I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15005

The Impact Evaluation of Vietnam's Escuela Nueva (New School) Program on Students' Cognitive and Non-cognitive Skills

Hai-Anh H. Dang
Paul Glewwe
Jongwook Lee
Khoa Vu

JANUARY 2022

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15005

The Impact Evaluation of Vietnam's Escuela Nueva (New School) Program on Students' Cognitive and Non-cognitive Skills

Hai-Anh H. Dang
World Bank, IZA, Indiana University, Vietnam National University and Vietnam's Academy of Social Sciences

Paul Glewwe
University of Minnesota

Jongwook Lee
Seoul National University

Khoa Vu
University of Minnesota

JANUARY 2022

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Impact Evaluation of Vietnam's Escuela Nueva (New School) Program on Students' Cognitive and Non-cognitive Skills*

This paper evaluates how Vietnam's Escuela Nueva (VNEN) program, an educational reform for primary schools supported by the World Bank, affected the cognitive (mathematics and Vietnamese) and non-cognitive (socioemotional) skills of students in that country. We use propensity score matching to estimate both short-term (1-3 years) and long-term (5-7 years) average treatment effects on the treated (ATT). We find that the impacts of VNEN on students' cognitive skills are relatively small in the short-term, and that they are larger for boys, ethnic minorities, and students in Northern Vietnam. The VNEN program modestly increased primary school students' non-cognitive skills in the short-term; these impacts on non-cognitive skills are sizable and significant for ethnic minority students, although there seems to be little gender difference. The long-term impacts are less precisely estimated, but they appear to fade away, showing little or no impact of the VNEN program on cognitive skills. There is little variation of long-term impacts by gender or geographical region, although the imprecision of the estimates for ethnic minority students does not allow us to rule out large long-term impacts on cognitive skills for those students. The program's impacts on non-cognitive skills also seem to have dissipated in the long-term.

JEL Classification: I2, O1

Keywords: VNEN, Vietnam Escuela Nueva, education, cognitive skills, non-cognitive skills, impact evaluation, propensity score matching, IV

Corresponding author:
Hai-Anh H. Dang
Data Production & Methods Unit Development Data Group
World Bank
1818 H St. N.W.
Washington, D.C. 20433
USA

E-mail: hdang@worldbank.org

* For comments received on previous versions of this paper, we would like to thank an anonymous reviewer from the RISE (Research to Improve Systems of Education) program and seminar participants at the University of Minnesota. We also thank Petra Todd for advice on matching estimation. Of course, we alone are responsible for any errors of omission or commission.

I. Introduction

Vietnam is often viewed as being a very strong performer in education, in part because of its surprisingly strong performance on the 2012 and 2015 PISA assessments (Dang and Glewwe, 2018). Yet both parents and policymakers in that country are concerned that Vietnam's education system has serious deficiencies. The Vietnamese government is particularly concerned that Vietnam's schools are not providing students the skills that the country will need for future economic growth and a higher standard of living.

In response to these concerns, Vietnam is undertaking a series of education reforms that are intended to improve and expand the skills that students acquire while in school. One program that Vietnam has recently implemented is the Vietnam *Escuela Nueva* (VNEN) program. This paper uses a propensity score matching methodology to assess the impact of that program, three years after it was first implemented, on the cognitive (mathematics and Vietnamese) and non-cognitive (socioemotional) skills of primary school students in Vietnam. For a subset of those students, the paper also estimates the long-term impacts by following those students as they transition from primary school (grades 1-5) to lower secondary school (grades 6-9).

An initial evaluation, by the World Bank (Parandekar et al., 2017), found some positive effects. This paper makes four further contributions. First, it clarifies the econometric methods used, and that the estimates are average treatment effects on the treated. Second, it presents results for population subgroups of interest, such as distinguishing between *Kinh* (ethnic Vietnamese) and ethnic minority students. Third, it develops an index of VNEN implementation that is based on the program's design, instead of using an index based on the observed differences between VNEN and non-VNEN schools. Fourth, for a subset of students it uses newly collected data to estimate the long-term impacts after they move from primary to lower secondary school.

We find that the impacts of VNEN on students' cognitive skills are modest at best. Averaging over all students, we find no large impacts of the VNEN program on mathematics and Vietnamese scores in any of the first three years of the program. While all estimates are positive, and some are statistically significant, they are generally small. There is some heterogeneity in estimated impacts. The VNEN program's impacts on cognitive skills are somewhat larger for boys than for girls, and for ethnic minority than for *Kinh* (ethnic Vietnamese) students. In Northern Vietnam, the program significantly increased both mathematics and Vietnamese scores, while in the Central and Southern areas of Vietnam the program had few significant impacts.

Turning to non-cognitive skills, the VNEN program seems to have modestly increased them among primary school students, at least in the short-term. Averaging over all students, the estimated short-term impacts are all positive, and although the estimate for intra-personal skills is not statistically significant, those for social skills and overall non-cognitive skills are significant at the 10% and 5% levels. Yet the estimated impacts are small: only 0.07 standard deviations for social skills and only 0.08 standard deviations for the overall index of non-cognitive skills.

Regarding heterogeneity, there seems to be little gender difference, but there are large differences between *Kinh* (ethnic Vietnamese) and ethnic minority students. Some aspects of VNEN were designed to help ethnic minority students, and this seems to have raised their non-cognitive skills. We estimate sizable and statistically significant increases in all three measures of non-cognitive skills for ethnic minority students, but much smaller and insignificant results for *Kinh* students. By region, the impacts for the North that are relatively large and statistically significant for all three indicators, which reflects the high proportion of ethnic minority students in the North. The impacts for the South are similar but significant only for social skills. In contrast, all estimated impacts for Central provinces are smaller and statistically insignificant.

Turning to long-term impacts, we find only small and statistically insignificant impacts of the VNEN program on math and Vietnamese scores when students were in grade 7 (two years after their last exposure to VNEN) and grade 9 (four years after their last exposure). Estimates by subgroups show little variation of impacts by gender or geographic regions, though imprecise estimates raise the possibility of sizeable impacts on ethnic minorities and in Northern Vietnam.

Long-term impacts on non-cognitive skills are less precisely estimated, but in almost all cases the estimates are statistically insignificant. For Vietnamese students as a whole, the estimated impacts are very small and statistically insignificant. Averaging over all three domains of non-cognitive skills, there are no significant impacts for ethnic Vietnamese (*Kinh*) or ethnic minority students, nor for boys or girls. While there is a significantly positive impact on average non-cognitive skills in the Southern region of Vietnam when students were in grade 7, there are no statistically significant impacts on average non-cognitive skills for any of Vietnam's three regions (North, Central and South) when they were in grade 9.

While the lack of a large impact on mathematics and Vietnamese test scores may be disappointing, one could argue that the main focus of the VNEN program was on other skills that could make students more productive workers when they are adults. Some observers worried that the program would reduce student learning, but that seems not to be the case, at least for mathematics and Vietnamese skills. The VNEN program appears to have increased students' non-cognitive skills, particularly ethnic minority students and students in the North and South (as opposed to the Central) regions of Vietnam. Yet almost all of these effects seem to have faded away four years after the students were exposed to the program.

The rest of this paper is organized as follows. The next section describes education in Vietnam, and the implementation of the VNEN program. Section III describes the design of the

evaluation and the data collected. Section IV presents the analytical framework. The results are presented in Section V. The final section summarizes and gives suggestions for future research.

II. Background on Education in Vietnam and Introduction of *Escuela Nueva* in Vietnam

A. Overview of Vietnam's Education System. Vietnamese culture has held education in high esteem for centuries, but for much of Vietnam's history formal education was reserved for the elite. This changed dramatically in the 20th century, when education – and schooling in particular – was deemed to be essential for the entire population. During the 1950s and 1960s, the Democratic Republic of Vietnam (DRV) in the northern half of Vietnam achieved massive increases in literacy and access to education, with estimates ranging between 2.5 and 10 million children and adults acquiring literacy, even amid war-time conditions (Dang, Hoang, and Nguyen, 2021). School enrollment rates also increased, albeit more slowly, in the southern half of Vietnam in the 1950s and 1960s. By the 1970s and the 1980s, however, very tight budgets hindered the provision of schooling in both the north and the south of the now-unified country, which particularly undermined the quality of education in Vietnam.

Vietnam's school enrollment rates declined in the late 1980s and early 1990s, especially at the secondary level, owing to both low school quality and incentives for rural households to use their children's labor after the *Doi Moi* reforms provided those households with individual plots of land (Glewwe and Jacoby, 1998). In many areas, teachers went months without being paid, leading most to seek additional employment and many to leave the education sector altogether. In response, Vietnamese authorities, schools, and households began to finance education through a complex mix of state subsidies and both formal and informal payments from households. By the mid-1990s, Vietnam's state finances had recovered, and political leaders had stabilized education sector financing, in part due to support from international donors.

Starting around 1990, Vietnam's economy improved dramatically, with annual GDP growth rates between 5 and 7 percent. This allowed Vietnam's government to invest heavily in education, although household payments continued to play a central role in education finance. By the late 1990s, primary and secondary enrollment rates had fully returned to levels attained in the mid 1980s, and since then enrollment rates have continued to increase at all levels.

Vietnam's education system includes primary school (grades 1-5, starting at age 6), lower secondary school (grades 6-9), and upper-secondary school (grades 10-12). It also has pre-primary education (for ages 3-5), secondary vocational training schools, and many different post-secondary institutions. In 2014, Vietnam had about 15,000 primary, 10,000 lower-secondary, and 2,300 upper-secondary schools. The education system operates through central and local government offices, in collaboration with parent-teacher associations and other stakeholder groups.

In 2013, about half of primary schools provided "full day" (6 hours) instruction; the other half provided only "half day" (3.5 hours) instruction, usually operating two shifts. One goal of the government is to extend full-day schooling to all schools, but this has proceeded slowly.

At first glance, Vietnamese children's time in school seems very low; it has one of the shortest school days, and one of the shortest school years, in the world. Yet in most areas of Vietnam, including rural areas, parents send their children to varying hours of "extra study" classes, though the time and resources devoted to this activity vary across provinces. While some schools offer these classes at no cost, the vast majority of parents pay fees for their children to attend them. The value of these classes, which focus on passing exams to advance to upper secondary and university study, has been questioned (World Bank 2014).

Examinations play a very important role in Vietnam's education system. Students must take an examination to graduate from upper secondary education, and their scores on this exam

can determine their admission to tertiary-level education. The demand for higher education often exceeds the supply, generating strong competition for admission to universities. This situation, coupled with a Confucian culture that stresses academic achievement, leads to a strong focus of teachers preparing students, and students studying, for these exams (Dang, 2013; Reed, 2018).

Vietnam's Ministry of Education and Training (MoET) has formal authority over the entire education system. It works with other line ministries to plan investments in education, and plays the lead role in education planning and in setting the content of curricula (London 2011). The Vietnamese government is keen to reform the current education system in order to equip students with the skills needed for a modern economy. It has issued various guidelines for encouraging students to develop their social skills, creativity, and self-learning ability, and that discourage teachers from using ineffective teaching practices such as passive learning and rigid memorization (Government of Vietnam, 2014). The government announced the "Fundamental and Comprehensive Reform of Education" in 2018. This has led to major revisions in the official curriculum, pedagogical methods, and teacher professional development to provide Vietnamese children the skills needed to effectively participate in Vietnam's economy and society. This new curriculum was implemented in grade 1 in the fall of 2020, and grades 2 and 6 in the fall of 2021.

B. The Implementation of *Escuela Nueva* in Vietnam.[1] Even before announcing these reforms, MoET implemented pilot programs that were designed to reform Vietnam's education system. Under a World Bank-supported project to improve primary education for disadvantaged children, in 2010 MoET implemented the *Escuela Nueva* program in 24 pilot schools. Encouraging results from this pilot generated great interest among teachers, parents, and students, and neighboring schools adopted the *Escuela Nueva* program on their own (Parandekar et al., 2017).

[1] This section is based on Parandekar et al. (2017).

The *Escuela Nueva* education program started in rural areas of Colombia in the 1970s. It gradually attracted international attention and has been adopted in many countries (Kline, 2002). In Vietnam it attracted interest due to dissatisfaction with several aspects of its education system. Vietnamese policy makers and international agencies identified challenges that were seen as barriers to Vietnam's continued economic growth. Most stakeholders agreed on the need to upgrade the skills of Vietnam's labor force, which require a strong education foundation. Both policy makers and business leaders were especially interested in improving students' non-cognitive (socioemotional) skills, such as creativity, independence, problem solving, and teamwork (Bodewig et al., 2014). These skills are widely considered as a required component of a well-rounded education, and for a productive 21st century worker, but they are rarely provided in schools that use traditional teaching and learning methods (World Bank, 2019). The *Escuela Nueva* program was of particular interest given its emphasis on developing these types of skills.

To undertake a quantitative analysis of the impact of the Vietnam *Escuela Nueva* (VNEN) program on students' skills, the starting point is an understanding of what constitutes a VNEN school. As explained below, not all VNEN schools fully adopted the VNEN model, and many non-VNEN schools operated in ways that had similarities to VNEN.

Generally speaking, VNEN schools have eight fundamental characteristics, three of which pertain to the organization of the classroom, and five of which concern pedagogy, which are summarized in Table 1. The first is the organization of classroom space. There must be adequate space for the teacher to move around among the students, and for students to move around to interact with each other. A related and very important characteristic of VNEN schools is arranging students' desks so that students can face each other and work in small groups, instead of having all students' desks face the front of the classroom.

A second fundamental characteristic of a VNEN school is the presence of six items, called "artifacts", in the classroom. The first is a *community map*, which shows the location of the school and each student's home, as well as local landmarks of interest. The second is a *classroom library*, which includes learning guides (described in the next paragraph) and other reading materials, such as reference material, story-books and magazines. If often consists of a set of shelves containing these reading materials. The third classroom artifact is a *learning corner*, which is an area of the classroom that contains didactic materials and serves as an area to store and display students' work. The fourth classroom artifact, which is similar to (and often located next to) the learning corner, is the *community corner*; it contains objects of importance to the local community, such as locally-grown crops, ethnic handicrafts and traditional, locally-made clothing; parents are invited to the classroom to help set it up. The fifth artifact is the *student mailbox*, which is a poster or chart with an envelope for each student that can be used for students to write notes to each other and is intended for students to develop their communication skills. Finally, the sixth classroom artifact is the *student government chart*, which shows which students are student government leaders, and which students belong to the committees that are part of the student government (described further below).

The third fundamental characteristic of VNEN schools, and the last one pertaining to the organization of and materials in the classroom, is the learning guides. They are similar to textbooks and in fact replace the standard textbooks. They also serve as workbooks in that they contain exercises and activities for students to do, and students write in them when doing these exercises and activities.[2] They are designed not only to improve students' academic skills but

[2] In principle, because students write in their learning guides, they can be used only once, and students can keep them. Yet in some schools teachers are not sure that they will get new learning guides every year, and so they ask students to use pencils when writing in them. These students' writing can be erased, so that the learning guides can be used by other students in the next year.

also their socio-emotional skills, as well as to help students better understand their own progress – and what actions to take to increase their progress – in becoming independent learners.

The five remaining fundamental characteristics of VNEN schools pertain to pedagogy. Thus the fourth fundamental characteristic is the teaching methods used. Teachers use "real life" objects (e.g. fruits or plants) as well as manmade objects in class discussions, and teachers find ways to relate classroom lessons to "real life". Teachers incorporate games in their lessons. They use "role-play" methods when teaching Vietnamese, and ask students to solve many problems when teaching math. Students are encouraged to ask questions, both of their teacher and of other students. They are also encouraged to study on their own and "find out new things by themselves". Finally, teachers provide assistance to individual students who need their help.

The fifth fundamental characteristic of VNEN schools is the use of group work. Students are organized into small groups, sitting together in desks that face each other and working together to complete assignments. Students work together in their groups on a variety of assignments, and the teacher spends most of his or her time working with individual groups or individual students, as opposed to lecturing to the whole class.

