Child Indicators Research (2020) 13:147–173
https://doi.org/10.1007/s12187-019-09673-z

The Effect of Sibship Size on Children's Outcomes: Evidence from Vietnam

Daniel Mont[1] · Cuong Viet Nguyen[2,3] · Anh Tran[4]

Accepted: 15 August 2019 / Published online: 20 August 2019

Abstract

The child quality-quantity debate pays extensive attention to the effect of sibship size on education, but less is known about other outcomes of children such as health and labor. In this study, we estimate the effect of sibship size on various outcomes of children including disability, education and the labor supply of children in Vietnam. We do not find significant effects of sibship size on disability and working status of children. However, we find that having an additional sibling increases the risk of dropping out of school. More specifically, having an additional sibling reduces the probability of schooling of the first-born children by 1.6 percentage points in families with at least a child. This effect is increasing to around 3.0 percentage points in families with at least three children. The effect of sibship size on the number of completed grades is also higher and more significant in large families.

Keywords Sibship size · Health · Disability · School enrolment · Employment · Vietnam

JEL Classification J12 · J13 · I12

✉ Cuong Viet Nguyen
nguyenvietcuong@tdtu.edu.vn

Daniel Mont
danielmont01@gmail.com

Anh Tran
trananh@indiana.edu

[1] Leonard Cheshire Disability and Inclusive Development Centre, University College London, London, UK

[2] Informetrics Research Group, Ton Duc Thang University, Ho Chi Minh City, Vietnam

[3] Faculty of Social Sciences and Humanities, Ton Duc Thang University, Ho Chi Minh City, Vietnam

[4] Indiana University Bloomington, 1315 E. 2nd St, R. 410J, Bloomington, IN 47405, USA

1 Introduction

Does an increase in the number of siblings lessen family per-capita investment in children's human capital, such as education and health? The tradeoff between number of children and investment in them, first explored by Becker (Becker 1960), is referred to in the economic literature as the 'quantity-quality' tradeoff. For a given family, economic resources invested in each child will be reduced as the number of children increases. As a results, children in a family with a large number of siblings can have lower outcomes such as health and education than those in a family with a small number of siblings (Becker 1960; Becker and Lewis 1973; Becker and Tomes 1976).

The question on the trade-off between child quantity and quality has received significant attention in economics as well as in other social sciences. Although theories predict the tradeoff between child quantity and child quality, current empirical findings do not uniformly support this prediction. A number of studies find a negative impact on sibship size on children's health (e.g., Cummings et al. 1994; Henderson et al. 2008; Rosenzweig and Zhang 2009), and on education (e.g., Belmont and Marolla 1973; Rosenzweig and Wolpin 1980; Booth and Kee 2009; Park and Chung 2012). However, other studies find a negligible and insignificant effect on health (e.g., Kruger et al. 2006; Peters et al. 2014), and on education (e.g., Lee 2008; Black et al. 2005; Angrist et al. 2010; Yucel and Yuan 2015).

The wide diversity of empirical results calls for more empirical studies to better understand the effects of the number of children on child outcomes. Thus our study seeks to contribute to the child quantity-quality debate by testing the causal effect of the number of siblings on different outcomes of children including disability, education and labor supply in Vietnam using data from the 2009 Vietnam Population and Housing Census. Our study is expected to make two contributions. Firstly, it provides new empirical findings for Vietnam. To our knowledge, there are few studies looking at the association between family size and children's outcomes in Vietnam. For example, using the 1994 Inter-Censal Demographic Survey, Anh et al. (1998) find a negative association between family size and children s school attendance and educational attainment in Vietnam. Jernström (2017) shows a negative effect of family size on height-for-age and weight-for-age of children using Young Lives data on Vietnamese children. Our study is the first attempt to examine the effect of the number of children on children's outcomes in Vietnam. Vietnam is a good setting for studying this question. Vietnam has a large population of nearly 95.5 million people in 2017, ranking it 13th in the world (GSO 2018). In the 1980s, the fertility rate was very high, at 4.2. To reduce the fertility rate, the government issued the first Decree on birth control in 1988 which encouraged parents to have maximum two children (Pham et al. 2012).[1] The fertility rate decreased quickly to 2.3 in 1999 (GSO 1999). Thus, the government have stopped policies to restrict the family size since 2003 (Pham et al. 2012). During the same period, Vietnam has achieved a great success in developing education (Dang and Glewwe 2018). If lower sibship size can increase child health and education, the decline fertility can explain partly the success in education of Vietnam.

[1] The two-child policy was strictly applied to civil servants and worker sin the public sector. People who violated the policy were not promoted and received a wage reduction (Pham et al. 2012).

Secondly, we examine the effect of sibship size on different outcomes of children including health, education, and child labor. Most previous studies focus on either health or education of children. In this study, we measure health by disability. Having a large number of children in a household may reduce the attention, nutrition, healthcare and other resources that parents can devote for each child. This reduction can increase the risk of infectious diseases and injuries, leading to child disabilities. Although a variety of indicators of child health are examined in the above studies, there are no studies looking at the relation between sibship size and disability among children. Currently, scientific and evidence-based information on disability issues is seriously lacking (WTO, 2011). The World Health Organization (WHO 2011) estimates that roughly 650 million individuals currently suffer from some form of a disability. Among them, about 200 million are children. There is a strong negative association between disability, education and employment in both developed and developing countries (e.g., see Filmer 2008; Mitra 2006; Mont et al. 2014). We also look at the effect of sibship size on child labor, which is an important issue for a transition country like Vietnam (e.g., see ILO, MOLISA and GSO 2014; Giang et al. 2017). By looking at different child outcomes, we can provide a more comprehensive finding on the effect of sibship size.

This paper is organized as follows. The second section provides a brief overview of the literature of child quality and quantity. The third section introduces the dataset. The fourth section describes sibship size and child outcomes in Vietnam. Then, the fifth section presents the estimation method. The sixth section discusses the empirical findings of the effect of sibship size on disability, education and labor supply. Lastly, the conclusions of the study are presented in the seventh section.

2 Literature Review

There are several theories which explain the 'quantity-quality' trade-off between child quality and quantity. According to the economic theory, as the number of children increases, the cost of maintaining a particular level of quality for each child by investing in their health, education, and other factors will increase. Given limited resources, investment in each child will be reduced (e.g., Becker 1960; Becker and Lewis 1973; Becker and Tomes 1976). Children in a family with a large number of siblings might have lower outcomes such as health and education than those in a family with a small number of siblings.

This quantity-quality theory is also known in sociology as a so-called resource dilution hypothesis. This hypothesis is based on the observation that parental resources are finite. Thus, as the size of a family increases, the amount of parental resources available for each child tends to be diluted. By the same token, increased sibship size decreases a child's wellbeing (Blake 1981, 1989; Downey 1995).

Another theory which explains the relationship between sibship size and child outcomes is the confluence model (Zajonc and Markus 1975; Zajonc 1976, 2001). This theory attempts to explain why on average IQ decreases with birth parity. The theory is that parents have less and less time to invest in each subsequent child, which hinders that child's intellectual development. Conversely the oldest children in the family have to take on a parental role with the younger children, and having to teach them solidifies their own intellectual development.

There are a large number of studies examining the effect of sibship size on children's outcomes, and the findings are ambiguous in the effect of the sibship size. Regarding the health outcomes, evidence supporting the quantity-health tradeoff theory is found in several studies. Children in larger families are found to have to a higher probability of injury or death from accidents (Cummings et al. 1994; Scholer et al. 1997; Schwartz et al. 2005). Henderson et al. (2008) find an increase in the number of siblings can lead to a decrease in weight-by-age of children in Indonesia. In Glick et al. (2007), the sibship size is found to have a negative effect on children's health, which is measured by height-by-age in Romania. Sarin (2004) and Rosenzweig and Zhang (2009) find a negative effect of family size on child health in India and China, respectively. However, Peters et al. (2014) do not find a statistically significant effect of sibship size on children's risk for diseases and the number of sick days children experienced in rural Bangladesh. Moreover, several medical studies find a negative relation between family size and children's diseases. For example, Kruger et al. (2006) show that children from large families are less likely to be obese. Rona et al. (1997) and Ponsonby et al. (1998) find children with a large number of siblings are less likely to suffer from asthma.

