The Decision to Invest in Child Quality over Quantity: Household Size and Household Investment in Education in Vietnam

Hai-Anh H. Dang and F. Halsey Rogers

During Vietnam's two decades of rapid economic growth, its fertility rate has fallen sharply at the same time that its educational attainment has risen rapidly—macro trends that are consistent with the hypothesis of a quantity-quality tradeoff in child-rearing. We investigate whether the micro-level evidence supports the hypothesis that Vietnamese parents are in fact making a tradeoff between quantity and "quality" of children. We present private tutoring—a widespread education phenomenon in Vietnam—as a new measure of household investment in children's quality, combining it with traditional measures of household education investments. To assess the quantity-quality tradeoff, we instrument for family size using the commune distance to the nearest family planning center. Our IV estimation results based on data from the Vietnam Household Living Standards Surveys (VHLSSs) and other sources show that rural families do indeed invest less in the education of school-age children who have larger numbers of siblings. This effect holds for several different indicators of educational investment and is robust to different definitions of family size, identification strategies, and model specifications that control for community characteristics as well as the distance to the city center. Finally, our estimation results suggest that private tutoring may be a better measure of quality-oriented household investments in education than traditional measures like enrollment, which are arguably less nuanced and less household-driven. JEL: I22, I28, J13, O15, O53, P36

Over the past four decades, there has been considerable study of the relationship between household choices on the quantity and quality of children, starting with the seminal studies by Becker (1960) and Becker and Lewis (1973). The

Hai-Anh H. Dang (corresponding author) is an economist with the Poverty and Inequality Unit, Development Research Group, World Bank; his email address is hdang@worldbank.org. F. Halsey Rogers is lead economist with the Global Education Practice, World Bank; his email address is hrogers@worldbank.org. We would like to thank the editor Andrew Foster, three anonymous referees, Mark Bray, Miriam Bruhn, Hanan Jacoby, Shahidur Khandker, Stuti Khemani, David McKenzie, Cem Mete, Cong Pham, Paul Schultz, and colleagues participating in the World Bank's Hewlett grant research program, and participants at the Population Association of America Meeting for helpful comments on earlier drafts of this paper. We would also like to thank the Hewlett Foundation for its generous support of this research (grant number 2005-6791). A supplemental appendix to this article is available at http://wber.oxfordjournals.org/.

THE WORLD BANK ECONOMIC REVIEW, VOL. 30, NO. 1, pp. 104–142 doi:10.1093/wber/lhv048
Advance Access Publication August 25, 2015

hypothesis driving the literature is that parents make tradeoffs between the number of children they bear and the "quality" of those children, which is shorthand for the amount of investment that parents make in their children's human capital. If this hypothesis is true, it has considerable implications for policies aimed at increasing economic growth and reducing poverty.[1] For example, this can motivate policy makers to work on policies that assist couples to avoid unwanted births or to subsidize birth control (Schultz 2008).

We investigate a different measure of household investment in their children in this paper, which is private tutoring—or extra classes—in mainstream subjects at schools that children are tested in. Private tutoring is now widespread in many countries, especially but not solely in East Asia,[2] and evidence indicates that it improves students' academic performance in some countries, including Germany, Israel, Japan, and Vietnam (Dang and Rogers 2008).[3] There has been considerable debate about tutoring among policymakers. One crucial question is whether widespread availability and use of private tutoring exacerbates or helps equalize social and income inequality (Bray 2009; Bray and Lykins 2012), a question that is relevant to both developing and developed countries.[4] Here, the link with demography is important: if use of tutoring is correlated with both smaller family size and higher family income, this heightens the risk that it could exacerbate inequality.

We make several conceptual and empirical contributions in this paper. Our conceptual contribution is to propose private tutoring as a new measure of household investment in their children's education quality in the context of the child quantity-quality tradeoff literature. Private tutoring may be an especially good measure of a household's decision to invest voluntarily in children's human capital—compared with enrollment, for example, which may also reflect exogenous factors such as compulsory schooling laws. Put differently, private tutoring

1. The empirical evidence on the correlation between household size and poverty appears inconclusive. For example, Lanjouw et al. (2004) argue that the common view that larger-sized households are poorer is sensitive to assumptions made about economies of scale in consumption.

2. Private tutoring (or supplementary education) is a widespread phenomenon, found in countries as diverse economically and geographically as Cambodia, the Arab Republic of Egypt, Japan, Kenya, Romania, Singapore, the United States, and the United Kingdom. A recent survey of the prevalence of tutoring in twenty-two developed and developing countries finds that in most of these countries, 25–90 percent of students at various levels of education are receiving or recently received private tutoring, and spending by households on private tutoring even rivals public sector education expenditures in some countries such as the Republic of Korea and Turkey (Dang and Rogers 2008).

3. Other recent studies that find tutoring to have positive on different measures of student academic performance include student test scores and academic performance in India (Banerjee et al. 2010) and the United States (Zimmer et al. 2010); but see Zhang (2013) for recent evidence that tutoring may benefit only certain student groups in China.

4. Given the rapid expansion of educational attainment around the developing world, the tradeoffs that households make between the quantity and quality of children may increasingly manifest themselves outside of the formal education system. For example, in a recent opinion piece in the *New York Times* on the widening inequality in the United States, the Nobel laureate Joseph Stiglitz (2013) calls for more "summer and extracurricular programs that enrich low-income students' skills" to help level the playing field between these students and their richer peers.

can capture the household's extra efforts to increase their children's human capital. In particular, in countries where the private-school sector is almost nonexistent (at least at the pre-tertiary school level) such as Vietnam, private tutoring represents a type of flexible household education investment, which is most likely to be the equivalent of household investment in private education in other contexts.[5] Very few, if any, existing studies offer such study of private tutoring seen in this light.

Furthermore, the existing literature on private tutoring focuses on examining this phenomenon on its own, rather than exploring its intertwined connection with regular school. We attempt to improve on this with an explicit investigation of this nexus. Theoretically, we (slightly) extend the standard Becker-Lewis quantity-quality tradeoff framework to provide further insights that can then guide our empirical analysis; empirically, we propose new measures that exploit both the *absolute* and *relative* differences between household investments in regular school and private tutoring. This combined approach thus provides new and original interpretations that appear not to have been attempted elsewhere.

We further make a threefold contribution with our empirical analysis. First, we improve on previous studies by providing the most comprehensive empirical investigation to date of different aspects of household investment in private tutoring for each child (i.e., at the child level). These include participation in tutoring, household monetary investment in tutoring, and time spent both in the short term (i.e., frequency of attending tutoring classes in one year) and in the long term (i.e., number of years attending tutoring classes) on tutoring. We also go one step beyond just looking at household investment in tutoring by considering the situation where households can make a joint decision on whether to enroll their children in school and to send them to tutoring classes.

Second, to identify the impacts of family size on household investment in private tutoring, we use as an instrument the distance from the household's commune to the nearest family planning center. In contrast to those used in most previous studies, this instrumental variable allows us to study the effects of family size for families with one child or more. Our results provide considerable support for the quantity-quality tradeoff in the Vietnamese context. Furthermore, the IV estimates of the impacts of family size are larger in magnitude than the uninstrumented results. These estimation results hold for several different measures of tutoring and are generally robust to different model specifications, identification strategies, and definitions of family size.

5. In this paper we focus on households' investment in their children rather than children's outcomes because doing so may provide a more direct test of the quantity-quality tradeoff hypothesis (see, for example, Caceres-Delpiano (2006) and Rosenzweig and Zhang (2009) for a similar approach). In the context of Vietnam, private tutoring as a new measure of the households' investment in the quality of their children appears more appropriate than traditional measures (such as education expenditures or private school attainment) for two reasons. First, Vietnam's education system is mostly public with more or less uniform tuition, and second, the market for private tutoring is well developed, with approximately 42 percent of children age 6–18 attending private tutoring in the past twelve months.

Finally, we explore the hypothesized child quantity-quality tradeoff in the context of rural Vietnam, a country that has undergone rapid change in fertility and educational attainment. The total fertility rate decreased steadily from 6 births per woman in the 1970s to 4 births per woman in the late 1980s and to just under 2 births per woman currently (World Bank 2014). Over the past two decades, the average number of years of schooling for the adult population has increased rapidly, from 4 in 1990 (Barro and Lee 2012) to 6.6 in 1998 and 8.1 in 2010 (VLSS 1998; VHLSS 2010).[6] The Government of Vietnam has paid much attention to family planning and has promulgated policies over the past fifty years encouraging (and in the case of government employees, requiring) families to restrict their number of children to one or two, but to our knowledge, our study is the first to investigate rigorously the quantity-quality tradeoff for this country.

Our estimation results indicate that each additional sibling reduces the rural household's investments in a child's schooling as measured through a variety of indicators: it reduces education expenditure and tutoring expenditure by 0.4 and 0.5 standard deviations, respectively; it decreases the child's probability of being enrolled in tutoring by 32 percentage points; it reduces the child's enrollment and tutoring index and tutoring attendance frequency by 0.34 and 0.49, respectively; and it cuts the average time spent on tutoring by 74 hours and 1.4 years of tutoring. With regard to the differences between tutoring and regular school, one more sibling reduces by 31 percentage points the probability of attending tutoring (unconditionally on whether the child is enrolled in school or not); reduces by D 243,000 the amount spent on education expenditure net of tutoring expenditure; and reduces by 8 percentage points and 20 percentage points, respectively, the share of tutoring expenditure in education expenditure and the share of years attending tutoring over completed years of schooling.

This paper has five sections. We provide a review of the literature in the next section, followed in section II by the data description and a description of family planning policies and the private tutoring context in Vietnam. Section III presents our theoretical and empirical framework of analysis and the instrumental variable, which is then followed by the estimation results in section IV and the conclusion in section V.

I. Empirical Literature: Testing the Quantity-Quality Tradeoff

Our paper straddles two strands of literature: the more established literature on the quantity-quality tradeoff and a smaller but growing number of studies on private tutoring. We briefly review the most relevant studies in this section.

One central and empirical challenge among the first literature, on the hypothesized quantity-quality tradeoff, is to address the endogeneity of family size

6. Unless otherwise noted, all estimates from the Vietnam Living Standards Surveys (VLSSs) and Vietnam Household Living Standards Surveys (VHLSSs) are authors' estimates.

convincingly in the data, since unobserved factors can affect both fertility and child human development outcomes. Different instrumental variables have been used and include unplanned (multiple) births (Rosenzweig and Wolpin 1980; Li, Zhang, and Zhu 2008), the gender mix of children combined with parental sex preference (Angrist and Evans 1998; Angrist, Lavy, and Schlosser 2010), and relaxation of government regulation on family size (Qian 2013). Despite these (and other) studies, the existing evidence on the quantity-quality tradeoff appears far from conclusive;[7] furthermore, while these identification strategies are useful, they cannot be applied in all contexts.

In the quantity-quality tradeoff framework proposed by Becker and Lewis (1973), a reduction in the costs of maternity care leads to changes in the relative price of quality and quantity of children and in the amount that parents choose to invest in their children. While no studies on the quantity-quality tradeoff appear to have used this insight to construct instruments, several studies in labor economics use variables related to family planning as instruments to identify the causal impacts of family size on female labor supply.[8] Instrumenting for fertility with state- and county-level indicators of abortion and family planning facilities and other variables, Klepinger, Lundberg, and Plotnick (1999) find that teenage childbearing has substantial negative effects on women's human capital and future labor market opportunities in the United States. Another US study by Bailey (2006) employs state-level variations in legislation on access to the contraceptive pill to instrument for fertility, and it also provides strong evidence for the impact of fertility on female labor force participation. More recently, Bloom et al. (2009) instrument for fertility with country-level abortion legislation in a panel of 97 countries over the period 1960–2000; they find that removing legal restrictions on abortion significantly reduces fertility and that a birth reduces a woman's labor supply by almost two years during her reproductive life.

We follow an identification strategy that is similar in spirit to that literature: we use the availability of family planning services as our instrument, which can reduce the cost of maternity care as well as the cost of controlling the quantity of children in general.[9] Specifically, in our test of the quantity-quality tradeoff

7. For example, Angrist, Lavy and Schlosser (2010) find no tradeoff in Israel; Lee (2008) finds a weak tradeoff in Korea that gets stronger with more children. In addition, conflicting results have been found for different countries including Brazil (e.g., Ponczek and Souza (2012) and Marteleto and de Souza (2012)), China (e.g., Li et al. (2008) and Qian (2013)), and Norway (Black, Devereux, and Salvanes (2005) and Black, Devereux, and Salvanes (2010)). See also Steelman et al. (2002) and Schultz (2008) for recent reviews.