A third component of VNEN pedagogy, and more generally a sixth fundamental characteristic of VNEN schools, is the use of assessments of student learning. VNEN pedagogy emphasizes formative assessment rather than summative assessment, which means that assessments are used to provide information on how students can improve their learning skills and how teachers can improve their teaching skills, as opposed to simply measuring how much students have learned. Formative assessment can be done both formally using tests and quizzes, but also informally through conversations of teachers with students and with other teachers. VNEN also encourages students to conduct assessments within their student groups. Examples

of this are: a) students can check each other's work in pairs; b) the group leader can assess the work of each student in his or her group; c) the group arrives at a consensus on the work done by each student; and d) the teacher can provide feedback on the project work of each group.

The fourth component of VNEN pedagogy, and thus the seventh fundamental characteristic of VNEN schools, is the establishment of student government within the classroom. The goals of this component are to increase the value that students place on service and on belonging to a community, and to develop responsibility and leadership skills. The underlying method is learning by doing; for example, the best way to learn how to be responsible is to be assigned a specific responsibility. In general, students take turns being group leaders, so that each student has the opportunity to be a leader. In addition, all students are assigned to committees that are to accomplish certain tasks, which provides the experience of working in groups and thus should develop their teamwork skills. Many different committees are formed (academics, community relations, arts and crafts, health and hygiene, sports, culture, etc.), which should allow each student to participate in a committee for which he or she has a particular talent or strength.

Finally, the fifth component of VNEN pedagogy, and thus the eighth fundamental characteristic of VNEN schools, is parental engagement. The goal is for parents to become more engaged in their children's education, both at home and at school. Examples of this are: a) inviting parents to the classroom to demonstrate or teach a skill; b) inviting parents to help make the six classroom artifacts described above; c) asking specific parents to help teachers teach a particular skill that they have; d) helping the teacher to communicate with students who speak an ethnic minority language; and e) discussing with teachers how to use the grants that were provided to VNEN schools to help them implement the VNEN program.

III. Impact Evaluation Design and Data

The World Bank not only assisted in the implementation of the VNEN program in Vietnam, but also collected data from 651 primary schools in order to conduct an initial evaluation, which is provided in Parandekar et al. (2017). We use the same data to conduct an analysis at the primary school level but, as explained in Section IV, we conduct a more thorough evaluation. In addition, we collected new data to follow a subset of the students in the World Bank study into lower secondary school in order to assess the long-term impacts of the VNEN program, five and seven years after the program started and two and four years after students were no longer in a VNEN primary school. This section describes the World Bank's overall evaluation design, and the data that were collected to conduct that evaluation.

A. Design of the World Bank's Evaluation of VNEN. The World Bank evaluation of the VNEN program used propensity score matching; randomized evaluation methods could not be used because schools were not randomly assigned to the VNEN program. The program started in the summer of 2012, before the start of the 2012-13 school year, and participation in the program varied across Vietnam's 63 provinces. Of the approximately 14,000 public primary schools in Vietnam, about 10% (1,447) adopted the VNEN program in 2012 or 2013.

The process by which schools participated in the VNEN program was the following. First, Vietnam's 63 provinces were divided into three groups: Priority I (most disadvantaged); Priority II (somewhat disadvantaged) and Priority III (least disadvantaged). Second, Vietnam's Ministry of Education and Training (MoET), in cooperation with the World Bank, developed a plan that that set the number of VNEN schools to be established in each province, with most going to Priority I provinces. Within each province, candidate VNEN schools were selected, giving priority to disadvantaged schools, which were defined as schools that had a high proport-

ion of ethnic minority students, were located in remote areas, and/or were located in communes with low income. Yet schools had to meet certain minimum requirements to be eligible for VNEN, such as having electricity (or an electric generator), having a desk and chair for each student, and operating a full school day (as opposed to a half day school). These requirements resulted in many disadvantaged schools being ineligible for VNEN. Each province, with input from the province-level Department of Education and Training, offered the VNEN program to the disadvantaged schools that met these criteria, with the number of offers per province being set by the number in the MoET plan. The schools offered the VNEN program did not have to accept it, but almost all did since the offer included additional funds and other types of support.

To evaluate the impact of this program on student outcomes in the 1,447 VNEN primary schools, a set of comparison schools that are similar to the VNEN schools is needed. The World Bank evaluation used propensity score matching (PSM) to select the comparison schools. More specifically, PSM was applied to the primary school census database, which contains data on school-level student academic performance, school facilities, infrastructure, and teaching staff qualifications for all primary schools in Vietnam (see Dang and Glewwe, 2018, for information on this database). The school characteristics used for this PSM matching were: size (number of students), percentage of poor students, percentage of ethnic minority students, percentage of weak-performing students in math, percentage of weak-performing students in Vietnamese, distance to the district bureau of education, level of parental support for students' education (as perceived by the school principal), number of the school's satellite campuses, number of computers, percentage of blackboards in good working condition, percentage of student desks in good working condition, whether the school meets national quality standards, percentage of teachers with upper secondary education, percentage of teachers with college training, location in

an urban area, and provincial dummy variables. Note that a much smaller subset of these variables was used in the prior process to select the 1,447 VNEN schools.

This procedure to select comparison schools was applied first in Priority I provinces, and then to schools in Priority II and III provinces, since the two sets of provinces are quite different. The following two-step procedure was used to select VNEN schools and comparable non-VNEN schools from the 1,104 VNEN schools and 4,451 non-VNEN schools in Priority I provinces:

Step 1: i) For the 5,555 schools in Priority I provinces, a probit model was estimated. The VNEN school variable was the dependent variable; the explanatory variables were the school census variables listed above. Estimates are in the first column in Appendix Table A.1. A propensity score (probability of being a VNEN school) was calculated for each school.

ii) After dropping 110 VNEN schools with propensity scores not in the region of common support, each of the remaining 994 VNEN schools was matched to the non-VNEN school (out of the 4,451) with the closest propensity score (nearest neighbor match). This yields 994 pairs of schools (one VNEN, one non-VNEN) with very similar propensity scores.

iii) Appendix Table A.2 shows that these 994 VNEN schools are not significantly different from the 994 matched non-VNEN schools for almost all matching variables. The only exceptions are dummy variables for provinces 3, 7 and 10. This table compares 35 variables, so finding three significantly different at the 5% level across the two types of schools is close to what one would see if there were no differences between the two types of schools.

Step 2: i) Next, 220 pairs of schools were randomly selected (simple random sampling) from the 994 pairs of schools identified in Step 1. These provide the 220 VNEN and 220 non-VNEN primary schools from Priority I provinces used in the analysis of this paper.

ii) As a further robustness check, the matching variables for these two groups of 220 schools were compared, and no significant difference exists between the two groups.

These two steps were also applied to the (combined) Priority II and III provinces to select 110 pairs of schools from those provinces' 263 VNEN and 9,402 non-VNEN primary schools. The second column in Appendix Table A.1 shows the probit regression that yields the propensity scores for this matching process. Table A.3 shows that the observable characteristics of these 110 VNEN and 110 non-VNEN schools are well matched: no difference is significant at the 5% level.

Data were to be collected from all of these 660 primary schools, but due to "logistical issues", the actual sample consists of 651 schools (323 VNEN and 328 non-VNEN).

The key assumption of propensity score matching is that the students in the 323 VNEN schools are statistically indistinguishable from those in the 323 non-VNEN schools, in terms of both observed and unobserved characteristics. In particular, after conditioning on schools' propensity scores, or more specifically after matching schools by their propensity scores, we assume that the (potential) student outcomes of interest are uncorrelated with the treatment status of students' schools (the VNEN school dummy variable). This is called the "ignorability" assumption.

B. Program Implementation and Data Collection. Data were collected from the 651 schools (323 VNEN, 328 non-VNEN) for three consecutive school years: 2013-14 (December, 2013), 2014-15 (December 2014) and 2015-16 (November, December, 2015, and May, 2016). Cross-sectional matching can be used to estimate the short-term effects of the VNEN program for each of these years.[3] For each year the length of exposure to the program can be calculated

[3] If the 2013-14 data had been collected before the program was implemented in the VNEN schools, one could relax the matching assumptions by combining matching with difference-in-differences methods (Heckman, Ichimura and Todd, 1997, 1998). Regrettably, this method cannot be used as data collection funds arrived too late: the first data were collected in December 2013, but the program started in the summer/fall of 2012 (Parandekar et al., 2017, p.14).

using the school principal questionnaire data, which record when the program started. This is shown in Table 2. Slightly over 80% (259) of VNEN school principals report that they started the program in the calendar year 2012, and all but one of the rest report that it started in 2013, so virtually all VNEN schools implemented the program before the first data collection in December of 2013. Even some non-VNEN school principals (25, which is about 8%) report starting the program in 2012 or 2013, and by 2015 88 non-VNEN school principals (27%) report that they had started the program. Principals also report when the first VNEN grant was received. Almost all VNEN principals report dates well before December of 2013: 23 report receiving grants in 2012, and 305 (93%) report receiving grants on or before May of 2013, seven months before the first data collection. Thus the first data collection round, December of 2013, is not baseline data.

The primary outcomes of interest for this paper are students' cognitive and non-cognitive skills. The cognitive skills measured were mathematics and Vietnamese (literature). The mathematics and Vietnamese assessments (tests) were developed by Vietnam's Ministry of Education and Training (MoET), assisted by an international testing expert, to measure the skills that the Vietnamese curriculum was designed to impart. The proposed items (test questions) were piloted in a pre-test, and IRT (Item Response Theory) analysis was used to select the items that performed best. The math and reading assessments were administered to 20 students per school in December of 2013, December of 2014, November and December of 2015, and May of 2016.

Students' non-cognitive skills were measured only at the end of the 2015-16 school year, in May of 2016. The 2015-16 parent questionnaire asked parents to assess the non-cognitive skills of their children. More specifically, they were asked to "compare your child's competency progress in the last three years with your expectation". For each of 21 "dimensions" (types of child behavior), parents were asked to rank their child's progress on a four-point scale relative to

their expectations: much higher; somewhat higher; same; a bit lower. These 21 dimensions can be aggregated into four categories: intra-personal (6 items); social (7 items); ethical (4 items); and community (3 items).[4] We focus on the overall skills (include all 21 items) and the two categories with the most items, intra-personal and social.

C. Sample for Evaluating the Long-term Effects of VNEN. This paper also estimates long-term impacts among lower secondary students of having attended a VNEN primary school. This required tracking students from (a subsample of) the 651 primary schools used in the World Bank study. In almost all cases, most students from a given primary school enroll in the same lower secondary school. It is possible that two or more of these 651 primary schools could send most of their students to the same lower secondary school, but this is rare since these 651 schools were spread over 51 provinces and about 300 districts, so that the students from these schools enrolled in almost the same number of lower secondary schools. Thus, to draw a subsample of a given number of lower secondary schools, one can draw a sample of the same number of primary schools from the World Bank study's 651 primary schools. Due to budget constraints, the target sample was 100 lower secondary schools, so a (sub)sample of 100 primary schools was drawn.

The (sub)sample of 100 primary schools was drawn from the 651 VNEN primary schools as follows. First, for the larger research project of which this paper is just one of several, a (new) sample of 140 primary schools, as well as this (sub)sample of 100 lower secondary schools, was drawn. Logistical and budget constraints limited the number of provinces from which these two samples could be drawn, so 23 provinces were randomly drawn to represent Vietnam as a whole. Of these 23 provinces, 17 were among the 51 provinces included in the World Bank VNEN

[4] Examples of these types of behavior are: "making plans and following through with them" (intra-personal); "sharing objects with siblings/friends" (social); "being sensitive to emotional needs of others" (ethical); and "love of nature (plants, animals, river, mountain, etc.)" (community).

study's 651 schools. These 17 provinces have 272 of these 651 primary schools. Of these 272 schools, 18 were dropped due to remoteness or having less than 10 students in grade 3 in 2012, leaving 254 schools (133 VNEN and 121 non-VNEN). Three of these 17 provinces had only one of the 254 schools; all three were dropped since using them would match a VNEN school with a non-VNEN school from a different province. In the remaining 14 provinces, a random sample of the remaining 251 primary schools was drawn with a probability of 100/251 to yield 100 primary schools from the World Bank study: 49 VNEN schools and 51 non-VNEN schools.

These 49 VNEN schools are generally representative of Vietnam's VNEN schools. To be specific, since the 23 provinces are representative of all of Vietnam, the VNEN schools in the 17 (of the 23) provinces with such schools in the World Bank study are representative of the VNEN schools in that study. Aside from dropping about 10% of VNEN schools that lacked common support, the VNEN schools in the World Bank study are representative of all VNEN schools in Vietnam. Except for one VNEN school in the World Bank study that we dropped since it was the only VNEN school in its province in that study (and the World Bank data had no matched non-VNEN school in that province), our sub-sample is a representative sample of the VNEN schools in that study, and so is close to a representative sample of all VNEN schools in Vietnam.

Finally, to obtain a sample of 100 lower secondary schools, the principals of these 100 primary schools (49 VNEN and 51 non-VNEN) were contacted and asked what lower secondary school the majority of their students enrolled in. This led to 99 lower secondary schools being selected (one was linked to two of the 100 selected primary schools). To provide student-level panel data, 30 students were randomly selected per lower secondary school: half of whom had participated in the World Bank study when they were in primary school and half of whom had not participated but were other students in the same 100 primary schools.

This led to a final sample of 99 lower secondary schools, with 2,959 lower secondary students. All 2,959 of these students are matched with a primary school in the World Bank study. Half of them were randomly drawn from those who participated in the World Bank primary school study; all of these (1,450) were matched to their records in that study using their student identification number, name, gender, and date of birth. Mathematics and Vietnamese test score data, and basic demographic data, were collected from these 2,959 lower secondary students, first in November of 2017 (soon after they entered grade 7), then in April of 2018 (one month before finishing grade 7), and finally in June of 2020 (one month before finishing grade 9).[5]

The data collected from these 99 lower secondary schools also include assessments of non-cognitive skills (collected using a student questionnaire) of 12 students per school for three different domains of those skills.[6] The first domain is students' self-perceptions of their competencies, appearance, behavior and overall self-worth, as measured using questions proposed by Harter (2012). We used principal components analysis to aggregate these self-perceptions into a general scale, and we expect that VNEN will raise these self-perceptions. The second is students' intrinsic orientation in the classroom, in particular whether they prefer to master concepts and skills on their own, and whether they have an intrinsic motivation to learn; this is measured by questions developed by Harter (1981). The third domain is the ability to cope when difficulties arise. It consists of several dimensions, such as seeking support from others, relying on oneself, and internalizing behavior (the last is a poor coping strategy). We used principal components

[5] The school year in Vietnam typically ends in May, but all lower secondary schools were closed between early February and early May of 2020 due to the COVID-19 pandemic. Therefore, the data were collected in June of 2021, after the schools had reopened. The 2019-20 school year ended in early July of 2020.

[6] Unlike the cognitive skill data, which were collected by testing 30 students per lower secondary school, the non-cognitive skills were obtained using questions to measure those skills that were administered as part of an interview. Interviews are time-consuming, so only 12 students were (randomly) selected per school to complete them. All 12 were selected from those who attended one of the 100 primary schools that fed into the 99 lower secondary schools; this selection was stratified by (November 2017) grade 7 test scores. Of these 12, 10 were students from whom data were collected in the World Bank study, and the other two were students who did not participate in that study.

analysis to aggregate the second and third domains, and to aggregate all three domains into an overall non-cognitive index. The data were collected at the end of grades 7 (2018) and 9 (2020).

By design, the 1,195 lower secondary students from whom we collected non-cognitive skill data are a random sample of the students in the World Bank study. Indeed, the fractions of students in VNEN schools in grade 5 in 2015-16 who are female, ethnic minorities, and have parents who cannot read or write Vietnamese are similar to those of these 1,195 students in grade 7 in 2017-18 who were in VNEN schools in grade 5. This also holds for the non-VNEN schools.[7]

Because the 100 primary schools are a random subsample of schools from the original 651 matched primary schools, the 49 VNEN schools should be, and indeed are, statistically indistinguishable from the 51 non-VNEN schools in terms of observed school characteristics, as shown in Appendix Table A.4. Student-level data for primary schools are also relatively well balanced. The analogous check for the 49 lower secondary schools that received students from these 49 VNEN primary schools, and the 50 lower secondary schools that received students from these 51 non-VNEN primary schools, is shown in Table A.5; while most variables have similar means across these two sets of lower secondary schools, the percentage of ethnic minority students is significantly higher (36.9% vs 16.7%) in the VNEN schools, and a joint significance test for all eight school-level variables in the table is significant at the 5% level (p-value of 0.04). For student-level variables, the level for our analysis, there are significant differences in students' ethnic minority status, and in their mothers' and fathers' ability to read and write in Vietnamese. The test for joint significance of the eight student-level variables has a p-value of 0.011.