Many existing studies focus on educational outcomes as a measure of children's quality. This natural focus on education occurs because education is an important dimension of human development and data on education are more often available from household and individual surveys. Similar to the health outcome, empirical findings on the effect of sibship size on education are not consistent. Several studies show that sibship size has a negative impact on educational attainment (e.g., Belmont and Marolla 1973; Rosenzweig and Wolpin 1980; Blake 1981; Booth and Kee 2009; Park and Chung 2012; Li et al. 2008; Rosenzweig and Zhang 2009). However, other studies find a negligible and insignificant effect on education (e.g., Lee 2008, Cáceres-Delpiano 2006; Black et al. 2005; Dayioğlu et al. 2009; Angrist et al. 2010; De Haan 2010; Yucel and Yuan 2015). Qian (2009) finds that having an additional child has a positive effect on the school enrollment of the first-born child in rural China.

Compared with health and education, the relation between sibship size and child labor has received less attention. There are few studies, which find positive effects of the number of siblings on the probability of work as well as the time that children spend on productive activities (e.g., Patrinos and Psacharopoulos 1997; Cigno and Rosati 2002; Deb and Rosati 2004; Ponczek and Souza 2011).

There are several studies which look at the relation between family or household size and children's outcomes in Vietnam. As mentioned in the previous section, Anh et al. (1998) and Jernström (2017) find a negative effect of family size on children's education and health. In other studies which look at children's outcomes, for example Nguyen (2016), Giang et al. (2017) and Arouri et al. (2019), family or household size is used as an explanatory variable. These studies show consistent findings that children in larger families tend to have lower education and health status. Giang et al. (2017) find that family size is significantly and positively correlated with the probability of working of children aged from 7 to 14.

Most studies in other countries focus on an outcome of children, either education or health. Moreover, there is little evidence on the effect of sibship size on child labor. For the case of Vietnam, to our knowledge, there are no studies which explicitly examine

the effect of sibling size on children's outcomes in Vietnam. Thus, this study aims to fill the research gap by looking at the effect of sibship size on more comprehensive outcomes of children in Vietnam, including school enrolment, disability and labor supply.

3 The Data Set

This study uses data from the 15% sample of the Vietnam Population and Housing Census (VPHC), which was conducted by the General Statistics Office of Vietnam in April 2009 with technical support from United Nations Population Fund (UNFPA).[2] The 2009 VPHC consists of two modules of questionnaires. The first module is used to collect basic demographic and housing data for the whole population of individuals and households throughout the country. Individual data includes age, gender, race and education of individuals.

The second module is used to collect more detailed information on individuals and households. For instance, the individual data contains information on not only basic demographics but also on education, employment, basic assets of the individuals and housing conditions of each household. In particular, this census contains some data on the disability of individuals. Lastly, this module was administered to a random sample of 15% of the population. Vietnam has 684 districts, and each district is divided into enumeration areas, which are hamlets in rural areas and wards in urban areas. Depending on the population size of a district and the average population of the enumeration areas in this district, a number of numeration areas were randomly selected for the 2009 VPHC. All of the households in the selected numeration areas were surveyed. The 15% sample is designed so that data can be representative at the district level. This sample covered 3,692,042 households with 14,177,590 individuals.

The sibship size is equal to the number of children and adults having the same biological mother. In the 2009 VPHC, women from 15 to 49 were asked about the total number of children that they had. Moreover, women were also asked about the number of children living as well as not living currently with them. Thus, households without any woman from 15 to 49 are deleted from our analysis. It means that families with single fathers are dropped. As mentioned, we use twins as the instrument for the sibling size. We can only identify twins who have the same biological mother using information from a module on fertility which was applied for woman from 15 to 49. In a family, twins or other multiples are those who were born in the same month of the same year by the same biological mother. In addition, households who have other children not living at home are also removed in order to have an accurate measure of sibship size. In other words, the sample is restricted to children in households in which the number of children born is equal to the number of children currently living in that same household.

In this study, children are restricted from 6 to 18 years old. Children older than 18 are dropped to avoid the possibility that children with disability are more likely to live with parents at older ages (or children without disability are less likely to live with

[2] UNFPA has assisted in testing, piloting, training and monitoring the survey process, including the design of the questionnaires,

parents). In the 2009 VPHC, there are no data on disability collected for children below 5 years old. In Vietnam, children start primary school from 6 years old. Thus we focus on children aged from 6.

As discussed in the fifth section, we will use the presence of twins as an instrument of sibship size and restrict the analysis on the sample of children born before the birth of twins. This estimation approach is widely used to ensure the exogeneity of the presence of twins (e.g., Cáceres-Delpiano 2006; Black et al. 2005; Angrist et al. 2010). In addition, twins are dropped from the sample since twins can be associated with poorer health.

The final data set consist of 2,081,903 individuals with the age range from 6 to 18. The number of boys and girls are 1,075,578 and 1,006,325, respectively. The number of children from Kinh is 1,755,903, and the number of children from the remaining ethnic minorities is 326,000. In this data set, around 7.9% of children have a sibship size of one, i.e., they do not have a sibling. Comparatively, 51.9% and 25.7% of children have a sibship size of two and three, respectively. The proportion of children with sibship size of four is 9.6%. Lastly, the proportion of children who have sibship size above four is only 4.8%.

4 Sibling Size and Children's Outcomes in Vietnam

Measuring school enrolment and employment is much easier than disability. Construction of an uncontroversial definition of disability is difficult. According to Haveman and Wolfe (2000), people with disabilities are those who have a limitation in either their mental or physical characteristics that can impair their normal daily and productive activities. A widely used method to measure disability in household surveys and censuses is to ask respondents about their difficulties in basic functional domains such as seeing, hearing, walking, remembering, self-care, and communication. This disability measurement is designed by the Washington Group on Disability Statistics, established by United Nations Statistical Division with the participation of over 100 National Statistical Offices and international agencies (Madans et al. 2010). This approach can provide an internationally comparable way to measure and collect data on disability. Compared with other self-reported data on disability, this approach provides a 'more comprehensive and inclusive measure of disability' (Schneider 2009).

The 2009 VPHC follows the Washington Group on Disability Statistics in measuring disability. In the 2009 VPHC, interviewees are asked about their difficulties in the four basic functions including seeing, hearing, walking, and remembering. There are four multiple exclusive responses which are as follows: (i) no difficulty, (ii) some difficulty, (iii) a lot of difficulty and (iv) cannot do at all. In the 2009 VPHC, there are no data on causes as well as the time of disabilities. In this study, we define a child with difficulties in a function domain if she or he reports one of (ii) some difficulty, (iii) a lot of difficulty, (iv) cannot do at all. We also tried to measure the impact of sibship size on disability defined by higher difficulty level (a lot of difficulty or cannot do at all). The results are very similar to those using the lower difficulty level. For interpretation in this paper, we will use the definition of disability using the lower difficulty level.

While this approach captures the majority of children with physical and sensory disabilities, as well as those with significant intellectual disabilities, it probably does miss a non-trivial number of children with developmental disabilities or psychosocial issues. That should be taken into account when interpreting the results (Cappa et al. 2015).

Table 1 presents outcomes of children aged 6 to 18 by their sibship size (the sibship size of a person is equal to the number of her or his siblings plus herself or himself). The proportion of children with difficulties in seeing and the proportion of children with difficulties in hearing are 0.74 and 0.38%, respectively. The proportion of children with difficulties in walking and the proportion of children with difficulties in remembering are 0.41 and 0.65%, respectively.

Regarding education outcomes, the 2009 VPHC contains information on education enrolment and the number of completed grades. There are no data on educational performance in school. Thus, we measure the educational attainment of children by school enrolment and the number of completed grades. Table 1 shows that the average number of completed grades is 6.0 and the schooling rate is 85.4%. Regarding the labor supply, we measure it by the working status 'whether a child has been working during the past 7 days for money earning'. It should be noted that there are no data on working collected for children below 15. Thus, in this study, the working status is analysed for the sample of individuals aged 15–18. It shows that 35.3% of individuals aged 15–18 were working.

