Preliminary draft for discussion and comments, please do not cite (2/25/20)

Getting to Grade 10 in Vietnam
Challenges from deprivation and a booming job market

Ian Coxhead,[a] Nguyen Vuong[a] and Phong Nguyen[b]

Abstract

Vietnam has enjoyed more than a generation of rapid economic growth led by growth of labor-intensive exports. Average incomes have risen very fast, but some communities, notably ethnic minorities, have experienced much slower growth. At the same time, disproportionately fast growth in low-skilled labor demand has reduced the rate of rise in the skill premium. We hypothesize that these conflicting influences help explain another puzzle, the relatively slow growth of educational progression to upper-secondary school. Slow and unevenly distributed increases in schooling are warning signs for the sustainability of future aggregate growth and for the distribution of growth gains.

Contrasting educational outcomes become especially evident as children reach legal working age. We use a new dataset on participation rates and scores in an exam taken by (mainly) 15-year olds to enter upper secondary school (Grade 10-12). Data on such tests provide good measures of educational outcomes, and compared with simple school attendance measures are better indicators of children's labor market aspirations and likely career paths. After merging with household and labor force survey data we analyze variation in participation rates and scores due to variation in demand and supply side factors. The data are drawn from less advanced provinces within Vietnam and as such, shed light on challenges of expanding educational development at the extensive margins of lower socioeconomic status and higher school levels, especially in areas with high ethnic minority populations.

Keywords: Education, test scores, school-work transition, ethnic minority, Vietnam

JEL codes: I24, I25, J24, O15, P36

---

[a] University of Wisconsin-Madison. Address for correspondence: ian.coxhead@wisc.edu

[b] Former director of Social and Environmental Statistics Division, General Statistics Office, Vietnam

The authors thank the staff of the Secondary Education Division of the Ministry of Education and Training and UNICEF-Hanoi (especially Le Anh Lan and Simone Vis) for making data available and facilitating collaborations. We also thank Mr. Nguyen Thanh Hai for outstanding support in data cleaning and management.

1. Introduction

Vietnam has made tremendous progress in increasing school enrolments and achievement. In about one generation it has achieved near-universal rates of primary and lower-secondary school enrolment, up from about 86% and 72% respectively in the early 1990s; upper-secondary enrolments have increased from 27% to over 75% in the same interval (Rolleston et al. 2013; Dang and Glewwe 2018). Gains in the stock of education (measured by years of schooling) are substantially higher than for comparable countries such as South Korea and Thailand at comparable income levels (Figure 1). Despite these aggregate gains, however, significant achievement and access gaps remain, especially between wealthy and poor populations, a distinction that is increasingly strongly correlated with ethnic majority and ethnic minority populations. Over these categories we see that educational flows (that is, the production of new graduates at each grade level) are quite unequal. Disparities widen rapidly after grade 9 (ages 14-15). From age 15 to 18, average school enrolments decline from over 90% to a little more than 50%, but the enrolment decline is much faster for children from poor and ethnic minority households. Moreover the available data suggest that schooling achievement gaps are closing very slowly, if at all.

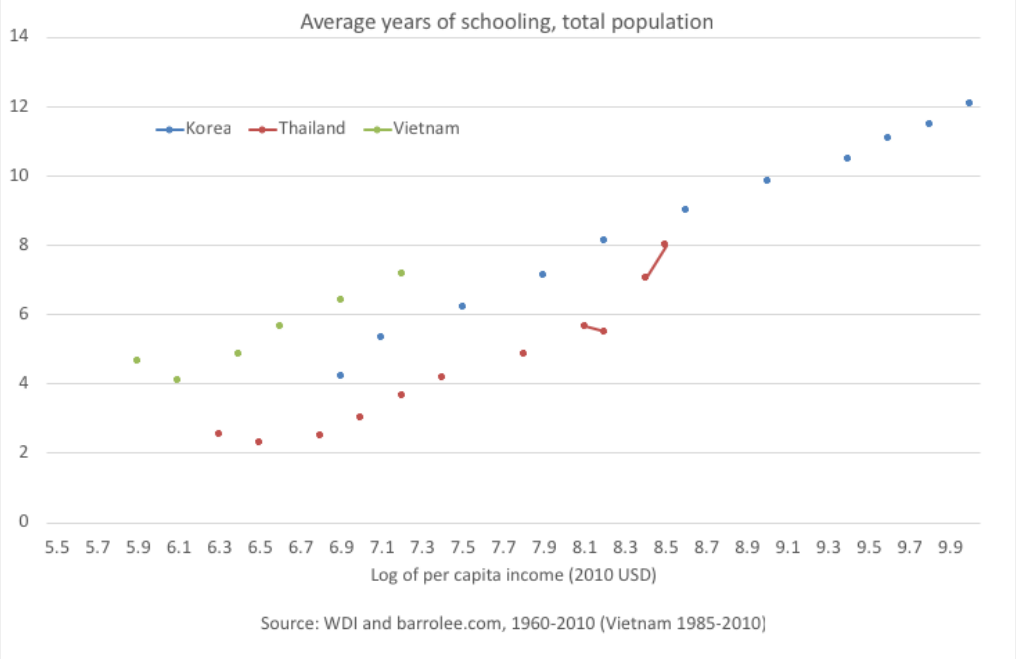

Figure 1: Average years of schooling, Vietnam, South Korea, and Thailand

Vietnam, like many countries at similar stages of development, offers few opportunities for adult education or for formal on-the-job training, so for most teenagers, the last year of schooling is the highest credential they will earn. This impacts their lifetime earnings, and opportunity gaps for the current generation may be transmitted to the next. This provides strong motivation for our

study. The decision to quit the school system has implications both for individual education stocks, which determine earnings, and for human capital accumulation in the economy as a whole, which affects future growth potential.

Demand for schooling in Vietnam is subject to income and substitution effects. Average incomes have grown very rapidly, but income growth is sustained by labor-intensive industries. Hence the opportunity cost of schooling (that is, the price of education) may drive an offsetting substitution effect. These two effects may apply with different force for different groups in the population, leading to heterogeneous patterns in educational decision-making.

Our goal in this research is to understand reasons for variation in educational attainment across groups in the population and to consider welfare and policy implications. Our central question is: how much variation in student school achievement and progress is due to income growth, how much to exogenous household or community factors such as ethnicity, and how much to "pull" from the labor market? Our purpose in asking these questions is to understand and inform educational policy, and especially the goal of educational deepening.

To answer this question, we use a unique dataset of individual test-takers and scores from the Vietnam 10th grade entrance examination (the "G10 exam"), a national exam to determine access to and placement in upper secondary schools (grades 10-12). We merge these with data from other sources, including the population census and national surveys of households and labor force, to obtain local-level controls for economic and social conditions and labor market activity. The data come from a subset of provinces that are poorer and more remote than average. As such—and also because the variable cost per student rises with school grade—they present an opportunity to study challenges to educational development at the extensive margin.

2. Globalization, growth and the labor market

To fix ideas, we begin with a stylized account of globalization in a small, low-income economy. Vietnam's experience with globalization has been the defining feature of its recent economic history. Exports of goods and services have grown from 7% of GDP in 1986 to over 100% in 2018; their real value has expanded at an average annual rate of more than 15%, three times faster than the low-middle income country average. Annual FDI inflows have averaged 6% of GDP, more than double the low-middle income average.[1] It follows that globalization has been a major driver of change in the domestic labor market.

Globalization—that is, increasing integration with global markets for products, services and factors—alters relative prices and raises returns on investment. The latter change draws in new

[1] Source: data.worldbank.org, accessed 10 February 2020.

investment, and the former helps to determine its sectoral destinations. New capital (and other productive attributes bundled with it) raises labor productivity. Those changes in turn help to determine the industries and occupations in which job growth is most active. Aggregate income increases, but the distribution of gains within the domestic economy is conditional on many factors, notably household assets (including labor and skills) and their mobility in response to changing sectoral structure.

Trade theory provides a useful lens through which to understand the main changes. Vietnam has specialized in tradable industries making intensive use of relatively low-skill labor, very often with complementary inputs of foreign investment.[2] Its import-substituting industries are far more capital-and technology-intensive than its exporters. This pattern of specialization has fueled a generation-long boom in non-farm job creation and GDP growth.

One implication of this pattern of investment and trade is that growth in demand for blue-collar labor has risen rapidly relative to that for skills.[3] The link from globalization to this structural change can be appreciated in stylized, static form with the help of a Lerner diagram (Deardorff 2002). Assume that the economy is endowed with skills (H) and blue-collar labor (L) and uses these to produce two goods of differing skill-intensity. In Figure 2, production technologies for the two goods are represented by unit isoquants $Q_i = 1/p_i$, each showing combinations of H and L compatible with production of one dollar's worth of output. In equilibrium, since factors are mobile between sectors, there is a single relative price of skills at which factor markets clear; this is shown as a unit isocost line with slope $w = w_H/w_L$ (its intercepts are $1/w_H$ and $1/w_L$). When factor markets clear, then the point on each unit isoquant with slope equal to w is the cost-minimizing input ratio $h_i = H_i/L_i$ in that industry. At the tangency, employment in each industry is found by measuring the corresponding distance along each axis.