[7] More specifically, in the World Bank study's VNEN schools the proportion of female students was 0.484 in grade 5 (2015-16) and 0.476 in grade 7 (2017-18), and the analogous figures for ethnic minorities were 0.367 and 0.369. Regarding mothers (fathers) who could read and write Vietnamese, the figures are 0.867 (0.921) in grade 5 and 0.859 (0.866) in grade 7. For non-VNEN schools, the proportion of students who were female was 0.535 in grade 5 and 0.509 in grade 7, and the analogous figures for ethnic minorities were 0.184 and 0.169. Regarding mothers (fathers) who could read and write Vietnamese, the figures are 0.955 (0.945) in grade 5 and 0.943 (0.910) in grade 7.

As with the primary school data, we again assume that these two sets of lower secondary schools are statistically indistinguishable in terms of their unobserved characteristics. That is, after matching schools based on the propensity scores of the primary schools that feed into them, or more strictly speaking after conditioning on these propensity scores, we assume that the (potential) outcomes of interest are uncorrelated with these schools' treatment status (the indicator variable for VNEN schools). Given the significant differences between these two sets of schools in students' ethnic minority status and of their mothers' and fathers' ability to read and write in Vietnamese, we control for these differences by adding these three variables to the regressions. This is a standard approach for randomized control trials when there are significant differences in observed variables between the treatment and control groups, and it should reduce the probability that this assumption does not hold with respect to the unobserved variables.

IV. Analytical Framework

This section explains the estimation methods used in this paper. To estimate the short-run impacts using the primary school data, the simplest approach is to use the matched sample of VNEN and non-VNEN schools to estimate the average treatment effect on the treated (ATT) impact of the VNEN program. Yet this may underestimate the impact since many control schools adopted some aspects of the VNEN program and, as explained below, many VNEN schools did not adopt all components of that program, so another approach is to use instrumental variables to estimate the impact of the *intensity* of adopting the VNEN program on student outcomes.

A. Estimation of the Short-term Average Treatment on the Treated. Given the use of nearest neighbor matching to produce matched pairs, if the propensity score matching (PSM) assumptions are met then the short-term ATT of the VNEN program can be estimated by the difference in the means of the outcome variables of interest between the VNEN and the non-

VNEN schools.[8] This can be done for all three years of data (2013-14, 2014-15 and 2015-16) for the 651 primary schools. Yet since some non-VNEN schools started to implement VNEN, as shown in Table 2, this will underestimate the short-term impact of the program.

A very convenient way to calculate the difference in the means between the two sets of schools, and thus to calculate ATT, is to use ordinary least squares (OLS) regression. One can simply regress the outcome variable of interest, the measures of students' cognitive and non-cognitive skills, on a constant term and a dummy variable for the VNEN schools, using the sample of matched pairs. This regression takes the following simple form:

$$Y = \alpha + \beta VNEN + u \qquad (1)$$

where Y is the student's cognitive or non-cognitive skill and $VNEN$ is a binary variable indicating a VNEN school. For convenience, Y is normalized to have a standard deviation of one. The OLS estimate of α is the sample mean of Y for the non-VNEN schools. The OLS estimate of β is the difference in the sample means of Y between the VNEN and non-VNEN schools; this is the PSM estimate of the ATT. The sum of the estimates of α and β is the mean of Y for the VNEN schools.

The statistical precision of the OLS estimate of equation (1) can be increased by adding exogenous variables that have explanatory power for students' skills that are not correlated with participation in the VNEN program. We add three such variables: an ethnic minority dummy variable, and two dummy variables for whether the student's mother and father can read and write Vietnamese).[9] Equation (1) can then be written as:

$$Y_i = \alpha + \beta VNEN_i + \boldsymbol{\gamma}'\boldsymbol{X}_i + u_i \quad (1')$$

[8] There are seven cases of a non-VNEN school being the "nearest neighbors" to two VNEN schools (so this applies to 14 VNEN schools). These seven schools receive a double weight when calculating the non-VNEN school means.
[9] Note that these three variables were not used to calculate the propensity scores used to match non-VNEN to VNEN primary schools. Although that school-level matching did use the proportion of ethnic minority students, the ethnic minority dummy variable in equation (1′) is at the student level, not the school level.

where $\boldsymbol{X}_i$ denotes these three variables, all of which have strong explanatory power.

Equation (1′) consistently estimates the short-term ATT when the standard assumptions for PSM estimation hold, but OLS estimates of the standard error of β are, in general, incorrect. One approach to obtain correct standard errors is to use boot-strapping methods, but Abadie and Imbens (2008) showed that this yields incorrect standard errors when nearest neighbor matching is used. Fortunately, Abadie and Spiess (2021) show that robust standard errors clustered at the matched pairs of schools provide correct inference for nearest neighbor matching.

B. Instrumental Variable Estimation of the Short-term Impact of the Intensity of VNEN Implementation. Even if the ignorability assumption holds, propensity score matching estimation will underestimate the short-term impact of the VNEN program because many of the non-VNEN schools started to adopt all or part of the VNEN program (see Table 2). Under-estimation can also occur if many VNEN schools did not fully implement that program. To accurately measure the extent of VNEN adoption, we constructed a VNEN implementation index for each year, for both VNEN and non-VNEN schools. The details are provided in the appendix.

While one could simply regress the outcomes of interest on a constant term and this index, this may lead to biased estimates of the short-term impact of VNEN implementation because the extent of adoption of that program within VNEN and non-VNEN schools could to be correlated with unobserved factors that directly affect the outcome variable. For example, a teacher in a VNEN school who is worried that VNEN pedagogy may reduce students' learning of math skills may not fully implement VNEN and may also undertake other, unobserved actions that directly affect students' mathematical skills. Yet if the assumptions of cross-sectional match-ing hold, then the VNEN dummy variable would be a valid instrument for the implementation index. Strictly speaking, however, because the VNEN dummy variable takes only two values, an

additional assumption is needed, which is that the impact of the implementation index is linear; this will allow a dummy variable instrument to estimate the short-term impact of this continuous variable for all values of that variable, that is for all degrees of implementation. Unfortunately, it is difficult to assess whether this linearity assumption is reasonable, so this should be kept in mind when interpreting the results of these instrumental variable estimates.

Figures 1, 2 and 3 show density plots of the VNEN implementation index for 2013-14, 2014-15 and 2015-16, respectively. In 2013-14, there is very little overlap between the VNEN (density function on the right) and non-VNEN (density function on the left) schools. Yet over time the VNEN distribution moves slightly to the right, while the non-VNEN distribution spreads out, moving to the right and also increasing its variance. This shows how, over time, many non-VNEN schools adopted, officially or unofficially, the aspects of the VNEN program.

C. Estimation of Long-term Treatment Effects. Finally, consider how the matched sample of 49 VNEN and 50 non-VNEN lower secondary schools can be used to estimate the long-term treatment effects of implementing the VNEN program in Vietnam's primary schools. The simplest approach is to run an OLS regression on the students in the sampled schools, using equation (1′); note that the three control variables added not only increase precision but also account for the lack of balance in these variables that was described in subsection III.C. The estimate of β is the PSM estimate of the average treatment effect on the treated (ATT).

Yet this estimate of β in equation (1′) suffers from two sources of bias. First, some lower secondary school teachers adopted some VNEN teaching methods, and the proportion who did differs for the two sets of lower secondary schools. Among the 49 lower secondary schools that received students from VNEN primary schools, 24% of the teachers report "voluntarily" using at least some VNEN teaching techniques, compared to only 6% of the teachers in the 50 lower

secondary schools that received students from non-VNEN primary schools. This could over-estimate the long-term impact of being in a VNEN primary school, assuming similar effects at the primary and lower secondary levels. Yet one could argue that the long-term impact of adopting VNEN at the primary level includes the higher rate of "voluntary" VNEN adoption at the lower secondary level. If one accepts this argument, then any impact from lower secondary school teachers using VNEN methods is not upward bias; the OLS estimate is simply a more complete estimate of the long-run impact of being in a VNEN primary school.

A second source of bias is in the opposite direction. OLS estimates will underestimate the long-run impact of the VNEN program because many of the non-VNEN primary schools adopted at least part of the VNEN program (see Table 2), and some VNEN schools did not fully adopt VNEN (see Figures 1, 2 and 3). This bias could be avoided by regressing Y on the VNEN implementation index described in subsection IV.B, but recall that the extent of VNEN adoption within VNEN and non-VNEN schools could be correlated with unobserved factors that directly affect the outcome variables. Yet if the assumptions for estimating equation (1′) hold, then the VNEN dummy variable is a valid instrument for the implementation index. However, one also must assume that the implementation index's impact is linear, as explained in subsection IV.B.

Figures 4, 5 and 6 present density plots of the VNEN implementation index in 2013-14, 2014-15 and 2015-16, respectively, for the 100 primary schools that correspond to the 99 lower secondary schools. They are very similar to Figures 1, 2 and 3, as one would expect since these 100 primary schools are effectively a random sample of the schools used in those figures. As before, over time, many non-VNEN schools adopted, officially or unofficially, at least some aspects of the VNEN program. To construct an average measure of VNEN implementation over these three years, we calculate a simple average value of the index over these three years.

V. Estimation Results

This section presents the estimation results, focusing on students' skills. The first and second subsections present estimates of the short-term impact of the VNEN program on students' mathematics and Vietnamese skills and on their non-cognitive (socioemotional) skills, respectively. The third and fourth subsections present estimates of the long-term estimates of the VNEN program on cognitive and non-cognitive skills, respectively.

A. Short-term Impacts on Students' Cognitive Skills. Table 3 presents propensity score matching estimates, that is OLS estimates of equation (1′), of the impact of the VNEN program on students' mathematics and Vietnamese test scores for the academic years 2013-14, 2014-15 and 2015-16. Students were tested twice in the 2015-16 school year, first in December of 2015 and again in May of 2016; the first set of estimates is the column labeled 2015 and the second set is in the column labeled 2016. The same students are followed over time, so they were in grade 3 in 2013-14, grade 4 in 2014-15 and grade 5 in 2015-16. For ease of interpretation, all test score variables are normalized to have a standard deviation of 1 (based on the distribution of test scores in non-VNEN schools). Recall that, as shown in Table 2, some non-VNEN schools started to adopt the VNEN program, so these estimates likely underestimate the impact of the program, especially in the 2015-16 school year, by which time about 27% of the non-VNEN schools report having adopted the VNEN program.

The results in Table 3 show only small impacts of the VNEN program on math and Vietnamese scores. All estimates are positive, and some are statistically significant, but the largest estimate, for Vietnamese in 2014, is an increase of only 0.11 standard deviations of the distribution of test scores (of non-VNEN schools). By 2016 the impacts for math and Vietnamese had dropped to 0.05 and 0.04 standard deviations, respectively, and neither is statistically significant.

Adoption of the VNEN program by about one fourth of non-VNEN schools, and some VNEN schools not fully implementing VNEN, could lead to underestimation of the program's impact. One way to avoid this bias is to regress the math and Vietnamese skill variables on the index for implementing the program. As explained in Section IV, using OLS could lead to biased estimates, so this index is instrumented using the VNEN dummy variable which, if the matching estimation assumptions hold, is uncorrelated with unobserved factors determining students' math and Vietnamese skills. The results are shown in Table 4. As in Table 3, all estimates are positive, and some are statistically significant, but after three years of the program the estimated impacts are modest and at best only weakly significant. The VNEN index is scaled so that it ranges from 0 (no implementation of VNEN) to 1 ("full" implementation). Thus the estimated impact of going from "no" to "full" implementation of VNEN is an increase of 0.21 standard deviations (significant at the 10% level) for math and 0.15 (not significant) for Vietnamese. Yet, as seen in figure 2, even in 2013, when only about 8% of non-VNEN schools report having adopted VNEN (see Table 2), very few VNEN schools are at the far left, the area where no VNEN methods are used. This implies that almost no schools went from "no" to "full" implementation of VNEN, so the above point estimates are likely to slightly exaggerate the impact of full adoption of VNEN.

Even if the average effects of adopting VNEN are small, there may be effects on certain population subgroups, although such estimates must be interpreted with caution because testing multiple hypotheses may produce "significant" impacts simply by chance when the number of subgroups is large. One obvious distinction is by student gender, so the first panel of Table 5 presents estimates of impacts on mathematics separately for girls and boys, and the first panel of Table 6 does the same for Vietnamese. The estimates for math in Table 5 are insignificant for girls, but they are significant for boys for the first three of the four years. The significant point

estimates for boys, however, are quite small – between 0.08 and 0.10 – and fall to 0.07 (and lose significance) by 2016. Turn to the impacts on Vietnamese in Table 6, the estimates for girls are small and mostly insignificant. In contrast, the estimated impacts for boys are significant in all years and range from 0.07 to 0.16 standard deviations, including a final estimate in 2016 of 0.10 standard deviations. Overall, there is little evidence of an impact on girls, but boys Vietnamese skills increased by a modest amount, with perhaps a smaller impact on their mathematics skills.

Perhaps the most important distinction to make is whether students are *Kinh* (ethnic Vietnamese) or members of one of Vietnam's many ethnic minority groups, which together constitute about 15% of Vietnam's population. Recall from Sections II and III that schools with more ethnic minority students were more likely to be invited to adopt the VNEN program, and the program itself has several aspects designed to encourage ethnic minority students, such as inviting parents to help teachers communicate with ethnic minority students, highlighting local ethnic handicrafts, and certain tailor-made lessons (for example, on life sciences) to suit the local context. The second panel of Table 5 provides estimates of the impact of the VNEN program on math skills, separately for *Kinh* and ethnic minority students, and the second panel of Table 6 does the same for Vietnamese skills. There is little evidence in Table 5 of a large differential impact of VNEN across these two groups of students; all estimates for both groups are small and statistically insignificant. In contrast, the estimated impacts in Table 6 for Vietnamese in 2013 and 2014 are significant at the 5% level for ethnic minority students, and these two point estimates approach a level that may be of policy interest: 0.17-0.18 standard deviations of the distribution of test scores. However, the impacts are much smaller (0.08 and 0.10) and statistically insignificant in 2015 and 2016. For *Kinh* students, there is a small (0.07 standard deviations) and significant impact in 2014, but the impacts for other years are even smaller, and

all are statistically insignificant. Overall, the VNEN program's components that are intended to help ethnic minority students had, at best, a modest impact on those students' Vietnamese skills.

A final way to divide the population is geographically, splitting Vietnam into North, Central and South. The results are shown in the third panels of Tables 5 and 6, and the impacts clearly vary for these three areas. For the Northern areas of Vietnam, the VNEN program significantly increased mathematics scores in all four years by 0.15 to 0.17 standard deviations, while in the Central and Southern areas of Vietnam the estimated impacts on math scores are much smaller, and all are statistically insignificant. Turning to Vietnamese scores (Table 6), the results are similar. All four estimates for the North are significant, although the impacts are smaller in the last two years (0.11 in 2015 and 0.12 in 2016) than in the first two years (0.26 in 2013 and 0.22 in 2014). The estimates for the Central and Southern provinces are all much smaller and statistically insignificant. Overall, the VNEN program seems to have increased both mathematics and Vietnamese test scores in the northern third of Vietnam, but had little to no effect in the central and southern provinces. This likely reflects the high percentage of ethnic minority students in the North (68%) compared to the Center (18%) and South (17%).

B. Short-term Impacts on Students' Non-Cognitive Skills. While the overall impact of VNEN on mathematics and Vietnamese test scores is modest at best, one could argue that the VNEN program's focus was on other skills that would make students more productive workers when they are adults. Some observers worried that the VNEN program would reduce students' learning of cognitive skills, but the estimates in the previous subsection show that this did not occur, at least for mathematics and Vietnamese skills. This subsection addresses the next question: Did the VNEN program increase students' non-cognitive (socio-emotional) skills? Note that these skills were measured only once during the 2015-16 school year.

