Sibship Size and Health Outcomes in Later Life among the Mexican Elderly

Takashi Yamashita

Department of Economics
Reed College

3203 SE Woodstock Blvd.
Portland, OR 97202-8199

e-mail: takashiy@reed.edu

April 2008

Abstract

This paper investigates whether the number of siblings is associated with health outcomes in the elderly population, using the Mexican Health and Aging Study (MHAS). The main questions the paper tries to answer are (i) whether the association between the number of siblings and health outcomes exists, and if so, (ii) whether such association remains significant even after controlling for childhood and adult characteristics, such as childhood socioeconomic status, educational attainment and work history. Empirical estimates suggest that sibship size matters in predicting incidence of cancer (for both men and women), respiratory illness, and stroke (for men) while it does not show any association for other illnesses. As the sibship size is positively related to adult height, it seems unlikely that parental resource dilution for a large number of children is the mechanism through which childhood circumstances affect health in later life. Instead, as the number of siblings appears to be mostly associated with incidence of disease with infectious etiology, early exposures to certain infectious agents may play a role in transmitting the family background to health outcomes later in life. These findings confirm results found in the epidemiological literature.

JEL Classification Codes: I1, J1, D1

Keywords: sibship size, sibling rivalry, life course models, pathway models, health status, aged

I. Introduction

Socioeconomic conditions during early-childhood years are shown to be related to health conditions in adult years and mortality later in life. Children from poorer households tend to be less healthy, and children with poorer health have disadvantages in education as they tend to end up with lower educational attainment (Case, Lubotsky, and Paxon, 2002). In turn, educational attainment is known to be strongly correlated with health status of adults (REF). Thus the health status of elderly individuals may have an origin in the family environment in which these individuals grew up.

One of the socioeconomic conditions that are thought to be related to children's well-being is the number of siblings. A large body of literature has documented that children born in large families tend to have lower educational attainment (e.g., Butcher and Case 1994) and somewhat poorer health (e.g., Garg and Murdoch 1998). Much attention has been paid to the association between sibship size and education, education and adult health, or childhood environment and adult health, but there have been relatively few studies that have addressed the specific question of how the number of sibling is related to health outcomes in later life.

In this paper, I link these two strands of the literature and examine the relationship between sibship size and health outcomes later in life. I use large-scale population-based survey data collected in Mexico. The data set, Mexican Health and Aging Study (MHAS), is suitable for this study, as it is one of a few survey data sets of the elderly population that contain information on the number of siblings born alive to mothers of the survey respondents. The survey also collected detailed information on family background and health and economic conditions of the respondent.

Uncovering the relationship between health and sibship size is important as findings here may provide insight into a mechanism of possible transmission of childhood environment to adult health. Despite the mounting empirical evidence on the relationship between childhood circumstances and health in later life, there is little consensus on the relative importance of mechanisms that lead from lower socioeconomic status in childhood to poor health in later life. If a family size is a proxy for parental socioeconomic status, then individuals who grew up in a large family may have poorer health overall. On the other hand, if sibship size is associated with some health conditions but not with others, and has little explanatory power for health in general, the mechanism of transmission that lead from childhood socioeconomic status to adult health may be traced to factors that are specific to certain illnesses.

I find that sibship size matters in predicting the incidence of some diseases while it does not show any association with other illnesses. The diseases that exhibit a relationship with the number of siblings are cancer, respiratory illness and stroke for men and cancer for women. As the sibship size is positively related to adult height, it seems unlikely that parental resource dilution for a large number of children is the mechanism through which childhood circumstances affect health in later life. Instead, because the number of siblings appears to be mostly associated with diseases with infectious etiology, early exposures to certain infectious agents may play a role in transmitting the family background to health outcomes later in life. Findings in this paper thus confirm results found in the epidemiological literature.

II. Sibship Size, Educational Attainment and Health

The positive association between schooling and health is one of the most robust patterns found across different countries and generations. Not as universal, the association between sibship size and education is also found in many different cultures, particularly in poorer parts of the world (Strauss and Thomas 1995). Confronted with these two associations, one cannot help thinking if sibship size is related to health outcomes later in life, and if so, whether the relationship is transmitted through education, or if the number of siblings would have an association with adult health independent of education.

Consistent with a large body of literature, a strong, positive relationship between health and education is also found in the Mexican data. Using data of individuals aged 50 and older from the MHAS, panel (a) of figure 1 plots the average of self-rated health (1=poor to 5=excellent) by the years of schooling for men and women. There is clearly a strong, almost linear, association between self-reported health and years of schooling: both men and women who have 13 or more years of education report their health "Good," nearly a full one point better than those with no schooling who rate their health as "Fair."

The association of educational attainment with sibship size is also present in the older cohort of Mexicans. Panel (b) of figure 1 presents the average years of schooling by the number of siblings by gender. There is clearly a negative, if not linear, relationship between the sibship size and average years of schooling. Men born into a family with 13 or more siblings on average have a one-year shorter schooling compared to those with three or fewer siblings. For women, the relationship is similar, as those with a large number of siblings (13 or more) have nearly two-years shorter schooling than those born into a small family (one to two siblings). The main question in this paper is whether there is a direct association of the sibship size with health outcomes beyond what is accounted for by education and other childhood and adulthood socioeconomic characteristics.

There are two main hypotheses to explain the correlation between childhood circumstances and adult health.[1] The life-course model emphasizes that childhood circumstances have long-

---

[1] There is a third explanation, namely the fetal-origin hypothesis, which links nutrition *in vitro* to health in later life. However, since we do not have data on nutrient intake of mothers during pregnancy, or the detailed birth place information, it is impossible to test whether *in vitro* malnutrition could play a role in the sample individuals' health.

term consequences in adult health, directly through illness and deficiency in important nutrients and partly by hindering educational attainment and opportunities in life. In contrast, the pathway model posits that the effects of childhood environment on adult health are through intermediate steps of income and socioeconomic characteristics and the observed correlation between lower socioeconomic status and health in adulthood is not directly attributable to childhood circumstances (e.g., Brunner et al. 2005, Hambleton at al. 2005, Marmot et al. 2001). According to this theory, childhood environment is important because it affects initial adult socioeconomic position but would not have independent impact on adult health beyond what can be explained by education and adult socioeconomic statues.

These two theories provide testable hypotheses on the relationship between sibship size and adult health. We could first examine whether or not sibship size and other childhood characteristics have association with health outcomes in a multivariate regression framework. We then analyze the data by adding further controls that may be associated with parental and childhood background, such as education, socioeconomic characteristics in adulthood and occupations. If we find association between sibship size and health in later life above and beyond education and adult characteristics, there may be an independent association of the number of siblings with health. If so, such relationships may offer a key to understanding how childhood environment is transmitted to health conditions later in life.

III. Empirical Framework

There is a negative and significant association of sibship size with educational attainment, as well as strong and significant associations of educational attainment with the self-reported health status of the Mexican elderly. The question is whether the number of siblings would have an independent relationship with health in adult life, beyond its effects through education.

Following Case et al (2005), I model a measure of adult health (h_A) as linear functions of vectors of age (**A**), parental education (**P**), socioeconomic conditions in childhood ($\mathbf{e_C}$), sibship size (**S**), educational attainment (**E**), and socioeconomic and labor market characteristics during adult life ($\mathbf{e_A}$):

$$h_A = \beta_0 + \mathbf{A}\boldsymbol{\beta}_\mathbf{A} + \mathbf{P}\boldsymbol{\beta}_\mathbf{P} + \mathbf{S}\boldsymbol{\beta}_\mathbf{S} + \mathbf{e}_\mathbf{C}\boldsymbol{\beta}_\mathbf{C} + \mathbf{E}\boldsymbol{\beta}_\mathbf{E} + \mathbf{e}_\mathbf{A}\boldsymbol{\beta}_\mathbf{A} + \varepsilon, \quad (1)$$

where ε is an error term.