We can choose units such that the price of the skill-intensive good (p_2) is a numeraire with value unity. The autarky relative price of the less skill-intensive good (p_1) is lower than its world price, so globalization raises it. As a result, the Q_1 moves closer to the origin. The original factor price ratio w is now no longer compatible with market clearing. To re-establish factor market equilibrium the relative price of L, the factor used intensively in industry 1, must rise. The new factor price w' reflects this shift. In other language, a rise in the relative price of the labor-intensive good has caused the skill premium to decline, since $w' < w$.

---

[2] Provide some production and trade data

[3] Demand for skills has also grown. But skilled jobs are concentrated in government and state-owned enterprises whose activity levels have grown far slower than private-sector industries. The latter are subject to crowding-out in domestic capital markets and have much lower capital and skill-intensity (Phan and Coxhead 2013).

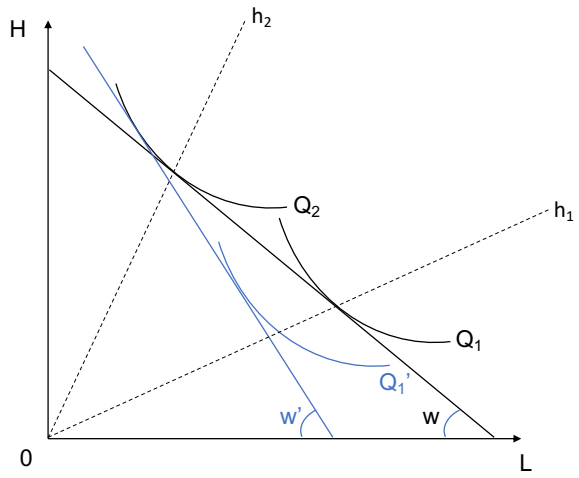

Figure 2: Effect of a rise in p_1/p_2 on the skill premium

Under full employment, both output and the quantity of each factor employed in 1 rises while in 2 they must decline. In Figure 3, point *E* shows the total endowments of *H* and *L* available to the two sectors, and the allocation of each factor to each sector is found by vector addition using the cost-minimizing factor proportions measured by h_i^j, where i =1,2 denotes sector and j = 0 in the initial equilibrium, j = 1 after the price change. The relative price change that raises demand for low-skill labor increases the skill-intensity of production in both sectors (another stylized fact in globalizing economies), but if *E* remains unchanged as shown, then the share of sector 1 in employment of both factors must also increase proportional to its output as measured by distance from the origin along the relevant ray, and that in sector 2 must fall. As drawn, for example, the share of *L* employed in sector 1 increases as a result of the price change, from $0L_1^0/0L^T$ to $0L_1^1/0L^T$.

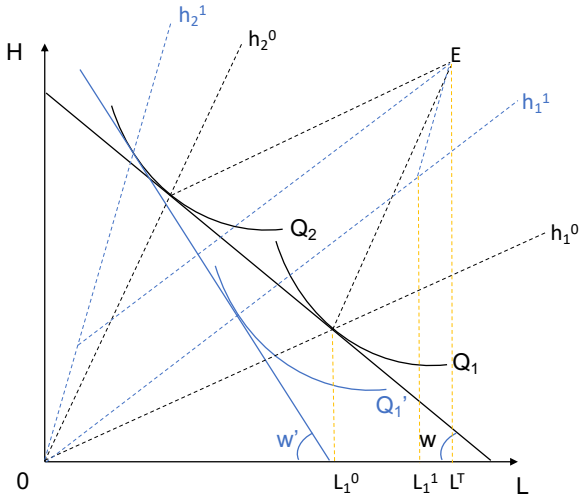

Figure 3: output and employment effects of the price rise

Of course, real-world adjustment may involve more conditions—notably, differential elasticities of factor supply, non-traded goods, adjustment times and costs, and transitions between informal and formal employment (Winters et al. 2004). In addition, we know globalization increases aggregate income due to gains from specialization and trade, along with the structural change shown in the diagram. Together, these results highlight the fundamental schooling decision problem. On one hand, income growth induces greater schooling by helping to relax credit and other constraints on educational spending by households. On the other hand, the lower skill premium and faster job growth in blue-collar occupations (i.e. in *L* rather than *H*) raises the opportunity cost of schooling. Predictably, these two forces will be felt with different intensity by different groups depending on the distribution of gains from trade and on how they calculate the expected *net* benefits of additional schooling.

The dynamics of the globalization-human capital nexus were first explored by Findlay and Kierzkowski (1983), in a model that integrated Heckscher-Ohlin specialization with models of human capital accumulation due to Mincer (1958), Shultz (1961) and Becker (1962). Their model used exogenous changes in relative product prices to drive educational decisions by individuals, taking account not only of skill-specific expected lifetime earnings but also of the direct and implicit (opportunity) costs of schooling. An insight from the Findlay-Kierzkowski model is that in a two-good economy the process is self-reinforcing over time. A rise in the relative price of the less skill-intensive good (which lowers the skill premium) induces a change in the composition of the labor force, and this in turn further increases the relative size of the less skill-intensive industry through Rybczinksi effects (Findlay and Kierzkowski 1983: 968-969).[4] In Figure 3, a more elastic supply of L than H would see the endowment point moving east-northeast over time from E. Changes in endowments and in relative factor prices are correlated.

An important qualification not shown in this model is that the economy also produces and consumes nontradable services (Lloyd 198X). Demand for services is typically income-elastic, and as such grows at least as rapidly as the aggregate economy. Therefore, growth may drive up nontradables' prices. Whether this second-round change raises or lowers the skill premium depends on the relative skill-intensity of nontradable production. In lower-income countries, services are dominated by lower-skill activities such as construction, sales, food and hospitality, and personal services, and these are largely supplied by small, often family-run firms with low levels of capitalization. In this case a rise in demand for services will raise relative demand for blue-collar labor even more than shown.[5]

---

[4] Findlay and Kierzkowski note that when all individuals are identical, the equilibrium difference in lifetime earnings between skilled and unskilled workers is zero: higher earnings of skilled workers are exactly offset by additional schooling costs that they incur. It follows that other things equal, a relative price change that favors unskilled labor will, over time, induce more children to quit school earlier.

[5] If nontradable production is of intermediate factor intensity relative to sectors 1 and 2 then an expansion in demand must reduce the skill premium. Suppose that the endowment point E represents factors available for tradable

Finally, it may be realistic to suppose that the supply of low-skilled labor is elastic due to underemployment in a "backstop" sector such as agriculture. World Bank data show employment in Vietnamese industry expanding from 12% of the labor force in 1991 to 26% in 2019, and that in services rising from 19% to 35% over the same interval. Agriculture's labor share fell from 69% to 39% and its share in GDP decreased from 40% to 15%, suggesting a substantial reallocation of low-skill workers. Suppose that the services production function uses only labor (or a combination of labor plus the entrepreneurial flair of owner-operators). Then an elastic labor supply means that service output can increase without significant price rises. In this case, the total net increase in low-skill labor demand is the sum of that in the expanding sector 1, minus that in the contracting sector 2, plus that in the expanding nontraded services sectors.

3. Related literature

Vietnam's experience with trade liberalization may be typical of that among globalizing lower-middle-income countries. Rapid employment growth concentrated in low-skill occupations poses a particular challenge when adolescents complete basic education and must choose between continuing in school or joining the labor market. These choices have lifetime implications for the current generation—lower skill workers earn much less over their working careers—but also for overall economic growth and for the distribution of gains from growth in future generations.

It is by now very well known that variation in cognitive and non-cognitive abilities in adolescence or even later in life originates very early, even in utero (Almond and Currie 2011a). Maternal, household and environmental conditions all play a potentially long-lasting role in young children's intellectual development and this in turn affects a child's performance upon entry to the formal education system (Almond and Currie 2011b). Thereafter, there is potential for a persistent widening of educational achievement gaps (in the absence of targeted interventions) as more advanced and better-motivated children, and those whose home conditions are more conducive to learning, both learn better and capture more educational resources, such as the attention and encouragement of teachers (Chiappori, HCEO). Thus variation in socioeconomic status, an indicator for a range of health and nutritional variables that contribute to early life and childhood development, is an important predictor of later-life outcomes.

Quantifying and understanding education gaps presents a challenge in early childhood unless children are subjected to specific tests of cognitive capacity. In adolescence and beyond, scores on standardized tests provide a more transparently comparable metric of educational attainment. Test scores are complementary with other frequently used measures such as total schooling years

---

production after demand for nontradable services (N) has been satisfied, so $E(H,L) = E(H^T–H_N, L^T–L_N)$. Then increased output of N will reduce E, moving it south and west of its original location. The new skill premium must be lower than its original value.

and schooling-for-age (SAGE) (Ray and Lancaster 2005), but can reduce measurement errors and other sources of noise associated with such quantitative measures. Test scores have the additional and distinctive feature that they provide a direct measure of knowledge acquisition or cognitive capacity, something that can only be very broadly inferred from data on schooling duration or grade progression since these depend more closely on school availability and quality (Hanushek and Woessmann 2012). Studies in other labor markets find that high-school test scores are robust predictors of labor market outcomes (cf. Neal and Johnson 1996). Finally, selection into test-taking is itself an important indicator of variation within the population being studied.