There is a correlation between the outcomes of children and their sibship size. The proportion of children with difficulties in seeing is lower for children with larger sibship size than those with lower sibship size. However, the proportion of children with difficulties in other domains tends to be lower for children with lower sibship size. It could be that vision problems are more noticeable when children go to school and have difficulties reading or seeing the board from the back of the class. This is actually consistent with other work from Vietnam that shows that while the presence of disabilities other than mild or moderate vision difficulties among adults is associated with poorer outcomes, those with more mild vision difficulties are more likely to have positive outcomes, quite possibly because they have jobs that demand use of their eyes in ways that lower skilled workers do not experience (Mont and Cuong 2011). Table 1 shows that children with a high number of siblings have lower education and higher employment rates than those with a low number of siblings.

Table 2 presents disability, education and employment disaggregated by other demographic characteristics. Boys are more likely to have a disability than girls. Boys also have lower education and higher working rates than girls. The Kinh majority group tends to have a lower rate of disability and higher education than ethnic minorities.[3] There is a strong correlation between children's outcomes and the mother's education.

[3] According our estimates from the 2009 VHPC, Vietnam consists of 54 ethnic groups, in which Kinh or Vietnamese accounts for 85% of the total population. Compared with ethnic minorities, Kinh people have higher income and living standards (e.g., Lanjouw et al. 2017).

Table 1 Disability, education and labor participation of children by sibship size

Sibship size	Rate of having difficulty in seeing (in percent)	Rate of having difficulty in hearing (in percent)	Rate of having difficulty in walking (in percent)	Rate of having difficulty in remembering (in percent)	The number of completed grades	School enrolment rate (in percent)	The working rate (in percent)
1	0.756	0.322	0.335	0.568	5.43	90.50	26.91
	(0.034)	(0.018)	(0.018)	(0.023)	(0.01)	(0.11)	(0.36)
2	0.777	0.291	0.333	0.524	6.02	90.41	27.00
	(0.021)	(0.007)	(0.007)	(0.010)	(0.01)	(0.06)	(0.20)
3	0.727	0.502	0.523	0.843	6.26	82.24	39.75
	(0.018)	(0.012)	(0.012)	(0.016)	(0.01)	(0.11)	(0.24)
4	0.723	0.520	0.506	0.794	6.03	74.53	49.23
	(0.029)	(0.021)	(0.020)	(0.025)	(0.01)	(0.18)	(0.34)
5	0.465	0.486	0.466	0.758	5.46	68.72	57.36
	(0.035)	(0.031)	(0.031)	(0.043)	(0.02)	(0.33)	(0.56)
6	0.545	0.583	0.473	0.806	4.68	61.97	67.00
	(0.060)	(0.048)	(0.040)	(0.066)	(0.04)	(0.45)	(0.69)
Total	0.743	0.382	0.406	0.649	6.00	85.43	35.32
	(0.015)	(0.006)	(0.006)	(0.008)	(0.01)	(0.08)	(0.20)
Number of obs.	2,081,903	2,081,903	2,081,903	2,081,903	2,081,903	2,081,903	2,081,903

The working rate is computed for individuals aged 15 to 18. Disability and education variables are computed for children aged 6 to 18

Standard errors in parentheses

Author's estimation from the 2009 VPHC

Table 2 Disability, education and labor participation of children by demography

Groups	Rate of having difficulty in seeing (in percent)	Rate of having difficulty in hearing (in percent)	Rate of having difficulty in walking (in percent)	Rate of having difficulty in remembering (in percent)	The number of completed grades	School enrolment rate (in percent)	The working rate (in percent)
Gender							
Girl	0.780	0.334	0.361	0.564	6.04	86.75	31.99
	(0.019)	(0.008)	(0.008)	(0.010)	(0.01)	(0.08)	(0.21)
Boy	0.708	0.427	0.448	0.728	5.96	84.20	38.36
	(0.016)	(0.008)	(0.008)	(0.011)	(0.01)	(0.09)	(0.21)
Age							
6–10	0.461	0.323	0.372	0.587	3.02	97.32	
	(0.012)	(0.009)	(0.009)	(0.013)	(0.00)	(0.04)	
11–14	0.863	0.414	0.406	0.668	7.07	88.56	
	(0.022)	(0.009)	(0.009)	(0.011)	(0.01)	(0.09)	
15–18	0.1055	0.433	0.475	0.730	8.52	60.24	35.32
	(0.029)	(0.012)	(0.013)	(0.016)	(0.01)	(0.18)	(0.20)
Ethnicity							
Ethnic minorities	0.490	0.561	0.528	0.733	5.12	75.36	65.83
	(0.017)	(0.019)	(0.015)	(0.022)	(0.02)	(0.19)	(0.37)
Kinh majority	0.790	0.349	0.384	0.633	6.16	87.29	30.44
	(0.018)	(0.006)	(0.007)	(0.009)	(0.00)	(0.08)	(0.20)
Highest education of mother							
No degree	0.557	0.467	0.459	0.774	5.30	74.03	56.07
	(0.014)	(0.011)	(0.010)	(0.015)	(0.01)	(0.13)	(0.25)
Primary education	0.709	0.362	0.401	0.630	6.30	89.56	30.54
	(0.016)	(0.008)	(0.008)	(0.011)	(0.00)	(0.06)	(0.21)
Lower-secondary	0.1148	0.293	0.328	0.477	6.67	94.93	14.09
	(0.048)	(0.014)	(0.015)	(0.018)	(0.01)	(0.07)	(0.23)
Post secondary	0.1455	0.206	0.279	0.379	5.98	98.67	3.20
	(0.094)	(0.019)	(0.022)	(0.025)	(0.01)	(0.05)	(0.16)
Total	0.743	0.382	0.406	0.649	6.00	85.43	35.32
	(0.015)	(0.006)	(0.006)	(0.008)	(0.01)	(0.08)	(0.20)
Number of obs.	2,081,903	2,081,903	2,081,903	2,081,903	2,081,903	2,081,903	2,081,903

The working rate is computed for individuals aged 15 to 18. Disability and education variables are computed for children aged 6 to 18

Standard errors in parentheses

Author's estimation from the 2009 VPHC

Children with mothers that have a lower level of education are more likely to have high disability, low education and high working rates. A correlation between mother's education and disability rates has also been found in Bangladesh and Bhutan (Mont et al. 2014; Khan et al. 2011).

5 Estimation Method

To measure the impact of the number of siblings on an individuals' disability, we start with a simple linear model:

$$Y_{i,j} = \beta_0 + Sibship_{i,j}\beta_1 + C_i\beta_2 + X_{i,j}\beta_3 + \varepsilon_{i,j}, \quad (1)$$

where: $Y_{i,j}$ is an outcome indicator of child i in family j, $\text{Sibship}_{i,j}$ is his or her sibship size, and C_i is a vector of child-level control variables, and $X_{i,j}$ is a vector of family-level control variables such as mother's education, that include gender, age, ethnicity, age and education of mother. We tend to use exogenous and pre-determined control variables, which are not affected by the number of children (Angrist and Pischke 2008). The use of control variables is to increase the estimation efficiency. We also tried the models without control variables, and the estimation results are very similar to the models with control variables. The summary statistics of these control variables are presented in Table 7 in the Appendix.

The main problem in estimating β_1 is the endogeneity of the sibship size. In other words, a child's outcome and the number of siblings might be jointly determined. Omitted variables in eq. (1), such as household economic characteristics, can affect both sibship size and children's outcomes; as a result, the estimation of the impact of sibship size can be biased. There can be a large number of unobserved variables, and thus we are not usually able to predict the bias direction of OLS estimators. For instance, parents who pay more attention to children tend to have few children and invest more in children's healthcare and nutrition at the same time. As a result, the sibship size can be negatively correlated with children's disability status. On the other hand, parents with poor health might have few children but also less investment in their children. In this case, the sibship size can be positively correlated with the children's disability. The endogeneity bias also affects the estimation of the effect of sibship size on education and employment of children.

To address this problem, we use instrumental-variable regressions to measure the impact of sibship size. A valid instrument should correlate with sibship size but not the error term in the disability equation. One popular instrument for sibship size in empirical studies is the presence of twins (Rosenzweig and Wolpin 1980; Schultz 2005; Cáceres-Delpiano 2006; Black et al. 2005; Angrist et al. 2010; De Haan 2010; Mogstad and Wiswall 2016). Twins provide an exogenous increase in the sibship size. Similar to empirical studies in this literature such as Black et al. (2005), Angrist et al. (2010), Mogstad and Wiswall (2016), we use the presence of later twins as the instrument of sibship size to measure the effect of sibship size on outcomes of children born before twins. It means that children who are born after twins are excluded. This is because the presence of twins is an exogenous shock to children born before the twins but not to children born after the twins. The final sample which is used for the estimation is not fully representative of the population. As a result, the effect estimated using the sample of children before twins might be different from the effect on children in the full sample. Thus, we should be cautious in interpreting the estimates for the whole population.