The first row of Table 7 shows propensity score matching estimates of the (ATT) impact of the VNEN program on intra-personal, social, and overall (all 21 dimensions) non-cognitive skills. All estimates are positive, and although the intra-personal skills estimate is insignificant, those for social skills and overall non-cognitive skills are significant at the 10% and 5% levels, respectively. Thus the VNEN program appears to have increased students' non-cognitive skills, although the estimated impacts are small: the impacts are only 0.05, 0.07 and 0.08 standard deviations for the intra-personal category, the social category, and the overall measure, respectively. The second panel of Table 7 presents instrumental variable estimates of the impact of the VNEN implementation index, in order to avoid underestimating the impact due to some non-VNEN schools adopting VNEN and some VNEN schools not fully adopting it. All estimated impacts are positive, and those for the social and combined skills are significant at the 10% level. All three are large enough to be of policy interest, ranging from 0.15 to 0.22 for a school switching from "no" to "full" VNEN implementation. Yet recall that most non-VNEN schools were not at "zero" even in 2013, so even full implementation does not, on average, raise non-cognitive skills by these amounts. Also, these estimates are similar in size to those for math and Vietnamese skills, so VNEN seems equally effective at raising both cognitive and non-cognitive skills.

As with cognitive skills, the VNEN program may have had different effects for different students. This is examined in Table 8. The first panel examines differences in effects for boys and girls. Although all three estimated impacts for boys, unlike two of the three for girls, are statistically significant, the differences between the two sets of estimates are not statistically significant, so there seems to be little or no gender difference.

The second panel of Table 8 presents separate impacts for *Kinh* (ethnic Vietnamese) and ethnic minority students. A priori, one would expect larger impacts for the latter since some

VNEN components were intended to help ethnic minority students. The results strongly support this outcome, with statistically significant improvements in all three skills for ethnic minority students, but much smaller and statistically insignificant results for *Kinh* students. These estimated impacts are large enough to be of policy interest; for example, the overall index estimate is an increase of 0.15 standard deviations of that variable's distribution.

Finally, the third panel in Table 8 presents separate results for the North, Central and South of Vietnam. The impacts for the North are modest at best, between 0.07 and 0.12 standard deviations, but two are statistically significant. The impacts for the South are slightly larger, between 0.09 and 0.15 standard deviations, and one, for social skills, is significant at the 10% level. In contrast, the impacts for Central provinces are smaller and statistically insignificant. One explanation for the large impact in the North is that about 68% of the sampled students in the North are ethnic minority students, while the proportions in the Central and South are only 18% and 17%, respectively, but this does not explain the large impact in the South.

C. Long-term Impacts on Cognitive Skills. The data collected from the 99 lower secondary schools near the beginning and end of the 2017-18 school year (November, 2017, and April, 2018) and near the end of the 2019-20 school year (June, 2020) included assessments of the cognitive skills of close to 3,000 students in those schools. These students were given tests of their mathematics and Vietnamese (literature) skills for each of these three rounds of data collection. The assessments (tests) used were developed by Vietnam's Ministry of Education and Training, assisted by an international testing expert, to measure the skills that the Vietnamese curriculum was designed to impart. The proposed items (test questions) were administered in a pre-test, and IRT (Item Response Theory) analysis was used to select the best-performing items. The final tests used for the November 2017 round of data collection included 35 questions

on the math test and 30 questions on the Vietnamese test. The analogous figures for the April 2018 round of data collection are 31 and 30, and for the June 2020 round they are 35 and 35.

Estimates of Equation (1′) are shown in Table 9, and instrumental variable (IV) estimates are shown in Table 10. Note that the need to control for students' ethnic minority status and their parents' ability to read and write Vietnamese reduces the sample to about 1,200 because the interviews that collected these data were done for only 12 students per school. The estimated long-term impacts of the VNEN program on mathematics skills, shown in Table 9, are slightly negative and statistically insignificant. For example, the largest estimated impact, in 2017 when the students were in grade 7, is -0.052 standard deviations of the distribution of the math test scores in the non-VNEN schools. For Vietnamese skills, the estimated impacts are all very close to zero and statistically insignificant. The small samples reduce the precision of the results, yet it is still possible to rule out large positive impacts. For example the upper bound of the 95% confidence interval for grade 7 math scores is 0.14, which implies that the actual impact is unlikely to be larger, and very likely to be much smaller. For Vietnamese the largest upper bound for a 95% confidence interval is 0.17. These results are not surprising, given that the short-term effects in Table 3 were not particularly large and seemed to decline over time.

The OLS estimates in Table 9 may be biased downward due to some of the non-VNEN schools adopting the VNEN program, and also due to some VNEN schools not fully implementing the VNEN program. This bias can be avoided by regressing the math and Vietnamese skill variables on the VNEN implementation index. As explained in Section IV, OLS regressions could lead to biased estimates, so this index is instrumented using the VNEN dummy variable which, if the assumptions of the matching estimation hold, is uncorrelated with all unobserved factors determining students' mathematics and Vietnamese skills.

Table 10 shows IV estimates of the impact of the VNEN index. The F-statistics for the instrument are very high, indicating no weak instrument problems. The estimated impacts on students' mathematics scores are all negative, but are too small to be statistically significant. The estimated impacts on Vietnamese scores are either close to zero or slightly negative, and all are far from statistically significant. While these estimates are imprecise due to the small sample, the associated confidence intervals often rule out very large effects. For example, the 95% confidence interval for grade 7 math (in 2017, 2018 is similar) ranges from -0.68 to +0.38 standard deviations of the distribution of test scores, and the 90% confidence interval ranges from -0.59 to + 0.30; despite these estimates' imprecision, they do rule out very large impacts. For Vietnamese, the upper bounds of the 90% and 95% confidence intervals are higher, but even here very large effects can be ruled out – e.g. for grade 9 (2020) the upper bound on the 90% confidence interval is 0.39 standard deviations – and the point estimates are essentially zero or slightly negative. Recall also that almost no VNEN primary schools went from "no" to "full" implementation of VNEN, so actual impacts are lower than these estimates.

The small and statistically insignificant estimates of the average long-term effects of the VNEN program may hide larger effects on some subgroups of the population. Tables 11 and 12 present OLS estimates of impacts on mathematics and Vietnamese, respectively, separately for boys and girls, ethnic minority and *Kinh* (ethnic Vietnamese) students, and the three different regions of Vietnam (North, Central and South). The results are easily summarized: for none of these groups is there a statistically significant long-term impact of the VNEN program on math skills, and for only two of the 21 estimates is there a significant impact on Vietnamese skills. The only pattern of interest is that there seems to be a large difference in the impacts of the program on *Kinh* and ethnic minority students: negative impacts for the former and positive

impacts for the latter. While these impacts are not significantly different from zero, in four of six cases (math in 2017 and 2018, and Vietnamese in 2017 and 2020) the coefficients for the *Kinh* students and ethnic minority students are significantly different from each other at the 10% (but not the 5%) level. Also, one cannot rule out sizeable impacts for ethnic minority students given the imprecision of the estimates. For example, the upper bounds for the 90% and 95% confidence intervals for grade 9 mathematics are 0.30 and 0.35, respectively; impacts in this range, which are unlikely but possible, would make a strong case for adopting this program.[10]

Overall, averaging over all students there is no evidence of any sizeable long-term impact of the VNEN program on students' mathematics and Vietnamese skills, although there may be sizeable impacts on ethnic minority students.

D. Long-term Impacts on Non-Cognitive Skills. The main benefit of the VNEN program may be its impact on non-cognitive (socioemotional) skills. Table 13 provides estimates of the long-term impact on those skills of being enrolled in a VNEN school while in primary school. Of the eight reported estimates (four measures of non-cognitive skills for two different years), all are small and statistically insignificant. While the small sizes reduce the precision of the results, the results for the aggregated non-cognitive skill variable rule out large effects. In particular, the 95% confidence interval in grade 9 ranges from -0.22 to +0.11 standard deviations, which rules out effects greater than 0.11 standard deviations Thus, we conclude that, on average, the VNEN program produced no long-lasting benefits in terms of students' non-cognitive skills. Table 14 reports the associated instrumental variable (IV) regressions, and again all eight estimates are small and statistically insignificant. Note in particular that the 95% confidence interval for the aggregate impact in grade 9 ranges from -0.61 to +0.31 standard

[10] The strong impact on Vietnamese scores in the northern areas of Vietnam likely reflects the impact of the program on ethnic minority students; about 68% of lower secondary students in those areas are ethnic minority students.

deviations, and the 90% confidence interval ranges from -0.54 to +0.23 standard deviations, which implies that sizeable impact (e.g. > 0.25 standard deviations) are unlikely. Moreover, since almost no VNEN primary schools changed from "no" to "full" implementation of VNEN, actual impacts are lower than these estimates.

Despite the lack of statistical significance for Vietnamese students as a whole, there may be significant impacts on specific subsets of students. This is examined in Tables 15, 16 and 17. Table 15 examines differences by gender. There are no significant effects of any kind for boys or girls in grade 7. In grade 9, there are no significant effects for girls, and only one marginally significant (10% level) negative impact for boys, on intrinsic orientation. The aggregate impacts in the last column of Table 15 show no significant effects, and except for one (insignificant) negative effect for girls in grade 9 the estimated aggregate effects are all very small. Overall, the (lack of) long-term impacts of the VNEN program show little or no differentiation by gender.

Table 16 focuses on differences between ethnic Vietnamese (*Kinh*) and ethnic minority students. For *Kinh* students, all eight estimates are small and none is statistically significant, so there is no evidence of any impact on these students. By grade 9, the 95% confidence interval for the aggregate index ranges from -0.26 to + 0.14, which rules out sizeable positive effects. For ethnic minority students, there is a marginally significant (10% level) negative effect when they were in grade 7, a reduction in their intrinsic orientation. For grade 9, there is a significantly positive impact on coping mechanisms, but again a negative impact on intrinsic orientation (both significant at the 5% level). Overall, there is some evidence of positive and negative impacts for ethnic minority students, but the estimated impact on aggregated non-cognitive skills over these three domains (last column of Table 16) is small and statistically insignificant, and the 95% confidence interval for aggregate skills in grade 9 rules out impacts over 0.14 standard deviations.

Finally, Table 17 considers differences by region. When students were in grade 7, there were no significant effects for the North and Central regions. However, for the South there are large (0.30 standard deviations) and significantly positive impacts on both students' self-perceptions and on the aggregate index (both at the 5% level). Yet, by grade 9, there are both signifycantly negative (intrinsic orientation, 1% level) and positive (coping, 5% level) impacts in the North, but for the aggregate index there is no overall effect. In the Central and South regions there are no significant effects of any kind. Thus, while there seem to be differential effects by region when the students were in grade 7, they dissipate by the time they are in grade 9. For both the North and Central regions, the 95% confidence interval for grade 9 aggregate skills rules out impacts greater than 0.14 standard deviations, although for the South the (more imprecisely estimated) confidence interval ranges from -0.27 to +0.40 standard deviations.

VI. Conclusion

Vietnam is often considered to be an education success story, in part because of its performance on the 2012 and 2015 PISA assessments, but Vietnamese policymakers and others are concerned that Vietnam's education system is not producing adequate skills that the country will need for future economic growth and, more generally, a higher quality of life. One program that Vietnam has recently implemented to address this concern is the Vietnam *Escuela Nueva* (VNEN) program. This paper has used propensity score matching to assess of the short- and long-term impacts of that program on the cognitive (mathematics and Vietnamese) and non-cognitive (socioemotional) skills of students in Vietnam. To our knowledge, this is the first rigorous quantitative evaluation of *Escuela Nueva* for any country.

The impact of VNEN on students' short-term cognitive skills are modest at best. Of the ATT estimates that aggregate over all students, the largest impact in any year is only 0.11

standard deviations of the distribution of test scores, and by the end of the third year (2016) the impacts on math and Vietnamese are only 0.05 and 0.04 standard deviations, respectively, and statistically insignificant. Even moving from "no" to "full" implementation of VNEN, an extreme scenario, would increase math and Vietnamese by only 0.21 and 0.15 standard deviations, respectively, and the former is significant only at the 10% level and the latter is insignificant. Disaggregating by gender, ethnic group and region shows at most a modest impact on the Vietnamese skills of boys and of ethnic minority students, and modest impacts on both math and Vietnamese skills in Northern Vietnam (which in this sample is mostly ethnic minority students).

Turning to short-term non-cognitive skills, ATT estimates averaging over all students are again modest for all measures of those skills, ranging from 0.05 to 0.08 standard deviations. Estimates of going from "no" to "full" VNEN implementation are larger, ranging from 0.15 to 0.22 standard deviations, but again few schools actually move from 0% to 100% implementation of the program. For specific subgroups of students, there is little or no difference in impacts by gender, but for ethnic minority students the program does seem to increase the aggregate index of these skills by 0.15 standard deviations. Regarding regions, there are modest impacts in the North (0.08 to 0.12 standard deviations) and the South (0.09 to 0.15).

Long-term impacts are even smaller, although less precisely estimated. For cognitive skills, when averaging over all students, by grade 9 (four years after exposure to the program ended) there is no discernable impact on either mathematics or Vietnamese skills, although one cannot rule out possibly large impacts for ethnic minority students. A similar result holds for long-term non-cognitive skills; there is little evidence of a sizeable impact on average skills, or on any particular group of students.

In summary, the VNEN program raised both cognitive and non-cognitive skills to some extent, particularly for ethnic minority students. There is no evidence of any harm to students' cognitive skills, at least in terms of their mathematics and Vietnamese skills. Yet, as with many programs, these impacts fade out over time, so that four years after the program had ended there is little evidence of any long-term effect, although the precision of these estimates is lower.

Future research on the VNEN program in Vietnam could take several directions. First, reasons for the variation of the short-term impacts on non-cognitive skills, especially in the Northern and Southern thirds of the country (and the lack of impact in the Central third of the country) are puzzling; further research to understand this variation could be of interest. Second, it would be very useful to estimate the impact of both cognitive and non-cognitive skills on labor productivity, to see whether the non-cognitive skills that the VNEN program was designed to increase are the same ones that make individuals more productive workers when they are adults. Third, after the World Bank data were collected, the Ministry of Education and Training (MoET) in Vietnam decided not to further expand the VNEN program, in part because of opposition from parents; the reasons for this opposition, and for MoET's decision not to go forward with VNEN, would be of interest in order to better understand education policy making in Vietnam.

References

Abadie, Alberto, and Guido Imbens. 2008. "On the Failure of the Bootstrap for Matching Estimators." *Econometrica* 76 (6): 1537–57.

Abadie, Alberto, and Jann Spiess. 2021. "Robust Post-Matching Inference", *Journal of the American Statistical Association.* https://doi.org/10.1080/01621459.2020.1840383

Bodewig, Christian, Reena Badiani-Magnusson, Kevin Macdonald, David Newhouse and Jan Rutkowski. 2014. *Skilling Up Vietnam: Preparing the Workforce for a Modern Market Economy*. Washington, D.C.: World Bank.

Dang, Hai-Anh. 2013. "Private Tutoring in Vietnam: A Review of Current Issues and Its Major Correlates". In Janice Aurini, Julian Dierkes and Scott Davis. (Eds.) *Out of the Shadows: The Global Intensification of Supplementary Education.* International Perspectives on Education and Society, Vol. 22: 95-128. Emerald Press.

Dang, Hai-Anh, and Paul Glewwe. 2018. "Well Begun, but Aiming Higher: A Review of Vietnam's Education Trends in the past 20 Years and Emerging Challenges". *Journal of Development Studies* 54(7):1171-1195.

Dang, Hai-Anh, Trung Hoang, and Ha Nguyen. 2021. "The Long-run and Gender-equalizing Impacts of School Access: Evidence from the First Indochina War". *Economic Development and Cultural Change,* 70(1): 453-484.

Glewwe, Paul, and Hanan Jacoby. 1998. "School Enrollment and Completion in Vietnam: An Investigation of Recent Trends". In D. Dollar, P. Glewwe and J. Litvack, eds., *Household Welfare and Vietnam's Transition*. World Bank Regional and Sectoral Studies. Washington, D.C.: World Bank.