8. Another thread of the quantity-quality tradeoff literature estimates the reduced-form impacts of family planning services instead (see, for example, Rosenzweig and Schultz (1985) and Joshi and Schultz (2013)). Recent studies that find that family planning-related variables have important impacts on fertility include DeGraff, Bilsborrow, and Guilkey (1997) for the Philippines, Miller (2010) for Columbia, and Portner, Beegle, and Christiaensen (2011) for Ethiopia.

9. Throughout this paper, we follow the literature by using the term "quality" of children to refer to the amount of human capital invested in them. Needless to say, this should not be taken as a value judgment about their worth as individuals. As noted earlier, however, higher human capital is associated with a host of other desirable development outcomes, at both the individual and societal levels.

hypothesis, we use the distance to the nearest family planning center at the commune level as an instrumental variable for the quantity of children.[10] Perhaps the greatest advantage of this instrument over other commonly used instruments such as twins and sibling sex composition is that the family-planning instrument allows us to analyze the impacts of family size on all of the children in the household (or the single child, if there is only one), while using either twins or children sex composition restricts analysis to a subset of these children.[11] We discuss this instrument further in section III.

Turning now to the second strand of literature, on private tutoring, few papers have investigated the correlation between household size and household educational investment in their children through private tutoring. To our knowledge, the exceptions are the two papers on Korea by Lee (2008) and Kang (2011), and the former touches only briefly on tutoring. Both of these papers share the same identification strategy, in that they use the sex of the first-born child as an instrument for family size,[12] but the former implements this analysis at the household level, while the latter does so at the level of the child. Lee (2008) finds a negative impact of larger family size on household investment in education in general and tutoring in particular, but Kang (2011) finds these negative impacts to be significant only for girls.

II. Data Description, Family Planning and Tutoring in Vietnam

Data Description

In this paper, we analyze data from three rounds (2002, 2006, and 2008) of the Vietnam Household Living Standards Surveys (VHLSSs). The VHLSSs are implemented by Vietnam's General Statistical Office (GSO) with technical assistance from the World Bank and cover around 9,200 households in approximately

10. Distance to services is often used as an instrument in the literature. For example, distance to college is used to identify the returns to education (Card 1995), distance to the tax registration office is used to identify the impact of tax registration on business profitability (McKenzie and Sakho 2010), and distance to the origins of the virus is used to estimate the response of sexual behavior to HIV prevalence rates in Africa (Oster 2012). Gibson and McKenzie (2007) provide a related review of household surveys' use of distances measured via global positioning systems (GPS).

11. Using twins as the instrument also requires a much larger estimation sample size; as a result, most previous studies that took this strategy have had to rely on population censuses.

12. The use of the sex of the first-born child as an IV has some limitations. First, it requires the assumption of son preference—which appears to be a weak IV, so that Kang (2011) has to rely on bound analysis to identify bounds of impacts of family size in the case of boys. Second, the assumption of son preference in turn requires the assumption that parents do not abort girls at their first childbearing; if they do, the sex of the first-born child is clearly not valid as an exogenous instrument. This concern is especially relevant to Vietnam, which has one of the highest abortion rates in the world (Henshaw, Singh, and Haas 1999). And finally, this identification approach may only work for families with more than one child; our study makes no such restriction on family size, investigating families with between one and seven children.

3,000 communes across the country in each round.[13] The surveys provide detailed information on household demographics, consumption, and education. The surveys also collect data on community infrastructure and facilities such as distances to schools or family planning facilities. Since 2002, the VHLSSs have been implemented biannually and have collected more data for rotating themes for each survey round; for example, the 2006 round focused on educational activities and tutoring. These surveys are widely used for education analysis by the government and the donor community in Vietnam.

Since only the 2002 round collected data on the distance to family planning for rural communes, we restrict our analysis to rural households in Vietnam. The VHLSSs' commune sample frame remains almost the same during the period 2002–08, which allows us to match the commune information from the 2002 survey round to most of the households in the 2006 and 2008 survey rounds.[14] However, we focus on the 2006 round of the VHLSSs for the outcome variables, since this round has the most detailed information on household investment in tutoring activities. We also supplement our analysis with data from another nationally representative survey (VHTS) focused on private tutoring that we fielded in 2008,[15] as well as data on teacher qualifications in the community from the primary school census (DFA) database.[16]

Since most children start their first grade at six years old, we restrict our analysis to children who are between six and eighteen years old.[17] To address concerns about grown-up children that have already moved away from home, we consider only children who are living at home and households where the total number of children born of the same mother is equal to the number of children living in the household. We define family size as consisting of children born of the same mother, but we also experiment with a more relaxed definition of family size that

13. A commune in Vietnam is roughly equal to a town and is the third administratively largest level (i.e., below the province and district levels) and higher than the village level. There are approximately 9,100 communes in the country (GSO 2012). The respondents for the community module of the VHLSSs are mostly the (deputy) head of the commune.

14. This matching process is complicated by the fact that there were administrative changes resulting in changes to administrative commune codes between 2002, 2006, and 2008. For around 150 communes, we have to rely on both commune and district names (in addition to province and district codes) for matching. We can match 96 percent of all of the communes in 2002 to those in 2006 and 2008 (i.e., we can match 2,808 communes out of 2,933 communes in 2002).

15. For details on this survey, see Dang and Glewwe (2009). We collaborated on designing the survey with other researchers, including Paul Glewwe (University of Minnesota), Seema Jayachandran (Northwestern University), and Jeffrey Waite (World Bank). The survey was administered by Vietnam's Government Statistics Office, using funding from the World Bank's Research Support Budget and the Hewlett Foundation.

16. This database is initiated and maintained by World Bank-supported projects. For a brief description on the history and objectives for the primary school census database, see Attfield and Vu (2013).

17. We also experimented with other age ranges such as ages 10–18 and 12–18. Estimation results (available upon request) are qualitatively very similar and even more statistically significant than those for the age range 6–18.

considers all children living together in the households, as well as other stricter definitions to be discussed later.

Overview of Family Planning in Vietnam[18]

Vietnam's family planning policy dates back to 1961 in the North of Vietnam, but it initially had limited success. Following the unification of Vietnam in 1975, policymakers responded to the faster growth of the population than the economy by setting a goal of lowering population growth rates to less than 2 percent. Subsequently, in 1988 the government adopted a policy restricting families to one to two children, which has largely remained in effect until now. The highlights of this policy include the universal and free provision of contraceptives and abortion services, incentives for families, and strict penalties for families with more than two children. Vietnam's approach to family planning policy closely follows that of one-child-per-family in China, but it is administered less rigorously (Goodkind 1995). This lack of rigor contributes to our analysis of the quantity-quality tradeoff, in fact, by expanding the range of variation of family size.[19]

An important administrative landmark for family planning—and one that is quite relevant to the discussion below of our instrument's validity—was the establishment of the ministry-level National Council of Population and Family Planning (NCPFP) in 1984. By the late 1980s, the NCPFP had established administrative offices and staff down to the commune level to ensure that their activities reached the whole population. Together with the official administrative apparatus, the NCPFP also built up a wide-reaching network of family planning volunteers, both at the village level and in most government agencies, to promote family planning policies.[20]

Background on Tutoring in Vietnam

The current education system in Vietnam has three levels: primary (grades one to five), secondary (grades six to nine for lower secondary sublevel and grades ten to twelve for upper secondary sublevel), and tertiary (post-secondary). Almost all schools in rural Vietnam are public schools and provided by the government. Vietnam has almost achieved universal primary education with 94 percent of Vietnamese children age 15–19 having completed primary education (VHLSS 2006). High-stakes examinations are widely used in the education system for

18. This section is mostly based on GDPFP (2011). See also Vu (1994) for discussion of family planning policies in earlier periods.

19. The family size penalties include fines, restrictions on promotion (or even demotions) for government employees, and denial of urban registration status. We attempted in an earlier draft to use households' exposure to the two-child-per-family policy as an instrument since the strictness with which it is applied varies with certain characteristics that can be largely exogenous to the family. However, it turned out that the policy was not implemented rigorously enough to make it a viable instrument.

20. In 2007, the NCPFP was merged into the Ministry of Health and renamed the General Department of Population and Family Planning (GDPFP).

TABLE 1. Reasons for Attending Private Tutoring Classes for Students Age 9–20 (Percent), Vietnam 2007

	Tutoring organized by school	Tutoring not organized by school
Prepare for examinations	47.2	41.7
Do not catch up with the class	12.9	14.4
Acquire skills for future employment	12.2	12.7
Like this subject	6.4	11.3
Parents too busy to take care	2.7*	1.6*
Poor quality lessons in school	2.7*	6.0*
Subjects not taught in mainstream classes	0.5*	1.5*
Others	15.4	10.9
Total	100	100
N	376	301

Note: *Fewer than 20 observations.

Source: Authors' analysis based on data from Vietnam Household and Tutoring Survey 2007–08.

performance evaluation, and performance on the exams determines whether students can obtain secondary-school degrees and gain admission to colleges/universities. The strict rationing at the tertiary level results in strong competition among high school students, which helps fuel the demand for private tutoring.

Private tutoring is such a major feature of the Vietnamese educational landscape that it is hotly debated, both in the media and during the Minister of Education's presentations to the National Assembly. Policymakers, educators, and parents fall into two main opinion camps—one arguing that private tutoring worsens educational outcomes and harms children, and the other that tutoring can improve the quality of education. The former group calls for a total ban on private tutoring, while the latter supports the (controlled) development of tutoring.[21]

Table 1 lists the reasons that students take private tutoring classes, according to data from the VHTS. Tutoring classes are divided into two categories: tutoring classes organized by the student's own school, and other tutoring classes. Across the two types of tutoring, the most important reason for taking tutoring is to prepare for examinations, which accounts for almost half of all responses (42–47 percent). Other commonly cited reasons given include to catch up with the class (13–14 percent), to acquire better skills for future employment (13 percent), and to pursue a subject that the student enjoys (6–11 percent). Other reasons, such as to get childcare, to compensate for poor-quality lessons in school, or to study subjects not taught in mainstream classes, account for a smaller proportion of all responses (1–6 percent each). The preeminence of exam preparation over other

21. See also Dang (2011, 2013) for more detailed discussions of the private tutoring phenomenon in Vietnam.

TABLE 2. Household Expenditure on Private Tutoring Classes by Consumption Quintiles, Vietnam 2006

	Poorest	Quintile 2	Quintile 3	Quintile 4	Richest	All Vietnam
Average household expenditure on tutoring in 2006 (D '000)	54.2	126.4	222.8	325.0	814.3	321.3
Distribution of household with exp. on private tutoring as percent of total expenditure in 2006						
0%	78.8	61.8	55.1	56.3	52.6	60.4
1%–5%	20.0	36.4	41.6	38.7	38.9	35.6
5%–10%	1.0*	1.5*	3.0	4.4	7.0	3.5
10% or higher	0.1*	0.3*	0.2*	0.6*	1.6*	0.6
Total	100	100	100	100	100	100
No. of households	1,278	1,269	1,263	1,290	1,198	6,298

Note: *Fewer than 20 observations.

Source: Authors' analysis based on data from Vietnam Household Living Standards Survey 2006.

reasons for taking tutoring classes reflects the importance of examinations in the school system in Vietnam.[22]

Richer households in Vietnam spend more on tutoring classes than do poorer households, as shown in table 2. Currently about 40 percent (=100−60.4) of households in Vietnam send their children to private lessons, and the majority of them (90 percent) spend between 1 percent and 5 percent of household expenditure on tutoring classes. The percentage of households with positive expenditures on tutoring classes is only 21 percent in the poorest (1st) consumption quintile but nearly doubles to 38 percent in the next richer quintile (2nd) and hovers around 35 percent in the top three quintiles (3rd to 5th). In terms of actual expenditure, the mean expenditure on tutoring classes by the wealthiest 20 percent of households is fifteen times higher than expenditure by the poorest 20 percent of households. And more expenditure on tutoring is found to increase student grade point average (GPA) ranking in Vietnam, with a larger influence for lower secondary students (Dang 2007, 2008).

Our calculation (not shown) using the 2006 VHLSS shows that the majority of children age 6–18 have at most three siblings, with 10 percent having no sibling, 48 percent having one sibling, 27 percent having two siblings, and 10 percent having three siblings; only five percent of these children have four siblings or more. Table 3 provides a first look at children age 6–18 that are currently enrolled in school that comprise our estimation sample, of whom 42 percent attended private tutoring in the past twelve months. They spent on average

22. For examining our hypothesis of the quantity-quality tradeoff, we are in fact assuming that sending children to tutoring classes are completely determined by parents. If corrupt teachers force tutoring on their own students beyond parental control (see, e.g., Bray 2009; Jayachandran 2014), household investment in tutoring would not provide valid evidence for this tradeoff. However, the results in table 1 suggest this concern is a minor one in the context of Vietnam.