Parameter estimates of (1) can be used to test whether the life-course model or the pathway model can explain better health outcomes later in life. If the pathway model is true, childhood circumstances affect adult outcomes through their effects on education, work and adult socioeconomic status. If so, we would expect estimates of coefficient related to siblings and childhood characteristic ($\boldsymbol{\beta}_\mathbf{S}, \boldsymbol{\beta}_\mathbf{C}$) to be zero, after controlling for education and adult socioeconomic characteristics. If we find significant association of siblings and childhood characteristics with health in adulthood independent of one's education and later-life characteristics, then childhood characteristics may have a direct impact on adult health.

Epidemiological research suggests that the relationship between the number of siblings and health outcomes is not necessarily negative. Sometimes large family size and later birth favor health, as early exposures to some infectious agents may only lead to mild symptoms, while a delay in exposure to adolescence may severely increase the risk of certain diseases. Therefore, the direction of association between adult health outcomes and sibship size may vary depending on the type of illness, and a single measure of health outcomes may not be appropriate to study the relationship between sibship size and health outcomes. I will thus study the relationship between family structure and various health outcomes by looking at the probability of diagnosis with as many conditions as contained in the data.

IV. Data and Variable Construction

I use the first wave of the Mexican Health and Aging Study (MHAS) conducted in 2001. The MHAS is a two-wave longitudinal study of Mexican men and women who were born prior to 1951. About 11,000 households with at least one age-eligible individual were selected for interviews. Spouses and partners of the sample individuals residing in the same household were also interviewed regardless of their age. The sample was designed to be nationally representative of the population aged 50 or older, while individuals residing in the Mexican states with high rates of out-migration to the United States were oversampled.

Data were collected on multiple domains of health as well as retrospective information including childhood health and living conditions, the U.S. migration history, and about parents. Parental information includes educational attainment of both mothers and fathers. In addition to a question on self-reported health status, the respondents are asked if they have been diagnosed by health professionals with certain illnesses such as cancer, hypertension, diabetes and arthritis. The data on siblings are the number of siblings born alive to the sample individuals' mothers and the number of siblings alive at the time of the interview. Figure 2 presents the distribution of the number of siblings estimated from the 2001 MHAS. While the distribution of sibship size is right-skewed, a relatively small fraction of the sample individuals are born to family with 12 or more siblings and it is most common to have five to seven siblings. Only about two-thirds of siblings survived to the time of interviews of the sample individuals (appendix table). Unfortunately, the MHAS does not ask questions on birth order or the sex composition of the siblings. Although birth order is considered to be an important determinant in educational outcomes (Black, Devereux and Salvanes 2005) and health (Karmaus and Botezan 2002), I am not able to study the effect of birth order due to the data limitation.

In the full model, I include the respondents' height and body mass index (BMI) to control for factors that may be related to physiological aspects of health but may not be captured by childhood environment or socio-economic status of sample individuals. While it is desirable to use measured height and weight to construct the BMI, height and weight are measured only for a small subset of respondents.[2] The MHAS, however, has collected self-reported height and weight from a larger number of sample individuals. Although reported height and weight can be used to construct the BMI, such numbers may be subject to reporting errors.[3] I address this measurement error problem by using the strategy similar to Antecol and Bedard (2006). More specifically, I regress measured true weight and height on reported weight and height and their squared values and age, age squared, years of education separately for men and women. I then use the coefficient estimates to predict true height and weight for a subset of sample individuals who reported valid height and weight figures but from whom measurements were not taken. I then calculate the BMI from these predicted height and weight figures. The predicted values of height and BMI are used in all regressions hereafter.[4]

After deleting observations with missing data,[5] the final analysis sample contains 2,515 men and 3,037 women, of ages 50 to 95 (men) and 50 to 98 (women) in 2001. Summary statistics of key variables used for analysis are presented in the appendix table. Outcome variables are years of education (in years) and a battery of binary variables of health measures indicating whether the respondents consider their health good or better and whether they have been diagnosed with

---

2 With valid responses on measured height and weight, the BMI can be calculated for only 16% of the male sample and 16.8% of the female sample.

3 Antecol and Bedard (2006) compare the US NHIS and NHANSE III and report the extent of underreporting.

4 Results do not change appreciably if the BMI constructed from self-reported weight and heights are used.

5 Respondents with proxy interviews are excluded from analysis. Further, observations are excluded if answers to one of the following questions are "don't know," refusal or missing: age, number of siblings born alive, years of schooling, and self-reported health.

certain conditions by health professionals. I group control variables depending on the timing of occurrence in the respondents' lives. For example, parental education (indicator variables for mother's and father's education) is mostly predetermined prior to birth; the presence of siblings may affect health at an early age as well as the indicator for access to in-house toilet at age 10 as well as experience of major health problems before age 10. Educational attainment is often a combination of parental and individual choice. Controls for adulthood and work history characteristics include a full set of dummies for marital status, occupation, self-employment status, current and past smoking and drinking status, and own- and spouse's living experience in the United States as well as the number of years worked. Since the survey-based data measure not actual incidence of a disease but a diagnosis of it, access to health insurance is an important determinant of one's knowledge of his or her health conditions. I therefore include a dummy variable whether one has health insurance in the set of adult SES variables. Dummy variables for quartiles of total household income and wealth are also included in the adult SES control. Finally, (predicted) anthropometric measures (height and BMI) are added, as they are considered to be important predictors of certain diseases.

Figure 3 illustrates simple relationships between years of education and nine health outcome measures for men and women. In some cases, clear relationships are discerned from the unconditional scatter plots, while in others, the relationship is not clear. For example, both men and women with higher education seem to have a higher prevalence of cancer, while they are less likely to have experienced a stroke. Men with longer years of schooling seem to exhibit a lower prevalence of liver/kidney infections, while the relationship among women is opposite but highly nonlinear. Figure 4 demonstrates the relationship between the number of siblings and the same health measures. Here the relationships are even less clear, while one may argue there may be positive relationships between the sibship size and arthritis, liver and kidney infections and heart attack.

These unconditional plots can be misleading in that they are not adjusted for family background, demographic and other adulthood characteristics that may be influencing the prevalence of these conditions. I therefore estimate the relationship between sibship size and health outcomes in a regression framework. It turns out that after controlling for adulthood characteristics, much of the apparent association of education and sibship size with health outcomes disappears. However, in certain diseases, the association remains strong, although the adulthood characteristics explain a large part of the variation in the data.

V. Estimation Results

(a) Relationship between Sibship Size and Education

First, I estimate the statistical association between the number of siblings and educational attainment in the data. Years of schooling are regressed on age, age squared, dummy variables for parental education and childhood characteristics, and the number of siblings. The results are reported in table 1. There is a strong and negative association between the sibship size and educational attainment of both men and women, while the relationship seems stronger for men. For a man with six siblings, an additional sibling will reduce his years of schooling by nearly one-third of a year, while for a comparable woman an additional sibling will shorten her education by one-quarter of a year. The estimates of the quadratic terms of sibling variables are jointly significant for both genders. While the association between the number of siblings and education seems strong, the variation in data explained by sibship size is rather small: the incremental contribution that the sibling variables make to R^2 is minimal, at 2.5 percent for men and 1.2 percent for women.

