evidence Using Household Data from Vietnam. *Journal of Development Studies*, 50(4): 520-529. https://doi.org/10.1080/00220388.2013.875535

Vu, TM. (2014b). One Male Offspring Preference: Evidence from Vietnam using a Split-Population Model. *Review of Economics of the Household*, *12*(4): 689-715. http://doi.org/10.1007/s11150-013-9183-z

Vu, TM. & Matsushige, H. (2016). Gender, Sibling Order, and Differences in the Quantity and Quality of Education: Evidence from Japanese Twins. *Asian Economic Journal, 30(2):* 147-170. http://doi.org/10.1111/asej.12088

Welch, J. R., Vincent, J. R., Auffhammer, M., Moya, P. F., Dobermann, A., & Dawe, D. (2010). Rice Yields in Tropical/Subtropical Asia Exhibit Large but Opposing Sensitivities to Minimum and Maximum Temperatures. *Proceedings of the National Academy of Sciences*, *107*(33), 14562–14567. http://doi.org/10.1073/pnas.1001222107

Appendix 1 Details on Vietnamese National Entrance Examinations to University and College until the examination system was reformed in 2015

The National Entrance Examinations to university and college are uniform nationally and are centrally organized by Vietnamese Ministry of Education and Training (MOET). Each year, MOET establishes the National Entrance Exams (to University and College) Council (NEEC) and invites key persons such as principals and deans of universities and colleges, and even state police, to organize all-in-one exams across the country.

As early as February, the MOET decides the test dates, major test sites, regulations and related issues and publishes all information in a guidebook for the exams. The e-version of the book was put online for free access, which is how I accessed most of the important information to identify codes used in the test score data. In 2009, the examinations for entrance to university were on July 4 and 5, while those for entrance to college were on July 15 and 16. The time for the examination was from 7:15 am to 10:15 am and 2:15 pm to 5:15 pm. Based on the information, 12th grade high school students and others who had previously failed the test registered three names of universities and majors by descending order of preference in their applications for the National Entrance Exams to University and College from mid-March to mid-April.

MOET decided in March 2009 on a solution called "cost saving for test takers." For example, test takers for university with family registration ("ho khau") in Nghe An, Ha Tinh, Quang Binh, and Quang Tri provinces who wanted to be in university located in Hanoi, should take the test in Vinh University. Test takers from Binh Dinh, Phu Yen, Gia Lai, Kon Tum, Quang Ngai, and Quang Nam province who desired a university located in Hanoi and/or Ho Chi Minh City should take the test in Quy Nhon. Finally, test takers from Ca Mau, Bac Lieu, Kien Giang, An Giang, Vinh Long, Hau Giang, Tra Vinh, Soc Trang, and Can Tho who wanted a university located in Ho Chi Minh city should take the test in Can Tho City.

In each test site, students were assigned to test rooms by both university faculty and alphabetical order; however, students were prohibited to organize into test rooms according to their residence origin (MOET, 2008). Examinees in the same test room would be in the same ambient environment. MOET decided what test takers can bring into the test room, and limited options to pen, pencil, compass, ruler, and calculator (without memory and without word composer) (MOET, 2008). MOET required test monitors to randomly assign the test seats prior to the appearance of test takers on the test sites and to re-allocate differently from the previous test subject for the next consecutive test subject.

The test questions prior to the testing day were a national top secret. Teachers involved in creating problems were secretly employed by the NEEUCC and put into an isolated area heavily controlled by state police force and without outside connectivity. Phone and any wireless signals were cancelled locally where the teachers were staying. The teachers only were released after two thirds of the test time in the corresponding subject had passed (MOET, 2008). However, the solutions for all the test subjects were released to the public after the last subject test ended.

Examinee identities are anonymous to graders. Test takers input their identity on their answer sheets for the test day subject. However, this part will be cut away by a council in charge of NEEUCC (who must not be test graders) and sealed in envelopes during the grading time. Test graders were anonymously recruited by another council belonging to NEEUCC and randomly assigned for grading based on a detailed guideline of how to score the test (MOET, 2008). Each answer sheet was graded twice in two independent rounds. The first grader was not allowed to write on the answer sheet, but to write on a grader sheet. MOET (2008) documented in detail how the score should be decided or re-graded if score difference between two graders was 0.5 point and above. Graders were often placed in positive working environments (sometimes with air conditioners) to reduce the influence of hot summer weather. Therefore, the quality of grading would be homogenous and independent from the surrounding environment, including weather at the grading sites.