Why do scores and participation rates vary? In the U.S. literature (e.g., Card and Rothstein 2007; Thomas 2004) there is much attention to differences based on race and ethnicity. Some of these are correlates of structural inequality: race is a strong predictor of differences in household income, parental schooling and other predictors of children's school performance. Others are attributable to discrimination. In any heterogeneous population, the correct mix of causes is both intrinsically important (we want to understand the reason for differences) but also of great value to policy. Finding the least-cost path to an acceptably low level of variation across sub-populations begins with a good diagnosis of the causes of that variation.

Empirically, our study is closely related to those in other countries where progression within the education system is conditional on scores achieved in a screening exam. Our data also come from an educational system in which entry to the school system depends not on the actual age of the child but on the calendar birth year. In that respect it has a lot in common with several recent studies of adolescent test scores and school progression in Norway, a country with an especially rigid set of school admission criteria (Solli 2017, Strøm 2004). Such studies draw attention to the important role played by SES in compensating for other disadvantages, such as late birthdate relative to cohort. However our study (unlike many from wealthier countries with relatively homogeneous populations) must also control for selection into test participation.

Other contributions to the comparative international literature on educational attainment tend to rely heavily on quantity measures like years of schooling (for example, Blanchard and Olney 2017), but these provide very little information as to quality of training. Some recent international studies use results from the PISA standardized tests (oecd.org/pisa). However, Vietnam's PISA data have been recognized as drawn from a non-random sample (Glewwe and Hai-Anh Dang 2018) and the organization itself no longer reports the most recent Vietnam test results.

Our study also connects to a worldwide literature that identifies ethnicity with differential access to and benefits from schooling (e.g. Baert and Cockx 2013). Ethnic variation is one of particular salience in Vietnam (Kozel 2014; Baulch et al. 2010).

Trends in Vietnam data

In Vietnam, recent economic growth has been concentrated in export-oriented processing industries such as garments, footwear and electronics. Employment and skill premium data are broadly consistent with the Stolper-Samuelson conclusion of the previous section that growth has increased relative demand for less skilled workers and lowered the skill premium, other things equal. In this section we summarize evidence from studies of these trends.

The Vietnam-US Bilateral Trade Agreement (USBTA) of 2001 brought about a sudden and substantial lowering of U.S. tariffs on Vietnam's exports. Using a measure of provincial exposure to export-increasing trade policy changes, Fukase (2013) found that the effect of the USBTA on relative wages in export-exposed provinces significantly mitigated a generalized national rise in the skill premium due to ongoing domestic economic reforms. In another study, McCaig and Pavcnik (2018) found that the USBTA brought about a substantial increase in formal employment in industries most closely related to trade and among younger workers, who are both more mobile and also more intensively employed in the expanding industries.

Internal migration responses to globalization in Vietnam have been substantial (Coxhead et al. 2019). Rapid expansion of low-skilled jobs appears to have raised the opportunity cost of schooling: in particular, globalization-related job creation, proxied by local intensity of jobs in foreign-invested firms, is seen to have had a significantly negative effect on high school attendance (Coxhead and Shrestha 2017). In short, the aggregate evidence points to a clear tension in schooling trends between the positive effects of income growth and the negative effects of a booming low-skill job market.[6]

Structural change in labor demand translates into differential changes in demand for skills, and these in turn are reflected in changes in the skill premium (Katz and Murphy 2002). During Vietnam's reform era since the early 1990s, returns to education rose rapidly, albeit from a low base. Much of this increase can be attributed to the relaxation of the command-economy wage grid, in which earnings were at best only weakly correlated with skills or labor productivity.

However, recent analyses show a peak, and in some cases a pronounced decline in returns to schooling after the mid-2000s. Quantitative studies by Doan and Gibson (2012), Doan et al. (2018), and Phan and Coxhead (2013) show only modest to zero increases during the 2000s. Phan and Coxhead (2013) find that returns to lower and upper secondary schooling peaked in 2002-04, and Doan et al. (2018) find declining overall returns to schooling after 2008.[7]

---

[6] In another uniquely Vietnamese dimension, low overall returns to private sector employment of skilled workers have been exacerbated by capital constraints due to crowding-out by state sector enterprises (Phan and Coxhead 2013).

[7] Phan and Coxhead also point out that the average return to schooling conceals significant variation by subgroups in the population. Jobs in state and government sectors, which are limited in number and are largely allocated

According to the latter estimates the return to an additional year of schooling in 2014 was only 5.7%, significantly down from its peak of almost 11% in 2008, and substantially lower than world averages. The returns to upper secondary schooling in particular are very low, with wage offers to grade 12 graduates barely above those to job-seekers with grade 9 credentials.[8]

Figure 5, using Labor Force Survey data, shows that the skill premium is especially low when comparing cohorts that are close in age to the current generation of school-leavers. Moving from left to right in the figure, the upper age limit of the reference group used to calculate returns to each level of schooling is progressively reduced until on the far right, the maximum age is 25. If students respond more to perceptions of the experience of those close to them in age, this is the most relevant comparison group. At that level, the premium for upper secondary school (Grade 12) over middle school (Grade 9) is effectively zero. Premia for non-college tertiary and college are higher, but not greatly so, and their role in expected earnings as seen by a 15-year old must be weighted by the probability of applying to and gaining admission to a tertiary institution, and conditional on admission, on completing the program of study.[9]

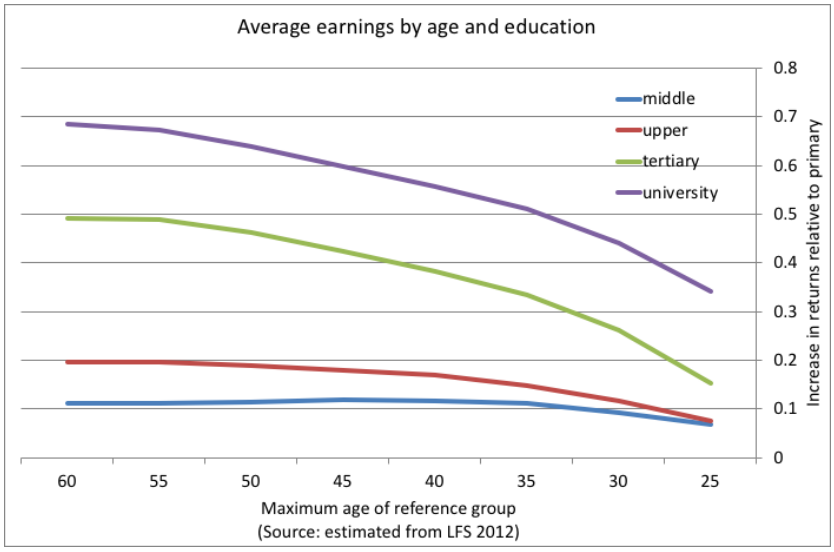

Figure 5: Average earnings by age and education, by maximum age of reference group

through non-market mechanisms, reward education at about double the rate of those in the private and self-employment sectors.

[8] These numbers are lower than the widely accepted world average of 10% (Psacharopoulos and Patrinos 2004), but are comparable with estimates from similar regional economies. Tangtipongtul (2015) estimated an average return to education in Thailand of about 13%, but this is highly convex, with estimated returns to primary schooling only 1.8%, and general secondary school only 5%; the majority of the labor force is educated to these levels or below. For Indonesia, overall returns are estimated in the medium-high single digits (Purnastuti et al. 2013; Newhouse and Suryadarma 2011; Coxhead 2014).

[9] Other dimensions of employment show similar differences and trends. In particular, Demombynes and Testaverde (2018) show no gain, and perhaps a decline, in formal wage employment (with contract) relative to informal, and a persistent difference in formality between ethnic majority and minority workers.

Much research attributes the observed decline in returns to schooling, in particular to college education, to the rising supply of high school and college graduates (e.g. Doan et al. 2018). However, comparable countries have experienced rapid growth in educational attainment without appreciable declines in returns. China, whose experience Vietnam's closely resembles in many ways, is one. From 1995 to 2014, college admissions in China increased more than sevenfold (Li et al. 2017) yet returns to college education *increased* (Gao and Smyth 2015). The most likely explanation is capital-skill complementarity: growth of skilled labor raises returns on capital investments, and the resulting increase in capital stocks raises skilled labor productivity—so increasing returns to education. This complementarity favors growth of industries that are intensive in capital and skills, a progression that is evident in China's recent history but much less visible thus far in Vietnam. In Vietnam, tertiary enrollments increased from 1998 to 2013 by a factor of 2.5 (Doan et al. 2018): a substantial gain, but far smaller than in China. So if returns to schooling have fallen in Vietnam despite a huge investment boom, then either the type of new investment was not conducive to capital-skill complementarities, and/or its impacts were diminished by other factors. In our view both the nature of Vietnam's globalization, taking advantage of its abundant low-skilled labor, and the persistence of command-economy policies in capital markets, have been contributing factors.