Government of Vietnam. 2014. *Resolution No. 44/NQ-CP*, June 09, 2014.

Harter, Susan. 1981. "A Scale of Intrinsic versus Extrinsic Orientation in the Classroom". Department of Psychology, University of Denver. Denver, CO.

Harter, Susan. 2012. "Self-Perception Profile for Children: Manual and Questionnaires (Grades 3-8)". Department of Psychology, University of Denver. Denver, CO.

Heckman, James, Hidehiko Ichimura, and Petra Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program." *Review of Economic Studies* 64 (4): 605–54.

Heckman, James, Hidehiko Ichimura, and Petra Todd. 1998. "Matching as an Econometric Evaluation Estimator." *Review of Economic Studies* 65 (2): 261–94.

Kline, Rachel. 2002. "A Model for Improving Rural Schools: *Escuela Nueva* in Colombia and Guatemala". *Current Issues in Comparative Education* 2(2):171-181.

London, Jonathan. 2011. *Education in Vietnam*. Singapore: Institute of Southeast Asian Studies.

McEwan, Patrick. 1998. "The Effectiveness of Multigrade Schools in Colombia". *International Journal of Educational Development*, 18(6): 435–452.

Parandekar, Suhas, Futoshi Yamauchi, Andrew Ragatz, Elisabeth Sedmik and Akiko Sawamoto. 2017. "Enhancing School Quality in Vietnam through Participative and Collaborative Learning: Vietnam *Escuela Nueva* Impact Evaluation Study." The World Bank, Washington, DC.

Reed, John. 2018. "Education in Vietnam: very good on paper". *Financial Times*. Accessed on March 28, 2021 at https://www.ft.com/content/da4387d0-aba8-11e8-8253-48106866cd8a.

World Bank. 2014. *Skilling up Vietnam: Preparing the workforce for a modern market economy*. Hanoi: The World Bank.

World Bank. 2019. *World Development Report 2019: The Changing Nature of Work.* Washington, DC: World Bank.

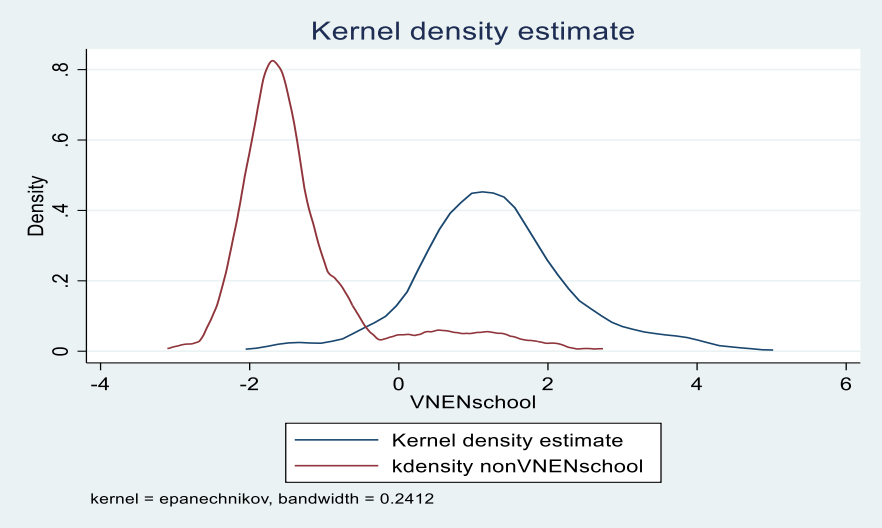

Figure 1: VNEN Implementation Index for 2013

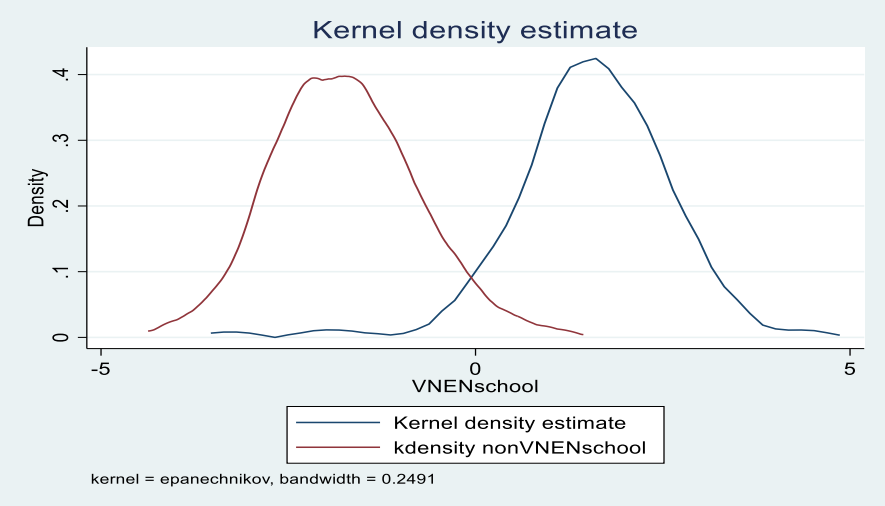

Figure 2: VNEN Implementation Index for 2014

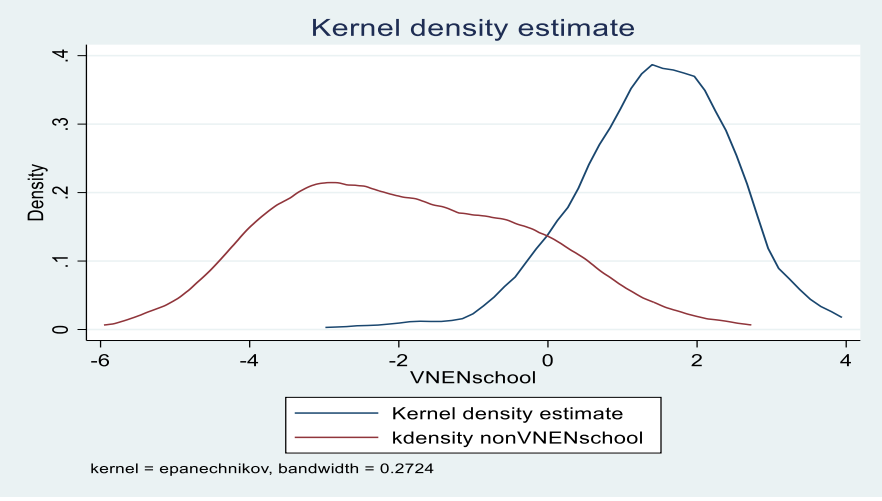

Figure 3: VNEN Implementation Index, 2015

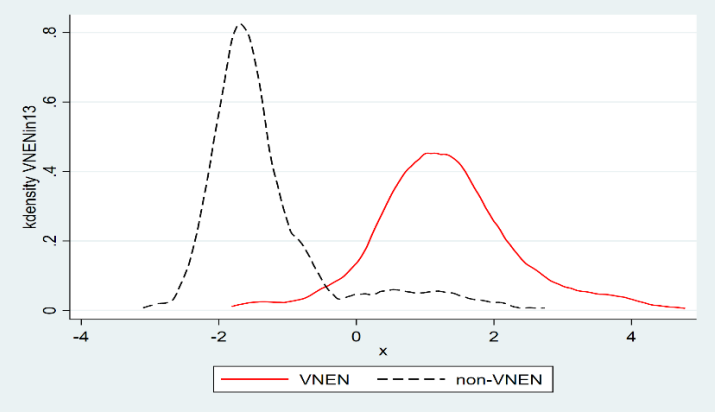

Figure 4: VNEN Implementation Index for 2013 for 100 Primary Schools Linked to 99 Lower Secondary Schools

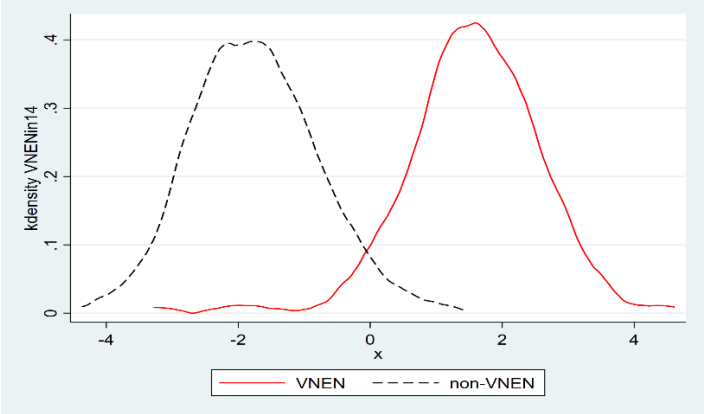

Figure 5: VNEN Implementation Index for 2014 for 100 Primary Schools Linked to 99 Lower Secondary Schools

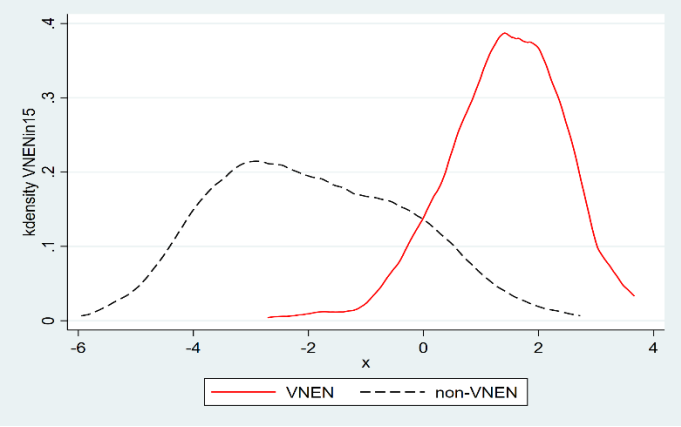

Figure 6: VNEN Implementation Index for 2015 for 100 Primary Schools Linked to 99 Lower Secondary Schools

Table 1: The Eight Fundamental Characteristics of VNEN Schools

Classroom Organization

1. Students sit in groups of 4 or 5, with desks arranged so that they face each other. There is enough room in the classroom for the teacher to move among the students, and the students can move among each other.
2. The classroom has six "artifacts":
    a) Community map, which shows the school and each student's home
    b) Classroom library
    c) Learning corner, which includes didactic materials and displays of students' work
    d) Community corner, which contains objects important to the community
    e) Student mailboxes, for students to write messages to each other
    f) Student government organizational chart
3. Learning guides, which are workbooks designed to help students become independent learners

Pedagogy

4. Teachers use "real life" objects in class discussions to relate classroom lessons to "real life". Teachers use games as part of their lessons. Students are encouraged to ask questions, both of the teacher and of other students.
5. Group work: Students are organized into small groups, sitting together in desks that face each other, to working together to complete assignments.
6. Use of formative assessments rather than summative assessments, which can be done both formally using tests and quizzes but also informally through conversations of teachers with students and with other teachers.
7. Establishment of student government, in order to increase the value that students place on service and on belonging to a community, and to develop responsibility and leadership skills, with each student having the opportunity to be a leader.
8. Parental engagement, in order for parents to become more involved in their children's education, both at home and at school.

Table 2: Date when School Started Implementing VNEN, from Principal Questionnaire

		VNEN schools	non- VNEN schools
Year when school first implemented the VNEN model (2015-16 principal questionnaire, questions Q1002b and Q1002c)	2012	259	7
	2013	61	18
	2014	1	29
	2015	0	34
	Never	0	242
Month and year that school received a VNEN school sub-grant (2013-14 principal questionnaire, question 106) This question was not asked for non-VNEN schools.	Before 2012	2	
	January 2012	2	
	February 2012		
	March 2012	6	
	April 2012	5	
	May 2012		
	June 2012		
	July 2012		
	August 2012		
	September 2012	1	
	October 2012	2	
	November 2012	1	
	December 2012		
	Month unknown 2012	4	
	January 2013	20	
	February 2013	15	
	March 2013	110	
	April 2013	110	
	May 2013	27	
	June 2013	7	
	July 2013	3	
	August 2013	2	
	September 2013	1	
	October 2013	1	
	Month unknown 2013	3	1

Table 3: Propensity Score Matching Estimates of the Impact of VNEN on Math and Vietnamese Skills, 2013-2016 (Regression Based Estimates)

VARIABLES	2013	2014	2015	2016
Mathematics				
VNEN dummy	0.058	0.060*	0.067**	0.051
	(0.041)	(0.034)	(0.033)	(0.038)
Minority	-0.594***	-0.610***	-0.624***	-0.584***
	(0.039)	(0.035)	(0.034)	(0.038)
Mother knows Vietnamese	0.417***	0.526***	0.501***	0.409***
	(0.055)	(0.049)	(0.035)	(0.037)
Father knows Vietnamese	0.306***	0.354***	0.321***	0.277***
	(0.040)	(0.036)	(0.028)	(0.029)
Constant	-0.395***	-0.552***	-0.503***	-0.400***
	(0.067)	(0.064)	(0.051)	(0.052)
Observations	11,630	11,605	11,387	11,218
Vietnamese				
VNEN dummy	0.096**	0.114***	0.033	0.039
	(0.040)	(0.035)	(0.034)	(0.032)
Minority	-0.557***	-0.504***	-0.450***	-0.430***
	(0.038)	(0.035)	(0.034)	(0.034)
Mother knows Vietnamese	0.591***	0.654***	0.700***	0.623***
	(0.058)	(0.055)	(0.052)	(0.044)
Father knows Vietnamese	0.295***	0.412***	0.347***	0.345***
	(0.039)	(0.040)	(0.033)	(0.032)
Constant	-0.544***	-0.753***	-0.756***	-0.697***
	(0.068)	(0.068)	(0.066)	(0.058)
Observations	11,630	11,595	11,387	11,218

Notes: Standard errors are clustered at the matched school pair level. The outcome variables are IRT latent values (2-parameter model) normalized using the non-VNEN students. All regressions are weighted in this and later tables are weighted, where weights equal 1 for all VNEN schools and equal the number of matches for non-VNEN schools (2 for seven non-VNEN schools and 1 for all other non-VNEN schools).

Table 4: Propensity Score Matching Estimates of the Impact of VNEN on Math and Vietnamese Skills, 2013-2016 (IV Regression Based Estimates)

VARIABLES	2013	2014	2015	2016
Mathematics				
VNEN Index	0.177	0.166*	0.243**	0.212*
	(0.124)	(0.097)	(0.101)	(0.116)
Minority	-0.593***	-0.611***	-0.615***	-0.575***
	(0.039)	(0.036)	(0.036)	(0.041)
Mother knows Vietnamese	0.417***	0.525***	0.501***	0.413***
	(0.055)	(0.050)	(0.037)	(0.039)
Father knows Vietnamese	0.308***	0.351***	0.307***	0.264***
	(0.040)	(0.036)	(0.029)	(0.032)
Constant	-0.437***	-0.591***	-0.616***	-0.508***
	(0.079)	(0.077)	(0.080)	(0.087)
Kleibergen-Paap F-stat	1,079.3	1,799.3	823.2	826.2
Observations	11,557	10,901	9,779	9,636
Vietnamese				
VNEN Index	0.289**	0.308***	0.163	0.149
	(0.121)	(0.099)	(0.109)	(0.104)
Minority	-0.555***	-0.505***	-0.451***	-0.424***
	(0.038)	(0.036)	(0.036)	(0.036)
Mother knows Vietnamese	0.592***	0.645***	0.711***	0.645***
	(0.057)	(0.056)	(0.056)	(0.046)
Father knows Vietnamese	0.296***	0.417***	0.325***	0.325***
	(0.039)	(0.040)	(0.033)	(0.033)
Constant	-0.613***	-0.838***	-0.837***	-0.775***
	(0.077)	(0.083)	(0.100)	(0.091)
Kleibergen-Paap F-stat	1,079.3	1,799.3	823.2	826.2
Observations	11,557	10,891	9,780	9,636

Notes: Standard errors are clustered at the matched school pair level. The outcome variables are IRT latent values (2-parameter model) normalized using the non-VNEN students.