After all grading is completed, the test scores were published and disseminated among news agencies online. Students could access some specific websites to check their own scores. However, the majority of students failed the first university exams in their three chosen preferences. Their scores needed then to be transferred to the second university in the list. Universities/colleges were required to consider fairly students with the name of university in the “second-choice” before moving to consider those with the name of university in the “third-choice.” However, these universities/colleges have an incentive to accept students who had a high score regardless of students’ choice order. For example, an institute might have rather accepted a “third-choice” student who had higher score than that of the “second-choice” student. As a result, the eligible “second-choice” student would have failed. Therefore, in 2009, Higher Education Department of MOET asked all these institutes to send and publish the microdata of test scores and choices to the department website unconditionally so that everyone might monitor this process and so that educators could perform some quality analysis.

Appendix 2 Distance between nearest weather stations and locations of schools and test sites

Variable	Mean	Std. Dev.
Among 460,424 test takers for university entrance exams		
Weather station and high school location at grade 10	37,564.54	35,537.00
Weather station and high school location at grade 11	37,460.81	38,971.54
Weather station and high school location at grade 12	37,413.50	45,087.50
Weather station and test sites	24,009.22	31,656.01
Among 103,592 test takers for both entrance exams		
Weather station and high school location at grade 10	37,280.08	26,951.91
Weather station and high school location at grade 11	37,119.38	26,930.70
Weather station and high school location at grade 12	37,003.34	26,918.60
Weather station and test sites (university)	14,992.37	19,925.94
Weather station and test sites (colleges)	18,684.87	17,843.84

Note:

Distance unit is in meters.

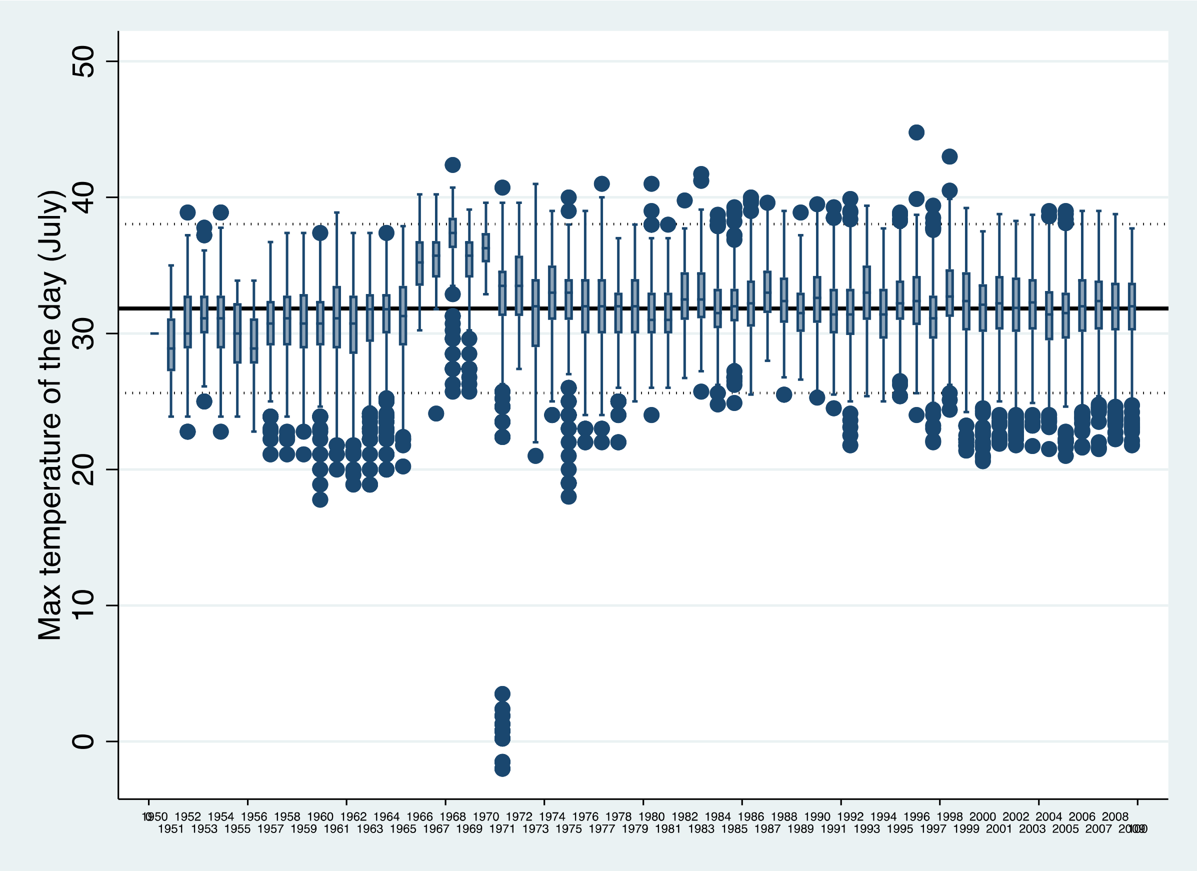
Appendix 3 Individuals in test sites correspondent with weather stations