The first stage regression in 2SLS is expressed as follows:

$$Sibship_{i,j} = \theta_0 + Twin_{i,j}\theta_1 + C_i\theta_2 + X_{i,j}\theta_3 + u_{i,j}, \quad (2)$$

Having twins can be considered as a random event that causes sibship size to increase. A recent concern with the randomness of twin is fertility treatments. Parents who used certain fertility techniques such as artificial insemination and vitro maturation might be more likely to have twins. However, this issue is not an issue for the case of Vietnam for at least two reasons. Firstly, fertility techniques are very expensive, and few people are able to use the services. Secondly, we use 'having a twin sibling' as the instrument to measure the effect on children born before twins. It means that parents have at least a child before having twin children. If parents already have at least a child, they are less likely to use fertility techniques. In this study, the proportion of children born before twins within the age range of 6 to 18 having twin siblings is 0.5%.

It should be noted that although dependent variables in eq. (1) are binary we estimate eq. (1) by applying two-stage least squares (2SLS) regression for the linear probability model, as there are no available estimators for a binary model with endogenous count variables. 2SLS estimators are consistent and can be applied for the binary model with count endogenous variables (e.g., see Angrist 2001; Angrist and Krueger 2001; Cáceres-Delpiano 2006; Angrist et al. 2010).

An important issue is non-linearity in the impact of children. Since we have a large data set, we are able to see whether the impact of increasing from one to two children is different from the impact of increasing from two to three children, and so on. Thus, we try different samples of children as follows. Firstly, we estimate the effect of the sibship size on the first-born children using the sample of first-born children who have at least a younger sibling (i.e., the sibship size is larger than one). The instrument is the presence of twins at the second birth. Secondly, we estimate the effect of the sibship size on the outcomes of the first-born and second-born children who have sibship size larger than two. In other words, the sample is limited to children in families with at least three siblings. The presence of twins at the third birth is the instrument for the sibship size. Thirdly, we also use the similar identification strategy. We use the presence of twins at the fourth birth to estimate the impact of sibship size on the first-born, second-born and third-born children who have the sibship size is larger than three (The sample is limited to children in families with at least four siblings). We are not able to go further to estimate the effect of sibship size for sibship size larger than four, since there are only a few observations of children with the sibship size larger than four, and the presences of twins at the fifth birth is very small.

6 Empirical Results

We start with the OLS regressions, which are presented in Tables 8, 9, 10, 11 in Appendix. The OLS regressions show children with larger sibship sizes tend to have higher rates of disability, lower education, and a higher probability of working. Since the OLS results can be biased, we will use the 2SLS results for interpretation. The first-stage regressions of sibship size on the instrumental variables and control variables are presented in Table 11 in the Appendix. As expected, the presence of twins increases sibship size strongly, and having the first-born son reduces sibship size. The t-statistic of twins and a firstborn's

gender is very high. Furthermore, we also perform a Cragg-Donald weak identification test of the instruments; the statistic is very high, indicating that the instruments are very strong.[4]

Table 3 presents the 2SLS regressions using the sample of first-born children. Compared with OLS, 2SLS produces smaller magnitude and less significant estimates of the effect of sibship size. There are no significant effects of the number of children on disability of the first-born children. This finding is similar to Peters et al. (2014), which do not find a statistically significant effect of sibship size on children's sickness in Bangladesh.

The effect on the number of completed grades has a negative sign but not significant. The point estimate of sibship size on the completed grades is also smaller than other studies such as Booth and Kee (2009) (the effect estimate at −0.024) and Rosenzweig and Zhang (2009) (the estimate at 0.23). The effect of sibship size on the working status is not statistically significant at the conventional level. However, there is a significant effect on the school enrolment. Having an additional sibling reduces the probability of schooling by 0.016 or 1.6 percentage points. This effect estimate is slightly lower than the effect estimate found in Park and Chung (2012), which find that the sibship size reduces the school enrollment rate by 0.029 in Bangladesh. The negative effect of sibship size on education enrollment in China that is found by Li et al. (2008) is −0.03. Although the point estimate of the effect on sibship size on school enrolment is lower in Vietnam than in Bangladesh (Park and Chung 2012) and China (Li et al. 2008), the differences are not statistically significant once the standard errors are taken into account.

Table 4 presents the 2SLS regressions using the sample of first-born and second-born children. The sample is limited to children in families with at least three siblings. Thus, the sample in this table is smaller than that in Table 3. The results are very similar to the effect of the first-born children. Only the effect of sibship size on school enrolment is significant. Having an additional sibling is estimated to reduce the probability of school enrolment by around 0.006 or 0.6 percentage points.

In Table 5, we still do not find significant effects of sibship size on disability and employment. The sample is limited to children in families with at least four siblings. Again, we find a larger (negative) effect of sibship size on education. We find a significant effect on the number of completed grades. For children who have at least four siblings, having an additional sibling reduces the number of completed grades by 0.14 (or around 23% of the mean of the number of completed grades). This effect is rather similar to the effect found in Rosenzweig and Zhang (2009) for the case of China. The effect of sibship size on school enrolment is also larger, at 0.03 or 3.0 percentage points. This implies an increasing effect of sibship size on education of children. When parents already have a large number of children, having an additional one can cause them more difficulties in affording education for their children.

In addition, we also include interactions between sibship size and other characteristics including gender, age, urbanity and the education level of the mother, to examine whether the effect of sibship size differs among these groups (Table 6). The instrumental variables for these interactions are interactions between twins and the characteristics of the individuals (sex and age of children, age and the education level of the

[4] We used command ivreg2 in Stata to estimate the 2SLS regression and performance tests. As a rule of thumb, if a F- statistic is under 10, the instruments might be weak (Staiger and Stock 1997).

Table 3 2SLS regression of children's outcomes on sibship size: sample of the first-born children

Explanatory variables	Having difficulty in Seeing	Having difficulty in Hearing	Having difficulty in Walking	Having difficulty in Remember-ing	The number of completed grades	School enrolment (yes = 1, no = 0)	Work (yes = 1, no = 0)
Sibship size	−0.000163	0.000223	−0.000473	−0.000122	−0.012144	−0.016052**	0.018403
	(0.001639)	(0.002752)	(0.001992)	(0.001097)	(0.030329)	(0.006798)	(0.017279)
Sex (male = 1, female = 0)	0.001365***	0.001448***	−0.001143***	0.001092***	−0.120831***	−0.029333***	0.055256***
	(0.000324)	(0.000551)	(0.000412)	(0.000230)	(0.006333)	(0.001413)	(0.004250)
Age	0.000107	0.000625***	0.000516***	0.000072	0.685070***	−0.036757***	0.071706***
	(0.000082)	(0.000140)	(0.000104)	(0.000061)	(0.001738)	(0.000397)	(0.001365)
Kinh majority (yes = 1, no = 0)	0.000934**	−0.002369***	0.000001	−0.001743***	0.352819***	0.032012***	−0.206919***
	(0.000435)	(0.000805)	(0.000544)	(0.000353)	(0.013059)	(0.002414)	(0.006472)
Age of mother	0.000112***	0.000456***	0.000331***	0.000058**	−0.011676***	−0.003089***	−0.002309***
	(0.000029)	(0.000057)	(0.000047)	(0.000027)	(0.000636)	(0.000131)	(0.000355)
Schooling years of mother	−0.000217***	0.000116	0.000613***	−0.000187***	0.103438***	0.024303***	−0.038146***
	(0.000073)	(0.000131)	(0.000101)	(0.000050)	(0.001633)	(0.000340)	(0.001049)
Constant	0.000220	−0.007787	−0.012232**	0.003732	−2.924168***	1.278455***	−0.358911***
	(0.004175)	(0.007097)	(0.005185)	(0.002841)	(0.078904)	(0.017743)	(0.051383)
Observations	870,769	870,769	870,769	870,769	823,375	823,375	299,165
R-squared	0.001	0.001	0.002	0.001	0.721	0.219	0.182