In explaining variation in years of schooling, maternal education appears to account for a great deal of differences within the data. Individuals born to mothers with some elementary school education have 1.7 years longer schooling than those with mothers with no formal education. The association is even bigger for individuals with mothers who have more than elementary school education. In addition, close to 60 percent of explained variation is accounted for by the two dummy variables indicating mother's education for both men and women. Such a large contribution by maternal education stands out against a relatively minor contribution of paternal education. The literature suggests that mother's education is an important determinant of children's health in some cultures (e.g., Thomas 1994, Desai and Alva 1998). The estimates from Mexican data confirms that maternal education is a crucial determinant in children's human capital investment.

(b) Relationship between Sibship Size and Health Outcomes

Next, I investigate the relationship between sibship size and self-reported health. Self-reported health has been shown to predict mortality and is known to be strongly correlated with one's socioeconomic status (see references cited in Case et al. 2005). I use the binary measure of self-reported health, assigning the value of 1 if a respondent reports his or her health is good, very good, or excellent and 0 otherwise and estimate its association with individual characteristics by logit. Table 2 reports the results of such regressions, all of which include age, age squared, and four indicators of parental education. Age is to control for age effect and parental education is largely pre-determined before birth. Column (1) adds the quadratic function of the number of siblings, and in column (2), conditions before age 10 are included to control for childhood socioeconomic status. Such controls include: indicators for having experienced a major health problem before age 10, having had a toilet in house at age 10, and the fraction of surviving siblings to the total number of siblings born alive. Column (3) adds own educational attainment, and the adulthood socioeconomic characteristics including indicators for income and wealth quartiles and work history controls are added in column (4). Finally, specification (5) includes predicted height and BMI. As additional variables are added sequentially, I report the measures of incremental contribution of a set of variables to pseudo-R^2, computed as a fraction of the total pseudo-R^2 for the full model (5).

In general, the association of sibship size with self-reported health is negative for the relevant range, and the relationship between education and health is positive, as expected. For both men and women, adult socioeconomic characteristics account for a large fraction of variation in self-reported health status. These similarities not withstanding, other aspects of the results for men and women are in stark contrast. For men (panel (a)), none of the sibling variables are significant either individually or jointly. Furthermore, adding the sibling variables contributes only 3.2 percent of the total pseudo-R^2 of the full model. Therefore, the number of siblings seems to account for very little in explaining men's self-reported health status.

While the relationship of self-reported health with sibship size is weak, the association of self-reported health with education levels is strong for men. Although the estimates of education coefficients become somewhat smaller as more controls are added to regression analysis, the magnitude of change is relatively small and the estimates remain highly significant. More concretely, a man with 12 years of schooling is 23 percentage points more likely to report his health "good" or better compared to men with no education (evaluated at the means of other variables). Adding controls for educational attainment also improves the fit of the model by increasing explained variation in data by more than one-quarter. Educational attainment is thus an important predictor of men's health independent of their past and present socioeconomic characteristics.

In contrast to the estimates obtained from the male sample, the estimates from the female sample highlight a different relationship between self-reported health and sibship size. In specification (3), the coefficient estimates of the sibling variables jointly become marginally significant (p-value: 0.052). In the full model with adulthood characteristics and anthropometric measures, however, the sibling variables lose statistical significance. In addition, the sibling variables seem to play a larger role as indicated by a large (relative to the male sample) incremental contribution of the sibling variables to the overall explained variation of the data.

With respect to education, its role in explaining women's health is much smaller compared to men's. Although the years of schooling variables are jointly highly significant in specification (3), they become insignificant as adult socioeconomic characteristics and anthropometric measures are added. Their incremental contribution to pseudo-R^2 is only marginally bigger (at 17.6 percent) than that of sibling variables (at 14.7 percent), while the adult socioeconomic characteristics explain more than one-third of the explained data variation.

Although the relationship between sibship size and adult health is weak, it is premature to conclude that childhood conditions operate through educational attainment or adult socioeconomic status. The first column of table 3 reports the coefficient estimates of an indicator variable whether sample individuals had a serious health problem before age 10 from the full model. Even after controlling for a full set of variables, this indicator has a large and significant association with self-reported health. In the regression of years of education on parental and childhood socioeconomic characteristics (not reported here), the estimate of the dummy variable indicating health problem before age 10 is not significant. The childhood health problem, therefore, seems to have an independent association with adult self-reported health.

In tables 4 and 5, I study the relationship among sibship size, education and health outcomes by focusing on diagnoses of specific diseases. While the regressions are estimated for all disease outcomes, only the results of outcomes are presented for which the siblings variables are independently or jointly significant at least at the 10 percent level for either sex. Columns (1), (2), (3) of these tables correspond to specifications in columns (1), (3), (5) of table 2. Again, the results are vastly different by gender as well as by types of illnesses. For men, the larger number of siblings is associated with a higher likelihood of cancer and stroke. For women, the larger number of siblings is associated with a higher likelihood of cancer but not with other conditions. Comparing estimates between genders in the full model (columns (3)), both the male and female samples yield similar estimates for cancer, but the estimates of the sibling variables are vastly different between men and women. These indicate different health production mechanisms at work for men and women.

Despite the apparent large magnitude of some estimates, particularly of stroke for the male sample, I emphasize here that the actual impact of sibling variables is quite small. For example, a man with 6 years of schooling is expected to have had a stroke with a probability of 0.0136 if he has four siblings, and this expected probability increases to 0.0172 with five siblings and to 0.0183 with six siblings. These predicted probabilities are not statistically different from each other. In addition, the marginal contribution of siblings variable to total psuedo-R^2 is small --- it ranges from 0.001 for respiratory illness of women to 0.186 for a stroke for men. These results indicate that the number of siblings has a small impact on one's overall health in old age.

When education and other childhood circumstance controls are added (columns (2)), the fit of regressions generally improves. Furthermore, the adulthood socioeconomic characteristics account for a substantial fraction of explained variation in the data in all conditions, ranging from 43 percent (stroke for men, respiratory illness for women) to 66 percent (stroke for women). Although the data and education and adult socioeconomic status do a much better job of explaining the variation in the data than the sibling variables, changes in the coefficient estimates on sibling controls are generally small. This implies that sibship size and educational attainment and adult socioeconomic status are correlated little in explaining variation in health outcomes and they seem to influence health outcomes separately.

While the relationships between sibship size and health conditions are generally positive, estimates of respiratory illness for men stand out as its association with the number of siblings is mildly negative for a range of the sibship size. For example, for a man with six years of schooling, predicted probability of having been diagnosed with respiratory illness declines from 0.0252 with one sibling, 0.0248 with two siblings, and to 0.246 with three siblings. For a man with 12 years of education, the comparable numbers are 0.0326, 0.0322, and 0.0320. In the epidemiology literature, a larger sibship size is associated with lower risk of atopic disorders (e.g., allergy and asthma), and immune maturation and development of resistance in response to early exposures to infectious agents are considered to be a factor (Karmaus and Botezan 2002). Association of larger sibship size with a lower risk of respiratory illness seems to confirm such findings, although the estimates here are jointly significant only at the 10% level.

In the second to last columns of table 3, I report estimates of the indicator variable for having had severe health problems before age 10. As before, I only report outcomes for which this variable are significant at the 10 percent level for either sex. Here, the results for men and women are more similar in that for both men and women, those who had severe health problems before age 10 are less likely to rate their health good, and a much higher likelihood of diagnosis with arthritis. For some conditions, childhood health conditions do persist beyond their influence through education.

(c) Possible Transmission Mechanism of Family Background to Health

While there is some evidence for the association of health outcomes later in life with sibship size, mechanisms through which the number of siblings early in life would manifest in health conditions later in life are not clear. In economics literature, one of the hypotheses proposed to explain association between sibship size and educational attainment is the resource dilution/liquidity constraint hypothesis. A large number of children are thought to dilute resources available for each child's education. Given imperfect capital markets, parents with many mouths to feed would have to stretch resources across many children. Previous studies indicate that liquidity constraints may be important in determining educational attainment. If the resource dilution/liquidity constraint is also important in explaining the relationship between certain health outcomes and sibship size, we would expect that the association would manifest stronger in health conditions that are correlated with nutrition status in early life.