Table 5: Propensity Score Matching Estimates of the Differential Impact of VNEN on Mathematics Skills, 2013-2016 (Regression Based Estimates)

	2013	2014	2015	2016
Gender				
Girl	0.115***	0.158***	0.067***	0.079***
	(0.024)	(0.025)	(0.025)	(0.026)
VNEN x Girl	0.021	0.030	0.052	0.038
	(0.044)	(0.037)	(0.035)	(0.039)
VNEN x Boy	0.097**	0.094**	0.084**	0.067
	(0.044)	(0.040)	(0.040)	(0.044)
Minority	-0.595***	-0.613***	-0.625***	-0.585***
	(0.039)	(0.035)	(0.034)	(0.038)
Mother knows Vietnamese	0.417***	0.525***	0.500***	0.409***
	(0.055)	(0.049)	(0.035)	(0.037)
Father knows Vietnamese	0.303***	0.349***	0.318***	0.274***
	(0.040)	(0.035)	(0.028)	(0.029)
Constant	-0.448***	-0.625***	-0.534***	-0.437***
	(0.069)	(0.066)	(0.054)	(0.054)
Minority				
VNEN x Minority	0.111	0.089	0.076	0.079
	(0.071)	(0.062)	(0.053)	(0.054)
VNEN x Kinh	0.027	0.043	0.062	0.034
	(0.047)	(0.040)	(0.040)	(0.048)
Minority	-0.637***	-0.634***	-0.631***	-0.607***
	(0.054)	(0.051)	(0.048)	(0.052)
Mother knows Vietnamese	0.418***	0.526***	0.501***	0.410***
	(0.055)	(0.049)	(0.035)	(0.037)
Father knows Vietnamese	0.306***	0.354***	0.321***	0.276***
	(0.040)	(0.036)	(0.028)	(0.029)
Constant	-0.380***	-0.544***	-0.501***	-0.392***
	(0.069)	(0.065)	(0.053)	(0.055)
Region				
Central	0.112*	0.080	0.019	0.034
	(0.063)	(0.057)	(0.052)	(0.057)
South	-0.259***	-0.240***	-0.226***	-0.329***
	(0.079)	(0.070)	(0.068)	(0.073)
VNEN x North	0.173**	0.155***	0.171***	0.145***
	(0.069)	(0.059)	(0.053)	(0.055)
VNEN x Central	-0.080	-0.014	0.007	-0.021
	(0.054)	(0.048)	(0.048)	(0.058)
VNEN x South	0.081	0.012	-0.026	-0.006
	(0.083)	(0.063)	(0.072)	(0.083)
Minority	-0.667***	-0.683***	-0.716***	-0.690***
	(0.047)	(0.041)	(0.038)	(0.041)
Mother knows Vietnamese	0.394***	0.502***	0.484***	0.387***
	(0.055)	(0.048)	(0.035)	(0.037)
Father knows Vietnamese	0.268***	0.315***	0.280***	0.227***
	(0.040)	(0.036)	(0.028)	(0.027)
Constant	-0.300***	-0.448***	-0.376***	-0.237***
	(0.087)	(0.079)	(0.066)	(0.063)
Observations	11,630	11,605	11,387	11,218

Notes: Standard errors are clustered at the matched school pair level. The outcome variables are IRT latent values (2-parameter model) normalized using the non-VNEN students.

Table 6: Propensity Score Matching Estimates of the Differential Impact of VNEN on Vietnamese (Literature) Skills, 2013-2016 (Regression Based Estimates)

	2013	2014	2015	2016
Gender				
Girl	0.328***	0.372***	0.367***	0.460***
	(0.024)	(0.025)	(0.023)	(0.025)
VNEN x Girl	0.059	0.076**	0.014	0.001
	(0.043)	(0.039)	(0.036)	(0.036)
VNEN x Boy	0.141***	0.161***	0.067*	0.096**
	(0.043)	(0.040)	(0.038)	(0.037)
Minority	-0.563***	-0.511***	-0.458***	-0.439***
	(0.038)	(0.035)	(0.034)	(0.034)
Mother knows Vietnamese	0.590***	0.653***	0.696***	0.619***
	(0.058)	(0.055)	(0.051)	(0.044)
Father knows Vietnamese	0.283***	0.400***	0.334***	0.329***
	(0.039)	(0.039)	(0.032)	(0.032)
Constant	-0.695***	-0.925***	-0.925***	-0.911***
	(0.069)	(0.070)	(0.067)	(0.060)
Ethnicity				
VNEN x Minority	0.165**	0.180**	0.072	0.090
	(0.077)	(0.072)	(0.069)	(0.062)
VNEN x Kinh	0.054	0.074**	0.010	0.009
	(0.040)	(0.033)	(0.033)	(0.035)
Minority	-0.614***	-0.559***	-0.482***	-0.472***
	(0.058)	(0.055)	(0.055)	(0.053)
Mother knows Vietnamese	0.591***	0.654***	0.700***	0.623***
	(0.057)	(0.055)	(0.051)	(0.043)
Father knows Vietnamese	0.295***	0.412***	0.347***	0.344***
	(0.039)	(0.040)	(0.033)	(0.032)
Constant	-0.524***	-0.734***	-0.745***	-0.682***
	(0.068)	(0.069)	(0.065)	(0.058)
Region				
Central	0.133**	-0.029	-0.088	-0.014
	(0.065)	(0.057)	(0.057)	(0.056)
South	-0.150**	-0.194***	-0.197***	-0.214***
	(0.076)	(0.074)	(0.066)	(0.060)
VNEN x North	0.256***	0.218***	0.110*	0.123**
	(0.073)	(0.063)	(0.063)	(0.058)
VNEN x Central	-0.074	0.002	-0.042	-0.069
	(0.048)	(0.043)	(0.046)	(0.048)
VNEN x South	0.092	0.095	0.002	0.059
	(0.079)	(0.070)	(0.065)	(0.058)
Minority	-0.618***	-0.603***	-0.555***	-0.518***
	(0.045)	(0.041)	(0.037)	(0.037)
Mother knows Vietnamese	0.574***	0.642***	0.694***	0.614***
	(0.058)	(0.055)	(0.052)	(0.044)
Father knows Vietnamese	0.266***	0.388***	0.320***	0.316***
	(0.039)	(0.040)	(0.033)	(0.032)
Constant	-0.502***	-0.629***	-0.609***	-0.578***
	(0.090)	(0.084)	(0.084)	(0.075)
Observations	11,630	11,595	11,387	11,218

Notes: Standard errors are clustered at the matched school pair level. The outcome variables are IRT latent values (2-parameter model) normalized using the non-VNEN students.

Table 7: Estimates of the Impact of VNEN on Non-cognitive (Socioemotional) Skills, 2015-2016

VARIABLES	Intra-personal	Social	Combined measure
OLS Regression			
VNEN dummy	0.054	0.068*	0.077**
	(0.036)	(0.035)	(0.037)
Minority	-0.209***	-0.177***	-0.220***
	(0.041)	(0.041)	(0.043)
Mother knows Vietnamese	0.237***	0.309***	0.305***
	(0.047)	(0.047)	(0.049)
Father knows Vietnamese	0.014	0.043	0.050
	(0.039)	(0.043)	(0.041)
Constant	-0.148**	-0.252***	-0.240***
	(0.061)	(0.062)	(0.064)
Observations	10,931	10,932	10,931
IV Regression			
VNEN Index	0.146	0.193*	0.220*
	(0.111)	(0.111)	(0.116)
Minority	-0.174***	-0.138***	-0.177***
	(0.043)	(0.044)	(0.046)
Mother knows Vietnamese	0.258***	0.335***	0.327***
	(0.050)	(0.049)	(0.052)
Father knows Vietnamese	-0.008	0.024	0.037
	(0.042)	(0.046)	(0.044)
Constant	-0.210**	-0.346***	-0.347***
	(0.090)	(0.091)	(0.094)
Kleibergen-Paap F-stat	822.9	822.9	822.9
Observations	9,404	9,405	9,404

Notes: Standard errors are clustered at the matched school pair level. The outcome variables are IRT latent values (2-parameter model) normalized using the non-VNEN students.

Table 8: Propensity Score Matching Estimates of the Differential Impact of VNEN on Non-cognitive Skills, 2015-2016 (Regression Based Estimates)

VARIABLES	Intra-personal	Social	Combined measure
Gender			
Girl	0.194***	0.161***	0.206***
	(0.026)	(0.028)	(0.027)
VNEN x Girl	0.048	0.055	0.070*
	(0.042)	(0.040)	(0.042)
VNEN x Boy	0.069*	0.087**	0.094**
	(0.041)	(0.042)	(0.043)
Minority	-0.212***	-0.180***	-0.224***
	(0.041)	(0.041)	(0.043)
Mother knows Vietnamese	0.234***	0.307***	0.302***
	(0.047)	(0.047)	(0.049)
Father knows Vietnamese	0.007	0.037	0.043
	(0.039)	(0.043)	(0.041)
Constant	-0.237***	-0.326***	-0.334***
	(0.063)	(0.065)	(0.066)
Ethnicity			
VNEN x Minority	0.112*	0.126*	0.147**
	(0.061)	(0.064)	(0.066)
VNEN x Kinh	0.020	0.034	0.036
	(0.041)	(0.038)	(0.041)
Minority	-0.256***	-0.225***	-0.278***
	(0.055)	(0.055)	(0.058)
Mother knows Vietnamese	0.237***	0.310***	0.306***
	(0.047)	(0.046)	(0.049)
Father knows Vietnamese	0.013	0.041	0.048
	(0.039)	(0.043)	(0.041)
Constant	-0.131**	-0.235***	-0.219***
	(0.062)	(0.063)	(0.065)
Region			
Central	0.202***	0.194***	0.206***
	(0.060)	(0.060)	(0.062)
South	0.452***	0.330***	0.417***
	(0.071)	(0.070)	(0.074)
VNEN x North	0.082	0.103*	0.122**
	(0.055)	(0.056)	(0.058)
VNEN x Central	0.025	-0.001	0.018
	(0.052)	(0.053)	(0.055)
VNEN x South	0.086	0.151*	0.130
	(0.077)	(0.077)	(0.083)
Minority	-0.054	-0.055	-0.079*
	(0.043)	(0.043)	(0.045)
Mother knows Vietnamese	0.252***	0.321***	0.320***
	(0.045)	(0.045)	(0.047)
Father knows Vietnamese	0.061	0.079*	0.093**
	(0.038)	(0.042)	(0.040)
Constant	-0.444***	-0.493***	-0.520***
	(0.068)	(0.070)	(0.071)
Observations	10,931	10,932	10,931

Notes: Standard errors are clustered at the matched school pair level. The outcome variables are IRT latent values (2-parameter model) normalized using the non-VNEN students.

Table 9: Propensity Score Matching Estimates of the Long-term Impact of VNEN on Math and Vietnamese Skills, 2017-2020 (Regression Based Estimates)

	Nov. 2017 (Grade 7)	April 2018 (Grade 7)	June 2020 (Grade 9)
Mathematics			
VNEN dummy	-0.052	-0.053	-0.076
	(0.097)	(0.094)	(0.119)
Minority	-0.451***	-0.617***	-0.639***
	(0.082)	(0.094)	(0.100)
Mother knows Vietnamese	0.243**	0.302***	0.164
	(0.097)	(0.084)	(0.105)
Father knows Vietnamese	0.348***	0.291***	0.128
	(0.075)	(0.064)	(0.093)
Constant	-0.434***	-0.405***	-0.105
	(0.129)	(0.118)	(0.153)
Observations	1,199	1,191	1,080
Vietnamese			
VNEN dummy	0.001	-0.016	-0.007
	(0.083)	(0.091)	(0.089)
Minority	-0.543***	-0.581***	-0.593***
	(0.086)	(0.081)	(0.108)
Mother knows Vietnamese	0.362***	0.271***	0.121
	(0.099)	(0.090)	(0.132)
Father knows Vietnamese	0.232***	0.145	-0.058
	(0.085)	(0.088)	(0.110)
Constant	-0.284**	-0.269**	0.071
	(0.119)	(0.125)	(0.159)
Observations	1,199	1,191	1,073

Notes: Standard errors are clustered at the matched school pair level. The outcome variables are IRT latent values (2-parameter model) normalized to have a mean of zero and a standard deviation of one using the non-VNEN students.

Table 10: Propensity Score Matching Estimates of the Long-term Impact of VNEN on Math and Vietnamese Skills, 2017-2020 (IV Regression Based Estimates)

	Nov. 2017 (Grade 7)	April 2018 (Grade 7)	June 2020 (Grade 9)
Mathematics			
VNEN Index	-0.148	-0.148	-0.215
	(0.271)	(0.262)	(0.335)
Minority	-0.451***	-0.618***	-0.639***
	(0.0815)	(0.0938)	(0.0996)
Mother knows Vietnamese	0.250***	0.309***	0.175*
	(0.0934)	(0.0825)	(0.103)
Father knows Vietnamese	0.347***	0.290***	0.134
	(0.0742)	(0.0634)	(0.0932)
Constant	-0.399**	-0.370**	-0.0626
	(0.167)	(0.157)	(0.192)
Kleibergen-Paap F-statistic	257.8	262.9	248.2
Observations	1,199	1,191	1,080
Vietnamese			
VNEN Index	0.002	-0.045	-0.020
	(0.232)	(0.256)	(0.250)
Minority	-0.543***	-0.581***	-0.593***
	(0.086)	(0.081)	(0.107)
Mother knows Vietnamese	0.362***	0.273***	0.122
	(0.099)	(0.091)	(0.132)
Father knows Vietnamese	0.232***	0.145*	-0.0569
	(0.085)	(0.087)	(0.109)
Constant	-0.284**	-0.259*	0.075
	(0.145)	(0.155)	(0.176)
Kleibergen-Paap F-statistic	257.8	255.7	246.8
Observations	1,199	1,191	1,073

Notes: The VNEN index is average value of VNEN index in 2013, 2014, and 2015. It is normalized so that it has a range from 0 to 1, where 0 corresponds to not adopting any aspects of the VNEN program and 1 indicates adoption of all aspects of the VNEN program. Standard errors are clustered at the matched school pair level. The outcome variables are IRT latent values (2-parameter model) normalized to have a mean of zero and a standard deviation of one using the non-VNEN students.

Table 11: Propensity Score Matching Estimates of the Differential Long-term Impacts of VNEN on Mathematics Skills, 2017-2020 (Regression Based Estimates)

	Nov. 2017 (Grade 7)	April 2018 (Grade 7)	June 2020 (Grade 9)
Gender			
Girl	-0.001	0.031	0.035
	(0.087)	(0.082)	(0.085)
VNEN x Girl	-0.066	-0.072	-0.081
	(0.110)	(0.111)	(0.128)
VNEN x Boy	-0.040	-0.034	-0.069
	(0.112)	(0.104)	(0.135)
Minority	-0.451***	-0.617***	-0.638***
	(0.082)	(0.094)	(0.101)
Mother knows Vietnamese	0.242**	0.304***	0.164
	(0.097)	(0.084)	(0.106)
Father knows Vietnamese	0.347***	0.290***	0.131
	(0.075)	(0.064)	(0.094)
Constant	-0.432***	-0.421***	-0.127
	(0.145)	(0.127)	(0.171)
Ethnicity			
VNEN x Minority	0.179	0.209	0.048
	(0.120)	(0.144)	(0.156)
VNEN x Kinh	-0.127	-0.137	-0.113
	(0.118)	(0.112)	(0.144)
Minority	-0.638***	-0.830***	-0.737***
	(0.115)	(0.140)	(0.156)
Mother knows Vietnamese	0.246**	0.305***	0.173
	(0.095)	(0.084)	(0.107)
Father knows Vietnamese	0.353***	0.296***	0.130
	(0.076)	(0.066)	(0.093)
Constant	-0.411***	-0.377***	-0.100
	(0.133)	(0.121)	(0.153)
Region			
Central	-0.031	0.136	-0.039
	(0.200)	(0.174)	(0.276)
South	-0.627***	-0.306	-0.340
	(0.202)	(0.206)	(0.269)
VNEN x North	0.073	0.214	0.052
	(0.180)	(0.167)	(0.276)
VNEN x Central	-0.115	-0.115	-0.059
	(0.113)	(0.110)	(0.143)
VNEN x South	0.0731	-0.0297	-0.300
	(0.145)	(0.164)	(0.201)
Minority	-0.544***	-0.678***	-0.718***
	(0.090)	(0.105)	(0.111)
Mother knows Vietnamese	0.203**	0.274***	0.130
	(0.086)	(0.083)	(0.091)
Father knows Vietnamese	0.324***	0.271***	0.089
	(0.077)	(0.067)	(0.098)
Constant	-0.238	-0.404**	0.050
	(0.219)	(0.196)	(0.292)
Observations	1,199	1,191	1,080

Notes: Standard errors are clustered at the matched school pair level. The outcome variables are IRT latent values (2-parameter model) normalized to have a mean of zero and standard deviation of one using the non-VNEN students.