Schooling trends and policy targets

The Vietnamese government's education strategy for 2011-2020 set out the following targets: "By 2020, the rate of primary school students and lower secondary school students of eligible age will reach 99% and 95% respectively; 80% of young people will reach education of upper secondary school level or equivalent at eligible age; 70% of disabled children will get schooling."

Conflicting trends and population heterogeneity present the designers of education policy with a particular problem, since it is no longer the case that "one size fits all" in measures to promote schooling. This heterogeneity is especially important in Vietnam, where the pace of economic growth is rapid but has not had uniform effects. For example, Figure 3 shows the proportion of children in school, by level of schooling, for the top and bottom quintiles of households by expenditure over a decade-long interval, 2006-2016. In the top quintile, school persistence and grade progression improved (from already high initial rates), especially at upper secondary and tertiary levels, while in the bottom quintile the change was far more modest. This differential response may be due to greater income effects among wealthier households; or labor demand growth causing higher increase in the opportunity cost of schooling; or institutional differences in rates of return to secondary and tertiary schooling (Phan and Coxhead 2013); or differences in ethnic composition, since ethnic minority groups are heavily over-represented in the lowest income quintile (Kozel 2014). Figure 4 confirms that by comparison with the *kinh* majority, educational progress among Vietnam's ethnic minority children has been solid up to about age 14, but sparse thereafter.

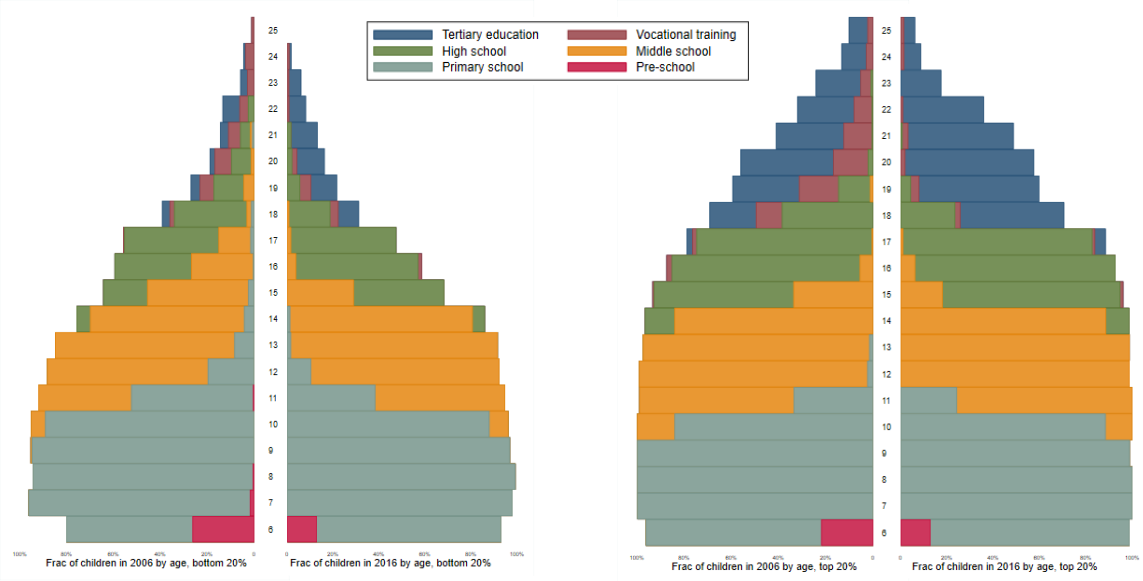

Figure 3: Enrolment rates by age and school type, 1st and 5th quintiles, 2006 and 2016
(Source of data: VHLSS)

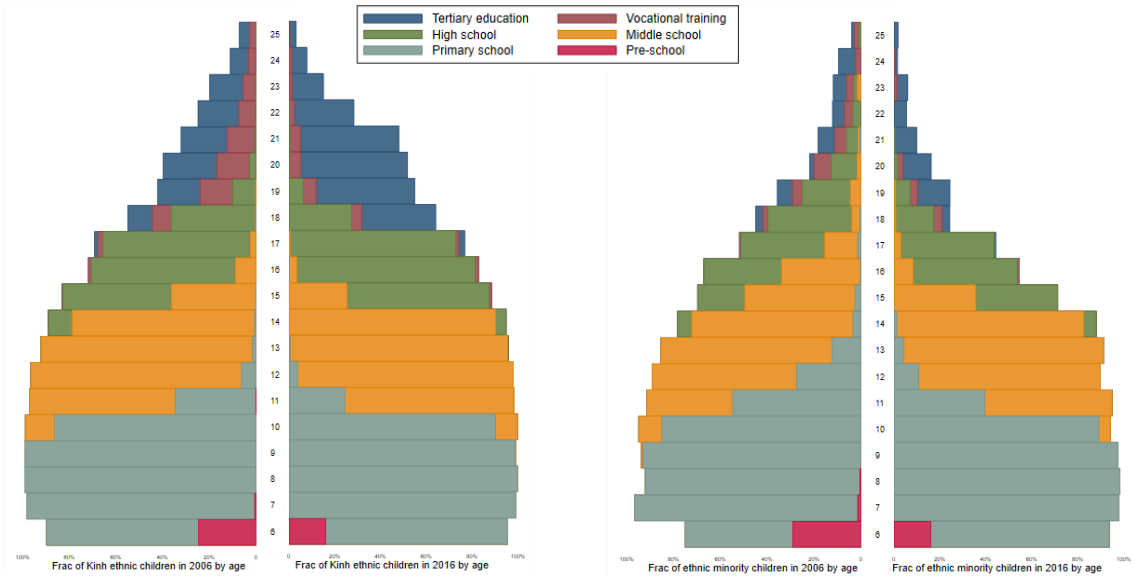

Figure 4: Enrolment rates by age and school type and ethnicity, 2006 and 2016
(Source of data: VHLSS)

That poorer quintiles and underrepresented minorities should be lagging in education is neither surprising nor unusual. Alongside these familiar dimensions of the schooling problem, however, Vietnam also displays some less common challenges associated with its very rapid growth and the reforms aimed at creating a "socialist market economy". A benchmark 2016 study from eight of the country's 63 provinces evaluated school attendance and progression and focused on the

"out of school children" (OOSC) phenomenon (UNICEF 2016). This study identified supply-side constraints including school access and quality, teacher shortages and training, transport and tuition costs, safety and security, and language barriers for ethnic minority students. But it also drew attention to demand-side issues, notably persistent poverty in some sub-populations; demand for child and adolescent labor, and disruptions associated with migration—both for children moving to cities (UNICEF 2017), and for "left-behind children" in mainly rural areas (Zhang et al. 2014; C.V. Nguyen 2016a).

Improving the supply of educational resources—schools, teachers, books and equipment, and services such as curriculum design—is fundamental and is the primary policy domain of public sector agencies with educational mandates. But schooling demand beyond the age of compulsory education is also subject to influence from labor markets, credit markets, and more. Employment, returns to skills, and incomes may all rise in the course of economic growth, but each of these changes will have a distinct (and not necessarily positive) influence on incentives to remain in school (Becker 1954). Moreover, each of these variables may have different effects on schooling when individuals are heterogeneous along some exogenous or differentially constrained dimension such as ethnicity, credit access or labor mobility.

4. Data

This project will make use of a unique data set with significant complementarities to existing data on schooling and labor market transitions in Vietnam. Students completing Grade 9 lower secondary school) who wish to continue to upper secondary and perhaps beyond must take a qualifying exam the Grade 10 entrance test. This test covers two core subjects (mathematics and literature) and may include other subjects at the discretion of provincial education authorities.[10] The total score is 50 points. High schools determine admission cutoff scores based on a variety of factors including school capacity and quality. We use the Grade 10 entry exam scores, merged at local geographic scale with data on from VHLSS, censuses, and labor force surveys to provide additional information on variables such as local incomes and distribution, population structure, and labor demand.

A test like Vietnam's G10 exam provides a reasonably clear indication of the candidate's intention to continue her studies, since there is no reason for taking the test other than to go on to 10th grade. As noted, however, an important qualification is that not all children who are eligible to take the G10 exam actually choose to do so.

[10] When test scores are the dependent variable, province fixed effects play an especially important role since the G10 exam is not nationally uniform. All exams are required to have math and literature sections, but individual provinces have the discretion to add other sections and to apply their own grading standards.

This dataset is unique, but it is not the only vehicle for studying returns to education or skills in Vietnam. In the early years of Vietnam's transition from command to mixed economy, nearly all studies of the educational implications of growth relied on data from the Vietnam Living Standards Survey (1993, 1998) and its successor, the Vietnam Household Living Standards Survey (VHLSS, 2002-present). Early studies found evidence of a strong income effect from domestic market liberalization, leading to withdrawal of children from the family/farm labor force and a preference for increased schooling, even among low-income households (e.g. Edmonds and Pacvnik 2005). Income effects may have been found to dominate in early studies in part because large-scale labor demand growth by foreign-invested firms and their domestic counterparts began in earnest only from about the year 2000.