Regression of working uses the sample of individuals aged 15 to 18, and regressions of disability and education variables use the sample of children aged 6 to 18. The sample is limited to children in families with at least two siblings

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Author's estimation from the 2009 VPHC

Table 4 2SLS regression of children's outcomes on sibship size: sample of the first-born and second-born children

Explanatory variables	Having difficulty in Seeing	Having difficulty in Hearing	Having difficulty in Walking	Having difficulty in Remember-ing	The number of completed grades	School enrolment (yes = 1, no = 0)	Work (yes = 1, no = 0)
Sibship size	−0.001478	−0.003920	−0.001527	−0.001361	−0.062048	−0.006022*	−0.011348
	(0.001664)	(0.002903)	(0.002116)	(0.001385)	(0.056978)	(0.003586)	(0.021064)
Sex (male = 1, female = 0)	0.003010***	0.004311***	0.000323	0.002567***	−0.211427***	−0.045703***	0.069156***
	(0.000293)	(0.000490)	(0.000345)	(0.000266)	(0.007369)	(0.001502)	(0.003064)
Age	0.000316***	0.000872***	0.000507***	0.000257***	0.617386***	−0.053678***	0.076197***
	(0.000066)	(0.000116)	(0.000084)	(0.000059)	(0.002160)	(0.000431)	(0.001204)
Kinh majority (yes = 1, no = 0)	0.000188	−0.002808***	−0.000290	−0.002060***	0.386390***	0.019144***	−0.186683***
	(0.000529)	(0.000974)	(0.000654)	(0.000484)	(0.021438)	(0.003827)	(0.007660)
Age of mother	−0.000055	0.000013	0.000038	−0.000099***	−0.013067***	−0.004068***	0.000941**
	(0.000035)	(0.000063)	(0.000045)	(0.000032)	(0.000908)	(0.000183)	(0.000372)
Schooling years of mother	0.000167**	0.000197	0.000483***	0.000068	0.167371***	0.033256***	−0.037523***
	(0.000077)	(0.000135)	(0.000097)	(0.000064)	(0.002719)	(0.000487)	(0.001082)
Constant	0.006837	0.017848*	0.002024	0.010575**	−2.264558***	1.448837***	−0.443165***
	(0.005487)	(0.009622)	(0.007011)	(0.004621)	(0.189036)	(0.035290)	(0.077612)
Observations	565,901	565,901	565,901	565,901	544,561	544,561	217,896
R-squared	0.001	0.001	0.001	0.001	0.605	0.262	0.158

Regression of working uses the sample of individuals aged 15 to 18, and regressions of disability and education variables use the sample of children aged 6 to 18. The sample is limited to children in families with at least three siblings

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Author's estimation from the 2009 VPHC

Table 5 2SLS regression of children's outcomes on sibship size: sample of the first-born, second-born and third-born children

Explanatory variables	Having difficulty in Seeing	Having difficulty in Hearing	Having difficulty in Walking	Having difficulty in Remember-ing	The number of completed grades	School enrolment (yes = 1, no = 0)	Work (yes = 1, no = 0)
Sibship size	0.004723	−0.001046	−0.000100	0.001802	−0.143042**	−0.029539**	0.011555
	(0.003305)	(0.003996)	(0.002855)	(0.002884)	(0.071790)	(0.014631)	(0.025574)
Sex (male = 1, female = 0)	0.002582***	0.004684***	0.000820*	0.002488***	−0.205842***	−0.039146***	0.065971***
	(0.000408)	(0.000682)	(0.000488)	(0.000381)	(0.011842)	(0.002366)	(0.004040)
Age	0.000223**	0.000943***	0.000512***	0.000230**	0.554973***	−0.064406***	0.073590***
	(0.000103)	(0.000145)	(0.000098)	(0.000090)	(0.002755)	(0.000528)	(0.001666)
Kinh majority (yes = 1, no = 0)	0.002167**	−0.001800	0.000222	−0.000862	0.483280***	0.009666*	−0.159457***
	(0.001016)	(0.001503)	(0.000993)	(0.000917)	(0.032350)	(0.005775)	(0.010110)
Age of mother	−0.000085*	−0.000291***	−0.000058	−0.000152***	−0.010907***	−0.002624***	0.000237
	(0.000051)	(0.000083)	(0.000058)	(0.000044)	(0.001532)	(0.000294)	(0.000542)
Schooling years of mother	0.000240*	0.000308*	0.000473***	0.000119	0.213073***	0.035165***	−0.032323***
	(0.000124)	(0.000184)	(0.000137)	(0.000107)	(0.003640)	(0.000666)	(0.001251)
Constant	−0.018078	0.017268	0.000049	−0.000643	−1.537834***	1.615656***	−0.477934***
	(0.013825)	(0.016774)	(0.011871)	(0.012109)	(0.303479)	(0.061909)	(0.118214)
Observations	253,229	253,229	253,229	253,229	244,358	244,358	96,386
R-squared	0.001	0.001	0.001	0.001	0.529	0.274	0.144

Regression of working uses the sample of individuals aged 15 to 18,and regressions of disability and education variables use the sample of children aged 6 to 18. The sample is limited to children in families with at least four siblings

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Author's estimation from the 2009 VPHC

mother).[5] In the 2SLS regressions, estimates of interaction terms are small and not statistically significant, suggesting that the effect of sibship size does not differ significantly across children with different characteristics.

Finally, the regressions show the effects of other control variables on children's outcomes. Boys are more likely to have disabilities than girls. Boys tend to have lower education and higher working rates than girls. Kinh children and those with more highly educated mothers have lower disability and working rates, and higher education than ethnic minority children and those with lowly-educated mothers.

7 Conclusion

This study provides evidence for the link between children's quantity and various outcomes including disability, education and working status. To estimate the effect of the number of children, we use the data from the 15% sample of the Vietnam Population and Housing Census in 2009. In addition to education and employment data, this census contains data on individuals' limitations in different human functionalities, including hearing, seeing, mobility, and cognition, which allows for a measure of disability consistent with the UN Statistical Division's recommendations on census questions for identifying people with disabilities. An advantage of the census is that it covers a large proportion of the population. In addition, there are a large number of observations that include twins; therefore, the presence of twins can be used as an instrumental variable for sibship size.

We do not find significant effects of sibship size on disability. It implies that the number of children does not have a large negative effect on the investment of parents in their children's health. From these results, however, no clear policy prescription emerges. However, we find a negative and significant effect on the education of children. Although this effect is not large, it tends to be increasing in the number of children. Having an additional sibling reduces the probability of schooling of the first-born children by 1.6 percentage points in the sample of children from families with at least a child. This effect is increasing to around 3.0 percentage points in the sample of children from families with at least three children. In the larger families, the effect of sibship size on the number of completed grades is also higher and more significant. This finding suggests that education of children in families with a large number of children should be paid more attention. Possible policy prescriptions would be to provide subsidies for school expenses to large families, however a side effect of that could be to provide an incentive for additional births. Given the overall expenses of having children, however, it is unlikely that effect would be significant.

The negative effect of sibship size on education implies that the decline in the fertility rate in 1980s and 1990s due to the family planning policy might contribute to the increase in educational attainment in Vietnam. In recent years, the proportion of households having more than two children has increased (Minh 2016). This trend might reduce the progress in education development in Vietnam.

[5] We also tried the two instruments (twins and sex of the firstborn) in the regressions with interactions. It means that instruments for interactions between sibship size and individual characteristics are interactions between the two instruments and individual characteristics. The estimation results are similar to those using one instrument of twins. Therefore, we report the estimates using one instrument 'presence of twins'.