TABLE 3. Summary Statistics for Children age 6–18, Vietnam 2006

Variable	Obs.	Mean	Std. Dev.	Min	Max
Enrollment in past 12 months	5012	0.87	0.33	0	1
Total education expenditure in past 12 months (D'000)	4248	583.83	745.71	0	20165
Completed years of schooling	5012	5.80	3.25	0	12
Private tutoring attendance in past 12 months	4125	0.42	0.49	0	1
Enrollment and private tutoring attendance in past 12 months (0 = not enrolled in school, 1 = enrolled in school but have no tutoring, 2 = enrolled in school and have tutoring)	5012	1.22	0.65	0	2
Expenditure on private tutoring in past 12 months (D'000)	4125	104.15	465.35	0	18000
Expenditure on private tutoring in past 12 months for those attending private tutoring (D'000)	1614	246.59	691.19	6	18000
Number of hours spent on private tutoring in past 12 months	4247	89.06	158.71	0	1728
Number of hours spent on private tutoring in past 12 months for those attending private tutoring	1624	215.43	183.61	2	1728
Tutoring attendance frequency (0 = no tutoring, 1 = tutoring either during school year or holidays/ break, 2 = tutoring during both school year and holidays/ break)	4248	0.65	0.77	0	2
Years attending private tutoring to date	4248	1.90	2.58	0	13
Number of siblings age 0–18	4248	1.58	1.04	0	7
Distance to family planning center	4248	8.56	9.78	0	80.5
Age	4248	11.90	3.20	6	18
Male	4248	0.50	0.50	0	1
Years before last grade in current school level	4248	1.67	1.23	0	4
Secondary school	4248	0.58	0.49	0	1
Mother age	4248	37.38	6.00	21	68
Female-headed household	4248	0.12	0.32	0	1
Head's years of schooling	4248	7.36	3.39	0	16
Ethnic majority group	4248	0.83	0.37	0	1
Total household expenditures	4248	19222	10209	2145	175393
Distance to primary school	4248	0.82	1.25	0	10
Distance to secondary school	4248	2.78	2.81	0	25
North East and West region	4248	0.16	0.37	0	1
North Central region	4248	0.19	0.39	0	1
South Central region	4248	0.09	0.29	0	1
Central Highlands region	4248	0.06	0.24	0	1
South East region	4248	0.09	0.29	0	1
Mekong River Delta region	4248	0.16	0.37	0	1

Note: All numbers are weighted using population weights.

Source: Authors' analysis based on data from Vietnam Household Living Standards Survey 2006.

D 104,150 (equivalent to $US 6)[23] and eighty-nine hours on these tutoring classes also in the past twelve months, and had attended tutoring for 1.9 years; for those that attended tutoring in the past twelve months, the corresponding

23. The exchange rate was D 15,994 for $US 1 in 2006 (World Bank 2014).

expenditure and hours spent on tutoring are D 246,590 and 215 hours. Most tutoring attendees (80 percent) take these classes organized by their school (VHLSS 2006).[24] Table 3 also shows that the children in our estimation sample have 1.6 siblings on average, are mostly in secondary school (58 percent), and live an average of 8.6 kilometers away from the nearest family planning center.

III. Framework of Analysis

Family Size, Private Tutoring, and Regular school

We present a simple theoretical model that builds on the standard quantity-quality tradeoff framework (Becker and Lewis 1973) for interpreting the interwoven connection between private tutoring and regular school. We note three main specific features with private tutoring, which provide the underlying assumptions behind our model. First, the existence of private tutoring depends on the mainstream education system and it does not stand alone as an independent educational activity;[25] second, it can offer lessons that are often much more flexible and informal than regular school; and third, compared to the public-subsidized regular school, private tutoring is more costly for the average household.

The household maximizes its utility function U(n, q, y)

$$\max U(n, q, y) \tag{1}$$

subject to its budget constraint

$$y + n(p_u e_u + p_r e_r) = I \tag{2}$$

where n is the number of children, q is their quality, y is the other (numeraire) good with its price set to 1, p_k is the price of household investment in (or expenditure on) their children's quality, for k = u or r, and I is household income. A child's quality is assumed to be equivalent to the total amount of public education (e_u) and private tutoring (e_r) that the household invests in the child:

$$q = e_u + e_r \tag{3}$$

We also assume further that regardless of consumer demand, there is a limit ($\bar{e}_u$) on the capacity of public schools to provide the quality of education desired by the household.[26]

24. See also table S1.1 in the online appendix for a breakdown of tutoring prevalence and expenditure by urban/ rural areas.

25. This supplementary aspect of private tutoring helps explain why it has been referred to as "shadow education" (Bray 2009) or "supplementary education" (Aurini et al. 2013).

26. Particularly in developing countries, the public education system is well known for its rigidity, lack of teacher incentives and accountability, and inadequate infrastructure (see Glewwe and Kremer (2006) for a recent review). In our model, this inelasticity of supply should hold at least in the short run.

$$e_u \leq \bar{e}_u \tag{4}$$

Examples of this limit can be the inability of public schools to provide more than, say, the basic reading skills in primary grades or a fixed number of hours of instruction, given short-run constraints on resources and capacities. We then make the standard assumptions that the number of children and the goods are nonnegative—that is $n \geq 0,\ q \geq 0,\ y \geq 0$. Our model extends the standard quantity-quality framework by introducing household tutoring consumption into the household utility function (1), the budget constraint (2), and the limit on public education consumption. Without these extensions (i.e., with $e_r = 0$ and $e_u \leq \infty$), the standard Becker-Lewis model results.

Assuming the marginal utilities of income (λ_1) is positive, the Kuhn-Tucker conditions for maximizing the utility function subject to the child quality function, the budget constraint, and the public education constraint yield the following results:

$$U_n - \lambda_1(p_u e_u + p_r e_r) = 0 \tag{5}$$

$$Ue_u - \lambda_1 n p_u - \lambda_2 = 0 \tag{6}$$

$$Ue_r - \lambda_1 n p_r = 0 \tag{7}$$

$$U_y - \lambda_1 = 0 \tag{8}$$

$$I - y - n(p_u e_u + p_r e_r) = 0 \tag{9}$$

$$\lambda_2(\bar{e}_u - e_u) = 0 \tag{10}$$

Equations (5) to (9) thus yield the same result as under the standard Becker-Lewis model: the shadow prices of the quality of children for either public education (np_u) or private tutoring (np_r) are proportional to the quantity of children; or, put differently, an increase in quality is more expensive if there are more children. Under this standard model, a reduction in quantity-related costs such as contraception costs would increase the shadow prices of quantity relative to quality and other goods, leading to smaller household size and better-quality children.

Furthermore, the different values of the marginal utility of relaxing the public education constraint (λ_2) offer the following results:

(i) If $\lambda_2 = 0$, then the typical household does not consume the maximum available quality of public education (i.e., $e_u < \bar{e}_u$). However, this case is likely to be the exception rather than the norm, since a Vietnamese child

that is currently in school typically has more than a 40 percent chance of attending private tutoring in the past year (table 3) and around half of these children resort to private tutoring besides their regular classes to better prepare for examinations (table 1).

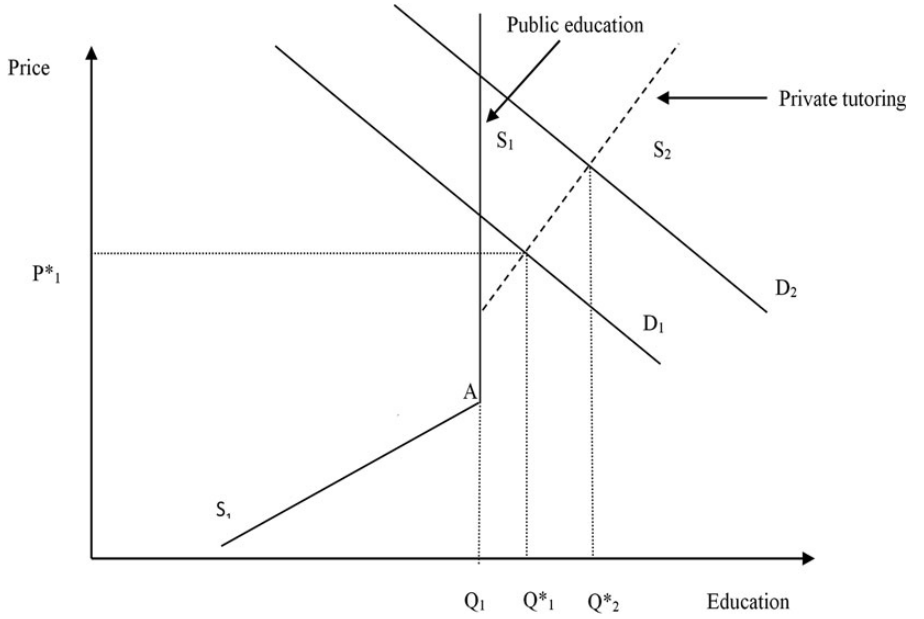
(ii) If $\lambda_2 > 0$, then the household consumes the maximum available quality of public education (i.e., $e_u = \bar{e}_u$), which has several important implications. First, to improve the quality of its children, the household's only option is to invest in tutoring; equivalently, since $e_u = \bar{e}_u$, private tutoring is the only choice variable for maximizing the household's utility function.[27] Second, when coupled with the standard result of quantity-quality tradeoff, this result leads to household demand for private tutoring that is more elastic to household size than the household's demand for public education is. The model can thus better capture the tradeoff of household investment in their children's education. In other words, our model indicates that households would cut down on tutoring consumption and increasingly shift their education expenses to the public subsidies as their family size grows. Finally, since private tutoring is more costly than regular education, relaxing the capacity constraint of public education—for example by providing more teacher time with students—can help reduce the demand for tutoring. This result comes from equation (9) where, given a fixed budget constraint, increasing e_u ($=\bar{e}_u$) would *ceteris paribus* result in a lower value of e_r. Analogously, for a better and fuller picture on the quantity-quality tradeoff, household investment in private tutoring should be examined together with investment in the regular school.

Figure 1 provides a graphical illustration for a typical household in case (ii) discussed above. The supply of education is represented by the supply curves S_1 (solid line) for public education and S_2 (dashed line) for private tutoring. The gradient of S_2 is flatter than the vertical segment of S_1 but steeper than the upward-sloping segment of S_1; these relationships represent, respectively, the fact that private tutoring can fill in the demand for education where the public education system cannot and that private tutoring is more expensive than public schooling. Since private tutoring is prevalent in Vietnam (as shown with tables 1 to 3), the average household would consume the maximum available quality of public education and also some private tutoring. Household demand for tutoring can be represented by a demand curve that lies higher and to the right of point A and that cuts across both the public education supply S_1 and private tutoring supply S_2.[28]

27. This result can generally apply to contexts where the household has no other choice besides public education, and already consumes the maximum available quality of public education. In such cases, household investment in public education would not respond to changes in family size.

28. For case (ii), households consume the maximal available quality of public education (Q1), and therefore we do not show the demand curve for public education in Figure 1.

FIGURE 1. Demand and Supply of Education with Private Tutoring



Source: Illustrations based on the theoretical model discussed in the text.

This graphical model helps illustrate our theoretical results. First, other things equal, since public education supply is inelastic after point A, family size would have little or no impact on the household's consumption of public education; consequently, household investment in private tutoring is a better measure of household quantity-quality tradeoff. Second, compared to a representative household with the demand curve D_1, the demand curve D_2 represents another household that is assumed to have stronger education preferences, which can be represented by a smaller family size according to our theoretical model.[29] Thus, the household with smaller family size would consume more private tutoring (Q_2^*) than the household with larger family size (Q_1^*). Finally, focusing on investigating private tutoring on its own rather than examining its intertwined relationship with regular school is equivalent to studying the dashed line S_2 in Figure 1 alone without taking into consideration its connection with the solid line S_1. This can result in an incomplete—or even potentially misleading—picture of private tutoring.

29. Other factors that shift the demand curve include household income, the price of substitute goods or the number of buyers on the market, or expectations about future returns to education.

These findings offer new interpretations of private tutoring as a new measure of household education investment.[30] We will validate these theoretical predictions empirically in later sections, after first discussing the empirical framework and the instrument.

