Does the Number of Siblings Affect Health? Evidence from Swedish Register Data

Anna Baranowska-Rataj[a)b)], Xavier de Luna[c)]

[a)] *Department of Sociology, Umeå University*

[b)] *Institute of Statistics and Demography, Warsaw School of Economics*

[c)] *Department of Statistics, Umeå School of Business and Economics*

Paper prepared for Population Association of America Annual Meeting

30 April -2 May 2015, San Diego, CA

Abstract

The aim of this paper is to examine the impact of the number of siblings on long term physical and mental health. The literature suggests that on the one hand, family size may dilute parental resources, but on the other hand, growing up in a large family may contribute to the development of the immune system and hence prevent diseases in adult life.

We use record linkage data from Swedish Registers that provide rich information on family structure, socio-economic background and health outcomes of family members. We compare the results of propensity score matching and instrumental variable models in order to examine causal effects of family size on long term health. Our results indicate that growing up in a large family does not have a detrimental effect on health.

BACKGROUND AND OBJECTIVES

The relationship between family size and child well-being has been a central topic in demographic, economic, sociological and epidemiological research (Steelman, Powell, Werum, & Carter, 2002). It has also attracted a lot of attention from policy-makers, who are concerned with the way that fertility and population growth affect life chances, human capital accumulation and hence economic development. In some countries, governments interpret a commonly observed association between family size and poor child outcomes as evidence for the need to introduce policy measures that discourage parents from having many children (Baez, 2008; Desai, 1995; Qian, 2009).

There are a number of mechanisms through which family size may negatively affect child well-being. Mothers of large families are likely to suffer from depression during pregnancy and afterwards (Kemppainen et al., 2000), which causes emotional and cognitive distress among their children. Moreover, since children come in close contact to each other, siblings raise probability of exposure to infectious agents and thus increase the risk of diseases (Altieri, Castro, Bermejo, & Hemminki, 2006). In large families, the parental financial support as well as personal attention are distributed across a larger number of siblings, and therefore the parental investments per each child may be lower (Becker & Lewis, 1973; Becker & Tomes, 1976; Blake, 1981; Downey, 1995). Therefore, children who have more siblings tend to live in more overcrowded accommodation, with greater exposures to early infections, and with access to a less adequate diet (Hart & Smith, 2003). If parents with more children devote less attention towards monitoring child activities and assuring that their children adopt a healthy life style, the family size may have a long-lasting effect on the quality and duration of children's lives (Grundy & Sloggett, 2003; Mucci et al., 2004; Van den Berg, Doblhammer, & Christensen, 2009).

However, not all of the health outcomes of children should be expected to worsen as the family size increases. For example, according to the hygiene hypothesis, the increasing efficiency of hygiene combined with declining family size reduces the contact with infections in early life which may have a harmful effect for the development of the immune system (Strachan, 2000). This in turn raises the risk of allergy and asthma. And indeed, many empirical studies confirm a negative association between being raised in large families and the risk of allergy and asthma, although it has not yet been established whether it is the increased exposure to infectious agents *per se* that generates this association (Karmaus & Botezan, 2002).

While theoretical models of the impact of family size on child well-being provide a lot of interesting hypotheses empirical research on this topic has been rather limited. The few existing studies confirm that in developing or low- and middle-income countries, parents with large families may indeed face

difficulties with providing their children with adequate conditions for leading a healthy life. The evidence confirming a negative association between family size and child health has been provided by Glick, Marini, and Sahn (2007) for Romania, Millimet and Wang (2011) for Indonesia, Rosenzweig and Zhang (2009) for China, Hatton and Martin (2010) for Britain in the 1930ies and Smith, Mineau, Garibotti, and Kerber (2009) for the State of Utah, USA. However, these studies concern very specific contexts of harsh economic conditions and limited access to welfare state support. This opens up the question of whether growing up in a large family may negatively affect child health in affluent societies, where parental resources are not dramatically scarce and may be complemented by social policies.

We address this question by providing evidence from Sweden, i.e. a country where the well-being of children is seen as a matter of public interest, and not just a matter of parents' concern (Sandin, 1995; Sandin, Sjoberg, & Sparrman, 2012). The Swedish welfare state not only provides favorable conditions for child development, but it also takes direct active measures to intervene whenever parents need support. It provides financial and in-kind benefits to secure child health and care needs. There is also a comprehensive institutional system of support for families, that encompasses among others parental education, family counseling, maternity centers, child care centers and youth clubs. Moreover, the provision of public health care services in Sweden is universal and comprehensive, which substantially eliminates socioeconomic inequalities in use of health care (Burström, 2002). It could be argued that in such a country, even if parents with large families fail to provide their children with adequate resources, the state may make up for it.

In this paper, we use linked register data that provide information on both family structure and health outcomes of family members for the whole Swedish population in order to carry out a systematic assessment of the impact of the number of siblings on individual physical and mental health. Since family size is a choice variable and cannot be regarded as randomly distributed in the population, we use propensity score matching (Oakes & Johnson, 2006) in order to analyze differences in health of individuals raised in small and large families. Given that this technique assures comparability of individuals under the assumption of unconfoundness, we additionally implement instrumental variable models, which yield identification of causal effects under the presence of unobserved characteristics that simultaneously affect family size and health.