Table 12: Propensity Score Matching Estimates of the Differential Long-term Impacts of VNEN on Vietnamese Skills, 2017-2020 (Regression Based Estimates)

	Nov. 2017 (Grade 7)	April 2018 (Grade 7)	June 2020 (Grade 9)
Gender			
Girl	0.211**	0.252***	0.272***
	(0.094)	(0.091)	(0.088)
VNEN x Girl	0.030	0.029	-0.046
	(0.100)	(0.102)	(0.088)
VNEN x Boy	-0.013	-0.04	0.044
	(0.104)	(0.114)	(0.121)
Minority	-0.537***	-0.574***	-0.584***
	(0.089)	(0.082)	(0.107)
Mother knows Vietnamese	0.379***	0.292***	0.125
	(0.100)	(0.091)	(0.134)
Father knows Vietnamese	0.238***	0.155*	-0.032
	(0.087)	(0.086)	(0.110)
Constant	-0.414***	-0.427***	-0.098
	(0.143)	(0.136)	(0.172)
Ethnicity			
VNEN x Minority	0.234*	0.092	0.292
	(0.124)	(0.130)	(0.198)
VNEN x Kinh	-0.075	-0.050	-0.098
	(0.099)	(0.112)	(0.094)
Minority	-0.732***	-0.669***	-0.831***
	(0.105)	(0.124)	(0.194)
Mother knows Vietnamese	0.365***	0.273***	0.142
	(0.100)	(0.090)	(0.131)
Father knows Vietnamese	0.237***	0.147	-0.053
	(0.086)	(0.090)	(0.112)
Constant	-0.260**	-0.259**	0.085
	(0.122)	(0.128)	(0.162)
Region			
Central	0.093	0.043	0.303
	(0.125)	(0.192)	(0.262)
South	-0.310**	-0.220	0.165
	(0.155)	(0.222)	(0.286)
VNEN x North	0.166	0.124	0.607**
	(0.155)	(0.195)	(0.279)
VNEN x Central	-0.030	-0.036	-0.101
	(0.097)	(0.108)	(0.095)
VNEN x South	-0.030	-0.070	-0.140
	(0.147)	(0.168)	(0.163)
Minority	-0.587***	-0.627***	-0.671***
	(0.109)	(0.098)	(0.090)
Mother knows Vietnamese	0.329***	0.246***	0.135
	(0.095)	(0.085)	(0.115)
Father knows Vietnamese	0.213**	0.130	-0.067
	(0.087)	(0.089)	(0.108)
Constant	-0.248	-0.221	-0.174
	(0.165)	(0.213)	(0.265)
Observations	1,199	1,191	1,073

Notes: Standard errors are clustered at the matched school pair level. The outcome variables are IRT latent values (2-para-meter model) normalized to have a mean of zero and a standard deviation of one using the non-VNEN students.

Table 13: Propensity Score Matching Estimates of the Long-term Impact of VNEN on Non-Cognitive Skills, 2018-2020 (Regression Based Estimates)

	Self-perception Profile	Intrinsic-Extrinsic Orientation	Coping	Aggregate non-cognitive skill
Grade 7 (April 2018)				
VNEN dummy	0.027	-0.073	0.073	0.009
	(0.066)	(0.075)	(0.076)	(0.074)
Minority	0.086	-0.270***	-0.261***	-0.215***
	(0.071)	(0.061)	(0.079)	(0.063)
Mother knows Vietnamese	-0.039	0.005	-0.023	-0.030
	(0.104)	(0.100)	(0.115)	(0.087)
Father knows Vietnamese	0.072	0.095	0.225**	0.194**
	(0.084)	(0.095)	(0.101)	(0.092)
Constant	0.029	-0.196	-0.113	-0.141
	(0.134)	(0.133)	(0.155)	(0.139)
Observations	1,195	1,195	1,195	1,195
Grade 9 (June 2020)				
VNEN dummy	0.011	-0.125	0.013	-0.054
	(0.068)	(0.081)	(0.099)	(0.084)
Minority	-0.135*	-0.418***	0.028	-0.288***
	(0.081)	(0.081)	(0.111)	(0.085)
Mother knows Vietnamese	0.072	0.129	0.184	0.198**
	(0.102)	(0.104)	(0.126)	(0.100)
Father knows Vietnamese	0.191	0.181*	0.170*	0.287**
	(0.120)	(0.100)	(0.097)	(0.115)
Constant	-0.304*	-0.059	-0.365**	-0.376**
	(0.158)	(0.142)	(0.169)	(0.152)
Observations	1,088	1,088	1,087	1,087

Notes: Standard errors are clustered at the matched school pair level. The outcome variables are normalized to have a mean of zero and a standard deviation of one using the non-VNEN students.

Table 14: Propensity Score Matching Estimates of the Long-term Impact of VNEN on Non-Cognitive Skills, 2018-2020 (IV Regression Based Estimates)

	Self-perception Profile	Intrinsic-Extrinsic Orientation	Coping	Aggregate non-cognitive skill
Grade 7 (April 2018)				
VNEN Index	0.075	-0.205	0.206	0.024
	(0.185)	(0.211)	(0.211)	(0.205)
Minority	0.087	-0.270***	-0.260***	-0.215***
	(0.070)	(0.060)	(0.077)	(0.063)
Mother knows Vietnamese	-0.042	0.014	-0.033	-0.031
	(0.103)	(0.099)	(0.115)	(0.085)
Father knows Vietnamese	0.072	0.094	0.226**	0.194**
	(0.084)	(0.095)	(0.100)	(0.092)
Constant	0.011	-0.147	-0.162	-0.147
	(0.153)	(0.162)	(0.182)	(0.170)
Kleibergen-Paap F-stat	256.5	256.5	256.5	256.5
Observations	1,195	1,195	1,195	1,195
Grade 9 (June 2020)				
VNEN Index	0.032	-0.354	0.037	-0.153
	(0.191)	(0.230)	(0.277)	(0.235)
Minority	-0.135*	-0.418***	0.028	-0.288***
	(0.081)	(0.082)	(0.110)	(0.085)
Mother knows Vietnamese	0.070	0.150	0.181	0.207**
	(0.100)	(0.110)	(0.124)	(0.099)
Father knows Vietnamese	0.191	0.181*	0.169*	0.287**
	(0.119)	(0.101)	(0.096)	(0.115)
Constant	-0.311*	0.019	-0.373*	-0.343**
	(0.174)	(0.161)	(0.190)	(0.166)
Kleibergen-Paap F-stat	250.2	250.2	250.8	250.8
Observations	1,088	1,088	1,087	1,087

The VNEN index is average value of VNEN index in 2013, 2014, and 2015. It is normalized so that it has a range from 0 to 1, where 0 corresponds to not adopting any aspects of the VNEN program and 1 indicates adoption of all aspects of the VNEN program. Standard errors are clustered at the matched school pair level. The outcome variables are normalized to have a mean of zero and a standard deviation of one using the non-VNEN students.

Table 15: Propensity Score Matching Estimates of the Long-term Impacts of VNEN on Non-Cognitive Skills by Gender, 2018-2020 (Regression Based Estimates)

	Self-perception Profile	Intrinsic-Extrinsic Orientation	Coping	Aggregate non-cognitive skill
Grade 7 (April 2018)				
Girl	-0.182**	-0.120*	0.089	-0.129
	(0.082)	(0.071)	(0.081)	(0.086)
VNEN x Girl	-0.012	-0.022	0.081	0.018
	(0.089)	(0.096)	(0.094)	(0.099)
VNEN x Boy	0.050	-0.127	0.072	-0.008
	(0.086)	(0.097)	(0.094)	(0.090)
Minority	0.083	-0.272***	-0.260***	-0.217***
	(0.072)	(0.060)	(0.078)	(0.064)
Mother knows Vietnamese	-0.056	-0.002	-0.015	-0.040
	(0.103)	(0.099)	(0.115)	(0.087)
Father knows Vietnamese	0.066	0.095	0.227**	0.192**
	(0.085)	(0.096)	(0.101)	(0.093)
Constant	0.144	-0.128	-0.169	-0.063
	(0.137)	(0.134)	(0.165)	(0.145)
Observations	1,195	1,195	1,195	1,195
Grade 9 (June 2020)				
Girl	-0.360***	-0.065	0.468***	-0.031
	(0.078)	(0.087)	(0.088)	(0.078)
VNEN x Girl	-0.048	-0.088	-0.145	-0.142
	(0.086)	(0.112)	(0.116)	(0.104)
VNEN x Boy	0.046	-0.162*	0.184	0.025
	(0.090)	(0.091)	(0.114)	(0.101)
Minority	-0.151*	-0.419***	0.041	-0.292***
	(0.083)	(0.081)	(0.110)	(0.085)
Mother knows Vietnamese	0.039	0.128	0.203	0.187*
	(0.099)	(0.103)	(0.126)	(0.098)
Father knows Vietnamese	0.170	0.183*	0.170*	0.276**
	(0.118)	(0.101)	(0.093)	(0.115)
Constant	-0.065	-0.026	-0.626***	-0.339**
	(0.162)	(0.140)	(0.170)	(0.160)
Observations	1,088	1,088	1,087	1,087

Notes: Standard errors are clustered at the matched school pair level. The outcome variables are normalized to have a mean of zero and a standard deviation of one using the non-VNEN students.

Table 16: Propensity Score Matching Estimates of the Long-term Impacts of VNEN on Non-Cognitive Skills by Ethnicity, 2018-2020 (Regression Based Estimates)

	Self-perception Profile	Intrinsic-Extrinsic Orientation	Coping	Aggregate non-cognitive skill
Grade 7 (April 2018)				
VNEN x Minority	0.120	-0.203*	0.182	0.040
	(0.121)	(0.107)	(0.146)	(0.122)
VNEN x Kinh	-0.003	-0.031	0.039	-0.001
	(0.076)	(0.089)	(0.084)	(0.084)
Minority	0.011	-0.164	-0.349***	-0.240**
	(0.100)	(0.105)	(0.123)	(0.114)
Mother knows Vietnamese	-0.037	0.002	-0.021	-0.029
	(0.104)	(0.098)	(0.114)	(0.087)
Father knows Vietnamese	0.074	0.092	0.227**	0.195**
	(0.085)	(0.095)	(0.101)	(0.092)
Constant	0.038	-0.208	-0.103	-0.138
	(0.135)	(0.133)	(0.153)	(0.139)
Observations	1,195	1,195	1,195	1,195
Grade 9 (June 2020)				
VNEN x Minority	-0.070	-0.268**	0.319**	-0.036
	(0.139)	(0.125)	(0.148)	(0.091)
VNEN x Kinh	0.036	-0.082	-0.079	-0.059
	(0.071)	(0.098)	(0.116)	(0.102)
Minority	-0.070	-0.302**	-0.217*	-0.303***
	(0.115)	(0.119)	(0.124)	(0.088)
Mother knows Vietnamese	0.069	0.123	0.198	0.199**
	(0.103)	(0.101)	(0.122)	(0.100)
Father knows Vietnamese	0.191	0.180*	0.173*	0.287**
	(0.120)	(0.101)	(0.097)	(0.115)
Constant	-0.310**	-0.070	-0.342**	-0.375**
	(0.155)	(0.142)	(0.166)	(0.153)
Observations	1,088	1,088	1,087	1,087

Notes: Standard errors are clustered at the matched school pair level. The outcome variables are normalized to have a mean of zero and a standard deviation of one using the non-VNEN students.

Table 17: Propensity Score Matching Estimates of the Long-term Impacts of VNEN on Non-Cognitive Skills by Region, 2018-2020 (Regression Based Estimates)

	Self-perception Profile	Intrinsic-Extrinsic Orientation	Coping	Aggregate non-cognitive skill
Grade 7 (April 2018)				
Central	0.200	-0.074	0.216	0.172
	(0.168)	(0.122)	(0.196)	(0.169)
South	0.085	-0.355**	0.138	-0.082
	(0.172)	(0.160)	(0.242)	(0.200)
VNEN x North	0.117	-0.160	0.048	0.002
	(0.184)	(0.132)	(0.213)	(0.182)
VNEN x Central	-0.042	-0.102	0.076	-0.046
	(0.079)	(0.094)	(0.086)	(0.089)
VNEN x South	0.299**	0.120	0.130	0.295**
	(0.118)	(0.136)	(0.197)	(0.147)
Minority	0.134	-0.290***	-0.175**	-0.160**
	(0.094)	(0.067)	(0.072)	(0.071)
Mother knows Vietnamese	-0.019	-0.007	-0.024	-0.025
	(0.104)	(0.098)	(0.117)	(0.087)
Father knows Vietnamese	0.082	0.089	0.230**	0.199**
	(0.085)	(0.097)	(0.102)	(0.092)
Constant	-0.170	-0.065	-0.314	-0.274
	(0.199)	(0.150)	(0.259)	(0.204)
Observations	1,195	1,195	1,195	1,195
Grade 9 (June 2020)				
Central	-0.141	-0.155	0.058	-0.139
	(0.149)	(0.113)	(0.154)	(0.126)
South	-0.187	-0.452***	0.129	-0.295
	(0.196)	(0.160)	(0.169)	(0.184)
VNEN x North	-0.208	-0.398***	0.423**	-0.135
	(0.164)	(0.143)	(0.188)	(0.143)
VNEN x Central	0.038	-0.135	-0.024	-0.063
	(0.081)	(0.100)	(0.123)	(0.107)
VNEN x South	0.095	0.214	-0.232	0.067
	(0.171)	(0.159)	(0.164)	(0.171)
Minority	-0.126	-0.408***	-0.078	-0.328***
	(0.093)	(0.079)	(0.105)	(0.089)
Mother knows Vietnamese	0.062	0.111	0.195	0.187*
	(0.101)	(0.100)	(0.122)	(0.102)
Father knows Vietnamese	0.193	0.192*	0.164*	0.291**
	(0.120)	(0.100)	(0.098)	(0.116)
Constant	-0.165	0.125	-0.414*	-0.219
	(0.199)	(0.183)	(0.235)	(0.202)
Observations	1,088	1,088	1,087	1,087

Notes: Standard errors are clustered at the matched school pair level. The outcome variables are normalized to have a mean of zero and a standard deviation of one using the non-VNEN students.

Appendix: Construction of the VNEN Implementation (Intensity) Index

We used the description of the VNEN program in Section II to construct an index of VNEN implementation that applies to both the VNEN and non-VNEN schools in the World Bank data. This is used to assess the short-term impact of the degree of VNEN implementation on students' cognitive and non-cognitive skills. The index gives each of the eight fundamental characteristics of the VNEN program equal weight by constructing subindices of these eight characteristics and then applying principal components analysis to the eight subindices to construct an overall VNEN index. The rest of this subsection explains how the eight subindices were created, and then combined to generate an overall index. This was done separately for each year, because the 2013-14 data are more limited than the 2014-15 and 2015-16 data.

Organization of classroom space. The two main classroom space requirements are to have enough space for students and teachers to move around the classroom, and for students' desks to be arranged in small groups (typically four students) with their desks facing each other instead of facing the front of the classroom. In all three years of the survey, teachers were asked to give the size of their classroom in square meters. They were also asked how many student learning groups were in their classrooms, and how many students per learning group, which can be used to calculate the number of students in the classroom. This is used to calculate the area (square meters per student) of the classroom, which can serve as an index of space in the classroom to implement VNEN. In the last two survey years (2014-15 and 2015-16), but not in the first year (2013-14), teachers were asked whether they arranged students desks into groups. Thus, for first year of the survey the organization of classroom space variable is measured by square meters per student in the classroom, while in the second and third years it will be

measured by applying principal components analysis to that variable and to a dummy variable indicating whether students' desks are arranged in groups with students facing each other.[11]

Classroom artifacts. The teacher questionnaire for all three years asks teachers about the presence of the six classroom artifacts described above. Principal components analysis is used to construct a scalar subindex for the presence of these artifacts in the classroom.