Empirical Framework

Our basic estimation equations are for child j, j = 1,..,J in household I, i = 1,..,N

$$E_{ij} = \alpha + \beta FamSize_i + \gamma X_{ij} + \varepsilon_{ij} \tag{11}$$

$$FamSize_i = \delta + \lambda DisFam + \phi X_{ij} + \eta_{ij}, \tag{12}$$

where, for the first equation, the dependent variable E_{ij} includes household education investment. The traditional measures for E_{ij} include school enrollment, educational expenditure, and completed years of schooling.[31] The new measures include private tutoring attendance, a combined school enrollment/tutoring index (which takes a value of 2 if enrolled in both school and tutoring, 1 if school only, and 0 if neither), frequency of tutoring attendance (which takes a value of 2 if enrolled in tutoring during both school year and holidays, 1 if either school year or holidays, and 0 if neither), expenditure on tutoring,[32] and the number of hours in the past year and the number of years to date spent on tutoring. Of these measures, only tutoring attendance and expenditure appear to have been used in previous studies on tutoring.

If some parents decide to choose fewer children and greater investment in each child, a smaller family size will be strongly correlated with unobserved parental devotion to their children, thus biasing estimates upward; however, the opposite holds if parents decide to choose both more children and greater investment in them at the same time. Thus, estimating equation (11) alone would provide biased estimates of the relationship between family size and household investment. The direction of bias appears to be an empirical issue and depends on parental

30. Some further extensions can be added to our theoretical model. For example, we can generalize by assuming a child endowment component in equation (3) as in Becker and Tomes (1976), or another extension is to assume that, instead of prices being fixed, the price of tutoring is a function of the price of regular school. These extensions, however, do not change the main results. Another extension is to assume that e_u and e_r are multiplicative up to $\bar{e}_u$ (the constraint on public education), and are additive beyond this value. This would correspond to private tutoring being complementary up to this value, and being substitute after this value. The latter case, however, appears to be the dominant case in Vietnam as discussed above.

31. For children that are currently in school, completed years of schooling is right-censored since we do not observe the final years of schooling for these children. Thus for such children (and our estimation sample), this variable represents a lower-bound estimate only.

32. For easier interpretation of results and because of the large number of zero observations, in our preferred specification we do not transform variables such as expenditures and hours spent on tutoring to logarithmic scale. Estimation results with the transformed variables are similar, however, and coefficients are slightly more statistically significant.

heterogeneity of preference; the IV model would help remove this bias and uncover the true impacts of family size on household investment. Thus, we jointly estimate equations (11) and (12) in an IV model using the commune-level distance to the nearest family planning center (*DisFam*) as the instrumental variable.

X_{ij} is a vector of child, household, community and school characteristics that include age, gender, school level, mother's age,[33] mother's age squared, gender of the household head, head's years of schooling, ethnicity, household expenditure, and distances to the nearest primary and secondary schools. A variable indicating the number of years that remains before the last grade in the current school level is also added, since this variable can capture the increasing intensity of tutoring investment as students progress through school (Dang 2007), but this variable is left out in the regression for the enrollment/tutoring index since it applies only to children currently enrolled in school.

For easier interpretation of results, we jointly estimate equations (11) and (12) for all the outcomes above using a 2SLS model, except for expenditure and hours spent on tutoring, where we use an IV-Tobit model instead and subsequently provide separate estimates for the marginal effects since a large number of children have zero values for these variables.[34] Let E_{ij}^* be the latent variable that represents the household's potential spending (or hours) on tutoring, the Tobit model for equation (11) has the form

$$E_{ij}^* = \alpha + \beta FamSize_i + \gamma X_{ij} + \varepsilon_{ij}, \tag{13}$$

where the relationship of the actual (E_{ij}) and latent (E_{ij}^*) spending on tutoring is given by $E_{ij} = 0$ if $E_{ij}^* \leq 0$ and $E_{ij} = E_{ij}^*$ if $E_{ij}^* > 0$.

Similarly, we can examine the marginal impacts of family size (or other explanatory variables) on either households' propensity to spend or households' actual (observed) spending on tutoring classes. While the former interpretation (shown in table 5) may be more relevant for forecasting the future, the latter (shown in table S1.3 in the online appendix S1, available at http://wber.oxfordjournals.org/) is more commonly used and focuses on household spending at present.[35] For our purposes, we will use the latter interpretation of the marginal effects.

33. There are more missing observations with father's age so we omit this variable.

34. While the number of years of tutoring can also be fitted in a Tobit model, we prefer to use the OLS model for better interpretation. Estimation results using an IV-Tobit provide very similar results.

35. The marginal impacts for household propensity to spend can be calculated as $\frac{\partial E(E_{ij}^* | FamSize_i, Z_{ij})}{\partial FamSize_i} = \beta$, and the marginal impacts for household actual spending can be calculated as $\frac{\partial E(E_{ij} | FamSize_i, Z_{ij})}{\partial FamSize_i} = \beta\Phi\left(\frac{\alpha + \beta FamSize_i + \gamma Z_{ij}}{\sigma}\right)$, where we also assume $\varepsilon_{ij} \sim N(0, \sigma^2)$ as in the OLS models. See, for example, Greene (2012) for more discussion on the marginal effects with the Tobit model.

Distance to Family Planning Center as Instrument

Our instrumental variable for family size is the distance to the nearest family planning center since it meets the exogeneity, relevance, and exclusion restriction conditions. In this section, we consider these three criteria in turn.

A major exogeneity-related concern with using public programs, including placement of family planning centers, as instruments is that these programs may have been established in response to local demand (Rosenzweig and Wolpin 1986). The evidence suggests, however, that such demand response is not an issue in Vietnam, where family planning services were already offered at the commune level and reached virtually the whole population by the late 1980s (Goodkind 1995; GDPFP 2011). While little data exist on the local conditions when family planning centers were set up, it is generally the case with most policy implementation in Vietnam that the central government sets the national policies but it is the local governments that ultimately decide exactly how these policies will be implemented.[36]

Indeed, the provincial governments were observed to be responsible for all work related to family planning and for mothers and children's health in general (Vu 1994), which should include the establishment of family planning centers. This is corroborated by an analysis of a survey of local governments' family planning efforts in fifteen provinces across Vietnam by San et al. (1999), which finds that effort strength is mostly driven by the quality of local governments' leadership and implementation ability, rather than local conditions such as geographical terrain or the level of economic development.[37]

Still, some variation of the location (and timing) of family planning center may stem from differences in local governments' resources: communes with more resources might have been more likely to build a family planning center earlier. We argue, however, that once this channel is controlled for in the regressions (as proxied for by commune infrastructure in several model specifications we examine later), the location of the family planning center is exogenous to each household's decision on number of children. While it is impossible to test directly for the instrument's exogeneity, we use a three-pronged approach as an extra precaution to ensure its validity.

First, we use the distance to family planning centers in 2002 to instrument for the impacts of family size on household investments in education four years later, in 2006. This approach can help reduce any contemporaneous correlation between the former and the latter.

Second, in one of the robustness checks, we will restrict our analysis to a subsample of cases in which the family planning centers had already been established

36. Scornet (2001) observes that local governments' strong autonomy in implementing family planning policies takes its root in the traditional decentralization of monarchical governments in the past. Kaufman et al. (1992) note that the local governments in China—which had a similar although stricter regulation on family size—were similarly responsible for setting up family planning clinics.

37. San et al. (1999) also provide some evidence that their selected 15 provinces share many characteristics of the overall functioning of the national family planning program.

earlier. If family planning centers were more likely to be established first in locations with stronger demand for family planning, older centers would be more effective in reducing family size and would consequently allow households to increase investment in their children's education. Thus an analysis showing similar impacts of family size for the sample with older centers compared to those for the overall sample would provide evidence for the instrument's exogeneity.[38]

Finally, if it were true that family planning centers were more likely to be first established in locations where households have larger family size, assuming a negative relationship between family size and household investment in their children, we would expect this endogenous placement of these centers to weaken the impacts of the instrument and thus bias estimates upward toward zero. Thus, our estimation results would provide conservative estimates of the extent of the tradeoff.[39]

In terms of the relevance criterion for the instrument, our review of the literature from other countries suggests that access to family planning facilities is highly relevant to household decisions on family size. Previous studies for Vietnam using data from the 1997 Demographic and Health Survey offer similar findings that increased access to family planning services increases contraceptive use (Thang and Anh 2002; Thang and Huong 2003) and reduces unintended pregnancy (Le et al. 2004). Our first-stage estimates turn out to show a consistently strong and negative impact of the distance to family planning center on family size.

For the exclusion restriction, there may be concerns that family planning centers directly affect the investment in children by explicitly promoting the idea of a quantity-quality tradeoff. But given the uniform presence in every commune of family planning workers (GDPFP 2011) who can provide interested households with detailed information on the benefits of family planning, family planning centers mostly serve as facilities that provide options for restricting family size to the desired number of children.[40] These centers focus on services related to providing contraceptives—such as insertion of intrauterine devices (IUDs),

38. This check does not hold in the opposite direction since older centers may also be effective through other channels that are uncorrelated with endogeneity of location (e.g., longer existence simply increases the chances families know about and use the services at these centers). Larger impacts for family size in the sample of older centers thus would not necessarily indicate violation of exogeneity.

39. An additional concern related to exogeneity is that families could have immigrated to their current commune, meaning that they were not necessarily constrained by the current distance to family planning center when making their decision on giving birth. However, this concern does not apply in our context: we restrict our analysis to rural families only, and fewer than 3 percent of the total population over five years of age move within or to rural areas in Vietnam between 1994 and 1999 (Dang, Tacoli, and Hoang 2003).

40. A reviewer pointed out that family planning centers' services may also possibly operate through family planning workers/volunteers. However, since these workers were already present in all the communes by 2001 (and most of the communes well before that in the late 1980s), any additional impacts brought about by the new workers that are associated with these centers are likely to be small. This is consistent with Do and Koenig (2007)'s finding that family planning outreach programs (including visits by family planning workers) do not have statistically significant impact on women's continued use of contraceptive methods. Other programs such as communications campaigns or economic incentives were most often employed by the government through channels (e.g., administrative measures as discussed earlier) that are not typically associated with the activities of family planning centers.

provision of condoms and oral contraceptives, menstrual regulation, and advice on family planning—as well as birth-related medical services and abortions (MOH 2001). In 2002, around one third of the population lived in communes that were within one kilometer of such a center. Thus, access to family planning facilities should affect the educational outcomes of interest only through family size, which satisfies the exclusion restriction.

Another possible objection to the validity of the exclusion restriction is that the distance to the nearest family planning center may be correlated with unobserved commune characteristics that also affect household investment in their children. For example, more remote, less developed communes may also be farther away from any family planning center. In such cases, any negative impacts of household sizes on the outcome variables as instrumented by availability of family planning might be caused by the negative correlation between the general development level of the commune and these outcomes (e.g., poorer communes may spend less on their children's tutoring classes).

We use two strategies to address this concern. The first is to consider a number of different specifications that test for the strength of this instrument as different commune characteristics are included in the regressions. If the instrument becomes weaker or loses its statistical significance, this means that it is strongly correlated with these commune characteristics (or other unobserved characteristics proxied for by these variables) and concerns about the exclusion restriction are justified. Our second strategy is to use an alternative identification that relies on the heteroskedasticity of the error terms (Lewbel 2012) rather than a regular instrumental variable.[41] Heteroskedasticity-based identification has been used for some time (see, e.g., Klein and Vella 2010). In particular, the Lewbel identification approach has been applied in various settings to examine the impacts of body weight on academic performance (Sabia 2007) or the effects of access to domestic and international markets on household consumption (Emran and Hou 2013). Due to its reliance on higher moments, this identification strategy is less reliable than the standard IV approach, but it can provide a qualitative robustness check on our estimation results.

We show estimation results for the first strategy in table 4, which tests for the strength of this instrument using several different specifications sequentially. (Full estimation results are shown in table S1.2 in the online appendix S1.) Model 1, the most basic model, includes only the instrument and the regional dummy variables. Model 2 adds the children's characteristics and their household characteristics, while model 3 adds to model 2 the distances to the nearest

41. Our standard IV identification strategy comes from the exclusion restriction that the coefficient on the distance to family planning center be zero in equation (1). However, Lewbel (2012) shows that, given the standard regularity condition on the data, we do not need to use this restriction for identification if the error terms are uncorrelated with the right-hand side variables and we can find a variable (or vector of variables) Z that is uncorrelated with the product of the two error terms, that is, $\text{cov}(Z, \varepsilon_{ij}\eta_{ij}) = 0$. In other words, we can use $(Z - \bar{Z})\eta_{ij}$ as the instrument for family size in equation (1), where the distance to family planning center is Z.