Station code	Time=0	Time=1
	Freq.	Freq.
488030		19
488060	701	312
488200	31,491	29,303
488230	165	3,314
488260	5,847	6,696
488400	319	
488450	1,877	1,564
488480	86	
488520	3,120	2,216
488550	5,460	6,439
488630	736	667
488660	164	253
488700	11,008	2,459
488750	1,357	
488770	1,566	480
489000	38,184	49,066
489140	1,511	11
489070		793
Test takers corresponding to “missing” stations	1,762	812
in percent	1.70	0.78
N	103,592	103,592

Notes:

Station names/locations:
488060 Son La; 488030 Lao Cai; 488200 Noibai Intl (Hanoi); 488230 Nam Dinh; 488260 Phu Lien (Hai Phong); 488400 Thanh Hoa; 488450 Vinh; 488480 Dong Hoi (Quang Binh); 488520 Phu Bai (Hue); 488550 Danang intl (Da Nang); 488630 Quang Ngai; 488660 Pleiku; 488700 Quy Nhon; 488750 Ban Me Thuot; 488770 Nha Trang; 489070 Rach Gia (Phan Thiet); 489000 Tansonnhat intl (Ho Chi Minh City); 489140 Ca Mau.

Appendix 4 Maximum temperature of the day in July during period 1950 – 2009



Notes:

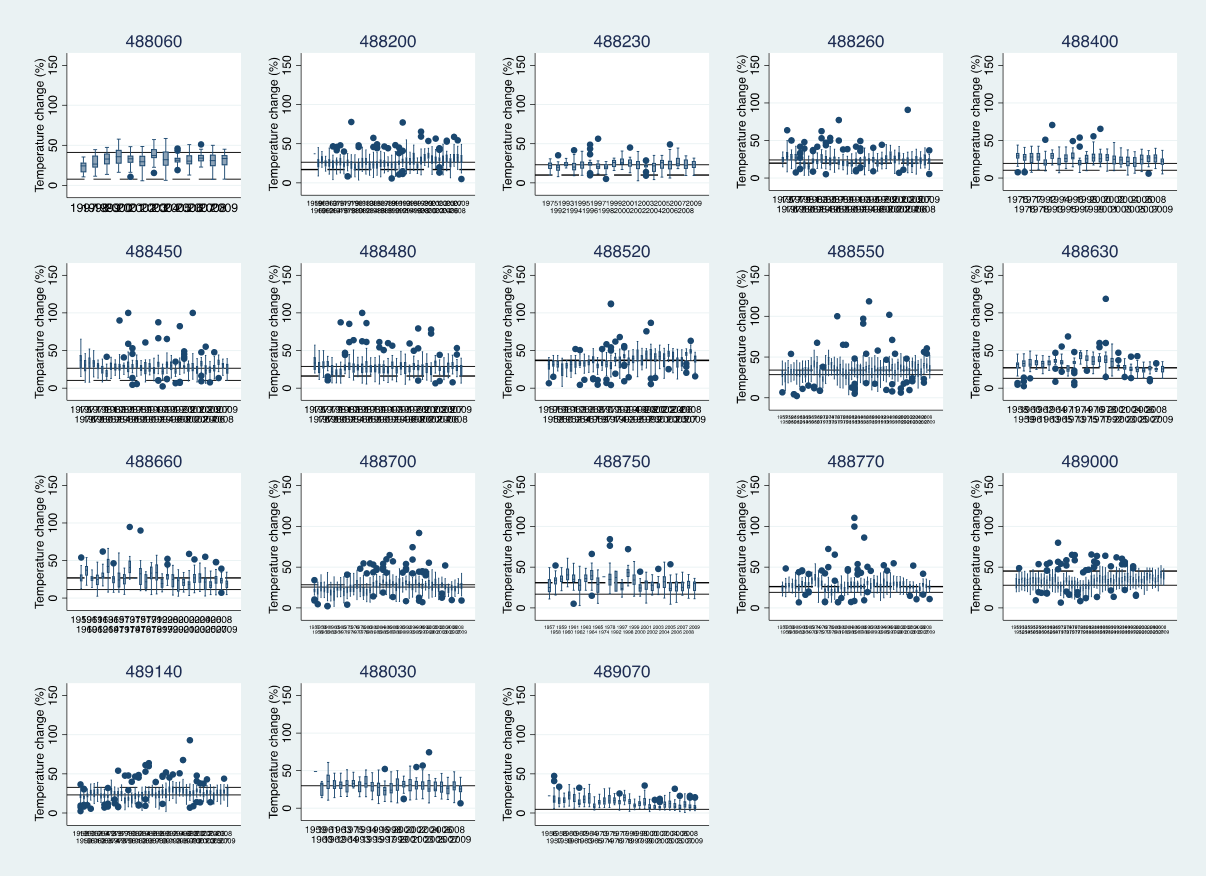
Reference lines

Mean of the period (solid bold line): 31.83737°C (89.307°F)

2-SD upper bound (dash line): 38.045514°C (100.482°F)

2-SD lower bound (dash line): 25.629226°C (78.133°F)

Appendix 5 Percentage of temperature change within the test days in comparison with the average of July in the past



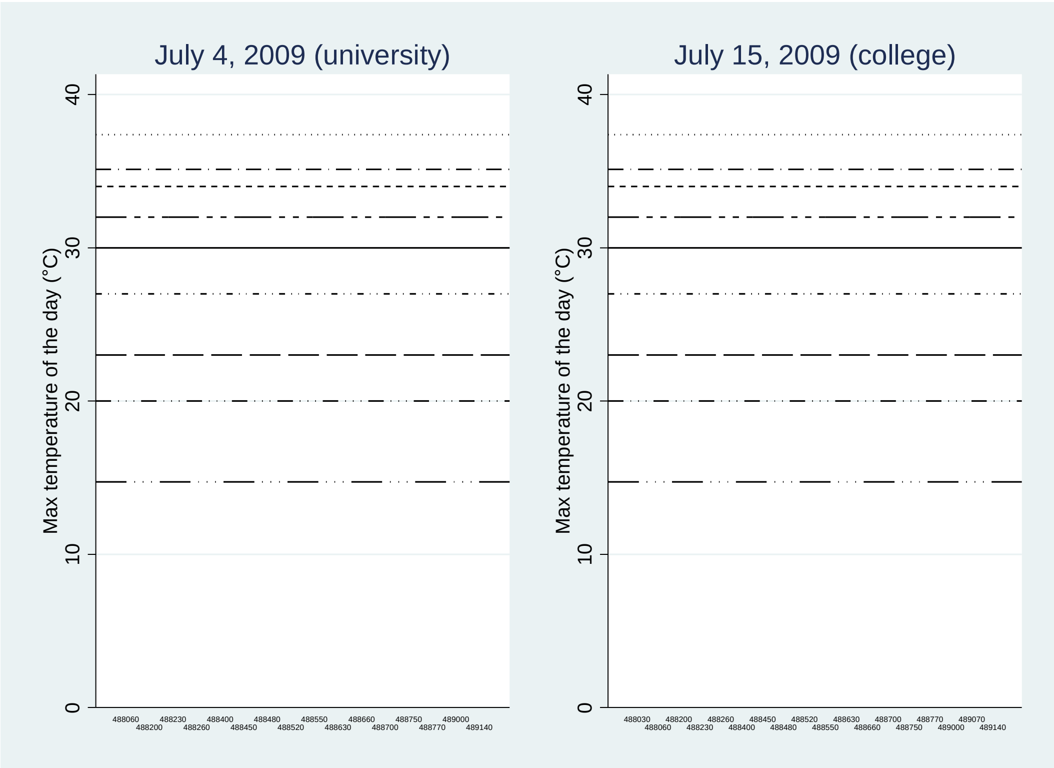
Notes:
Station names/locations:

488060 Son La; 488030 Lao Cai; 488200 Noibai Intl (Hanoi); 488230 Nam Dinh; 488260 Phu Lien (Hai Phong); 488400 Thanh Hoa; 488450 Vinh; 488480 Dong Hoi (Quang Binh); 488520 Phu Bai (Hue); 488550 Danang intl (Da Nang); 488630 Quang Ngai; 488660 Pleiku; 488700 Quy Nhon; 488750 Ban Me Thuot; 488770 Nha Trang; 489070 Rach Gia (Phan Thiet); 489000 Tansonnhat intl (Ho Chi Minh City); 489140 Ca Mau.

Reference lines:

- Dash lines: Percentage of temperature change within the day on university entrance exams July 4, 2009, at weather stations corresponding with test locations.
- Solid lines: Percentage of temperature change within the day on college entrance exams July 15, 2009, at weather stations corresponding with test locations.

Appendix 6 Maximum temperature at weather stations corresponding to test venues on two test days



Notes:
1950 – 2009 reference lines for max temperature of the day (any month)

- 50% of lowest days in the period (solid bold line): 30°C (86°F)
- 1% of lowest days (Long-dash 3-dots): 14.72222°C (58.5°F)
- 5% of lowest days (Dash 3-dots): 20°C (68°F)
- 10% of lowest days (Long-dash): 23°C (73.4°F)
- 25% of lowest days (Short-dash dot-dot): 27°C (80.6°F)
- 75% of lowest days (Long-dash short-dash): 32°C (89.6°F)
- 90% of lowest days (Short-dash): 34°C (93.2°F)
- 95% of lowest days (Dash-dot): 35.11111°C (95.2°F)
- 99% of lowest days (Dot): 37.38889°C (99.3°F)

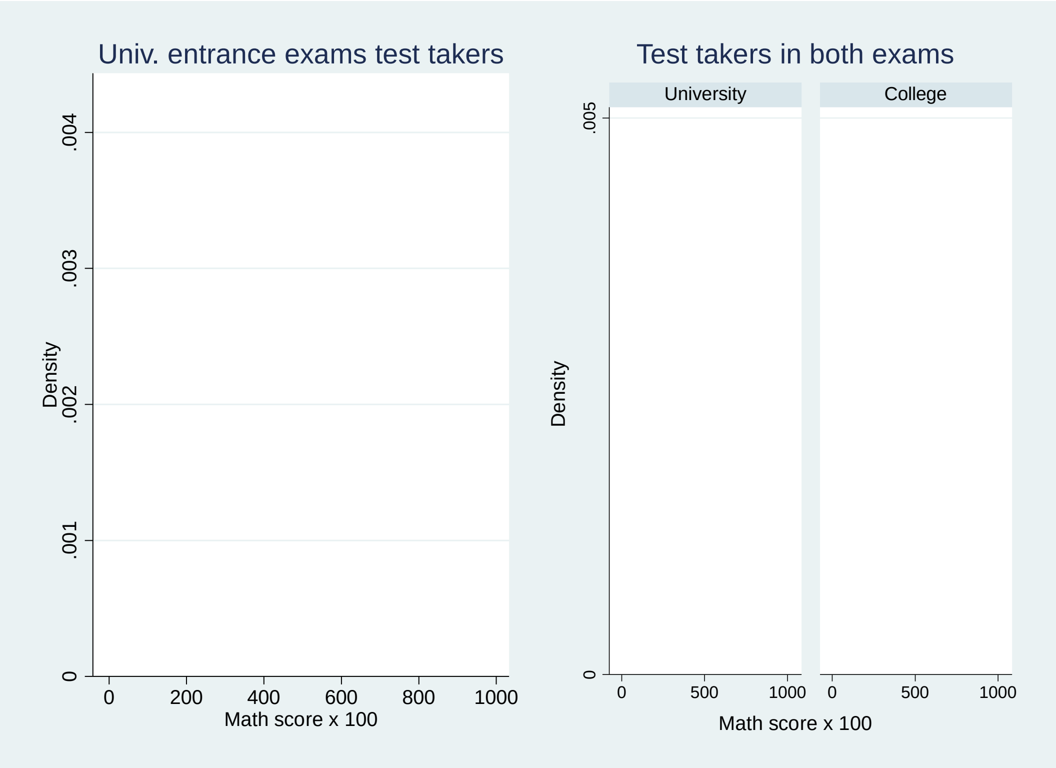
Dummies:

- (0-24.9%) Under_27°C=1 if max temperature of the test day at test site < 27°C
- (25-49.9%) 27_29.9°C=1 if max temperature of the test day at test site < 30°C but ≥ 27°C
- (50-74.9%) 30_31.9°C=1 if max temperature of the test day at test site < 32°C but ≥ 30°C
- (75- 89.9%) 32_33.9°C=1 if max temperature of the test day at test site < 34°C but ≥ 32°C
- (90- 94.9%) 34_37.37°C=1 if max temperature of the test day at test site < 37.38°C but ≥ 34°C

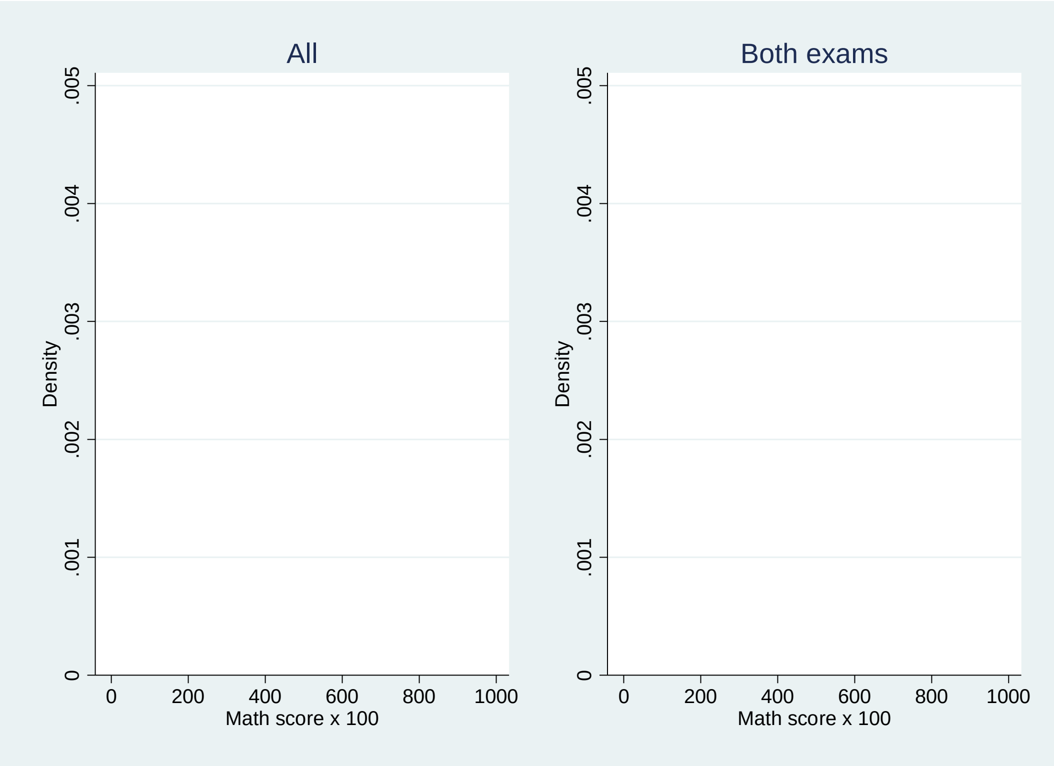
Station names/locations:

488060 Son La; 488030 Lao Cai; 488200 Noibai Intl (Hanoi); 488230 Nam Dinh; 488260 Phu Lien (Hai Phong); 488400 Thanh Hoa; 488450 Vinh; 488480 Dong Hoi (Quang Binh); 488520 Phu Bai (Hue); 488550 Danang intl (Da Nang); 488630 Quang Ngai; 488660 Pleiku; 488700 Quy Nhon; 488750 Ban Me Thuot; 488770 Nha Trang; 489070 Rach Gia (Phan Thiet); 489000 Tansonnhat intl (Ho Chi Minh City); 489140 Ca Mau.