Table 6 2SLS regression of the number of school enrollment

Explanatory variables	Dependent variable is school enrollment				
	Model 1	Model 2	Model 3	Model 4	Model 5
Sample of first born children from households with two or more children					
Sibship size	−0.00371	−0.00257	0.01841	−0.01160	−0.02974
	(0.00924)	(0.02629)	(0.02382)	(0.04715)	(0.01963)
Sibship size * Male	−0.02333				
	(0.01984)				
Sibship size * Age		−0.00108			
		(0.00243)			
Sibship size * Kinh majority			−0.03790		
			(0.02501)		
Sibship size * Mother's age				−0.00012	
				(0.00140)	
Sibship size * Mother's schooling years					0.00168
					(0.00201)
Sample of first and second born children from households with three or more children					
Sibship size	−0.00586	−0.00211	−0.03703	−0.15301*	0.00618
	(0.01324)	(0.03744)	(0.03077)	(0.08227)	(0.02746)
Sibship size * Male	−0.00034				
	(0.02063)				
Sibship size * Age		−0.00029			
		(0.00320)			
Sibship size * Kinh majority			0.03531		
			(0.03270)		
Sibship size * Mother's age				0.00397	

Table 6 (continued)

Explanatory variables	Dependent variable is school enrollment				
	Model 1	Model 2	Model 3	Model 4	Model 5
				(0.00320)	
Sibship size * Mother's schooling years					−0.00181
					(0.00345)
Sample of first, second and third born children from households with four or more children					
Sibship size	−0.04572**	0.01945	−0.03595	0.06037	−0.03428
	(0.01842)	(0.04832)	(0.02844)	(0.11105)	(0.02498)
Sibship size * Male	0.03784				
	(0.02758)				
Sibship size * Age		−0.00360			
		(0.00387)			
Sibship size * Kinh majority			0.00953		
			(0.03299)		
Sibship size * Mother's age				−0.00234	
				(0.00296)	
Sibship size * Mother's schooling years					0.00097
					(0.00386)

This table presents the coefficient estimates of sibship size and its interactions with other explanatory variables. The instrumental variable for sibship size is the presence of twins. Instrumental variables for interactions between sibship size and an explanatory variable such as age and gender are interactions between the presence of twins with the explanatory variable. The explanatory variables are the same as the regressions in previous tables. Full regressions are not reported in this paper

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Author's estimation from the 2009 VPHC

Although we do not find a significant effect of sibship size on child disability, we find a conformational finding that mother's education is linked to the presence of disability in children. Higher levels of mother's education may lead to better nutrition and health as well as more early childhood interventions that can decrease the prevalence of disability. This provides evidence not only for an increased focus on children's education, but also educational programs for mothers offered through the health or social protection systems.

A limitation of the data set used in this study is that there are no detailed data on health status and educational performance. Although sibship size has no significant effects on disability of children, it might have a significant effect on other health outcomes such as height and weight. There are few household surveys in Vietnam which contain data on health status, educational performance, cognitive and non-cognitive skills of children. However, these surveys have a small sample size, and cannot be used to construction the instrumental variable of 'twins'. Exploring the effect of sibship size on other outcomes is out of scope of this study but certainly very important for future studies.

Appendix

Table 7 Summary statistics of independent variables

Variables	Type	Mean	Std. Dev.	Min	Max
Dependent variables					
Having difficulty in seeing	Binary	0.0074	0.0859	0	1
Having difficulty in hearing	Binary	0.0038	0.0617	0	1
Having difficulty in walking	Binary	0.0041	0.0636	0	1
Having difficulty in remembering	Binary	0.0065	0.0803	0	1
The number of completed grades	Discrete	5.9976	2.7932	0	12
School enrolment (yes = 1, no = 0)	Binary	0.3532	0.4780	0	1
Work (yes = 1, no = 0)	Binary	0.8543	0.3528	0	1
Explanatory variables					
Sex (male = 1; female = 0)	Binary	0.5168	0.4997	0	1
Age	Discrete	11.814	3.635	6	18
Kinh (kinh = 1; ethnic minorities = 0)	Binary	0.8434	0.3634	0	1
Mother's age	Discrete	36.865	5.448	21	49
Mother's education years	Discrete	7.0975	3.7195	0	16
Observations		2,081,903			

Author's estimation from the 2009 VPHC

Table 8 OLS regression of children's outcomes on sibship size: sample of the first-born children

Explanatory variables	Having difficulty in Seeing	Having difficulty in Hearing	Having difficulty in Walking	Having difficulty in Remember-ing	The number of completed grades	School enrolment (yes = 1, no = 0)	Work (yes = 1, no = 0)
Sibship size	0.001974***	0.001949***	−0.000051	0.001515***	−0.238924***	−0.053536***	0.054753***
	(0.000161)	(0.000280)	(0.000195)	(0.000143)	(0.004787)	(0.000859)	(0.001487)
Sex (male = 1, female = 0)	0.001705***	0.001722***	−0.001076***	0.001352***	−0.158410***	−0.035544***	0.063321***
	(0.000184)	(0.000347)	(0.000274)	(0.000159)	(0.003853)	(0.000838)	(0.001877)
Age	0.000009	0.000546***	0.000497***	−0.000003	0.695810***	−0.034981***	0.069608***
	(0.000036)	(0.000068)	(0.000053)	(0.000035)	(0.000956)	(0.000231)	(0.000935)
Kinh majority (yes = 1, no = 0)	0.001389***	−0.002001***	0.000091	−0.001394***	0.302871***	0.023757***	−0.197008***
	(0.000255)	(0.000538)	(0.000338)	(0.000262)	(0.010480)	(0.001879)	(0.004452)
Age of mother	0.000121***	0.000463***	0.000333***	0.000065**	−0.012800***	−0.003275***	−0.001907***
	(0.000029)	(0.000056)	(0.000046)	(0.000027)	(0.000617)	(0.000125)	(0.000297)
Schooling years of mother	−0.000130***	0.000186***	0.000631***	−0.000121***	0.093800***	0.022710***	−0.036065***
	(0.000031)	(0.000069)	(0.000059)	(0.000025)	(0.000952)	(0.000179)	(0.000353)
Constant	−0.005155***	−0.012131***	−0.013294***	−0.000387	−2.348059***	1.373679***	−0.460647***
	(0.000798)	(0.001608)	(0.001245)	(0.000729)	(0.020254)	(0.004237)	(0.017536)
Observations	870,769	870,769	870,769	870,769	823,375	823,375	299,165
R-squared	0.001	0.001	0.002	0.001	0.724	0.223	0.185

Regression of working uses the sample of individuals aged 15 to 18,and regressions of disability and education variables use the sample of children aged 6 to 18

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Author's estimation from the 2009 VPHC

Table 9 OLS regression of children's outcomes on sibship size: sample of the first-born and second-born children

Explanatory variables	Having difficulty in Seeing	Having difficulty in Hearing	Having difficulty in Walking	Having difficulty in Remember-ing	The number of completed grades	School enrolment (yes = 1, no = 0)	Work (yes = 1, no = 0)
Sibship size	−0.000124	−0.000377	−0.000252	0.000093	−0.265784***	−0.041578***	0.035396***
	(0.000181)	(0.000339)	(0.000218)	(0.000180)	(0.007413)	(0.001247)	(0.001849)
Sex (male = 1, female = 0)	0.003095***	0.004533***	0.000402	0.002658***	−0.224552***	−0.047994***	0.072890***
	(0.000275)	(0.000460)	(0.000320)	(0.000253)	(0.006249)	(0.001336)	(0.002568)
Age	0.000277***	0.000769***	0.000470***	0.000215***	0.623250***	−0.052655***	0.075037***
	(0.000046)	(0.000082)	(0.000058)	(0.000044)	(0.001420)	(0.000303)	(0.001098)
Kinh majority (yes = 1, no = 0)	0.000525	−0.001926***	0.000027	−0.001699***	0.334470***	0.010081***	−0.173515***
	(0.000339)	(0.000681)	(0.000391)	(0.000347)	(0.015669)	(0.002754)	(0.004961)
Age of mother	−0.000060*	−0.000000	0.000033	−0.000105***	−0.012329***	−0.003939***	0.000888**
	(0.000034)	(0.000062)	(0.000044)	(0.000032)	(0.000874)	(0.000179)	(0.000369)
Schooling years of mother	0.000216***	0.000325***	0.000529***	0.000120***	0.159817***	0.031937***	−0.035390***
	(0.000049)	(0.000086)	(0.000062)	(0.000041)	(0.001630)	(0.000288)	(0.000518)
Constant	0.002420**	0.006286***	−0.002133	0.005832***	−1.595774***	1.565585***	−0.607967***
	(0.001126)	(0.002071)	(0.001381)	(0.001118)	(0.035875)	(0.006890)	(0.021773)
Observations	565,901	565,901	565,901	565,901	544,561	544,561	217,896
R-squared	0.001	0.001	0.001	0.001	0.608	0.265	0.163