TABLE 4. Impacts of Distance to Family Planning Center on Number of Siblings Age 6–18, Vietnam 2006 (First-Stage Regressions)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Distance to family planning center	0.009*** (3.60)	0.007*** (2.95)	0.007*** (2.87)	0.007*** (2.85)	0.007*** (2.86)	0.006*** (2.58)	0.006*** (2.61)
Additional control variables							
Regional dummy variables	Y	Y	Y	Y	Y	Y	Y
Individual & household characteristics		Y	Y	Y	Y	Y	Y
Distances to school			Y				Y
Community infrastructure				Y			Y
Distance to health facilities					Y		Y
Share of commune population working in agriculture						Y	Y
R2	0.12	0.25	0.23	0.23	0.23	0.23	0.24
N	6309	5413	4248	4178	4294	4294	4168

Notes: *p< .1, **p<0.05, ***p<0.01; robust t statistics in parentheses account for clustering at the household level. All regressions control for regional dummy variables, which include the following regions: Northeast and Northwest, North Central, South Central Coast, Central Highlands, South East, and Mekong River Delta. The reference category is the Red River Delta. All household expenditures are in million Vietnamese dong.

Source: Authors' analysis based on data from Vietnam Household Living Standards Surveys 2002 and 2006.

primary and secondary schools and includes the variables we use for the subsequent second-stage regressions. Model 4 then adds to model 3 basic commune characteristics such as distances to the nearest paved road, public transportation, and the post office, which are expected to proxy for the general level of economic development of the commune.

Next, to net out any effects that access to community health care has on family size (for example, inadequate health care may reduce family size through high child mortality rates), model 5 adds to model 3 the distance to the nearest health facilities.

Given the low-technology production techniques typically used in agriculture, rural farming households in Vietnam have had to rely for the most part on manpower for their farm work, giving them an incentive to have more children. Furthermore, government employees may be subject to a stricter enforcement of the one-to-two children rule than are farming households. Thus, in the IV terminology (see, e.g., Angrist and Pischke (2009)), farming households may be the population subgroup that is affected differently by the distance to the family planning center than other population subgroups.

To address this issue, in model 6 we add to model 3 a variable indicating the share of the commune population working in agriculture. If this addition changes significantly the estimated coefficient on the instrument, this result would suggest that the estimated impact of the distance to the family planning center on family size in model 3 is influenced by the farming-oriented occupation structure in the commune rather than the costs of family planning. Finally, model 7 includes all the variables from models 1 through 6.

The results in table 4 show that the distance to family planning center has a positive and strongly statistically significant impact on family size, as expected.[42] Importantly, except in the case of model 1 (which is clearly too simplistic), the magnitude of the estimated coefficient on the distance to the family planning center is almost identical in all the models at around 0.007; this magnitude indicates that a child living 10 kilometers further away from a family planning center would have 0.07 more siblings on average. The consistency of the point estimates suggests both the strong relevance and robustness of this instrument. Since most of the additional variables in models 4 to 6 are statistically insignificant, to keep our models parsimonious, we will use the variables in model 3 in subsequent regressions. In a later section on robustness analysis, we explore different specifications to further assess the validity of this instrument.

IV. Estimation Results

We investigate the impacts of family size on private tutoring alone in the next section, before turning to examining these impacts in the intertwined relationship with regular school.

Impacts of Family Size on Household Education Investment

Table 5 provides the instrumented regressions of the impacts of family size on household education investment; the uninstrumented coefficients on family size are also provided at the bottom of this table for comparison. The instrumented regressions shown in table 5 indicate that a quantity-quality tradeoff exists in Vietnam: all of the instrumented estimated coefficients on family size have a negative sign (as do all the uninstrumented estimated coefficients). While the instrumented coefficients on family size are not statistically significant for school enrollment and completed years of schooling, we can use the point estimates for a rough comparison with the results of previous studies. For example, the ratios for the instrumented coefficient over the uninstrumented coefficient in the regression for these variables (specifications 1 and 3) are around two and fall within a range of corresponding estimates by Li et al. (2008) and Qian (2013) for China;

42. The t-statistics for model 3 are equivalent to an F-statistic of 8.6, which is slightly below the value of 8.96 for a strong IV suggested by Stock and Yogo (2005). Note, however, that Stock and Yogo's critical values rely on the assumption of independently and identically distributed (iid) errors, whereas our F-statistic is obtained from a cluster-robust regression that is robust to heteroskedastic errors. Without this cluster-robust option, the F-statistic for model 3 is much higher at 22.6.

TABLE 5. Impacts of Family Size on Educational Investment for Children Age 6–18, Vietnam 2006

	Spec. 1	Spec. 2	Spec. 3	Spec. 4	Spec. 5	Spec. 6	Spec. 7	Spec. 8	Spec. 9
Instrumented Regressions	Enrollment	Total education expenditure	Completed years of schooling	Tutoring attendance	Enrollment & Tutoring attendance	Tutoring attendance frequency	Tutoring expenditure	Tutoring hours	Years attending tutoring
Number of siblings age 0–18	−0.072	−0.308**	−0.589	−0.318**	−0.337**	−0.488**	−573.957*	−188.425	−1.424**
	(−1.04)	(−2.02)	(−1.50)	(−2.17)	(−2.27)	(−2.45)	(−1.94)	(−1.51)	(−2.34)
Age	−0.033***	0.118***	0.783***	0.010**	−0.026***	0.017**	28.066***	9.061***	0.245***
	(−16.42)	(13.76)	(75.24)	(1.98)	(−7.10)	(2.32)	(3.38)	(2.61)	(9.88)
Male	−0.038**	−0.084***	−0.236***	−0.085***	−0.124***	−0.138***	−166.057**	−56.664**	−0.365***
	(−2.27)	(−2.64)	(−2.62)	(−2.75)	(−3.57)	(−3.10)	(−2.38)	(−2.10)	(−2.61)
Years before last grade in current school level		0.045***		−0.006		−0.023*	−6.165	−12.973**	−0.016
		(5.06)		(−0.71)		(−1.83)	(−0.39)	(−2.12)	(−0.44)
Secondary school		−0.359***		0.018		0.063	−41.929	−3.793	−0.176
		(−7.30)		(0.61)		(1.48)	(−0.81)	(−0.18)	(−1.28)
Mother age	0.048**	0.084	0.334**	0.111**	0.148***	0.148**	203.311*	54.731	0.433*
	(1.97)	(1.50)	(2.39)	(2.10)	(2.89)	(2.05)	(1.89)	(1.20)	(1.95)
Mother age squared	−0.001**	−0.001	−0.004**	−0.001**	−0.002***	−0.002**	−2.563*	−0.680	−0.006**
	(−2.01)	(−1.52)	(−2.39)	(−2.11)	(−2.91)	(−2.09)	(−1.90)	(−1.19)	(−1.99)
Female-headed household	−0.038	−0.018	−0.137	−0.044	−0.076	−0.069	−91.760	−27.225	−0.150
	(−1.43)	(−0.29)	(−0.91)	(−0.73)	(−1.23)	(−0.83)	(−0.84)	(−0.57)	(−0.58)
Head's years of schooling	0.009*	0.005	0.062**	−0.006	0.007	−0.000	−9.533	0.974	−0.012
	(1.75)	(0.53)	(2.25)	(−0.66)	(0.67)	(−0.04)	(−0.55)	(0.13)	(−0.32)
Ethnic majority group	0.010	0.069	0.200	0.091	0.080	0.096	189.625	127.766**	0.218
	(0.33)	(1.15)	(1.20)	(1.45)	(1.17)	(1.12)	(1.40)	(2.42)	(0.80)
Total household expenditures	0.004***	0.016***	0.022***	0.007***	0.010***	0.012***	0.017**	4.112***	0.034***
	(4.05)	(4.65)	(4.05)	(3.76)	(5.04)	(4.33)	(2.52)	(2.74)	(3.94)
Distance to primary school	0.003	0.006	0.055**	0.012	0.009	0.006	26.266	9.945	0.028
	(0.73)	(0.62)	(2.15)	(1.22)	(0.88)	(0.45)	(1.51)	(1.44)	(0.66)

Distance to secondary school	−0.002	−0.004	−0.031**	−0.003	−0.005	−0.006	−10.533	−4.657*	−0.035**
	(−0.81)	(−0.95)	(−2.52)	(−0.80)	(−1.35)	(−1.14)	(−1.39)	(−1.69)	(−2.17)
Constant	0.396	−2.134***	−9.903***	−1.221*	−0.745	−1.465	−3736.470**	−950.852	−5.988**
	(1.26)	(−2.85)	(−5.38)	(−1.72)	(−1.12)	(−1.50)	(−2.38)	(−1.55)	(−1.97)
Model	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	IV-Tobit	IV-Tobit	2SLS
F/ Chi^2 test	32.49	46.85	862.88	44.50	46.67	47.28	40.31	501.57	50.25
Log likelihood							−19019	−18262	
N	5012	4125	5012	4125	5012	4248	4125	4247	4248
Number of left-censored obs.							2511	2623	
Non-Instrumented Regressions	−0.038***	−66.390***	−0.240***	−0.043***	−0.085***	−0.083***	−79.516***	−46.051***	−0.233***
	(−6.50)	(−8.06)	(−7.60)	(−5.21)	(−7.91)	(−7.18)	(−3.66)	(−5.58)	(−6.19)

Notes: *p< .1, **p<0.05, ***p<0.01; robust t statistics in parentheses account for clustering at the household level. All regressions control for regional dummy variables, which include the following regions: Northeast and Northwest, North Central, South Central Coast, Central Highlands, South East, and Mekong River Delta. The reference category is the Red River Delta. Total household expenditure is net of education expenditure and tutoring expenditure respectively for the specifications of these outcomes. All household expenditures are in million Vietnamese dong, except for the expenditure variables in the Tutoring specification. For instrumented regressions, the instrumental variable is the distance from the commune to the nearest family planning center.

Source: Authors' analysis based on data from Vietnam Household Living Standards Surveys 2002 and 2006.

the former study finds the instrumented coefficients to range from 0 to 1.5 times the uninstrumented coefficients, but the latter study finds this ratio to be as large as 15 times.

The instrumented coefficients on family size are, however, statistically significant for all the tutoring variables except for tutoring hours. The instrumented coefficients on number of siblings have much larger absolute magnitude than the uninstrumented coefficients, ranging from four (enrollment and tutoring attendance index) to seven times (tutoring expenditure or attendance) as large as their uninstrumented counterparts, which points to the downward bias (in absolute magnitude) of the latter. Thus, both the stronger statistical significance and larger magnitudes for the former are consistent with our earlier theoretical discussion of private tutoring as a more elastic and refined measure of household educational investment than traditional measures.[43]

Controlling for other characteristics, each additional sibling results in reduced investments in a child's schooling: reductions in education expenditure and tutoring expenditure respectively by 0.4 standard deviations (or equivalently, a reduction of D 308,246) and 0.5 standard deviations (or D 211,087; see the online appendix S1 table S1.3); a decrease of 32 percentage points in his or her probability of being enrolled in tutoring; and a drop of 0.34 in the child's enrollment and tutoring index and 0.49 in the tutoring attendance frequency. One more sibling also leads to the child spending seventy-four fewer hours and 1.4 fewer years on tutoring, although the estimated coefficient on tutoring hours is no longer statistically significant.

Estimation results also indicate that, ceteris paribus, older children are less likely to enroll in school but more likely to attend tutoring, while boys are less likely either to enroll in school or attend tutoring.[44] Children that are farther

43. Since we control for the commune-level distances to school, the uninstrumented regression results that we presented (at the bottom of table 5) are identical to estimates using an OLS model with commune random effects. As suggested by a reviewer, we also estimate an OLS model with commune fixed effects and between-commune OLS (with variables aggregated at the commune level) for comparison. Estimation results are provided in tables S1.4 and S1.5 in the online appendix, where the former's estimated coefficients are smaller in magnitudes than the latter's, which are in turn smaller than those of the IV estimates. This suggests that the between-commune OLS estimates are less biased than the FE estimates, and appears consistent with the bias caused by the endogeneity of family size—which occurs at the household level. In particular, the FE estimates are the commune-fixed effects estimates, which rely on the variation of a small number (at most three) households in a commune for identification. Thus, the FE estimates can be severely biased. On the other hand, the between-commune OLS would first average out this variation (bias) in a commune in constructing the commune-aggregated variables, then rely on the variation between different communes (more than 1500) for identification. Thus, while estimates are still biased, these would be to a lesser extent than those from the FE estimates.