Whereas the VHLSS data used in most other studies have minimal panel components, Mergoupis et al. (2018) used a longitudinal survey of rural households (Vietnam Access to Resources Household Survey, VAHRS) to evaluate schooling decisions in the context of labor market trends. They found no relationship between local labor markets and schooling decisions. However, VAHRS is designed to study rural asset markets, and its indicator of labor market activity is wage offers in agricultural tasks such as harvesting. Agriculture, while still a large employer, is no longer the primary source of jobs for school-leavers, and wages for seasonal work like harvesting may not convey an accurate labor market signal.

Beginning in the early 2000s, an important new dataset has spawned an increasingly rich literature on schooling. Vietnam is one of four countries in the Young Lives (YL) survey (younglives.org.uk), a longitudinal study of child poverty with a strong focus on education. YL collects detailed data on a panel of children that it has followed since 2002. It has enabled numerous in-depth studies linking individual and family conditions to educational outcomes (Rolleston et al. 2013). Rolleston and Iyer (2019), for example, use YL data to confirm "meritocratic progression" by finding that good test scores in Grade 5 predict progression to Grade 10. C.V. Nguyen (2016a) examines outcomes for children "left behind" by parents who migrate for work. Ethnicity accounts for a large share of the income and education gap in Vietnam, and for this reason is much studied (Baulch et al. 2010; Truong 2011; Glewwe et al 2015; Arouri et al 2019; H.T.M. Nguyen 2019); the most detailed studies in this literature also rely on YL data. One conclusion that emerges consistently from these is that while individual and household conditions matter for all children's educational decisions and outcomes, external factors including parental education and economic conditions, peer and school effects, and language barriers are of relatively greater importance to children from minority backgrounds (Glewwe et al. 2015; Nguyen 2019). This conclusion, if robust in nationally representative data, suggests that measures aimed at improving education for ethnic minority children should focus more on their circumstances and less on providing specialized schools or curricula, as at present.

The dataset we will use is in some respects complementary to YL. We lack detailed information on individuals but will ultimately have more complete geographical coverage.[11] This will mean greater capacity to exploit local variation in labor markets, school access and quality, and community-level characteristics. Extant YL studies reveal very little about labor market influences on school decisions, yet there is clearly a need to focus on opportunity cost since a child of 15 is able legally to enter full-time employment. The ethnicity gap in labor market opportunities is of particular salience given Vietnam's dynamically growing economy. If demand for low-skilled workers is growing fast but ethnic minority children have been historically less able or willing to migrate for work, how is their schooling affected? If rapid growth is causing changes in the skill premium in employment, how will that interact with ethnicity-based differences in expected returns to schooling? If rapid economic growth and liberalization of domestic policies creates new opportunities for entrepreneurial activities by the self-employed or family businesses, how does *that* interact with ethnicity and other exogenous differences?

Constructing the data

We want to understand why some children taking the G10 test do well and others poorly, and why a fraction of children do not participate at all. The outcome variables of greatest interest are:

- The propensity of eligible children to take the exam; and
- Scores earned by those who take the exam.

The MOET test score records cannot be matched individually with records from surveys such as VHLSS, but the G10 exam datasets record both the student's location and the lower-secondary school they attended as well as the upper-secondary school where they take the exam—that is, their intended Grade 10 school. This is sufficient to merge with data from other surveys at provincial or district (sub-provincial) level.

The project of assembling test score data is ongoing. We currently have records for several years in just four provinces (Ho Chi Minh City, Dong Thap, Ninh Thuan and Lao Cai), so this paper reports "early harvest" results. Our analytical methods are constrained by a paucity of complementary information on individual test-takers. In this paper we work mainly with

---

[11] While the YL data form a very rich panel, they have some drawbacks. First, ethnic minority representation is 12% of the YL sample, but 16% of Vietnam's youth population (a related concern is *which* ethnic minority groups are represented, and whether they, in turn, adequately capture variation among Vietnam's 57 ethnic groups). Second, YL oversamples populations close to major cities. Sixty-three percent of the YL sample comes from provinces whose capitals are no more than 1.5 hours' drive from the center of the closest major city—either Ho Chi Minh City, Hanoi, or Da Nang respectively (these are our own calculations based on data reported in https://www.younglives.org.uk/sites/www.younglives.org.uk/files/Vietnam-School-Survey_Summary.pdf, accessed 10 July 2019.) Vietnam's true population distribution is considerably more remote, on average, than this.

individual observations aggregated to district level, in addition to some (limited) analysis using individual test score records.

It is very important to note that our data are drawn from relatively poor provinces and as such are not nationally representative. Rather, they embody characteristics analogous to the "last mile" problem of delivery costs that are increasingly convex in distance. At the extensive margin of development, people live in more remote locations, are more sparsely distributed, and are typically less wealthy, more rural, and more agriculture-dependent. These populations may also be less well-connected in terms of services, information flows (internet and telecommunications backbone), and migrant networks. Moreover the task of keeping children in school beyond legal working age is more challenging than that at lower grades. From the student's point of view, the direct and implicit costs of getting an education of a given quality are higher than for inframarginal populations in regional centers and cities.

Participation rate

We have test score data only on children who have taken the G10 exam. We first want to know what fraction of the eligible population they are. We use the 2014 Intercensal Survey of Population and Housing to compute the size of each cohort of test-taking age in our dataset. We then use a count of test participants as the numerator, and the census count as the denominator to calculate the G10 participation rate, or fraction of children taking the test in a given year.

Table 1 summarizes district-level results for three provinces.[12] Provincial average rates range from 42% to 55%, and girls take the test at a much higher rate than boys in two provinces while in a third, Lao Cai, their rate is much lower. Within provinces, district-level averages display a very wide range in Lao Cai, and smaller yet still notable ranges in the other provinces. Lao Cai is remarkable in that the range of district average test-taking rates is 21%-74% for boys, and 11%-73% for girls.

[12] The provinces in our dataset are poorer and more rural than the national average, so these test participation rates are also somewhat below the national average rate.

Table 1: Test-taking rates in 3 provinces

Province (# dist)/birth cohort	Sample	Mean	S.D.	Min	Max
Lao Cai (9)	All	0.418	0.213	0.147	0.691
1999	Male	0.466	0.195	0.207	0.740
	Female	0.380	0.238	0.105	0.733
Ninh Thuan (7)	All	0.591	0.166	0.415	0.897
2000	Male	0.469	0.181	0.237	0.778
	Female	0.750	0.175	0.527	1.033
Dong Thap (12)	All	0.547	0.067	0.450	0.649
2000	Male	0.495	0.049	0.400	0.577
	Female	0.616	0.125	0.431	0.826

Household income and other controls

Because the district-level data set is still small (and in particular because there is as yet little year-to-year variation), at this stage we have only a few variables to capture local economic conditions. At district level we seek to control for household socioeconomic status (SES). Our assumption, necessarily, is that within-location variation is less than that between locations, so that a district-wide measure of income is meaningful for individual test-takers. We use district-level income and poverty data merged from VHLSS,[13] as well as location and year fixed effects to sweep out other unobserved variation.

Table 2 shows SES correlates of the G10 exam participation rate at district level, obtained using OLS. About two-thirds of the variation in the participation rate is attributable to local economic conditions alone. This result provides an interesting counterpoint to findings in many YL studies, in which the explanatory variables of interest, such as parental education and child health and nutrition status, are predominantly correlates (or even, perhaps, consequences) of community-level wellbeing. Since education policies target communities rather than individuals, there is a strong case for the approach taken in our study.

[13] VHLSS covers only a sample of districts in each province. These have been expanded to a full set of district-level SES data using small-area estimation techniques by C.V. Nguyen (2016b).

Table 2: SES correlates of G10 exam participation rate

	Dependent variable: district average test participation rate				
VARIABLES	(1)	(2)	(3)	(4)	(5)
Head count index	-0.729***				
	(0.123)				
Log of per capita expenditure		0.403***			
		(0.063)			
Poverty gap index			-1.888***		
			(0.393)		
Poverty severity index				-3.788***	
				(0.929)	
Fraction of ethnic minorities					-0.461***
					(0.076)
Constant	0.788***	-3.462***	0.728***	0.684***	0.763***
	(0.073)	(0.604)	(0.076)	(0.077)	(0.068)
Year FE	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes
Observations	70	70	70	70	70
R-squared	0.697	0.690	0.677	0.666	0.736

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

An important point relating to Table 2 is that the correlation between poverty and the ethnic minority share in district population is very high (about 0.9 for headcount poverty). This reflects the widening inter-ethnic disparity in economic growth rates during Vietnam's boom years. In 2016, ethnic minority (i.e., non-Kinh and Hoa) groups made up about 15% of Vietnam's population but accounted for 75% of the poor population (World Bank 2018).

G10 test scores

Individual test scores are a composite of results for components of the G10 exam and sum to 50. Figure 6 summarizes some individual test score distributions over relevant characteristics. There is wide variation within provinces by ethnicity and location as well (in some) by gender. These distributions also highlight differences *between* provinces in locational and gender-based disparities (one province has negligible minority population).