Regression of working uses the sample of individuals aged 15 to 18,and regressions of disability and education variables use the sample of children aged 6 to 18

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Author's estimation from the 2009 VPHC

Table 10 OLS regression of children's outcomes on sibship size: sample of the first-born, second-born and third-born children

Explanatory variables	Having difficulty in Seeing	Having difficulty in Hearing	Having difficulty in Walking	Having difficulty in Remember-ing	The number of completed grades	School enrolment (yes = 1, no = 0)	Work (yes = 1, no = 0)
Sibship size	−0.000319	0.000002	−0.000572**	0.000358	−0.251663***	−0.029199***	0.024369***
	(0.000248)	(0.000472)	(0.000271)	(0.000255)	(0.011583)	(0.001880)	(0.002513)
Sex (male = 1, female = 0)	0.002436***	0.004714***	0.000806*	0.002446***	−0.209010***	−0.039136***	0.066283***
	(0.000397)	(0.000676)	(0.000481)	(0.000373)	(0.011617)	(0.002326)	(0.004001)
Age	0.000337***	0.000919***	0.000522***	0.000263***	0.557341***	−0.064414***	0.073364***
	(0.000068)	(0.000111)	(0.000074)	(0.000058)	(0.002269)	(0.000423)	(0.001610)
Kinh majority (yes = 1, no = 0)	0.000786	−0.001513	0.000093	−0.001257***	0.452901***	0.009761**	−0.155487***
	(0.000495)	(0.001000)	(0.000613)	(0.000486)	(0.025042)	(0.004190)	(0.006575)
Age of mother	−0.000050	−0.000299***	−0.000055	−0.000142***	−0.010164***	−0.002626***	0.000194
	(0.000047)	(0.000079)	(0.000054)	(0.000041)	(0.001443)	(0.000279)	(0.000538)
Schooling years of mother	0.000083	0.000341**	0.000458***	0.000074	0.209619***	0.035176***	−0.031830***
	(0.000073)	(0.000141)	(0.000108)	(0.000065)	(0.002833)	(0.000479)	(0.000791)
Constant	0.002826*	0.012922***	0.002005	0.005346***	−1.084799***	1.614241***	−0.534685***
	(0.001705)	(0.003040)	(0.001921)	(0.001728)	(0.068583)	(0.012173)	(0.032925)
Observations	253,229	253,229	253,229	253,229	244,358	244,358	96,386
R-squared	0.001	0.001	0.001	0.001	0.530	0.274	0.145

Regression of working uses the sample of individuals aged 15 to 18,and regressions of disability and education variables use the sample of children aged 6 to 18

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Author's estimation from the 2009 VPHC

Table 11 OLS first-stage regressions of sibship size

Explanatory variables	Dependent variable is sibship size					
	Sample of first born children aged 6–18 from families with at least two children	Sample of first born children aged 15–18 from families with at least two children	Sample of first and second born children aged 6–18 from families with at least three children	Sample of first and second born children aged 15–18 from families with at least three children	Sample of first, second and third born children aged 6–18 from families with at least four children	Sample of first, second and third born children aged 15–18 from families with at least four children
Having twin siblings at the second birth	0.8234***	0.7517***				
	(0.0074)	(0.0149)				
Having twin siblings at the third birth			0.8387***	0.8013***		
			(0.0123)	(0.0193)		
Having twin siblings at the fourth birth					0.9228***	0.9025***
					(0.0265)	(0.0380)
Sex (male = 1, female = 0)	−0.1590***	−0.2218***	−0.0626***	−0.0797***	−0.0290***	−0.0248***
	(0.0017)	(0.0033)	(0.0022)	(0.0038)	(0.0037)	(0.0061)
Age	0.0462***	0.0578***	0.0291***	0.0249***	0.0226***	0.0179***
	(0.0004)	(0.0015)	(0.0005)	(0.0015)	(0.0007)	(0.0022)
Kinh majority (yes = 1, no = 0)	−0.2142***	−0.2735***	−0.2505***	−0.2834***	−0.2725***	−0.3086***
	(0.0063)	(0.0097)	(0.0075)	(0.0108)	(0.0106)	(0.0145)
Age of mother	−0.0045***	−0.0112***	0.0037***	0.0010*	0.0068***	0.0032***
	(0.0002)	(0.0005)	(0.0003)	(0.0006)	(0.0006)	(0.0009)
	−0.0408***	−0.0573***	−0.0361***	−0.0456***	−0.0314***	−0.0386***

Table 11 (continued)

Explanatory variables	Dependent variable is sibship size					
	Sample of first born children aged 6–18 from families with at least two children	Sample of first born children aged 15–18 from families with at least two children	Sample of first and second born children aged 6–18 from families with at least three children	Sample of first and second born children aged 15–18 from families with at least three children	Sample of first, second and third born children aged 6–18 from families with at least four children	Sample of first, second and third born children aged 15–18 from families with at least four children
Schooling years of mother	(0.0004)	(0.0006)	(0.0006)	(0.0009)	(0.0010)	(0.0014)
Constant	2.5143***	2.7980***	3.2608***	3.5246***	4.1456***	4.4207***
	(0.0083)	(0.0273)	(0.0110)	(0.0307)	(0.0187)	(0.0471)
Observations	870,769	299,165	565,901	217,896	253,229	96,386
R-squared	0.149	0.131	0.112	0.095	0.100	0.093
Cragg-Donald weak identification test	9923	2897	5614	1513	3047	974

This table reports the first-stage results of a 2SLS estimation of the effect of the number of children on the probability of a child having a disability. In this first stage, the dependent variable is the number of siblings and the instrumental variable is a dummy variable indicating whether there are twins

Regression of working uses the sample of individuals aged 15 to 18, and regressions of disability and education variables use the sample of children aged 6 to 18

Robust standard errors are in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Author's estimation from the 2009 VPHC

References

Angrist, J. (2001). Estimation of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice. *Journal of Business & Economic Statistics, 29*(1), 1–28.

Angrist, J., & Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives, 15*(4), 69–85.

Angrist, J., & Pischke, J. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Angrist, J., Lavy, V., & Schlosser, A. (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics, 28*(4), 773–823.

Anh, T. S., Knodel, J., Lam, D., & Friedman, J. (1998). Family size and children's education in Vietnam. *Demography, 35*(1), 57–70.

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

Becker, G. S. (1960). An economic analysis of fertility. In G. S. Becker (Ed.), *Demographic and Economic Change in Developed Countries*. Princeton: Princeton University press.

Becker, G. S., & Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy, 81*(2), 279–288.

Becker, G. S., & Tomes, N. (1976). Child endowments and the quantity and quality of children. *Journal of Political Economy, 84*(4), 143–162.

Belmont, L., & Marolla, F. A. (1973). Birth order, family size, and intelligence. *Science, 182*(4117), 1096–1101.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics, 120*(2), 669–700.

Blake, J. (1981). Family size and the quality of children. *Demography, 18*(4), 421–442.

Blake, J. (1989). *Family size and achievement*. Berkeley: University of California Press.

Booth, A. L., & Kee, H. J. (2009). Birth order matters: The effect of family size and birth order on educational attainment. *Journal of Population Economics, 22*(2), 367–397.

Cáceres-Delpiano, J. (2006). The impacts of family size on investment in child quality. *Journal of Human Resources, 41*(4), 738–754.

Cappa, C., Petrowski, N., & Njelesani, J. (2015). Navigating the landscape of child disability measurement: A review of available data collection instruments. *ALTER-European Journal of Disability Research/Revue Européenne de Recherche sur le Handicap, 9*(4), 317–330.

Cigno, A., & Rosati, F. C. (2002). Child labour education and nutrition in rural India. *Pacific Economic Review, 7*(1), 65–83.

Cummings, P., Theis, M. K., Mueller, B. A., & Rivara, F. P. (1994). Infant injury death in Washington state, 1981 through 1990. *Archives of Pediatrics & Adolescent Medicine, 148*(10), 1021–1026.

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

Dayioğlu, M., Kirdar, M. G., & Tansel, A. (2009). Impact of sibship size, birth order and sex composition on school enrolment in urban Turkey. *Oxford Bulletin of Economics and Statistics, 71*(3), 399–426.

De Haan, M. (2010). Birth order, family size and educational attainment. *Economics of Education Review, 29*(4), 576–588.