To answer this question, I explore a relationship between sibship size and a measure of childhood investment in health: height. Figure 5 plots the average height of the MHAS sample individuals against the number of siblings. This relationship is robust even in a multivariate regression framework. Table 6 presents estimates from regressions including a full set of demographic, educational, parental and childhood background controls by sex. For both men and women, an increase in the number of siblings is strongly and significantly associated with taller statue for both men and women. Since a number of epidemiological studies have documented that those from lower social classes or those who experienced economic difficulties in their childhood tend to be shorter (e.g., Kuh and Wadsworth 1989, Silventoinen et al. 1999), we would expect a negative association between sibship size and adult height if the resource dilution/liquidity constraint hypothesis were true. However, the finding here strongly rejects that claim.

Instead of a large sibship size requiring parents to stretch resources over many children, which in turn results in unfavorable health outcomes later in life, the estimates here seem to support the view that early exposures to infectious pathogens are at play from the sibship size to certain health outcomes. Birth order and sibship size have been associated with diseases considered to have an infectious etiology, such as allergies and asthma (Karmaus et al. 2001, Karmaus and Botezan 2002), certain cancers (Westergaard et al. 1997, Chang et al. 2004) and periodontal disease (Mucchi et al. 2004). For instance, to the extent in which tooth loss and periodontal disease are caused by oral bacteria, risk of exposure may increase with a larger sibship size and crowded living conditions in the childhood home. Using data of Swedish twins, Mucchi et al. (2004) report increased risk of tooth loss with the increasing number of siblings. On the other hand, a larger family size may favor health in later life. Infection in early childhood with Epstein-Barr virus is usually asymptomatic. However, if infection is delayed from childhood to adolescence, it can substantially increase the risk of young-adult Hodgkin's Lymphoma. Late birth order and large family size therefore favor earlier exposure to infectious agents. Consistent with this hypothesis, incidence of Hodgkin's Lymphoma is associated with small sibship size and low birth order. I also find that the Mexican men with more siblings appear to have a lower likelihood of having had respiratory illness, although the relationship is weak and valid only for the number of siblings smaller than four. Other health conditions that have displayed association with sibship size may have infectious etiology such as some cancers. Even a stroke, for which sibling variables had the strongest association for the male sample, is considered to have an origin in bacterial infection (Dunne 2004). Although evidence presented here is far from conclusive, statistical associations between the sibship size and certain diseases may provide insight into possible mechanism of transmission.

VI. Conclusions

This paper has presented evidence of how the number of siblings is related to certain health conditions later in life, using the population-based survey data of the Mexican elderly. The sibship size is strongly correlated with one's educational attainment, and one's years of schooling is associated with health conditions among Mexican men and women age 50 and older. However, after controlling for childhood as well as adult socioeconomic characteristics, the association between the number of siblings and health conditions remained strong only for a couple of illnesses. Even with respect to diseases that may have association with sibship size, the overall relationship is small. Rather than the number of siblings, one's adult socioeconomic conditions appear the most important determinants of health conditions later in life.

Of health conditions and illnesses for which sibship size matters, many seem to have infectious etiology. Coupled with the lack of evidence of under-nutrition in large families arising from resource dilution, early exposures to infectious agents may be the culprit of transmission mechanism of early childhood socioeconomic characteristics to adult health. If that is the case, however, the findings in this paper pose more questions than it has provided answers. If the number of siblings matter through early exposure to infectious agents and immune maturation, why do we see the results so vastly different between men and women? Would boys and girls react differently to different types of infectious agents? Further research is needed to explain these unanswered questions.

Furthermore, if early exposures to infectious agents are important, the birth order of a child would be crucial in determining health outcomes in later life. However, as we do not have information on birth order in MHAS, further analysis is limited by the data limitation. On the other hand, there exist data sets in the United States that contain information on birth order, sibship size, and health outcomes in adult and elderly population (for example, Health and

Retirement Study and the 1979 National Longitudinal Study of Youth). Exploring these data sets would be a promising avenue for further research to advance our knowledge of the relationship between early childhood environment and health outcomes later in life.

References

Antecol, H., Bedard, K., 2006. Unhealthy assimilation: why do immigrants converge to American health status levels? Demography 43 (2), 337-360.

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. Quarterly Journal of Economics 120 (2), 669-700.

Brunner, E., Shipley, M.J., Blane, D., Smith, G.D., Marmot, M. G., 1999. When does cardiovascular risk start? Past and present socioeconomic circumstances and risk factors in adulthood. Journal of Epidemiology and Community Health 53 (12), 757-764.

Butcher, K., Case, A. 1994. The effects of sibling sex composition on women's education and earnings. Quarterly Journal of Economics 104 (4), 531-563.

Case, A., Lubotsky, D., Paxon, C., 2002. Socioeconomic status and health in childhood: the origin of the gradient. American Economic Review 92 (5), 1308-1334.

Case, A., Fertig, A., Paxon, C., 2005. The lasting impact of childhood health and circumstances. Journal of Health Economics 24 (2), 365-389.

Chang, E., Montgomery S.M. Richiardi, L., Ehlin, A., Ekbom, A. Lambe M., 2004. Number of siblings and risk of Hodgkin's Lymphoma. Cancer Epidemiology, Biomarkers & Prevention 13 (7), 1236-1243.

Desai, S., Alva, S., 1998. Maternal education and child health: is there a strong causal relationship? *Demography* 35 (1), 71-81.

Garg, A., Morduch, J., 1998. Sibling rivalry and the gender gap: evidence from child health outcomes in Ghana. Journal of Population Economics 11 (4), 471-493.

Dunne, M., 2004. Infectious Agents and Cardiovascular Disease. The Infectious Etiology of Chronic Diseases: Defining the Relationship, Enhancing the Research, and Mitigating the Effects. National Academy Press, Washington.

Hambleton, I.R., Clarke, K., Broome, H. L., Fraser, H.S., Brathwaite, F., Hennis, A.J., 2005. Historical and current predictors of self-reported health status among elderly persons in Barbados. Pan American Journal of Public Health 17(5/6), 342-352.

Hart, C.L., Smith, G.D., 2003. Relation between number of siblings and adult mortality and stroke risk: 25 year follow up of men in the Collaborative study. Journal of Epidemiology and Community Health 57 (5), 385-391.

Karmaus, W., Arshad, H., Mattes, J., 2001. Does the sibling effect have its origin in utero? Investigating birth order, cord blood immunoglobulin E concentration, and allergic sensitization at age 4 years. American Journal of Epidemiology 154 (10), 909-915.

Karmaus, W., Botezan, C., 2002. Does a higher number of siblings protect against the development of allergy and asthma? A review. Journal of Epidemiology and Community Health 56 (3), 209-217.

Kuh, D., Wadsworth, M., 1989. Parental height: childhood environment and subsequent adult height in a national birth cohort. International Journal of Epidemiology 18 (3), 663-668.

Marmot, M., Shipley, M., Brunner, E., Hemingway, H, 2001. Relative contribution of early life and adult socioeconomic factors to adult morbidity in the Whitehall II study. Journal of Epidemiology and Community Health 55(5), 301-307.

Mucci, L.A., Hsieh, C.C., Williams, P.L., Dickman, P.W., Björkman, L., Pedersen, N.L., 2004. Birth order, sibship size and housing density in relation to tooth loss and periodontal disease: a cohort study among Swedish twins. American Journal of Epidemiology 159(5), 499506-915.