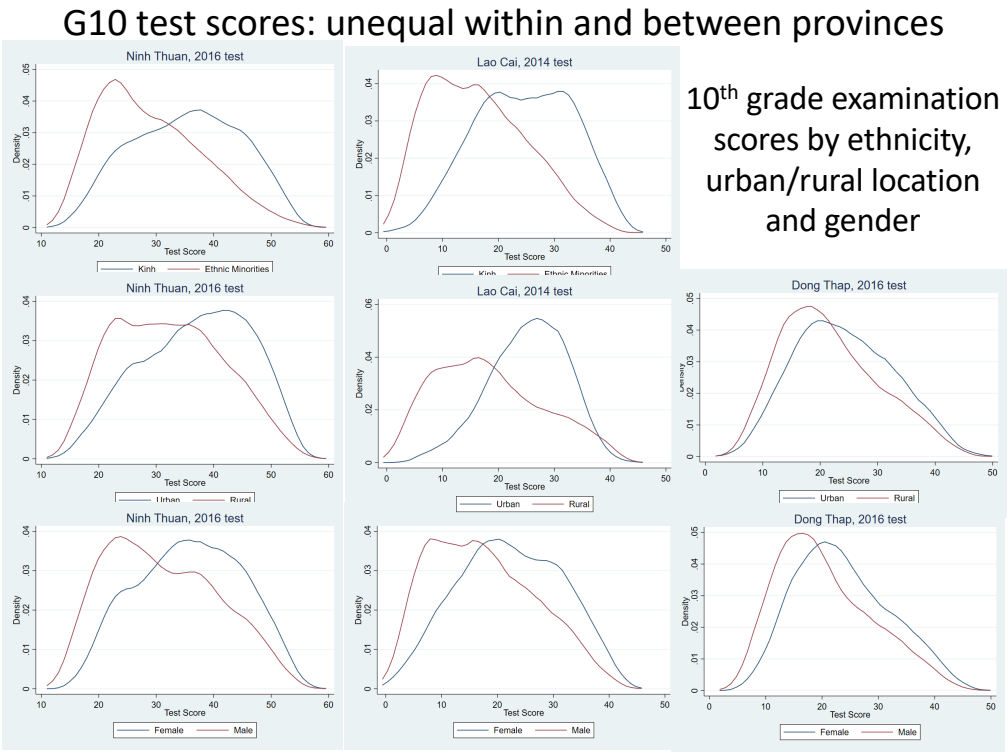

Figure 6: G10 test score distributions for by ethnicity, urban/rural, and gender

Labor market variables

The literature on school attainment in developing countries contains surprisingly few studies addressing demand-side phenomena. Atkin (2016) found the establishment of *maquiladoras* (foreign-invested factories producing for export) in Mexico had a negative effect on school enrolment among teenaged children in the same localities. Coxhead and Shrestha (2017) found that district-level FDI employment shares in Vietnam robustly predicted propensity for teenaged children to be out of school. In both cases the "treatment" variable should be interpreted less as a measure of direct employment effects and more as indicating an elevated overall level of labor demand. In Vietnam, employment by foreign-invested firms is a very small share of total employment other than in a few provinces in the industrial belt surrounding Ho Chi Minh City. The LFS shows that more than 90% of young school leavers aged 15 and 16 are employed instead in household businesses. However these businesses, most of which produce a variety of services such as construction, trade and transport, retail and personal services, flourish mainly in areas experiencing large injections of new investment.

Stimuli to employment may both increase incomes and raise the opportunity cost of schooling. Coxhead and Shrestha (2017) pursued a quantity-based approach using FDI employment shares.

Another approach to labor demand is through the influence of skill premia. A higher skill premium indicates higher relative return to investment in education and should be positively associated with test participation. Of course, labor mobility may mean that earnings per hour for a given skill level do not vary much over districts.[14] However, the structure of employment by skills should be an important ex ante influence on children's decisions. A location with a relatively high share of skilled jobs should give the skill premium greater weight in that location relative to one in which there are very few such jobs. As seen in Section 2, the skill premium and the skilled share of the labor force should be positively correlated in the long run since a higher premium incentivizes greater educational investment. We use two definitions of the skill premium, defined as the ratio of skilled to unskilled earnings: one where skilled is defined as having a high school (12^{th} grade) diploma or higher educational attainment, and another in which the cutoff is at 9^{th} grade completion or higher.

We expect that greater low-skill labor demand, other things equal, should be negatively associated with test participation while a higher skill premium should have a positive association. The influence of participation rates on test scores is an open question. If higher participation rates are the result of students at the lower margin of cognitive ability selecting into high school, test scores and participation rates should be negatively correlated. One the other hand, if the marginal test-taker is a student of relatively high ability deciding to remain in school rather than join the labor force, the association between scores and participation rates should be positive. The difference is of policy interest since it may imply potential for education policy to either widen or narrow the achievement gap between more and less well-off students.[15]

5. Determinants of test participation and test scores

Our goal is to test hypotheses about the contribution of different factors to variation in test participation and scores, and in particular to distinguish the influence of demand-side labor market trends from supply-side conditions such as socioeconomic status and ethnicity.

Participation rates

Table 3 reports results of OLS regressions with district average test participation rates as dependent variables. The right-hand side includes controls for household income or poverty as well as province and year fixed effects. The variable of interest is the skill premium. This is strongly correlated with per capita incomes (r=0.4) so there is a problem of multicollinearity in

[14] For wage workers with contract in 2014, Demombynes and Testaverde (2018) report regional data with a range from VND21,000/hr in the Mekong Delta to VND26,000/hr in the Central Highlands. The rural average was VND20,000/hr and the urban average VND27,000/hr.

[15] Clark et al. (2007) find that participation in U.S. college entrance examinations is positively selected. Participation rates are higher at better schools, and within schools, better students are more likely to take the test.

models containing both variables. But skill premia are not correlated with measures of poverty or poverty severity. With a reasonable degree of precision given the sample size, the results indicate that a higher skill premium induces higher test participation, with the effect for the more conservative (12th grade) skill premium definition about double that of the 9th-grade definition. In terms of magnitudes, a one standard deviation increase in the 9th grade skill premium increases G10 participation by about four percentage points. Conversely, a compression of the skill premium, as seen in Vietnam at least during the mid-2000s, discourages students from continuing their studies beyond 9th grade.

The final column of Table 3 uses the ethnic minority fraction in district population in place of income or poverty controls. The well-known correspondence between ethnic minority status and poverty in Vietnam (World Bank 2018) is immediately on display. Children in districts with higher ethnic minority concentrations are far less likely to take the G10 exam, no matter which definition of the skill premium we use. The fact that minority status and measures of socioeconomic status or poverty are so highly correlated that they cannot both be included in a single regression model speaks eloquently to the differential experience of this part of the Vietnamese population.

Test scores

Does an increase in the G10 participation rate mean that more children of lower cognitive ability (or poorer middle-school preparation) are taking the test? Or does it indicate that higher-ability children are opting to remain in education and thereby (presumably) acquiring greater skill rather than succumb to the shorter-term lure of increasingly well-paid blue-collar jobs? If the former, then we should worry that measures to encourage G10 test-taking may simply set more lower-ability children up for failure without adding appreciably to the stock of skilled workers. If the latter, then promoting test participation may make a positive contribution to skills accumulation both at individual and at aggregate levels.

Table 3: Socioeconomic status, skill premium and G10 exam participation rate

	Test participation rate				
VARIABLES	(1)	(2)	(3)	(4)	(5)
Panel A					
Skill premium of over-9th grader	0.060	0.011	0.086**	0.102**	0.083**
	(0.041)	(0.046)	(0.042)	(0.042)	(0.037)
Head count poverty	-0.623***				
	(0.142)				
Log of per capita expenditure		0.390***			
		(0.084)			
Poverty gap index			-1.474***		
			(0.432)		
Poverty severity index				-2.752***	
				(0.994)	
Fraction of ethnic minorities					-0.393***
					(0.079)
Constant	0.687***	-3.344***	0.593***	0.532***	0.648***
	(0.099)	(0.785)	(0.098)	(0.097)	(0.083)
Observations	70	70	70	70	70
R-squared	0.542	0.555	0.494	0.465	0.570
Panel B					
Skill premium of over-12th grader	0.145**	0.068	0.183**	0.203***	0.164**
	(0.072)	(0.077)	(0.074)	(0.076)	(0.069)
Head count poverty	-0.653***				
	(0.126)				
Log of per capita expenditure		0.374***			
		(0.071)			
Poverty gap index			-1.664***		
			(0.389)		
Poverty severity index				-3.315***	
				(0.904)	
Fraction of ethnic minorities					-0.417***
					(0.075)
Constant	0.694***	-3.207***	0.621***	0.573***	0.668***
	(0.085)	(0.670)	(0.085)	(0.084)	(0.077)
Observations	70	70	70	70	70
R-squared	0.555	0.560	0.508	0.476	0.573
Year FE	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Summary statistics for 2SLS

Variable	Obs	Mean	Std. Dev.	Min	Max
Score	66	25.19	5.93	8.82	38.76
Fraction of ethnic minorities	66	0.22	0.32	0.00	1.00
Test taking rate	66	0.57	0.15	0.15	1.01
Wage premium (High school diploma)	66	1.32	0.54	0.16	3.12
Ratio of HS diploma holders to non-holders (25-35 yo)	66	0.48	0.43	0.06	3.24

Table 5: 2SLS estimates of 10th grade entrance score

VARIABLES	(1) Test participation	(2) score
Fraction of ethnic minorities	-0.417***	-1.474
	(0.088)	(5.233)
Wage premium (HS diploma)	0.029	
	(0.033)	
HS diploma/Non-HS diploma, 25-35yo	0.073**	
	(0.036)	
Test taking rate		24.550**
		(10.716)
Constant	0.620***	8.117
	(0.092)	(8.319)
Province FE	Yes	Yes
Year FE	Yes	Yes
First stage F-stat	10.10	
Sargan's *J* test	0.201	
Wu-Hausman F-stats	5.066**	
Observations	66	66
R-squared	0.549	0.800

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

We address this question in a two stage least squares model with test scores as the dependent variable of interest. Summary statistics of variables are in Table 4, and estimates are in Table 5. The instrumental variables approach allows for non-random selection into test-taking. In the first stage, we instrument the participation rate using the 12th grade wage premium and the ratio of skilled to unskilled workers in the district labor force. In the second stage, the instrumented

participation rate enters as a control in the test score equation. Appendix table A-1 provides alternative results with other instruments; the estimates of interest are essentially unchanged.