Deb, P., & Rosati, C. (2004), Estimating the effect of fertility decisions on child labor and schooling. Understanding Children's Work Programme Working Paper. Available at SSRN: https://ssrn.com/abstract=1780272, Estimating the Effect of Fertility Decisions on Child Labor and Schooling

Downey, D. B. (1995). When bigger is not better: Family size, parental resources, and children's educational performance. *American Sociological Review, 60*, 746–761.

Filmer, D. (2008). Disability, poverty, and schooling in developing countries: Results from 14 household surveys. *The World Bank Economic Review, 22*(1), 141–163.

Giang, L. T., Nguyen, C. V., Tran, T. Q., & Thieu, V. (2017). Does firm agglomeration matter to labor and education of local children? Evidence in Vietnam. *Child Indicators Research, 10*(4), 1015–1041.

Glick, P. J., Marini, A., & Sahn, D. E. (2007). Estimating the consequences of unintended fertility for child health and education in Romania: An analysis using twins data. *Oxford Bulletin of Economics and Statistics, 69*(5), 667–691.

GSO. (1999). *Population and housing census in 1999*. Hanoi: General Statistics Office.

GSO. (2018). *Statistical year book 2017, general statistical Office of Vietnam (GSO)*. Hanoi: Publisher of Statistics.

Haveman, R., & Wolfe, B. (2000). The economics of disability and disability policy. *Handbook of Health Economics, 1*, 995–1051.

Henderson, D. J., Millimet, D. L., Parmeter, C. F., & Wang, L. (2008). Fertility and the health of children: A nonparametric investigation. In *Modelling and Evaluating Treatment Effects in Econometrics* (pp. 167–195). Emerald Group publishing limited.

ILO, MOLISA and GSO. (2014). Viet Nam national child labour survey 2012 Main findings. In *Joint report from International Labour Organization (ILO), Ministry of Labour, Invalids and Social Affairs of Vietnam (MOLISA) and General Statistics Office of Vietnam (GSO)*. Hanoi.

Jernström, L. (2017). *The marginal effects of family size on Children's health in Vietnam–does the effect of siblings differ by parity?* Master Thesis: Faculty of Social Sciences, University of Helsinki, Finland.

Khan, N. Z., Ferdous, S., Mahbub, M., Mobarak, R., Parveen, M., Mont, D., & Durkin, M. (2011). The impact of maternal and household characteristics on childhood impairments and disabilities in Bangladesh. *Bangladesh Journal of Child Health, 35*(2), 41–48.

Kruger, R., Kruger, H. S., & Macintyre, U. E. (2006). The determinants of overweight and obesity among 10- to 15-year-old schoolchildren in the north West Province, South Africa–the THUSA BANA (transition and health during urbanisation of south Africans; BANA, children) study. *Public Health and Nutrition, 9*(3), 351–358.

Lanjouw, P., Marra, M., & Nguyen, C. (2017). Vietnam's evolving poverty index map: Patterns and implications for policy. *Social Indicators Research, 133*(1), 93–118.

Lee, J. (2008). Sibling size and investment in children's education: An Asian instrument. *Journal of Population Economics, 21*(4), 855–875.

Li, H., Zhang, J., & Zhu, Y. (2008). The quantity-quality trade-off of children in a developing country: Identification using Chinese twins. *Demography, 45*(1), 223–243.

Madans, J. H., Loeb, M., & Altman, B. (2010). Measuring disability and monitoring the UN convention on the rights of persons with disabilities: The work of the Washington group on disability statistics. *BMC Public Health, 11*(Suppl 4), 1–8.

Minh, T. 2016. "Tỷ Lệ Sinh Con Thứ 3 Tiếp Tục Tăng", Sai Gon Giai Phong newspaper, 30/7/2016. Available at: http://www.sggp.org.vn/ty-le-sinh-con-thu-3-tro-len-tiep-tuc-tang-26330.html. Accessed 20 June 2017

Mitra, S. (2006). The capability approach and disability. *Journal of disability policy studies, 16*(4), 236–247.

Mogstad, M., & Wiswall, M. (2016). Testing the quantity–quality model of fertility: Estimation using unrestricted family size models. *Quantitative Economics, 7*(1), 157–192.

Mont, D., & Cuong, N. V. (2011). Disability and poverty in Vietnam. *The World Bank Economic Review, 25*(2), 323–359.

Mont, D., Cappa, C., Bahadur, C. N., Dorji, C., Hingst, G., & Khan, N. (2014). Poverty, maternal education and child disability: Evidence from Bhutan. *Disability and International Development, 3*, 13–20.

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

Park, C., & Chung, W. (2012). Sibship size, birth order, and children's education in developing countries: Evidence from Bangladesh. *Hitotsubashi Journal of Economics, 53*, 1–23.

Patrinos, H. A., & Psacharopoulos, G. (1997). Family size, schooling and child labor in Peru–an empirical analysis. *Journal of Population Economics, 10*(4), 387–405.

Peters, C., Rees, D. I., & Hernández-Julián, R. (2014). The trade-off between family size and child health in rural Bangladesh. *Eastern Economic Journal, 40*(1), 71–95.

Pham, B. N., Hill, P. S., Hall, W., & Rao, C. (2012). The evolution of population policy in Viet Nam. *Asia-Pacific Population Journal, 27*(2), 41–56.

Ponczek, V., & Souza, A. (2011). New evidence of the causal effect of family size on child quality in a developing country. *Journal of Human Resources, 47*(1), 64–106.

Ponsonby, A. L., Couper, D., Dwyer, T., & Carmichael, A. (1998). Cross sectional study of the relation between sibling number and asthma, hay fever, and eczema. *Archives of Disease in Childhood, 79*(4), 328–333.

Qian, N. (2009). Quantity-quality and the one child policy. In *The only-child disadvantage in school enrollment in rural China, NBER working papers no. 14973*. Massachusetts: National Bureau of economic research.

Rona, R. J., Duran-Tauleria, E., & Chinn, S. (1997). Family size, atopic disorders in parents, asthma in children, and ethnicity. *Journal of Allergy Clinical Immunology, 99*(4), 454–460.

Rosenzweig, M. R., & Wolpin, K. I. (1980). *Testing the quantity-quality fertility model: The use of twins as a natural experiment* (pp. 227–240). XLVIII: Econometrica.

Rosenzweig, M. R., & Zhang, J. (2009). Do population control policies induce more human capital investment? Twins, birth weight and China's "one-child" policy. *The Review of Economic Studies, 76*(3), 1149–1174.

Sarin, A. (2004). Are children from smaller families healthier? Examining the causal effects of on child welfare. Paper presented at the population association meetings in Boston, April 3. Harris Policy School, University of Chicago.

Schneider, M. (2009). The difference a word makes: Responding to questions on 'disability' and 'difficulty' in South Africa. *Disability and Rehabilitation, 31*(1), 42–50.

Scholer, S. J., Mitchel, E. F., & Ray, W. A. (1997). Predictors of injury mortality in early childhood. *Pediatrics, 100*(3), 342–347.

Schultz, T. P. (2005). Effects of fertility decline on family well being: Opportunities for evaluating population programs. Working Paper, http://www.econ.yale.edu/~pschultz/.

Schwartz, S., Eidelman, A. I., Zeidan, A., Applebaum, D., & Raveh, D. (2005). Childhood accidents: The relationship of family size to incidence, supervision, and rapidly seeking medical care. *Israeli Medical Association Journal, 7*(9), 558–563.

Staiger, D., & Stock, J. (1997). Instrumental variables regression with weak instruments. *Econometrica, 65*(3), 557–586.

WHO. 2011. *World Report On Disability*, WHO press, World Health Organization, 20 avenue Appia, 1211 Geneva 27, Switzerland.

Yucel, D., & Yuan, A. V. (2015). Do siblings matter? The effect of siblings on socio-emotional development and educational aspirations among early adolescents. *Child Indicators Research, 8*(3), 671–697.

Zajonc, R. B. (1976). Family configuration and intelligence. *Science, 19*, 227–236.

Zajonc, R. B. (2001). The family dynamics of intellectual development. *American Psychologist, 56*(6/7), 490–496.

Zajonc, R. B., & Markus, G. (1975). Birth order and intellectual development. *Psychological Review, 82*, 74–88.