Silventoinen, K., Lahelma, E., Rahkonen, O., 1999. Social background, adult body-height and health. International Journal of Epidemiology 28 (5), 911-918.

Strauss, J., Thomas, D., 1995. Human resources : empirical modeling of household and family decisions” Srinivasan, T. N. and Behrman, J. (eds.), Handbook of Development Economics, Vol. 3A, New York and Oxford: Elsevier Science, North Holland, 1883-2023.

Thomas, D., 1994. Life father, like son; like mother like daughter: Parental resources and child height. Journal of Human Resources 29 (4), 950-988.

Westergaard, T., Melbye, M., Pedersen, J.B., Frisch, M., Olsen, J.H., Andersen, P.K., 1997. Birth order, sibship size and risk of Hodgkin’s disease in children and young adults: a population-based study of 31 million person years. International Journal of Cancer 72 (6), 977-981.

Table 1 OLS estimates of Education and Sibship Size and Family Background, by Gender

	Dep Var: years of schooling		
	all	men	women
# of siblings born alive/10	–0.281	–0.092	–0.045
	(0.748)	(1.364)	(0.856)
# of siblings born alive squared/100	–0.306	–0.440	–0.340
	(0.393)	(0.691)	(0.459)
Mother's ed = some elementary	1.739**	1.786**	1.765**
	(0.239)	(0.432)	(0.275)
= more than elementary	2.259**	2.109**	2.325**
	(0.382)	(0.677)	(0.456)
Father's ed = some elementary	0.938**	0.989*	0.798**
	(0.225)	(0.370)	(0.255)
= more than elementary	2.755**	2.821**	2.721**
	(0.327)	(0.552)	(0.389)
Had a toilet in the house when age 10	2.576**	3.085**	2.199**
	(0.210)	(0.385)	(0.231)
Experienced a major health problem before age 10	0.017	0.461	–0.201
	(0.213)	(0.374)	(0.252)
P-value for joint significance of siblings variables	0.0002	0.0077	0.042
Incremental contribution of variables to R^2 as a fraction of total R^2			
Mother's education	0.583	0.561	0.597
Father's education	0.122	0.097	0.139
The number of siblings	0.021	0.025	0.012
Childhood environment	0.191	0.222	0.161
R^2	0.380	0.378	0.401
No. of observations	5552	2515	3037

Regressions also control for age and age squared. Total sample weight is used for all regressions.

* significant at the 5% level
** significant at the 1% level

Table 2 Logit Estimates of Self-Reported Health Status, by Gender

	Dep. Var: self-reported health = good, very good, excellent				
	(1)	(2)	(3)	(4)	(5)
			(a) Male		
# of siblings born alive/10	–0.359	–0.332	–0.319	–0.359	–0.392
	(0.781)	(0.764)	(0.713)	(0.674)	(0.655)
# of siblings squared/100	0.019	0.000	0.045	0.050	0.079
	(0.426)	(0.420)	(0.389)	(0.398)	(0.385)
p-value for joint significance	0.365	0.348	0.555	0.472	0.502
Yrs of schooling			0.206	0.209	0.213
			(0.108)	(0.111)	(0.111)
Yrs of schooling squared/10			–0.120	–0.197	–0.206
			(0.186)	(0.185)	(0.184)
Yrs of schooling cubed/100			0.041	0.077	0.081
			(0.078)	(0.076)	(0.075)
p-value for joint significance			<0.0001	0.003	0.002
Incremental contribution to R^2	0.020	0.035	0.289	0.295	0.031
			(b) Female		
# of siblings born alive/10	–0.580	–0.525	–0.451	–0.399	–0.478
	(0.536)	(0.535)	(0.538)	(0.519)	(0.522)
# of siblings squared/100	0.464	0.469	0.480	0.409	0.438
	(0.308)	(0.309)	(0.305)	(0.294)	(0.294)
p-value for joint significance	0.189	0.132	0.051	0.128	0.136
Yrs of schooling			–0.175	–0.164	–0.180
			(0.100)	(0.104)	(0.104)
Yrs of schooling squared/10			0.469*	0.405*	0.405*
			(0.183)	(0.191)	(0.191)
Yrs of schooling cubed/100			–0.186*	–0.158	–0.157
			(0.079)	(0.082)	(0.081)
p-value for joint significance			0.0002	0.032	0.237
Incremental contribution to R^2	0.026	0.147	0.176	0.367	0.064

Controlled for:					
Sibling variables	√	√	√	√	√
Childhood environment		√	√	√	√
Years of Schooling			√	√	√
Adult SES, etc.				√	√
BMI and height					√

Note: The number of observations are 2,515 and 3,036 for the male and female samples, respectively. All regressions include age, age squared, and indicator variables for parental education.

* significant at the 5% level

Table 3 Lasting Effects of Childhood Health, by Gender

	SR health ≥ good	hypertension	heart attack	arthritis
	(a) Male			
Experienced a health problem before 10	–0.474*	0.527*	0.192	0.532*
	(0.233)	(0.211)	(0.413)	(0.243)
Incremental contribution to R^2	0.028	0.030	0.005	0.060
	(b) Female			
Experienced a health problem before 10	–0.480*	0.110	0.917**	0.554**
	(0.208)	(0.217)	(0.351)	(0.196)
Incremental contribution to R^2	0.066	0.007	0.111	0.068

Note: The coefficient estimates are from the logit regressions of the full specification (i.e., model (5) in table 2).

* significant at the 5% level
** significant at the 1% level

Table 4 Logit estimates of relationship between Selected Disease Outcomes and Family Background, Male

Type of disease/health problem	cancer			respiratory illness			stroke		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Sibship Size									
# of siblings born alive/10	–0.460	0.525	3.474*	–0.303	–0.348	–0.077	9.721**	9.600**	11.671**
	(2.640)	(2.142)	(1.570)	(1.181)	(1.154)	(1.105)	(3.376)	(3.444)	(3.977)
# of siblings squared/100	0.699	0.417	–0.902	0.485	0.513	0.469	–8.458**	–8.329**	–9.839**
	(1.098)	(0.927)	(0.752)	(0.584)	(0.570)	(0.568)	(3.086)	(3.079)	(3.445)
p-value for joint significance	0.107	0.098	0.006	0.222	0.193	0.100	0.015	0.020	0.013
Educational Level									
Yrs of schooling		–0.656	–1.011*		–0.260	–0.179		0.086	0.060
		(0.414)	(0.404)		(0.212)	(0.179)		(0.313)	(0.305)
Yrs of schooling squared/10		1.435**	1.876**		0.549	0.431		–0.327	–0.377
		(0.552)	(0.623)		(0.386)	(0.312)		(0.541)	(0.466)
Yrs of schooling cubed/100		–0.602**	–0.774**		–0.283	–0.220		0.164	0.179
		(0.197)	(0.247)		(0.180)	(0.135)		(0.223)	(0.183)
p-value for joint significance		0.000	0.008		0.289	0.282		0.718	0.188
Incremental contribution to R^2	0.024	0.141	0.648	0.027	0.042	0.459	0.186	0.014	0.429

Note: Columns (1), (2), (3) correspond to columns (1), (3), (5) of table 2, respectively. Incremental contribution to pseudo-R^2 of column (1) is the fraction of pseudo-R^2 that sibling variables contribute. For (2) and (3), incremental contributions are calculated for the education variables and all adult socioeconomic status variables, respectively.