Diagnostics on the first-stage regressions confirm the relevance and validity of the instruments. The null hypothesis of an exogeneous participation rate is rejected. The test participation rate has a significantly positive influence on test scores. A higher ethnic minority population in the district has a significant negative association with test participation, as before. These results support the idea that measures to encourage G10 participation are likely to attract more candidates of higher ability rather than of lower ability. The participation rate variable is scaled from 0 to 1, so a one percentage point increase in this rate raises test scores by 0.246 points. The mean participation rate is 0.57 with a standard deviation of 0.15, so a one standard deviation increase in the participation rate raises test scores by 3.7 points, or 0.6 of a standard deviation.

Finally, we note once again the special circumstances of ethnic minority districts. In these, the propensity to take the G10 exam is significantly lower than in *kinh*-majority districts. However our results with the current data indicate that once differences in test participation are controlled, and taking account of province and year fixed effects, there is no significant ethnicity-based difference in test scores.

Discussion

The estimation results in Tables 3 and 5 provide useful information in their own right and also create optimism that stronger and richer results could be obtainable with a larger data set. Of course, the size and representativeness of our dataset is only one of several potentially important challenges to the validity and accuracy of the reported results. We conclude this section with a brief discussion of some other threats to be considered and perhaps addressed.

Omitted variables. We have minimal information on individual children, we are required to assume that individual differences in G10 exam participation and scores are randomly distributed after controlling for known characteristics (ethnicity, age, G9 school and local social and economic characteristics, including economic and labor market conditions). In this way our study is complementary to YL, which contains a rich set of individual characteristics but covers a small sample with limited cross-sectional variation.

Endogenous selection into lower-secondary schools. Some children attend schools outside their hometown in order to take advantage of better educational opportunities. If this is widespread then estimates of community-level influences over outcome variables will exhibit bias. The test score datasets record whether a child is a migrant to the area where they attend school, so there is some capacity to control for this form of selection.

Endogenous investment in schools and educational services: wealthier communities naturally lobby to win high-quality schools and teachers. If this is widespread and unobserved in the data,

then we will overestimate the effects of community characteristics. Conversely, if there is deliberate "catch-up" investment in deprived areas (and if that investment is effective) then we may underestimate the influence of community characteristics. In both cases, information on specific investments will help. School-level data from MOET will help minimize this source of bias.

6. Conclusions and policy discussion

Vietnam's economy has emerged into middle income after a generation of very rapid growth. That growth has been anchored on globalization, in which the country has exploited its comparative advantage in low-skill labor-intensive activities. The pattern of growth is not unambiguously positive. While higher per capita incomes (on average) facilitate and encourage continued schooling, the labor market offers instant rewards to school-leavers from a very young age. Moreover, the pattern of income gains and job market opportunities appears to be unevenly distributed, so for some sub-populations, labor market effects might dominate income effects. More work with richer data is required to quantify and understand these variations.

In this study we examine newly available data on education-labor market interactions from several less-privileged provinces within Vietnam. These data are not nationally representative; rather, they are drawn from provinces at the extensive margin of educational development. The populations of those provinces are poorer and ethnic minority groups are more prevalent than in the nation as a whole. In locations such as these, the marginal cost of providing additional educational opportunities, especially at upper secondary level, is likely to be high. This in turn is a factor inhibiting both poverty reduction and intergenerational mobility, and as such places an additional burden on policy.

We find that there is considerable variation in propensity to attempt the entrance examination to Grade 10, the gateway grade to rise above blue-collar labor market status. This variation runs in predictable patterns across SES and ethnicity. We also find, however, that opportunity cost in the form of skill premia and related variables in the labor market play a significant role. Looking further, we find that after instrumenting, the test participation rate is positively associated with test scores. This seems to indicate positive selection into test participation. One narrative consistent with this finding is that measures to raise the participation rate will tend to pull in children whose ability qualifies them to do well in upper secondary school but who might otherwise have opted out of education after 9th grade. If robust, this is a positive sign for individual human capital accumulation at both individual and aggregate scales.

The idea of measures to raise test participation rates brings us to education policy. The Education Law of 2019, a major revision of previous statutes, made education to Grade 9 free in the public school system and in doing so embodied the goal of universal education to that level. Clearly the

intent of the law is to promote further educational deepening, so progression to Grade 10 (and beyond) has become a focal point of education policy.

Rankings from a recent World Economic Forum report on "Readiness for the future of production" (WEF 2018) highlight the challenges that Vietnam faces in transitioning to globalization. On an indicator called "drivers of production," among 100 countries Vietnam is ranked 13th on "global trade and investment," between Australia and France. On "human capital," however, it is ranked 70th, between Sri Lanka and Georgia, and on "technology and innovation" it is ranked 90th, between Paraguay and Cameroon. The human capital index notably includes low rankings for Vietnam on key components such as mean years of schooling (rank: 74), quality of universities (75), quality of vocational training (80), and on-the-job training (74).

How to achieve educational deepening, at what cost and with what aggregate or individual benefits is a compelling set of questions for Vietnam.[16] At present, education policy reforms are focused on supply-side innovations and distributional equity. The former include school construction, teaching quality, and curriculum reform. These are undoubtedly important areas. Curriculum reform in particular has the potential to alter the balance of costs and returns to upper secondary schooling, especially as the current curriculum at that level is tightly focused on academic work in preparation for the college entrance examination, rather than on preparation for labor market entry per se.

Looking beyond one country, Vietnam has some special features, but it is also broadly representative of developing countries that arrived late to globalization and now (or will soon) confront the need to move beyond production based on natural resources, low-skill labor and largely foreign capital. Indonesia, Bangladesh, Cambodia and other countries now face similar human capital investment challenges to Vietnam. Lessons from one country will benefit others.

[16] See Goldin (1998) for a strikingly similar account of U.S. secondary education development, 1910-1940.

References

Almond, D., and J. Currie, 2011a. Killing me softly: the fetal origins hypothesis. *Journal of Economic Perspectives* 25(3): 153-172.

Almond, D., and J. Currie, 2011b. Human capital development before age 5. In *Handbook of Labor Economics* vol.4b: 1315-1486.

Arouri, M., A. Ben-Youssef, and C.V. Nguyen, 2019. Ethnic and racial disparities in children's education: comparative evidence from Ethiopia, India, Peru, and Viet Nam. *Children and Youth Services Review* 100: 503-514.

Baert, S., and B. Cockx, 2013. Pure ethnic gaps in educational attainment and school to work transitions: when do they arise? *Economics of Education Review* 36: 276-294.

Baulch, B., T.M.N. Hoa, T.T. Phuong, and T.P. Hung, 2010. Ethnic minority poverty in Vietnam. Working Paper No. 169. Manchester: Chronic Poverty Research Centre.

Becker, G., 1954. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, New York: Columbia University Press.

Coxhead, I., 2014. Indonesia: returns to occupations, education and ability during a resource export boom. In P-C Athukorala, A.A. Patunru, and B. Resosudarmo (eds), *Trade, development and Political Economy in East Asia: Essays in Honor of Hal Hill*. Singapore: ISEAS.

Coxhead, I., and R. Shrestha, 2017. Globalization and school-work choices in an emerging economy: Vietnam. *Asian Economic Papers* 16(2): 1-18.

Coxhead, I.; Nguyen Viet Cuong, and Vu Hoang Linh, 2019. Internal migration in Vietnam, 2000-2012. In A.Y.C. Liu and X. Meng (eds), *Rural-Urban Migration in Vietna*m, Springer: 67-96.

Dang, H.-A.H., and P. Glewwe, 2018. Well begun, but aiming higher: a review of Vietnam's educational trends in the past 20 years and emerging challenges. *Journal of Development Studies* 54 (7), 1171–1195.

Deardorff, A., 2002. Introduction to the Lerner diagram. http://www-personal.umich.edu/~alandear/writings/Lerner.pdf, accessed January 31, 2020.