44. We also experiment with using the distance to family center as the instrument for the number of male or female siblings, however, this instrument is statistically significant only in the first-stage regressions for the number of brothers, with qualitatively similar second-stage estimation results (not shown). While this result may indicate a degree of son preference in Vietnam, and it is consistent with previous studies (see, e.g., Phai et al. 1996; Belanger 2002), it may also suggest sex-selective abortion at the same time. Deeper analysis for intra-household gender differences would require better (and more than one) instruments than currently available. Thus, we leave this to further research.

from the last grade in their current school level are, as expected, less likely to have tutoring, but the coefficient on this variable is mostly statistically insignificant except in the case of tutoring hours. Older mothers and richer households invest more in their children's tutoring, but the quadratic term on mothers' age is negative, indicating that the marginal effect of age declines and eventually turns negative.

Robustness Checks

We further test the robustness of estimation results and provide them in table S1.6 in the online appendix S1. In the previous section (table 4), we have provided evidence against the concern that distance to the nearest family planning center may be proxying for other important unobserved commune characteristics. However, we test for this possibility again by including as control variables in the equation of interest some commune-level variables such as commune infrastructure, the distance to health facilities, and the share of the commune population working in agriculture. Since our estimation sample is restricted to rural households, to examine the hypothesis—albeit in an indirect way—that urban households spend more on tutoring, we also include the distance from the commune to the nearest major city in Vietnam.[45] Estimation results are largely qualitatively similar.[46]

Our previous study (Dang 2007) shows that communes with higher levels of education spend more on tutoring and argues that this impact can come from both the demand side (e.g., children have peer pressure to study harder or beneficial interaction with well-educated adults) and the supply side (e.g., communities with higher educational levels may be able to supply more tutors). We thus add to our equation of interest either the share of the commune adult population with upper secondary education or higher or a set of commune-averaged variables calculated from the primary school census (DFA) database including the shares of teachers with upper secondary education, upper secondary education plus two more years of additional training, two-year teacher training college education, four-year teacher training college education, and student-teacher ratios. These variables are expected to capture respectively the levels of commune education and the teacher and school quality in the commune.[47] Again, the estimation results are similar to those in our base specification.

45. These cities are Hanoi and Haiphong in northern Vietnam, Danang in central Vietnam, and Cantho and Ho Chi Minh in southern Vietnam. We also experiment with using the distance to the provincial city instead of the distance to these major cities and obtain similar, albeit slightly statistically weaker, results.

46. The only exception is the model specification with all the commune infrastructure and distances variables (row 1), but even in that case, magnitudes are similar but the coefficients have less statistical significance. This is perhaps unsurprising: the model is over-fitted, with all the distance variables statistically insignificant in both the first-stage regressions (as shown in table S1.2 in the online appendix) and second-stage regressions (not shown).

47. Detailed estimation results are provided in tables S1.7 and S1.8 in the online appendix.

While we have reduced some contemporaneous correlation between the distance to the nearest family planning center and household investment in their children by using values for the former in 2002 and the latter in 2006 in our regressions, this gap of four years may not be enough, given that households make their tutoring investments only when children at least six years old.[48] While a family planning center built in 2002 will have had no impact on parents' decision to give birth to the children who are at least six years old in 2006, the impact of the family planning center on family size in this case will come through the household decision on the number of younger siblings for these children and, subsequently, on total family size. Nevertheless, to examine this case, we restrict our estimation sample to the cases where the family planning center was already operating by 1997, which reduces the estimation sample by more than half.[49] Our results are for the most part qualitatively similar, except that the effects on education and tutoring expenditure now lose their statistical significance (though they keep their negative signs), while the effects on hours and years spent on tutoring become even more statistically significant.

In addition, we also implement other robustness checks including using the Lewbel heteroskedasticity-based IV model, and experiment with dropping out the outliers in the distance to the family planning center. Estimation results are, however, qualitatively similar. More detailed discussion of these results and other checks is provided in the working version of this paper (Dang and Rogers 2013).

Further/Heterogeneity Analysis

Estimation results thus far support the negative relationship between family size and household investment in tutoring classes. This subsection delves deeper into this result to provide heterogeneity analysis with, among other factors, different definitions of family size as well as subsets of the population. Estimation results are shown in table 6.

Different definitions of family size. Could our estimation results be sensitive to how we define family size? We provide further analysis based on different definitions of family size. First, we restrict the number of siblings to not more than three (row 1, table 6), to test whether the main result is driven by unusually large family sizes. Second, we extend the definition of family size from the children born of the same mother to all the children living in an extended family (row 2), which would perhaps be more consistent with an altruistic model in which

48. As predicted by the Becker-Lewis model, it is total family size that affects the quality-quantity tradeoff. Thus, the distance to the family planning center is still a relevant instrument as long as it can predict total family size.

49. There are a number of missing observations for the year a family planning center was set up, and the distances to school variables are not significant in these specifications, thus we left them out for larger sample sizes and more accurate estimates. As discussed in the previous section, the similarity in impacts of household size for the full sample and the sample with older family planning centers indicates that the locations of family planning centers are effectively independent of household size.

TABLE 6. Further/ Heterogeneity Analysis

No		Spec. 1 Enrollment	Spec. 2 Total education expenditure	Spec. 3 Completed years of schooling	Spec. 4 Tutoring attendance	Spec. 5 Enrollment & Tutoring attendance	Spec. 6 Tutoring attendance frequency	Spec. 7 Tutoring expenditure	Spec. 8 Tutoring hours	Spec. 9 Years attending tutoring
	Various definitions for family size									
1	Number of siblings age 0–18 less than or equal to 3	−0.110	−520.659	−1.001	−0.630*	−0.609**	−0.876**	−1158.703*	−321.754	−2.450*
		(−0.89)	(−1.63)	(−1.33)	(−1.88)	(−2.03)	(−2.05)	(−1.75)	(−1.43)	(−1.92)
	N	4750	3934	4750	3937	4750	4054	3937	4053	4054
2	Number of siblings age 0–18, relaxed definition	−0.136	−436.722**	−0.745	−0.474**	−0.541**	−0.767**	−902.347**	−347.998**	−2.283**
		(−1.37)	(−1.96)	(−1.41)	(−2.13)	(−2.15)	(−2.33)	(−2.03)	(−1.96)	(−2.28)
	N	7000	5540	7000	5550	7000	5704	5550	5703	5704
3	Number of siblings age 6–18	−0.115	−457.807*	−0.914	−0.461**	−0.523**	−0.729**	−846.219*	−283.496	−2.132**
		(−1.04)	(−1.89)	(−1.41)	(−2.08)	(−2.06)	(−2.27)	(−1.91)	(−1.54)	(−2.22)
	N	5015	4125	5015	4128	5015	4251	4128	4250	4251
	Birth order									
4	Birth order index added to the control variables	−0.025	−0.372*	−0.095	−0.433**	−0.429*	−0.688**	−933.382*	−262.425	−1.560*
		(−0.23)	(−1.68)	(−0.18)	(−2.11)	(−1.85)	(−2.21)	(−1.94)	(−1.41)	(−1.84)
	N	3880	3289	3880	3292	3880	3396	3292	3395	3396
	School quality									

(Continued)

TABLE 6. Continued

No		Spec. 1	Spec. 2	Spec. 3	Spec. 4	Spec. 5	Spec. 6	Spec. 7	Spec. 8	Spec. 9
		Enrollment	Total education expenditure	Completed years of schooling	Tutoring attendance	Enrollment & Tutoring attendance	Tutoring attendance frequency	Tutoring expenditure	Tutoring hours	Years attending tutoring
5	Estimation sample being restricted to the school considered to have good or excellent quality by parents	N/A	−177.341 (−1.17)	−0.790** (−2.38)	−0.306* (−1.94)	−0.288** (−1.97)	−0.565** (−2.34)	−602.280* (−1.85)	−150.720 (−1.23)	−1.293** (−1.97)
	N		2149	2215	2150	2215	2215	2150	2214	2215
	Outcomes in 2008									
6	All outcome variables in 2008	−0.215* (−1.90)	−413.753 (−1.13)	−0.073 (−0.15)	−0.519* (−1.91)	−0.576** (−2.28)	N/A	−1222.416** (−2.10)	N/A	N/A
	N	6030	4678	6030	4678	6030		4678		
	Model	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	IV-Tobit	IV-Tobit	2SLS

Notes: *p< .1, **p<0.05, ***p<0.01; robust t statistics in parentheses account for clustering at the household level. Unless otherwise noted, each cell provides the estimated coefficient on the number of siblings age 0–18 from a separation regression that controls for the same explanatory variables in the corresponding specification in table 5. All regressions control for regional dummy variables, which include the following regions: Northeast and Northwest, North Central, South Central Coast, Central Highlands, South East, and Mekong River Delta. The reference category is the Red River Delta. Total household expenditure is net of education expenditure and tutoring expenditure respectively for the specifications of these outcomes. All household expenditures are in million Vietnamese dong, except for the expenditure variables in the Tutoring specification. All regressions are estimated with IV method, where the instrumental variable is the distance from the commune to the nearest family planning center.

Source: Authors' analysis based on data from Vietnam Household Living Standards Surveys 2002, 2006, and 2008.

resources are shared within the extended family (e.g., Alger and Weibull 2010; Schwarze and Winkelmann 2011). An altruistic model may be an equally valid model in the context of Vietnam, where Confucian culture remains strong (Huu Ngoc 1996; Tran 2001). Third, we restrict the number of siblings to age 6–18 only (row 3), hypothesizing that the quantity-quality tradeoff will be stronger because households have to invest more in school-age children than in younger ones. Reassuringly, estimates are both larger in magnitude and have slightly stronger statistical significance when we use the more general definition of family size (row 2) and restrict the analysis to school-age siblings (row 3).

Birth order. Beyond the impacts of family size, the birth order of a child can also influence his or her parents' resource allocation in different directions. For example, first-born children may enjoy more parental time and investment due to their unique timing position (Price 2008; de Haan 2010), but younger siblings may benefit more if parents' earnings (Parish and Willis 1993) or child-rearing experience increase over the life cycle. Since birth order is closely related with family size (e.g., a child in a higher birth order is more likely to be in a larger family), we construct a birth order index suggested by Booth and Kee (2009) that is purged of family-size effect. This index is defined as $p/((n + 1)/2)$, where p is the child's birth order, and n the total number of children in the family. We add this birth-order index to our equation of interest (row 4) and find that coefficients become larger (in absolute value) but estimation results are qualitatively similar.[50]

Perception of school quality. We turn next to the role of school quality in influencing parents to send their children to tutoring lessons. Only a small proportion of households in Vietnam cite poor school quality as the reason for enrolling their children in tutoring classes (table 1), but other studies suggest that the opposite holds in other countries (Kim and Lee 2010; Bray and Lykins 2012). To examine the hypothesis that the negative impacts of family size may possibly not hold for children enrolled in high-quality schools, we restrict our estimation sample to children going to schools perceived by their parents as being of high quality, and we find estimation results for tutoring outcomes to be very similar, except that the impact of household size on education expenditure now loses its statistical significance (row 5).

50. The Pearson correlation coefficient with family size decreases from 0.49 for birth order to −0.08 with this index, which indicates that family size effect is largely netted out. We also try another birth order index suggested by Ejrnaes and Portner (2004) but find similar results. Because certain cultures, especially in Asia, may prefer sons over daughters, older sons may be more favored than their younger female siblings. We also try interacting this birth-order index with the male variable, but this interaction variable is not significant either. However, we do not have census data, and the birth order we have is for those children that are currently living in the household only. Thus we do not rule out the possibility that birth orders may have a (weak) impact on our results.

Outcomes for younger cohorts in 2008. Recent studies find that the quantity-quality tradeoff holds for younger but not older cohorts in Norway (Black, Devereux, and Salvanes 2010), turns from positive to no effect and then negative during the 1977–2009 period in Brazil (Marteleto and de Souza 2012), and changes from positive for older cohorts to negative for younger cohorts in urban areas in Indonesia (Maralani 2008). To investigate whether this tradeoff applies to younger cohorts in Vietnam, we rerun the same regressions using the 2008 round of the VHLSSs for children in the same age range (6–18).[51] While the 2008 data collect fewer variables on tutoring, our estimation results on the available indicators provide broadly qualitatively similar results (row 6), except that the effect on education expenditure is no longer statistically significant, while the effect on enrollment is statistically significant at the 10 percent level, and the effect on tutoring expenditure becomes stronger both in magnitude and statistical significance.[52]

Impacts of Family Size on Tutoring Investment Versus Traditional Measures

The regressions in tables 5 and 6 consider measures of household investment in tutoring only using equations (11) and (12). As discussed with our theoretical model, private tutoring should also be examined in its relationship with regular school. To operationalize this hypothesis, we can rewrite equation (11) slightly differently

$$E_{ijk} = \alpha_k + \beta_k FamSize_i + \gamma_k X_{ij} + \mu_{ik} + \vartheta_{ijk}, \tag{14}$$

where k indexes the different types of household investment in education such as education expenditure or private tutoring expenditure. The error term ε_{ij} is broken into two components that vary by household investment type: μ_{ik} and ϑ_{ijk}, which respectively represent unobserved household effects (e.g., household tastes for their children's education across different types of education investments) and the child idiosyncratic error term.