* significant at the 5% level
** significant at the 1% level

Table 5 Logit estimates of relationship between Selected Disease Outcomes and Family Background, Female

Type of disease/health problem	cancer			respiratory illness			stroke		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Sibship Size									
# of siblings born alive/10	2.906*	2.889*	3.798**	–0.166	–0.393	0.017	–1.513	–1.648	–0.752
	(1.320)	(1.286)	(1.428)	(0.988)	(0.955)	(0.981)	(1.400)	(1.380)	(1.216)
# of siblings squared/100	–1.815*	–1.818*	–2.320*	0.064	0.168	–0.001	0.657	0.722	0.332
	(0.763)	(0.789)	(0.988)	(0.592)	(0.594)	(0.579)	(0.997)	(0.709)	(0.735)
p-value for joint significance	0.065	0.066	0.024	0.979	0.893	0.999	0.541	0.458	0.775
Educational Level									
Yrs of schooling		–0.466	–0.459		0.211	0.202		0.179	0.160
		(0.294)	(0.246)		(0.205)	(0.184)		(0.267)	(0.242)
Yrs of schooling squared/10		0.615	0.618		–0.397	–0.326		–0.468	–0.431
		(0.514)	(0.411)		(0.360)	(0.332)		(0.494)	(0.423)
Yrs of schooling cubed/100		–0.208	–0.212		0.164	0.152		0.176	0.124
		(0.205)	(0.167)		(0.147)	(0.140)		(0.212)	(0.179)
p-value for joint significance		0.161	0.207		0.741	0.429		0.178	0.152
Incremental contribution to R^2	0.061	0.102	0.535	0.001	0.027	0.765	0.041	0.058	0.664

Note: See note of table 4.

* significant at the 5% level
** significant at the 1% level

Table 6 OLS estimates of Adult Height and Sibship Size and Education, by Gender

	men	women
# of siblings born alive/10	2.548	3.216*
	(1.802)	(1.499)
# of siblings born alive squared/100	–0.510	–0.939
	(0.981)	(0.829)
Years of education	0.924**	0.754**
	(0.269)	(0.227)
Years of education^2/10	–0.887*	–0.072
	(0.454)	(0.406)
Years of education^3/100	0.319	–0.075
	(0.194)	(0.175)
P-value for significance test of siblings variables	0.017	0.003
Adjusted R^2	0.008	0.239
No. of observations	2,515	3,037

Regressions also control for age and age squared, a fraction of surviving siblings, indicators for parental education, whether had a toiled in the house and experienced a major health problem before age 10. Total sample weight is used for all regressions.

* significant at the 5% level
** significant at the 1% level

Figure 1 Association between Sibship Size, Years of Schooling and Self-Reported Health Status, by Gender

Figure 2 Distribution of the Number of Siblings Born Alive, by Gender

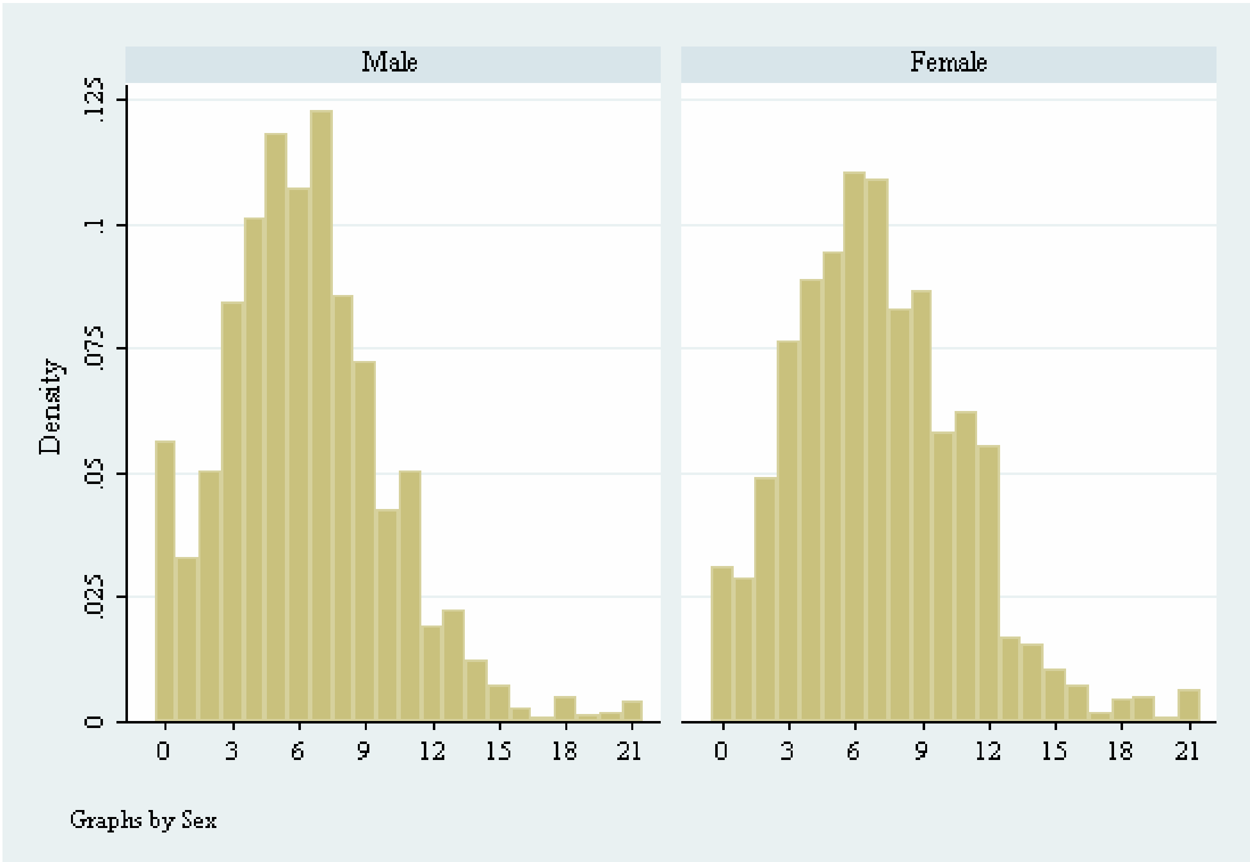
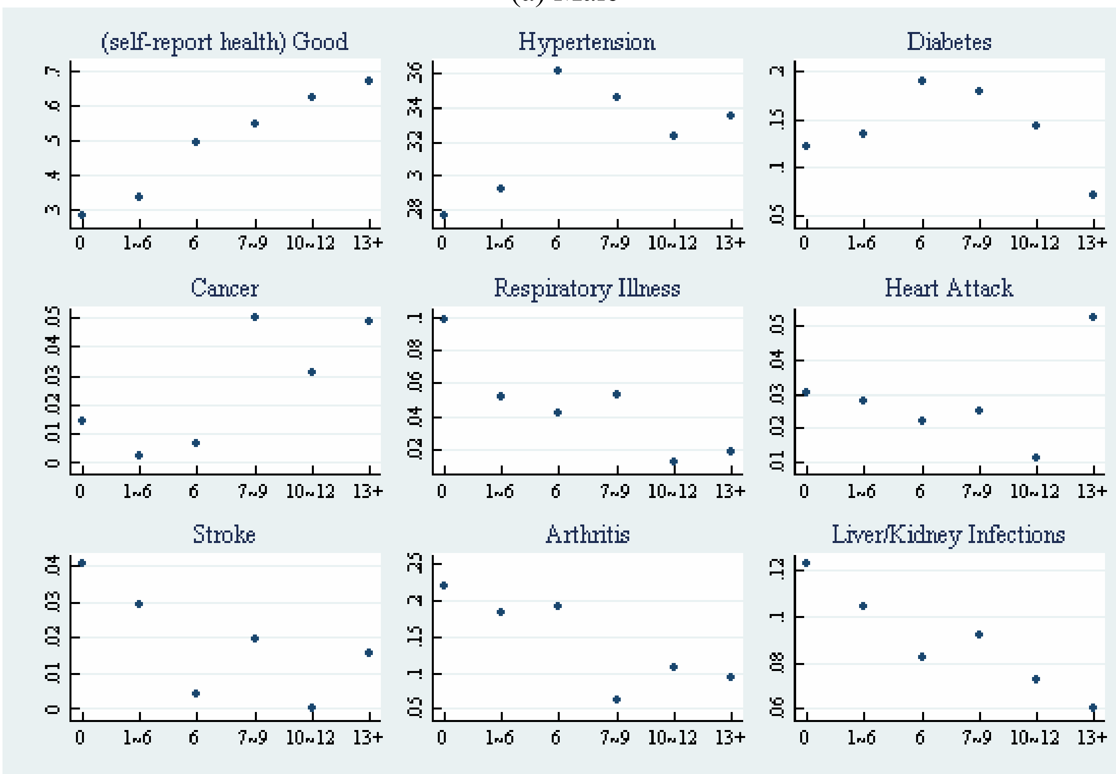