Doan, T., and J. Gibson, 2012. Return to education in Vietnam during the recent transformation. *International Journal of Education Economics and Development* 3(4): 314–329

Doan, T., Q.T. Tuyen, and L. Quan, 2018. Lost in transition? declining returns to education in Vietnam. *European Journal of Development Research* 30(2): 195-216.

Edmonds, E., and N. Pavcnik, 2005. The effect of trade liberalization on child labor. *Journal of International Economics* 65(2): 401-419.

Fukase, E., 2013. Export Liberalization, Job Creation and the Skill Premium: Evidence from the US-Vietnam Bilateral Trade Agreement." World Development 41: 317–37

Gao, W., and R. Smyth, 2014. Education expansion and returns to schooling in urban China, 2001-2010: evidence from three waves of the China Urban Labor Survey. *Journal of the Asia-Pacific Economy* 20(2): 178-201.

Glewwe, P., Q. Chen, and B. Katare, 2015. What determines learning among kinh and ethnic minority students in Vietnam? An analysis of the Round 2 Young Lives Data. *Asia & the Pacific Policy Studies* 2 (3): 494–516.

Goldin, C., 1998. America's graduation from high school: the evolution and spread of secondary education in the twentieth century. *Journal of Economic History* 58(2): 345-374.

Kozel, V., ed., 2014. *Well begun but not yet done: Progress and emerging challenges for poverty reduction in Vietnam*. World Bank.

Li, H.; P. Loyalka, S. Rozelle, and B. Wu, 2017. Human capital and China's future growth. *Journal of Economic Perspectives* 31(1): 25-48.

McCaig, B., and N. Pavcnik, 2018. Export markets and labor allocation in a low-income country. American Economic Review 108(7): 1899-1941.

Mergoupis, T.; V. Phan, and J.G. Sessions, 2018. Puzzle me this? The Vietnamese reverse gender education gap. WIDER Working Paper No. 2018/116.

Neal, D.A., and W.R. Johnson, 1996. The role of premarket factors in black-white wage differences. *Journal of Political Economy* 104(5): 869-895.

Newhouse, D., and D. Suryadarma. 2011. The value of vocational education: high school type and labor market outcomes in Indonesia. *World Bank Economic Review* 25 (2): 296–322.

Nguyen, C.V., 2016a. Does parental migration really benefit left-behind children? Comparative evidence from Ethiopia, India, Peru and Vietnam. *Social Science and Medicine* 153: 230-239.

Nguyen, C.V., 2016b. Poverty and inequality estimates using the 2014 Vietnam Household Living Standard Survey and the 2014 Intercensal Population and Housing Survey. Unpublished paper.

Nguyen, H.T.M., 2019. Ethnic gaps in child education outcomes in Vietnam: an investigation using Young Lives data. *Education Economics* 27(1): 93-111.

Phan, D. and I. Coxhead, 2013. Long-run costs of piecemeal reform: wage inequality and returns to education in Vietnam. *Journal of Comparative Economics* 41(4): 1106-1122.

Purnastuti, L.; P. Miller, and R. Salim, 2013. Declining rates of return to education: evidence for Indonesia'. *Bulletin of Indonesian Economic Studies* 49 (2): 213–36.

Ray, R., and G. Lancaster. 2005. The impact of children's work on schooling: multi-country evidence. *International Labor Review* 144(2):189-210.

Rolleston, C., and P. Iyer, 2019. Beyond the basics: access and equity in the expansion of post-compulsory schooling in Vietnam. *International J. of Educational Development* 66: 223-233.

Rolleston, C., Z. James, L. Pasquier-Doumer and N.T.M.T. Tran, 2013. Making progress: report of the Young Lives School Survey in Vietnam, Working Paper 100, Young Lives Project.

Solli, I.F., 2017. Left behind by birth month. *Education Economics* 25(4): 323-346.

Strøm, B., 2004. Student achievement and birth effects. Manuscript, Norwegian University of Science and Technology.

Tangtipongtul, K., 2015. Rates of return to schooling in Thailand. *Asian Development Review* 32(2): 38-64.

Truong, H.C., 2011. Eliminating inter-ethnic inequalities? Assessing impacts of education policies on ethnic minority children in Vietnam, Working Paper No. 69, Young Lives Project.

UNICEF and Ho Chi Minh City People's Committee, 2017. Situation analysis of children in Ho Chi Minh City, Vietnam 2017. Hanoi: UNICEF.

UNICEF and Ministry of Education and Training, 2016. Out-of-school children in Vietnam: A country study. Hanoi: UNICEF.

Winters, L.A., N. McCulloch, and A. McKay, 2004. Trade liberalization and poverty: the evidence so far. *Journal of Economic Literature* 42(1): 72-115.

World Bank, 2018. Climbing the ladder: poverty reduction and shared prosperity in Vietnam. Vietnam Poverty and Shared Prosperity Update Report 2018.

World Economic Forum, 2018. *Readiness for the Future of Production Report 2018.* http://www3.weforum.org/docs/FOP_Readiness_Report_2018.pdf.

Zhang, H.; J.R. Behrman, C. S. Fan, X. Wei, and J. Zhang, 2014. Does parental absence reduce cognitive achievements? Evidence from rural China. *Journal of Development Economics* 111:181-195.

Appendix: Calculating district-level wage premia from Labor Force Survey data

The LFS samples individuals of working age and then weights each observation to make it representative at district level. We use a two-level definition of skills, high (H) and low (L) (in estimation we define H two ways: schooling above 9th grade, and schooling above 12th grade).

Consider $s_i \in \{H, L\}$ is the skill and there are two types of skills, H and L. With ϖ_{ij} and y_{ij} as the sampling weight and wage of person i in district j, we calculate the district-level wage for high skilled labor as

$$w_{Hj} = \frac{\sum_i y_{ij}\varpi_{ij}.1\{s_{ij} = H\}}{\sum_i \varpi_{ij}.1\{s_{ij} = H\}},$$

And the wage for low skilled labor as

$$w_{Lj} = \frac{\sum_i y_{ij}\varpi_{ij}.1\{s_{ij} = L\}}{\sum_i \varpi_{ij}.1\{s_{ij} = L\}}.$$

The fraction of each type of labor is $s_{Hj} = \frac{\sum_i \varpi_{ij}.1\{s_{ij}=H\}}{\sum_i \varpi_{ij}}$ and $s_{Lj} = \frac{\sum_i \varpi_{ij}.1\{s_{ij}=L\}}{\sum_i \varpi_{ij}}$.

The skill premium p_j is the ratio of the representative wage for H divided by that for L, with each weighted by the share of labor of that type in district j:

$$p_j = \frac{w_{Hj}}{w_{Lj}}.\frac{s_{Hj}}{s_{Lj}} = \frac{\frac{\sum_i y_{ij}\varpi_{ij}.1\{s_{ij} = H\}}{\sum_i \varpi_{ij}.1\{s_{ij} = H\}}.\frac{\sum_i \varpi_{ij}.1\{s_{ij} = H\}}{\sum_i \varpi_{ij}}}{\frac{\sum_i y_{ij}\varpi_{ij}.1\{s_{ij} = L\}}{\sum_i \varpi_{ij}.1\{s_{ij} = L\}}.\frac{\sum_i \varpi_{ij}.1\{s_{ij} = L\}}{\sum_i \varpi_{ij}}}$$

Cancelling out unnecessary terms, we have

$$p_j = \frac{\sum_i y_{ij}\varpi_{ij}.1\{s_{ij} = H\}}{\sum_i y_{ij}\varpi_{ij}.1\{s_{ij} = L\}}.$$

Because the distribution of the labor force by skills is unique to each district, the skill premium is also unique even if wages by skill do not vary much over districts. Our goal in weighting by labor force shares is to capture not only the difference in returns by skill, but also the *ex ante* probability of working at a given skill level. So a district with a very high premium but very few skilled jobs (low s_H) may send a different labor market signal to schoolchildren compared with a district with the same premium but a much larger s_H.

Table A-1: Complete 2SLS results

VARIABLES	2SLS Test taking rate	2SLS Test score	2SLS Test taking rate	2SLS Test score	2SLS Test taking rate	2SLS Test score	OLS Test score
Fraction of ethnic minorities	-0.396***	-1.455	-0.393***	-1.979	-0.422***	0.853	-9.348***
	(0.080)	(5.351)	(0.079)	(5.212)	(0.076)	(6.735)	(1.746)
Skill premium Cut-off: 9th grade	0.067		0.082**				
	(0.071)		(0.037)				
Skill premium Cut-off: 12th grade	0.032				0.142**		
	(0.135)				(0.071)		
Test-taking rate		24.592**		23.449**		29.618**	7.398***
		(10.987)		(10.709)		(13.972)	(2.354)
Constant	0.650***	8.084	0.649***	8.955	0.680***	4.255	21.184***
	(0.083)	(8.522)	(0.083)	(8.305)	(0.077)	(10.802)	(2.173)
First stage F-stats	10.02		11.88		11.56		
Sargan's *J* test	0.922						
Wu-Hausman stats	4.815**		4.095**		6.40**		
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	66	66	66	66	66	66	66
R-squared	0.547	0.800	0.547	0.812	0.540	0.736	0.895

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1