Appendix 7 Math score distributions



Appendix 8 Math score distribution in university entrance exams by all test takers and by testers who also took college entrance exams



Appendix 9 Descriptive statistics for examinees at entrance exams to university

Variable	Description	Mean	Std. Dev.
Math z-core	Math z-score	–0.0007	0.9999
Math raw score	Math raw score x 100	286.5650	165.5105
Gender	=1 if female, 0 if otherwise	0.4864	0.4998
Under_27°C	=1 if max temperature on the test day-site was below 27°C, 0 if otherwise	0.0067	0.0813
27_29.9°C	=1 if max temperature on the test day-site was below 30°C but ≥27°C, 0 if otherwise	0.4187	0.4933
30_31.9°C	=1 if max temperature on the test day-site was below 32°C but ≥30°C, 0 if otherwise	0.0440	0.2051
32_33.9°C	=1 if max temperature on the test day-site was below 34°C but ≥32°C, 0 if otherwise	0.3802	0.4854
34_37.37°C	=1 if max temperature on the test day-site was below 37.38°C but ≥34°C, 0 if otherwise	0.1505	0.3576
Heavy Rain_Test Day	=1 if rainfall was over 199mm on test day-site	0.0010	0.0317
Temperature Shock_Test Day	Percentage of temperature changes from minimum to maximum temperature divided by minimum temperature at test day-site	28.4662	12.2894
Drought_Grade 10	Probability of having a month without rain in high-school-time where students studied their grade 10	0.0108	0.0454
Drought_Grade 11	Probability of having a month without rain in high-school-time where students studied their grade 11	0.0021	0.0212
Drought_Grade 12	Probability of having a month without rain in in high-school-time where students studied their grade 12	0.0002	0.0079
Drought	First principal component from Drought_Grade 10 to Drought_Grade 12	0.0001	1.0948
Average Drought	Average number of months without rain during high-school-year	0.0044	0.0175
Drought_School	First principal component from 3 year high-school-time without rain for months corresponding with schooling month calendar	0.0004	1.2088
Drought_No School	First principal component from 3 year high-school-time without rain for months corresponding with summer vacation (June to August)	–0.0006	1.2187
Average Heavy Rain	Average number of months in 3 year high-school-time experienced a least one rainfall over 199mm	1.3705	2.3035
Heavy Rain_High School	First principal component from 3 year high-school-time experienced at least one rainfall over 199mm	0.0004	1.6943
Heavy Rain_No School	First principal component from 3 year high-school-time experienced at least one rainfall over 199mm among months overlapping with schooling summer vacation (June to August)	0.0003	1.5339
Heavy Rain_School	First principal component from 3 year high-school-time experienced at least a rainfall over 199mm among months overlapping with schooling month calendar	0.0003	1.5216

Note

N=460,424

Appendix 10 Descriptive statistics for examinees for both entrance exams to university and college

	Time = 0 (July 4, 2009)		Time = 1 (July 15, 2009)	
Variable	Mean	Std.Dev.	Mean	Std.Dev.
Math z-core	–0.1081	0.8421	0.2614	0.9397
Math raw score	268.7874	139.3873	572.9132	192.2742
Gender	0.4813	0.4997	0.4813	0.4997
Under_27°C	0.0084	0.0910	0.0024	0.0494
27_29.9°C	0.3782	0.4849	0.4878	0.4999
30_31.9°C	0.0189	0.1363	0.0284	0.1660
32_33.9°C	0.4054	0.4910	0.1663	0.3723
34_37.37°C	0.1891	0.3916	0.3150	0.4645
Temperature Shock_Test Day	30.7151	12.5032	26.8097	3.1692
Heavy Rain_Test Day	0.0016	0.0398	0.0165	0.1275

Note:

N=103,592.