Figure 3 Years of Schooling and Selected Disease Measures, by Gender

(a) Male



(b) Female

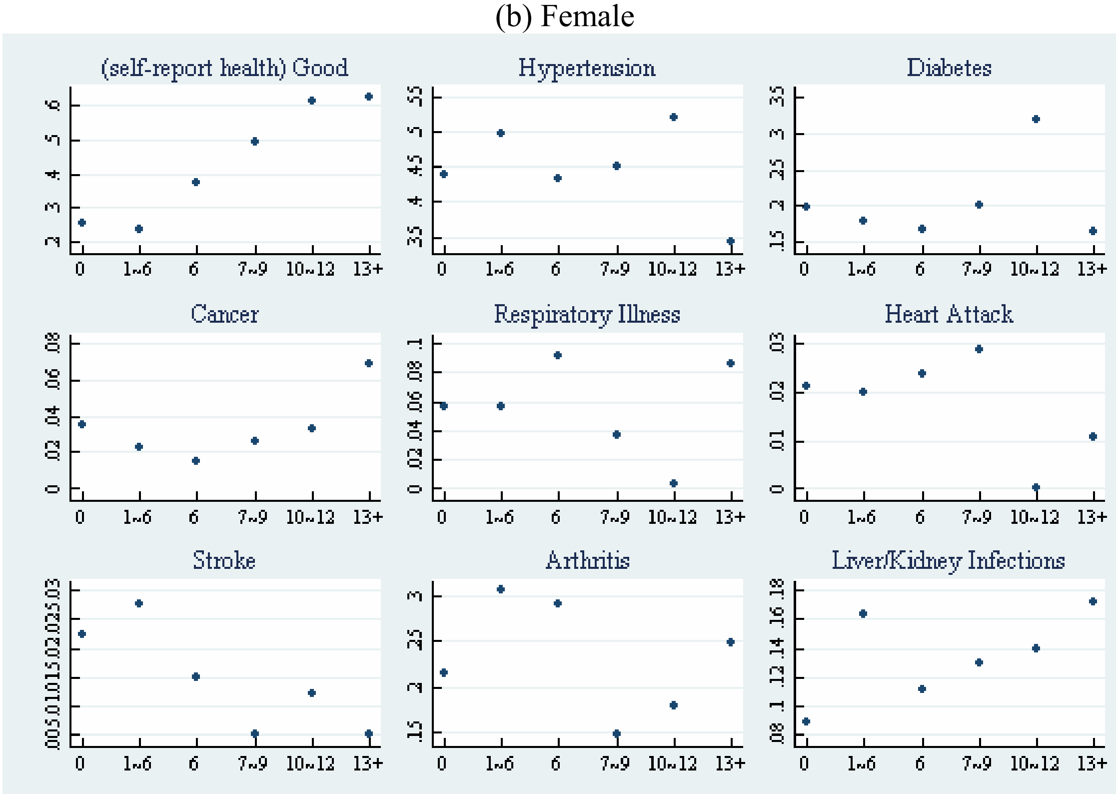
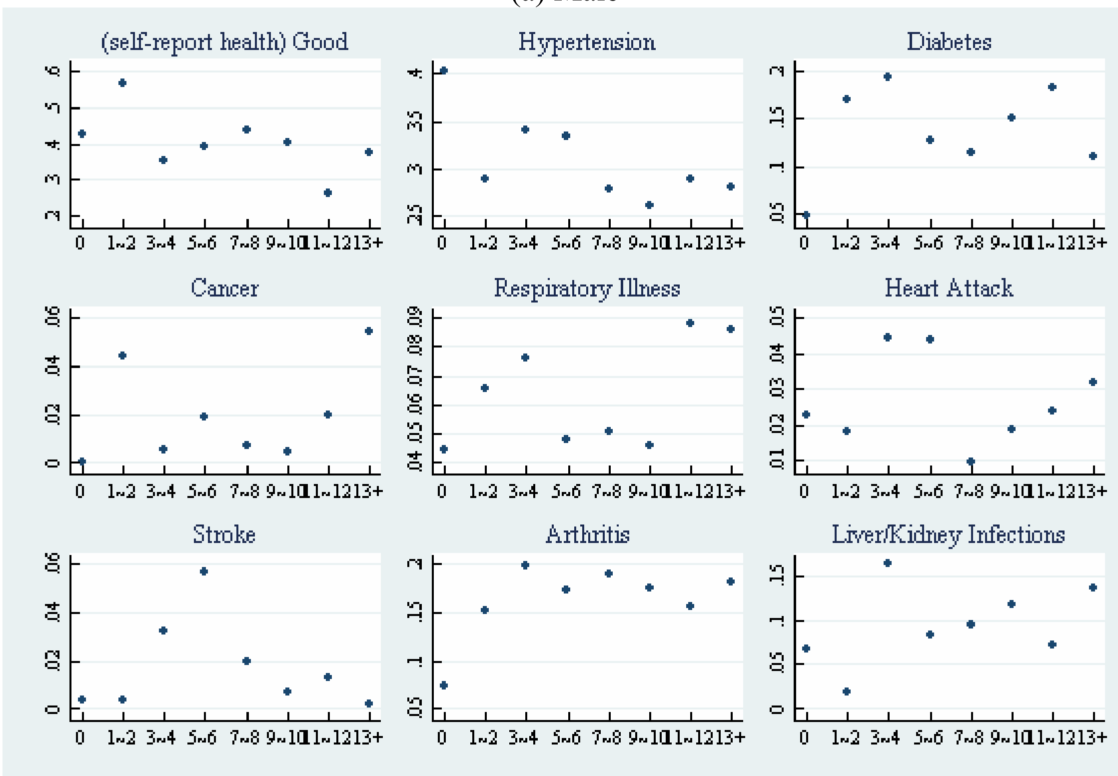