If we assume that households have the same preference over investment in regular school and tutoring (i.e., μ_{ik} being the same for these two investment

51. As in a previous robustness check regression (table 1.6, row 5), because the distances to school variables are not significant in this specification, we left them out to allow larger sample sizes and greater precision of estimates.

52. Since IV estimates may refer to the unobserved subset of the population that reacts to distance to the family planning center—which is known as the Local Average Treatment Effects (LATE) (see, for example, Imbens and Angrist 1994; Angrist and Pischke 2009)—one concern arises that our previous estimation results may apply only to these households, which may comprise a small share of the total. However, various additional estimation results such as restricting the estimation sample to better-off households in the richer three consumption quintiles and others (see table S1.9 in the online appendix) indicate that a substantial share of the population (i.e., half or more) appears to be influenced by this IV. Restricting the estimation sample to households in the poorest three consumption quintiles provides qualitatively similar but less statistically significant results (see table S1.10 in the online appendix). Also see Dang and Rogers (2013) for further discussion.

types), we can in fact difference out the unobserved household effects by considering the *absolute* difference of these two investments

$$\Delta E_{ij,ln} = \Delta\alpha_{ln} + \Delta\beta_{ln} FamSize_i + \Delta\gamma_{ln} X_{ij,ln} + \Delta\vartheta_{ij,ln} \tag{15a}$$

or, equivalently,

$$\Delta E_{ij,ln} = \alpha_a + \beta_a FamSize_i + \gamma_a X_{ij} + \vartheta_{a,ij}, \tag{15b}$$

where $\Delta E_{ij,ln} \equiv E_{ijl} - E_{ijn}$, with l and n being, respectively, the investment in regular school and tutoring, and the coefficients in equations (15a) rewritten for convenience of presentation (e.g., $\Delta\beta_{ln} \equiv \beta_a$). Similarly, we can consider the *relative* difference of these two investments

$$\frac{E_{ijl}}{E_{ijn}} = \Delta\alpha_{ln} + \Delta\beta_{ln} FamSize_{i,ln} + \Delta\gamma_{ln} X_{ij,ln} + \Delta\vartheta_{ij,ln} \tag{16a}$$

or

$$\frac{E_{ijl}}{E_{ijn}} = \alpha_r + \beta_r FamSize_i + \gamma_r X_{ij} + \vartheta_{r,ij}, \tag{16b}$$

where we have $\dfrac{\beta_l}{\beta_n} \equiv \Delta\beta_{ln} = \beta_r$ instead.[53]

Given this assumption of similar household preference over education investment types, we can simply estimate the impacts of family size on the difference between household investment in private tutoring and regular school with OLS method. However, if households have different preferences between tutoring and regular school, the unobserved household effects μ_{ik} cannot be differenced out and we would need to instrument for family size with the distance to the family planning center in estimating these equations. While it may not seem unreasonable to think that the assumption of similar preference can hold in certain contexts, we believe that this assumption may not hold for the average household in Vietnam given the diverse opinions frequently raised on tutoring in the local media. It thus appears that the uninstrumented regressions would, similar to the results shown in table 5, offer estimates of the impacts of family size that are biased upward toward zero.

Still, for comparison purposes we estimate equations (14) and (15) by both OLS and IV methods and provide estimation results in table 7, where the OLS results are shown at the bottom of this table. For the absolute differences, we

53. We derive equation (16a) by rewriting the dependent variable in equation (1) in log format before taking the ratios of the two investments, and then removing the log format of the ratio of the two investments for easier interpretation. Another way to think about this ratio is tutoring investment standardized by investment in regular schooling.

TABLE 7. Impacts of Family Size on Private Tutoring Versus Regular School for Children Age 6–18, Vietnam 2006

	Spec. 1	Spec. 2	Spec. 3	Spec. 4
Instrumented Regressions	Tutoring attendance (with nonattendance including both enrollment and non-enrollment)	Education expenditure net of tutoring expenditure	Share of tutoring expenditure in education expenditure	Share of years attending tutoring over completed years of schooling
Number of siblings age 0–18	−0.311**	−0.243**	−0.077*	−0.203**
	(−2.23)	(−2.02)	(−1.84)	(−2.12)
Age	0.009*	0.101***	0.002	−0.002
	(1.74)	(12.49)	(1.06)	(−0.48)
Male	−0.087***	−0.050*	−0.027***	−0.061***
	(−2.84)	(−1.93)	(−2.87)	(−2.83)
Years before last grade in current school level	−0.008	0.044***	−0.003	0.003
	(−0.95)	(6.74)	(−1.12)	(0.48)
Secondary school	0.020	−0.315***	−0.012	−0.044**
	(0.70)	(−6.79)	(−1.39)	(−2.30)
Mother age	0.112**	0.056	0.027*	0.064*
	(2.20)	(1.25)	(1.79)	(1.82)
Mother age squared	−0.001**	−0.001	−0.000*	−0.001*
	(−2.20)	(−1.28)	(−1.82)	(−1.88)
Female-headed household	−0.041	−0.026	0.007	−0.027
	(−0.71)	(−0.54)	(0.38)	(−0.69)
Head's years of schooling	−0.005	0.001	−0.001	−0.001
	(−0.64)	(0.19)	(−0.47)	(−0.24)
Ethnic majority group	0.085	0.063	0.026	0.073*
	(1.42)	(1.26)	(1.52)	(1.71)
Total household expenditures	0.007***	0.012***	0.003***	0.005***
	(3.59)	(6.12)	(4.23)	(3.74)
Distance to primary school	0.010	0.004	0.005*	0.005
	(1.02)	(0.45)	(1.81)	(0.85)
Distance to secondary school	−0.003	0.001	−0.002**	−0.005**
	(−0.71)	(0.21)	(−2.45)	(−2.21)
Constant	−1.273*	−1.585***	−0.264	−0.445

	(−1.85)	(−2.61)	(−1.28)	(−0.92)
Model	2SLS	2SLS	2SLS	2SLS
F test	39.98	45.59	31.68	37.61
N	4248	4125	4091	4248
Mean of dependent variable	0.41	0.47	0.11	0.30
Non-Instrumented Regressions	−0.045***	−0.050***	−0.014***	−0.034***
	(−5.32)	(−7.22)	(−5.13)	(−5.51)

Notes: *p< .1, **p<0.05, ***p<0.01; robust t statistics in parentheses account for clustering at the household level. All regressions control for regional dummy variables, which include the following regions: Northeast and Northwest, North Central, South Central Coast, Central Highlands, South East, and Mekong River Delta. The reference category is the Red River Delta. Total household expenditure is net of education expenditure and tutoring expenditure, respectively, for the specifications of these outcomes. All household expenditures are in million Vietnamese dong, except for the expenditure variables in the Tutoring specification. All regressions are estimated with IV method, where the instrumental variable is the distance from the commune to the nearest family planning center.

Source: Authors' analysis based on data from Vietnam Household Living Standards Surveys 2002 and 2006.

consider a dummy variable that is 1 if the child attended tutoring in the past twelve months and 0 otherwise (i.e., equivalent to subtracting the school enrollment variable from the enrollment and tutoring attendance variable), and education expenditure net of tutoring expenditure (i.e., equivalent to subtracting tutoring expenditure from total education expenditure). For the relative differences, we consider two share variables: tutoring expenditure over total education expenditure and years of tutoring over completed years of schooling.

The IV estimated coefficients on family size are negative and statistically significant at the 5 percent level for all these variables, except for the share of tutoring expenditure over education expenditure, which is significant at the 10 percent level. These results indicate that one more sibling reduces the probability of attending tutoring (unconditional on whether the child is enrolled in school or not) by 31 percentage points; reduces education expenditure net of tutoring expenditure by D 243,000; and reduces the two share variables by 8 percentage points and 20 percentage points, respectively. These estimated coefficients are roughly five or six times larger in absolute magnitude than the uninstrumented regression coefficients.

These estimation results thus validate our theoretical discussion that household demand for tutoring is more elastic to changes in family size than are other traditional measures and that tutoring investment merits more attention as a new measure of household education investment.

V. Conclusion

We find in this paper that families invest less in the education of school-age children who have larger numbers of siblings. Using the distance to the nearest family planning center as the instrument to identify the impacts of family size on household investment, the instrumented number of siblings has a strongly negative effect on education investment, and the estimated coefficient is much larger (in absolute value) than in the original uninstrumented regressions. This effect is robust across different indicators of educational investment—including the general education expenditure on the child, frequency of tutoring attendance, and expenditure and hours spent on tutoring—as well as with different specifications and definitions of family size.

Our results provide evidence that parents in Vietnam are indeed making a child quality-quantity tradeoff. The results suggest further that by lowering the relative cost of child quality and encouraging families to invest in quality, the availability of family planning services has increased investment in education in Vietnam. Finally, the analysis suggests that, compared with traditional indicators like enrollment, data on tutoring may be a more illuminating indicator of parents' willingness to invest in the quality of education of their children. Indeed, the hypothesized quantity-quality tradeoff appears much more strongly in the tutoring-based measures than in the simple enrollment decision, which may be a coarser indicator of the household's desire to invest in human capital. These results suggest the need for more research into these quality-oriented measures of

schooling investment, which could be examined in other contexts—besides the quantity-quality tradeoff model—that are broadly related to education efficiency and human capital enrichment.

References

Alger, I., and J. W. Weibull. 2010. "Kinship, Incentives, and Evolution." *American Economic Review* 100: 1725–58.

Angrist, J. D., and W. N. Evan. 1998. "Children and Their Parent's Labor Supply: Evidence from Exogenous Variation in Family Size." *American Economic Review* 80 (3): 313–36.

Angrist, J. D., and J.-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. New Jersey: Princeton University Press.

Angrist, J., V. Lavy, and A. Schlosser. 2010. "Multiple Experiments for the Causal Link between the Quantity and Quality of Children." *Journal of Labor Economics* 24 (4): 773–824.

Attfield, I., and B. T. Vu. 2013. "A Rising Tide of Primary School Standards—The Role of Data Systems in Improving Equitable Access for All to Quality Education in Vietnam." *International Journal of Education Development* 33: 74–87.

Bailey, M. J. 2006. "More Power to the Pill: The Impact of Contraceptive Freedom on Women's Lifecycle Labor Supply." *Quarterly Journal of Economics* 121 (1): 289–320.

Banerjee, A. V., R. Banerji, E. Duflo, R. Glennerster, and S. Khemani. 2010. "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India." *American Economic Journal: Economic Policy* 2 (1): 1–30.

Barro, R. J., and J.-W. Lee. 2012. "A New Data Set of Educational Attainment in the World, 1950–2010." Working paper. Department of Economics, Harvard University.

Becker, G. S. 1960. "An Economic Analysis of Fertility." In Gary S. Becker, eds. *Demographic and Economic Change in Developed Countries*. New Jersey: Princeton University Press.

Becker, G. S., and H. G. Lewis. 1973. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy* 81: S279–288.

Becker, G. S., and N. Tomes. 1976. "Child Endowments and the Quantity and Quality of Children." *Journal of Political Economy* 84 (4): S143–162.

Belanger, D. 2002. "Son Preference in a Rural Village in North Vietnam." *Studies in Family Planning* 33 (4): 321–34.

Black, S. E., P. J. Devereux, and K. G. Salvanes. 2005. "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education." *Quarterly Journal of Economics* 120 (2): 669–700.

———. 2010. "Small Family, Smart Family? Family Size and IQ Scores of Young Men." *Journal of Human Resources* 45 (1): 33–58.

Bloom, D. E., D. Canning, G. Fink, and J. E. Finlay. 2009. "Fertility, Female Labor Force Participation, and the Demographic Dividend." *Journal of Economic Growth* 14: 79–101.

Booth, A. L., and H. J. Kee. 2009. "Birth Order Matters: The Effect of Family Size and Birth Order on Educational Attainment." *Journal of Population Economics* 22: 367–97.

Bray, M. 2009. *Confronting the Shadow Education System. What Government Policies for What Private Tutoring*. International Institute for Educational Planning. Paris: UNESCO.

Bray, M., and C. Lykins. 2012. *Shadow Education: Private Supplementary Tutoring and Its Implications for Policy Makers in Asia*. Manila: Asia Development Bank.