Learning guides. The teacher questionnaire for the first two years (2013-14 and 2014-15) asks teachers how many learning guides are available in the classroom for math, Vietnamese and science, and, for each subject, whether there is a shortage of guides. It also asks for the percentage of students who had these learning guides at the beginning of the school year. Principal components analysis can be used to construct a scalar subindex for the presence of teaching guides in the classroom. The only information on the availability of learning guides in the third year (2015-16) is whether the teacher uses them (yes or no question), which is an indirect indicator of their availability since they can be used only if they are available.

Teaching methods. Recall that VNEN emphasizes teaching methods that: use "real life" as well as manmade objects in class discussions; finding ways to relate classroom lessons to "real life"; use games as part of their lessons, such as using "role-play" methods when teaching Vietnamese; ask students to solve many problems when teaching math; encouraging students to ask questions; encourage students to study on their own and "find out new things by themselves"; and provide assistance to individual students. The teacher questionnaire for all three years asks teachers about most of these types of pedagogy, including assigning exercises to

[11] Principal components analysis is used to summarize the variance in a set of variables by generating "principal components" that are weighted averages of those variables. It is useful for summarizing the overall variance in a set of variables that measure similar phenomena and so should be highly correlated with each other. The first principal component is the weighted average of those variables with the largest variance out of all possible weighted averages that satisfy the condition that the sum of the squared weights equals one. Intuitively, it contains the highest possible variance among all the variables in the set that can be captured in a single variable. The more highly correlated these variables are, the larger the proportion of their overall variance that can be captured in the first principal component.

do in class, encouraging students to ask questions, and letting students study on their own. Principal components analysis is used to construct a scalar subindex for these teaching methods.

Group work. The teacher questionnaire for all three years asks teachers about the teaching methods they use. Regarding group work, it asks whether they divide students to work in small groups, and whether they put students into groups for discussions. The 2014-15 and 2015-16 questionnaires also asked how frequently the teacher tried new methods for group work. Principal components analysis is used to construct a scalar subindex for the use of group work.

Assessment methods. The 2013-14 teacher questionnaire collects only limited information on assessment methods. It asks how often teachers use "routine" marking, and how often they use "periodical" marking, to assess their students' performance. The former refers to informal tests or quizzes to regularly check on student learning, and so are closer to formative assessments, while the latter are used in the middle or the end of school terms to assess what students have learned over longer periods of time, and so are closer to summative assessments. It also asks teachers how frequently they assess students' "non-cognitive" skills. The 2014-15 and 2015-16 teacher questionnaires collect additional assessment information, in particular whether: a) students check each other's work in pairs; b) the group leader assesses each student's work; c) the group arrives at a consensus on each student's work; and d) the teacher provides feedback on each group's project work. Principal components analysis is used to construct a scalar subindex for the use of these types of assessment methods.

Student government. The 2013-14 teacher questionnaire has only basic information on student government, specifically whether the classroom has a student committee organizing chart, and how often it is used. The 2014-15 and 2015-16 questionnaires collect more detail; they ask which of seven different types of student committees exist, how active the student

government officers are, and some questions about how student government is run. Principal components analysis is used to construct a subindex for these student government activities.

Parental Engagement. For each year, the teacher questionnaire asks about parental engagement in school activities. All three questionnaires ask teachers: a) how often they invite parents to discuss their child's situation; and b) how often they invite parents to participate in classroom teaching activities. The 2014-15 and 2015-16 teacher questionnaires also ask teachers how often they: a) invite parents to demonstrate or teach a particular skill they have; b) invite parents to help decorate the classroom; c) invite parents to help with general teaching; d) invite parents to communicate with students who speak an ethnic minority language; and e) discuss with parents the use the grants provided to VNEN schools. A scalar subindex of these types of parental engagement their children's school is constructed using principal components analysis.

Appendix Table A.1: Estimation Results from the Probit Matching Regression Using Primary School Census Data, Vietnam 2010-2011

	Priority I Provinces	Priority II & III Provinces
School size	-0.000*	-0.000
	(0.000)	(0.000)
Percentage of poor students	-0.001*	0.002
	(0.001)	(0.002)
Percentage of ethnic minority students	-0.000	0.001
	(0.001)	(0.001)
Percentage of weak-performing students in math	-0.002	-0.014
	(0.013)	(0.021)
Percentage of weak-performing students in Vietnamese	-0.016	-0.018
	(0.013)	(0.024)
Distance to district bureau of education	-0.008***	-0.013***
	(0.002)	(0.003)
Parental support for children's education	0.155***	0.116*
	(0.042)	(0.062)
Number of satellite campuses	-0.043***	-0.034
	(0.012)	(0.024)
Number of computers	0.005	0.001
	(0.003)	(0.004)
Percentage of blackboards in good working condition	-0.018	0.042
	(0.104)	(0.143)
Percentage of student seats in good working condition	0.151**	0.035
	(0.060)	(0.086)
School meets National Standards I or higher	0.192***	0.129*
	(0.053)	(0.073)
Percentage of teachers with upper secondary education	0.226	0.438
	(0.149)	(0.297)
Percentage of teachers with college training	0.292**	0.660***
	(0.115)	(0.175)
Urban	-0.092	-0.196
	(0.219)	(0.283)
Constant	-0.744***	-3.943***
	(0.224)	(0.537)
Pseudo- R2	0.08	0.15
Log likelihood	-2548	-1119
N	5555	9694

Note: *p< .1, **p<0.05, ***p<0.01; standard errors in parentheses. All regressions control for province dummy variables.

Appendix Table A.2: Post-matching Balance Tests for Priority I Provinces, Vietnam 2010-2011

	Mean		Percent of bias	t-test	P-value
	VNEN Schools	Non-VNEN Schools			
School size	261.50	261.97	-0.2	-0.05	0.96
Percentage of poor students	18.57	20.45	-6.7	-1.50	0.13
Percentage of ethnic minority students	50.25	52.59	-5.4	-1.18	0.24
Percentage of weak-performing students in math	2.10	1.95	4.6	1.16	0.25
Percentage of weak-performing students in Vietnamese	2.02	1.92	3.0	0.75	0.45
Distance to district bureau of education	13.65	13.17	3.5	0.85	0.40
Parental support for children's education	1.61	1.60	2.0	0.45	0.66
Number of satellite campuses	1.85	1.78	3.0	0.71	0.48
Number of computers	5.81	6.06	-3.2	-0.70	0.49
Percentage of blackboards in good working condition	0.95	0.95	-0.5	-0.12	0.91
Percentage of student seats in good working condition	0.76	0.76	0.5	0.12	0.91
School meets National Standards I or higher	0.63	0.66	-5.2	-1.17	0.24
Percentage of teachers with upper secondary education	0.86	0.87	-6.3	-1.48	0.14
Percentage of teachers with college training	0.58	0.58	2.6	0.60	0.55
Urban	0.01	0.01	1.1	0.24	0.81
Province 1	0.05	0.04	4.7	1.10	0.27
Province 2	0.07	0.09	-6.5	-1.25	0.21
Province 3	0.06	0.08	-11.2	-2.19	0.03
Province 4	0.07	0.08	-3.5	-0.67	0.51
Province 5	0.06	0.05	4.4	1.00	0.32
Province 6	0.05	0.06	-3.3	-0.67	0.50
Province 7	0.05	0.03	12.2	2.78	0.01
Province 8	0.09	0.09	-1.3	-0.31	0.76
Province 9	0.07	0.07	1.1	0.27	0.79
Province 10	0.03	0.01	9.5	2.44	0.02
Province 11	0.04	0.05	-5.4	-1.22	0.22
Province 12	0.02	0.02	1.8	0.46	0.64
Province 13	0.03	0.03	1.8	0.41	0.68
Province 14	0.06	0.05	5.9	1.29	0.20
Province 15	0.07	0.08	-4.7	-1.00	0.32
Province 16	0.03	0.04	-6.4	-1.25	0.21
Province 17	0.03	0.02	4.3	1.03	0.30
Province 18	0.04	0.05	-1.4	-0.33	0.74
Province 19	0.03	0.03	-1.7	-0.38	0.70
Province 20	0.04	0.04	4.2	1.03	0.30

These tests are based on the sample of 994 VNEN schools and 994 non-VNEN schools.

Appendix Table A.3: Post-matching Balance Tests for Priority II and III Provinces, Vietnam 2010-2011

	Mean		Percent of bias	t-test	P-value
	VNEN Schools	Non-VNEN Schools			
School size	344.02	366.89	-8.6	-1.08	0.28
Percentage of poor students	13.14	10.38	14.1	1.59	0.11
Percentage of ethnic minority students	20.12	15.27	15.7	1.73	0.08
Percentage of weak-performing students in math	1.64	1.38	11.2	1.50	0.13
Percentage of weak-performing students in Vietnamese	1.51	1.33	8.7	1.15	0.25
Distance to district bureau of education	10.73	9.75	9.4	1.22	0.22
Parental support for children's education	1.74	1.75	-2.3	-0.27	0.79
Number of satellite campuses	1.11	1.05	4.5	0.48	0.63
Number of computers	8.55	9.19	-5.9	-0.74	0.46
Percentage of blackboards in good working condition	0.94	0.95	-5.8	-0.66	0.51
Percentage of student seats in good working condition	0.71	0.73	-5.6	-0.62	0.54
School meets National Standards I or higher	0.62	0.60	3.1	0.36	0.72
Percentage of teachers with upper secondary education	0.92	0.93	-9.8	-1.24	0.22
Percentage of teachers with college training	0.68	0.71	-14.1	-1.70	0.09
Urban	0.03	0.03	-1.3	-0.26	0.79
Province 1	0.00	0.00	0.0	0.00	1.00
Province 2	0.00	0.00	3.3	1.00	0.32
Province 3	0.00	0.00	3.1	1.00	0.32
Province 4	0.00	0.01	-7.6	-1.00	0.32
Province 5	0.00	0.00	4.1	1.00	0.32
Province 6	0.00	0.01	-3.0	-0.58	0.56
Province 7	0.00	0.01	-6.0	-1.00	0.32
Province 8	0.00	0.00	4.0	1.00	0.32
Province 9	0.03	0.02	10.3	1.08	0.28
Province 10	0.02	0.02	0.0	0.00	1.00
Province 11	0.03	0.05	-12.8	-1.31	0.19
Province 12	0.06	0.04	8.1	0.80	0.42
Province 13	0.06	0.07	-5.9	-0.54	0.59
Province 14	0.00	0.01	-3.7	-0.58	0.56
Province 15	0.06	0.05	4.0	0.39	0.70
Province 16	0.00	0.00	4.0	1.00	0.32
Province 17	0.05	0.05	-2.0	-0.20	0.84
Province 18	0.02	0.02	2.4	0.34	0.74
Province 19	0.00	0.00	2.9	1.00	0.32
Province 20	0.05	0.05	4.2	0.40	0.69
Province 21	0.03	0.04	-2.4	-0.23	0.82
Province 22	0.00	0.00	4.7	1.00	0.32
Province 23	0.05	0.04	4.2	0.42	0.68
Province 24	0.06	0.05	4.0	0.38	0.71
Province 25	0.03	0.05	-10.1	-0.91	0.36
Province 26	0.03	0.02	6.8	0.84	0.40
Province 27	0.00	0.01	-4.9	-1.00	0.32
Province 28	0.05	0.05	-4.3	-0.40	0.69
Province 29	0.00	0.01	-3.0	-0.58	0.56
Province 30	0.00	0.01	-8.4	-1.00	0.32
Province 31	0.05	0.07	-9.8	-0.90	0.37
Province 32	0.05	0.06	-5.9	-0.55	0.58
Province 33	0.00	0.00	4.0	1.00	0.32
Province 34	0.00	0.00	0.0	N/A	N/A
Province 35	0.00	0.00	2.8	1.00	0.32
Province 36	0.05	0.06	-3.8	-0.39	0.70
Province 37	0.00	0.00	3.3	1.00	0.32
Province 38	0.00	0.00	3.3	1.00	0.32
Province 39	0.00	0.00	0.0	0.00	1.00
Province 40	0.00	0.01	-8.1	-1.00	0.32
Province 41	0.04	0.02	14.1	1.63	0.10
Province 42	0.05	0.03	10.3	1.07	0.29
Province 43	0.05	0.05	-4.6	-0.40	0.69

These tests are based on the sample of 110 VNEN schools and 110 non-VNEN schools.

Appendix Table A.4: Balance Tests for VNEN and Non-VNEN Schools (Primary level)

	Mean		
	VNEN schools	**Non-VNEN schools**	**Difference**
School-Level Variables	(N = 49)	(N = 51)	
School size	301.14	335.8	-34.64
Percentage of poor students	14.1	13.1	1.02
Percentage of ethnic minority students	36.2	21.0	15.2*
Percentage of weak-performing students in math	1.39	1.44	-0.05
Percentage of weak-performing students in Vietnamese	1.26	1.28	-0.01
Distance to district bureau of education	13.9	11.4	2.5
Parental support for children's education	1.73	1.88	-0.14*
Number of satellite campuses	1.06	1.10	-0.04
Number of computers	8.18	9.67	-1.48
Percentage of blackboards in good working condition	0.94	0.92	0.02
Percentage of student seats in good working condition	0.72	0.71	0.01
School meets National Standard I or higher	0.65	0.66	-0.01
Percentage of teachers with upper secondary education	0.89	0.91	-0.02
Percentage of teachers with college training	0.61	0.72	-0.11**
Propensity score	0.14	0.15	-0.01
F-statistic for joint significance			1.04
P-value for joint significance			0.42
Student-Level Variables	(N = 1,165)	(N = 1,193)	
Female	0.505	0.522	-0.016
Ethnic minority	0.372	0.199	0.173**
Birth order	1.908	1.999	-0.091
Mother can read and write in Vietnamese	0.875	0.925	-0.051*
Father can read and write in Vietnamese	0.876	0.918	-0.043*
Mother with upper secondary degree or higher	0.352	0.341	0.011
Father with upper secondary degree or higher	0.420	0.427	-0.007
Wealth index	3.40	3.516	-0.116
F-statistics for joint significance			1.67
P-value for joint significance			0.115

Note: *** $p < 0.01$; ** $p < 0.05$; * < 0.01 (bootstrapped with 500 replications). Sample includes students in 100 primary schools that are later linked with 99 lower secondary schools in the RISE study. Variable is at the individual level from the 2013 survey conducted by the World Bank.

Appendix Table A.5: Balance Tests for VNEN and Non-VNEN Schools (Lower Secondary)

	Mean		
	VNEN schools	**Non-VNEN schools**	**Difference**
School-Level Variables	(N = 49)	(N = 50)	
Class size	32.6	33.4	-0.75
Percentage of ethnic minority students	36.9	16.7	20.2***
Percentage of female students	47.6	51.0	3.42
Percentage of students helping with parent's business or working for extra income	13.1	9.5	3.62
Average meals per day	3.00	2.94	0.06
How often do parents ask about child's studying?	2.82	2.85	-0.03
How often do parents ensure that child set times to study at home?	3.12	3.25	-0.13*
How often do parents ensure that child get study done at home?	2.0	2.09	-0.10
F-statistics for joint significance			2.13
P-value for joint significance			0.04
Student-Level Variables	(N = 588)	(N = 611)	
Female	0.476	0.509	-0.033
Ethnic minority	0.369	0.169	0.200**
Birth order	1.893	1.933	-0.04
Mother can read and write in Vietnamese	0.859	0.943	-0.084***
Father can read and write in Vietnamese	0.866	0.910	-0.044**
Mother with upper secondary degree or higher	0.227	0.196	0.030
Father with upper secondary degree or higher	0.261	0.249	0.012
Wealth index	11.91	11.90	0.005
F-statistics for joint significance			2.66
P-value for joint significance			0.011

Note: *** $p < 0.01$; ** $p < 0.05$; * < 0.01 (bootstrapped with 500 replications). Sample includes students in 99 lower secondary schools in the RISE study. Variable is at the individual level from the 2017 student and parent questionnaires conducted by RISE.