Figure 4 Sibship Size and Selected Disease Measures, by Gender

(a) Male



(b) Female

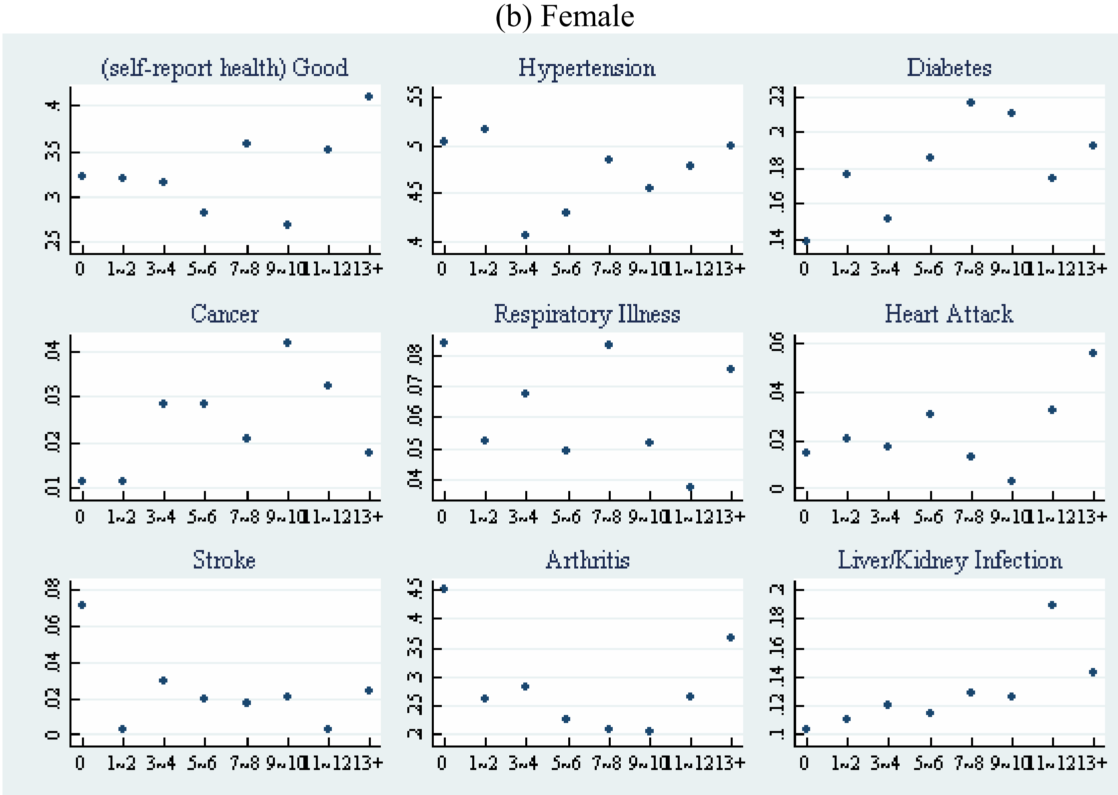
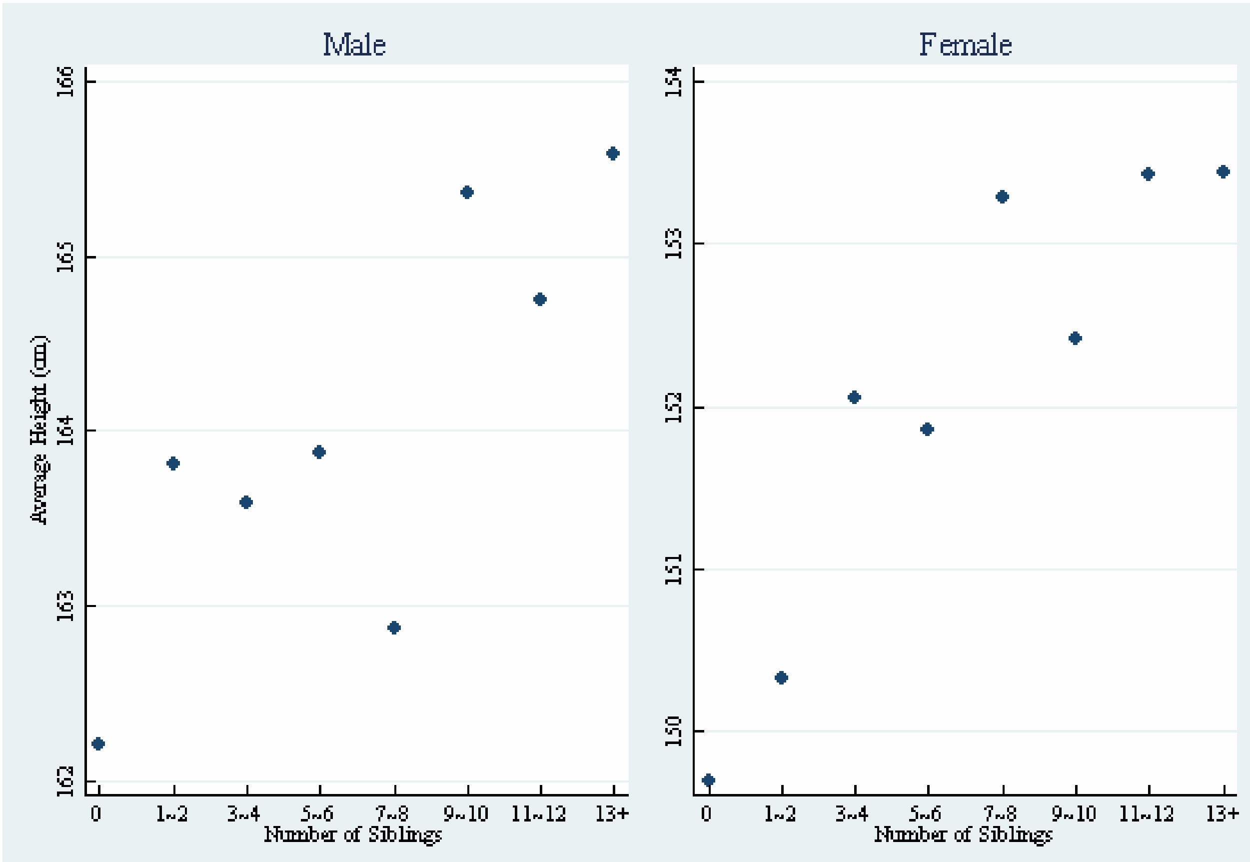


Figure 5 Sibship Size and Average Height, by Gender



Appendix Table Means and Standard Deviations of Key Variables, by Gender

	All	Male	Female
Age	63.88	65.54	62.45
	(9.29)	(9.08)	(9.23)
# of siblings born alive	6.68	6.29	7.03
	(3.79)	(3.67)	(3.87)
Proportion of surviving siblings	0.623	0.629	0.617
	(0.286)	(0.284)	(0.287)
Years of Schooling	4.16	4.31	4.04
	(4.11)	(4.43)	(3.80)
Mother's ed = some elementary	0.242	0.226	0.256
	(0.428)	(0.418)	(0.437)
= more than elementary	0.117	0.110	0.122
	(0.321)	(0.313)	(0.328)
Father's ed = some elementary	0.269	0.282	0.258
	(0.444)	(0.450)	(0.438)
= more than elementary	0.154	0.143	0.164
	(0.361)	(0.351)	(0.370)
Had a toilet in the house when age 10	0.310	0.294	0.324
	(0.463)	(0.456)	(0.468)
Experienced a major health problem before age 10	0.129	0.113	0.142
	(0.335)	(0.317)	(0.349)
Self-reported health = good, very good, or excellent	0.358	0.402	0.321
	(0.479)	(0.490)	(0.467)
Diagnosed with hypertension	0.390	0.309	0.459
	(0.488)	(0.462)	(0.498)
with diabetes	0.166	0.142	0.186
	(0.372)	(0.349)	(0.389)
with cancer	0.021	0.015	0.027
	(0.145)	(0.122)	(0.161)
with respiratory illness	0.060	0.060	0.061
	(0.238)	(0.237)	(0.239)
with a heart attack	0.025	0.029	0.022
	(0.156)	(0.166)	(0.145)
with a stroke	0.022	0.025	0.019
	(0.147)	(0.157)	(0.139)
with arthritis/rheumatism	0.216	0.173	0.252
	(0.411)	(0.379)	(0.434)
with liver or kidney infection	0.116	0.100	0.129
	(0.320)	(0.300)	(0.336)
No. of observations	5552	2515	3037

Author's calculation from the MHAS. Sampling weight is used to calculate means and standard deviations (parentheses).