Caceres-Delpiano, J. 2006. "The Impacts of Family Size on Investment in Child Quality." *Journal of Human Resources* 41 (4): 738–54.

Card, D. 1995. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." In L. N. Christofides, E. K. Grant, and R. Swidinsky, eds., *Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp*. University of Toronto Press: Toronto.

Dang, H.-A. 2007. "The Determinants and Impact of Private Tutoring Classes in Vietnam." *Economics of Education Review* 26 (6): 684–99.

———. 2008. *Private Tutoring in Vietnam: An Investigation of its Causes and Impacts with Policy Implications*, VDM Verlag Dr. Mueller Publishing House: Saarbrucken, Germany.

———. 2011. "A Bird's-Eye View of the Private Tutoring Phenomenon in Vietnam." International Institute for Asian Studies Newsletter *(Leiden, the Netherlands),* 56 (1): 26–27.

———. 2013. "Private Tutoring in Vietnam: A Review of Current Issues and Its Major Correlates." In J. Aurini, J. Dierkes, and S. Davis, eds., *Out of the Shadows: The Global Intensification of Supplementary Education.* United Kingdom: Emerald Press.

Dang, H.-A., and H. Rogers. 2008. "The Growing Phenomenon of Private Tutoring: Does It Deepen Human Capital, Widen Inequalities, or Waste Resources?" *World Bank Research Observer* 23 (2): 161–200.

———. 2013. "The Decision to Invest in Child Quality over Quantity: Household Size and Household Investment in Education in Vietnam." *World Bank Policy Research Working Paper* No. 6487.

Dang, H.-A., and P. Glewwe. 2009. "An Analysis of Learning Outcomes for Vietnam." *Working paper*.

Dang, N. A., C. Tacoli, and T. X. Hoang. 2003. "Migration in Vietnam A Review of Information on Current Trends and Patterns, and Their Policy Implications." Paper presented at the Regional Conference on Migration, Development and Pro-Poor Policy Choices in Dhaka, Bangladesh.

Do, M. P., and M. A. Koenig. 2007. "Effect of Family Planning Services on Modern Contraceptive Method Continuation in Vietnam." *Journal of Biosocial Science* 39 (2): 201–20.

DeGraff, D. S., R. E. Bilsborrow, and D. K. Guilkey. 1997. "Community Level Determinants of Contraceptive Use in the Philippines: A Structural Analysis." *Demography* 34 (3): 385–98.

de Haan, M. 2010. "Birth Order, Family Size and Educational Attainment." *Economics of Education Review* 29 (4): 576–88.

Ejrnaes, M., and C. C. Portner. 2004. "'Birth-Order and the Intrahousehold Allocation of Time and Education." *Review of Economics and Statistics* 86: 1008–19.

Emran, M. S., and Z. Hou. 2013. "Access to Markets and Rural Poverty: Evidence from Household Consumption in China." *Review of Economics and Statistics* 95 (2): 682–97.

General Department of Population and Family Planning (GDPFP). 2011. *Công Tác Dân Số- Kế Hoạch Hóa Gia Đình Việt Nam: 50 Năm Xây Dựng và Phát Triển (1961-2011).* (Population and Family Planning Activities in Vietnam: 50 Years of Development (1961–2011). Hanoi, Vietnam: Ministry of Health.

General Statistical Office (GSO). 2012. *Statistical Yearbook of Vietnam 2011*. Hanoi: Statistical Publishing House.

Gibson, J., and D. McKenzie. 2007. "Using Global Positioning Systems in Household Surveys for Better Economics and Better Policy." *World Bank Research Observer*, 22(2): 217–41.

Glewwe, P., and M. Kremer. 2006. "School, Teachers, and Education Outcomes in Developing Countries." In Eric A Hanushek, and Finis Welch, eds. *Handbook of the Economics of Education*. Amsterdam: North Holland.

Goodkind, D. M. 1995. "Vietnam's One-or-Two-Child Policy in Action." *Population and Development Review*, 21(1): 85–111.

Greene, W. H. 2012. *Econometric Analysis*, 7th Edition. New Jersey: Prentice Hall.

Henshaw, S. K., S. Singh, and T. Haas. 1999. "The Incidence of Abortion Worldwide." *International Family Planning Perspectives* 25: S30–S38.

Imbens, G. W., and J. D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62 (2): 467–75.

Jayachandran, S. 2014. "Incentives to Teach Badly? After-School Tutoring in Developing Countries." *Journal of Development Economics* 108: 190–205.

Joshi, S., and T. Paul Schultz. 2013. "Family Planning and Women's and Children's Health: Long-Term Consequences of an Outreach Program in Matlab, Bangladesh." *Demography*, 50: 149–80.

Kang, C. 2011. "Family Size and Educational Investments in Children: Evidence from Private Tutoring Expenditures in South Korea." *Oxford Bulletin of Economics and Statistics*, 73(1): 59–78.

Kaufman, J., Z. Zhirong, Q. Xinjian, and Z. Yang. 1992. "The Creation of Family Planning Service Stations in China." *International Family Planning Perspectives* 18 (1): 18–23.

Kim, S., and J.-H. Lee. 2010. "Private Tutoring and Demand for Education in South Korea." *Economic Development and Cultural Change* 58 (2): 259–96.

Klein, R., and F. Vella. 2010. "Estimating a Class of Triangular Simultaneous Equations Models without Exclusion Restrictions." *Journal of Econometrics* 154: 154–64.

Klepinger, D., S. Lundberg, and R. Plotnick. 1999. "How Does Adolescent Fertility Affect the Human Capital and Wages of Young Women?" *Journal of Human Resources* 34 (3): 421–48.

Lanjouw, J. O., P. Lanjouw, B. Milanovic, and S. Paternostro. 2004. "Relative Price Shifts, Economies of Scale and Poverty during Economic Transition." *Economics of Transition* 12, 509–36.

Le, L. C., R. Magnani, J. Rice, I. Speizer, and W. Bertrand. 2004. "Reassessing the Level of Unintended Pregnancy and Its Correlates in Vietnam." *Studies in Family Planning* 35 (1): 15–26.

Lee, J. 2008. "Sibling Size and Investment in Children's Education: An Asian Instrument." *Journal of Population Economics* 21: 855–75.

Lewbel, A. 2012. "Using Heteroscedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models." *Journal of Business & Economic Statistics* 30 (1): 67–80.

Li, H., J. Zhang, and Y. Zhu. 2008. "The Quantity-Quality Trade-off of Children in a Developing Country: Identification Using Chinese Twins." *Demography* 45 (1): 223–43.

Maralani, V. 2008. "The Changing Relationship between Family Size and Educational Attainment over the Course of Socioeconomic Development: Evidence from Indonesia." *Demography*, 45, 693–717.

Marteleto, L. J., and L. R. de Souza. 2012. "The Changing Impact of Family Size on Adolescents' Schooling: Assessing the Exogenous Variation in Fertility Using Twins in Brazil." *Demography* 49: 1453–77.

McKenzie, D., and Y. S. Sakho. 2010. "Does It Pay Firms to Register for Taxes? The Impact of Formality on Firm Profitability." *Journal of Development Economics* 91 (1): 15–24.

Miller, G. 2010. "Contraception as Development: New Evidence from Family Planning in Columbia." *Economic Journal* 120: 709–36.

Ministry of Health (MOH). 2001. "Quyết Định 385/2001/QĐ-BYT về Việc Ban Hành Qui Định Nhiệm Vụ Kỹ Thuật Trong Lĩnh Vực Chăm Sóc Sức Khỏe Sinh Sản Tại Các Cơ Sở Y Tế" (Decision 385/2001/QĐ-BYT on Provision of Technical Guidelines on Mothers' Birth Health at Health Facilities). Accessed on August 23 2013 at http://thuvienphapluat.vn/archive/Quyet-dinh-385-2001-QD-BYT-nhiem-vu-ky-thuat-trong-cham-soc-suc-khoe-sinh-san-tai-co-so-y-te-vb93175.aspx.

Ngoc, H. 1996. *Sketches for a Portrait of Vietnamese Culture*. Hanoi: The Gioi Publisher.

Oster, E. 2012. "HIV and Sexual Behavior Change: Why not Africa?" *Journal of Health Economics* 31: 35– 49.

Parish, W. L., and R. J. Willis. 1993. "Daughters. Education, and Family Budgets: Taiwan Experiences." *Journal of Human Resources* 28 (4): 863–98.

Phai, N. V., J. Knodel, M. V. Cam, and H. Xuyen. 1996. "Fertility and Family Planning in Vietnam: Evidence from the 1994 Intercensal Demographic Survey." *Studies in Family Planning* 27 (1): 1–17.

Ponczek, V., and A. P. Souza. 2012. "New Evidence of the Causal Effect of Family Size on Child Quality in a Developing Country." *Journal of Human Resources* 47 (1): 64–106.

Portner, C. C., K. Beegle, and L. Christiaensen. 2011. "Family Planning and Fertility Estimating Program Effects Using Cross-Sectional Data." *Policy Research Working Paper 5812*. Washington DC: World Bank.

Price, J. 2008. "Parent-Child Quality Time Does Birth Order Matter?" *Journal of Human Resources* 43 (1): 240–65.

Qian, N. 2013. "Quantity-Quality and the One Child Policy:The Only-Child Disadvantage in School Enrollment in Rural China." *Working paper, Yale University*.

Rosenzweig, M. R., and K. I. Wolpin. 1980. "Life Cycle Labor Supply and Fertility: Causal Inferences from Household Models." *Journal of Political Economy* 88 (2): 328–48.

———. 1986. "Evaluating the Effects of Optimally Distributed Public Programs: Child Health and Family Planning Interventions." *American Economic Review* 76 (3): 470–82.

Rosenzweig, M. R., and T. P. Schultz. 1985. "The Demand for and Supply of Births: Fertility and Its Life Cycle Consequences." *American Economic Review* 75 (5): 992–1015.

Rosenzweig, M. R., and J. Zhang. 2009. "Do Population Control Policies Induce More Human Capital Investment? Twins, Birth Weight and China's "One-Child" Policy." *Review of Economic Studies* 76: 1149–74.

Sabia, J. J. 2007. "The Effect of Body Weight on Adolescent Academic Performance." *Southern Economic Journal* 73 (4): 871–900.

San, P. B., J. A. Ross, N. L. Phuong, and N. D. Vinh. 1999. "Measuring Family Planning Program Effort at the Provincial Level: A Vietnam Application" *International Family Planning Perspectives* 25 (1): 4–9.

Schultz, T. P. 2008. "Population Policies, Fertility, Women's Human Capital and Child's Quality." In T Paul Schultz, and John Strauss. (Eds). *Handbook of Development Economics*. Vol 4. North Holand: Elsevier.

Schwarze, J., and R. Winkelmann. 2011. "Happiness and Altruism within the Extended Family." *Journal of Population Economics* 24: 1033–51.

Scornet, C. 2001. "An Example of Coercive Fertility Reduction, as Seen in the Region of the Red River Delta in Viet Nam." *Population: An English Selection* 13 (2): 101–34.

Steelman, L. C., B. Powell, R. Werum, and S. Carter. 2002. "Reconsidering the Effects of Sibling Configuration: Recent Advances and Challenges." *Annual Review of Sociology* 28: 243–69.

Stiglitz, J. E. 2013. "Equal Opportunity, Our National Myth." *New York Times*, 16 February 2013.

Stock, J. H., and M. Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." In D. W. K. Andrews, and J. H. Stock, eds., *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge: Cambridge University Press.

Thang, N. M., and D. N. Anh. 2002. "Accessibility and Use of Contraceptives in Vietnam." *International Family Planning Perspectives* 28 (4): 214–19.

Thang, N. M., and V. T. Huong. 2003. "Changes in Contraceptive Use in Vietnam." *Journal of Biosocial Science* 35 (4): 527–43.

Tran, N. T. 2001. *Tìm Về Bản Sắc Văn Hóa Việt Nam*. (*Discovering the Identity of Vietnamese Culture: Typological-Systematic Views*.) Ho Chi Minh City, Vietnam: Ho Chi Minh City Publishing House.

World Bank. 2014. World Bank Development Indicators Online.

Vu, Q. N. 1994. "Family Planning Programme in Viet Nam." *Vietnam Social Sciences* 39: 3–20.

Zhang, Y. 2013. "Does Private Tutoring Improve Students' National College Entrance Exam Performance?—A Case Study from Jinan, China." *Economics of Education Review* 32: 1–28.

Zimmer, R., L. Hamiltona, and R. Christina. 2010. "After-school Tutoring in the Context of no Child Left Behind: Effectiveness of Two Programs in the Pittsburgh Public Schools." *Economics of Education Review* 29 (1): 18–28.