The contribution of this paper to the literature on the health consequences of growing up in a large family is fourfold. First, our study takes an interdisciplinary and integrative approach to the analysis of health outcomes. Following recent suggestions from demographic research (Harris 2010), it bridges insights from both biomedical and social sciences and it considers a broad range of health

outcomes related to both physical and mental health. Second, it uses high quality register data with sufficiently large samples for studying even rare diseases such as those related to mental health. Our data are also detailed and accurate, and do not suffer from recall bias or ex-post rationalizations as it is the case for retrospective or self-reported data. Third, our paper provides evidence for the links between family size and health outcomes in a context of a developed country, one of the richest in the world, which has adopted a policy focus towards eliminating social inequalities in health and other aspects of quality of life. Finally, instead of describing associations between family size and individual health outcomes, we adopt a research design that aim at drawing causal inferences.

STUDY DESIGN

Data

In this paper, we take advantage of a unique database made available at the Umeå SIMSAM Data Lab, which links individual records from Swedish registers that cover the whole Swedish population over a number of decades. These data provide detailed longitudinal information on health of both parents and their children. In the Prescribed Drug Registry, we can identify prescribed and discharged medicines classified according to the Anatomical Therapeutic Chemical (ATC) classification system. Based on the data available in these registers it is possible to measure the incidence of receiving the prescriptions. The quality of these registers is very high according to validation studies (Wettermark et al., 2007) and can be used as an estimate of the occurrence of certain diseases in the population (Kramers, 2003). The data from these registers is also linked to the administrative registers from Statistics of Sweden, which provide information on demographic events (most importantly, births of siblings) and potential confounders (such as parental background or health outcomes at birth) that affect both family size and health of family members. Our data cover all the relevant health outcomes not just for the children, but also for their parents. This is important because many diseases are transmitted genetically, and previous generations' health status may affect family size choices. Hence, controlling for parents' health outcomes may be crucial for causal inference on the impact of family size on child health. The structure of the sample used in the analysis and the descriptive characteristics of the variables used are described in Table A1 in the Annex.

Our data have their limitations, as well. While register data are more detailed and accurate then self-reported data, they reflect the use of medicines that were prescribed and dispensed by the doctors. This means that illnesses that were not diagnosed due to restrictions in access to health care are not reflected in the data. However, it could be argued that in the Swedish context, such a problem should not bias the results of the analysis, because unlike in many other countries, the provision of public health care services in Sweden is universal and comprehensive. However, during the 1990ies

some market-oriented reforms have been introduced in the Swedish health care, including increases in user fees for visits to the doctor and filling in prescriptions (Burström, 2002). Preventive care is provided free of charge for children under 20 years, though (Mossialos, Dixon, Figueras, & Kutzin, 2002). Still, some of these changes in the organization and delivery of health care may have affected the patterns of utilization of health care in the Swedish population. Interestingly, recent research based on linked data from the Prescribed Drug Registry and the Survey of Living Conditions which provides self-reported measures of health suggests that affordability of medicines does not prevent drug utilization in Sweden (Nordin, Dackehag, & Gerdtham, 2013; Weitoft, Rosen, Ericsson, & Ljung, 2008). Most socioeconomic differences in drug use patterns captured in the Prescribed Drug Registry appear to be of the same magnitude as those found in studies on disease incidence and prevalence, meaning that drugs are prescribed and dispensed according to need. However, there are some important exceptions: antibiotics, sildenafil, hormone replacement drugs, anti-migraine drugs and angiotensin receptor blockers are used more often by some socioeconomic subgroups than it could be expected from disease incidence and prevalence data. We take this evidence into account when interpreting our findings according to the ATC groups of drugs analyzed in this paper.

Methods

Testing the hypothesis on trade-off between family size and child health is challenging because decisions to have another child may be driven by the same factors which simultaneously have impact on the offspring's health, and these factors may be difficult to observe and measure directly in the data. For example, some parents (defined by some characteristics) may prefer to have smaller families and simultaneously invest more time and attention in health of each child that they have. Hence, even in the absence of true causality, there may be a strong negative correlation between family size and child health outcomes. This means that the results from standard regression models might lead to misleading conclusions. Therefore, previous research on the relationship between family size and child well-being has been criticized on methodological grounds as it has applied methods that reveal associations rather than causal effects (Black et al. 2005).

The standard approach to infer the effects of causes is to conduct controlled randomized experiments, but such experimental designs are frequently (in particular in our case) infeasible for technical and/or ethical reasons. An increasingly popular design in observational studies, where a rich sets of observed characteristics is available, is propensity score matching. The latter has only recently been adopted in research on the consequences of family size choices (Arpino & Aassve, 2013). In lieu of opportunities for carrying out randomized experiments, social scientists have also started to utilize quasi experiments for causal inference. In this strand of literature, an instrumental

variable model is a commonly chosen approach (Moffitt, 2005; Rassen, Brookhart, Glynn, Mittleman, & Schneeweiss, 2009). Each of these two approaches has its advantages and limitations. In this paper, the impact of the number of siblings on child health is examined with a propensity score matching design and the robustness of our results is checked with instrumental variable models.

Propensity score matching is formalized with the framework of potential outcomes, which was adapted from the context of randomized experiments to studies drawing on observational data by (Rubin, 1974). Within this framework, in order to make the treatment variable comparable with the one defined in the instrumental variable design, we compare health outcomes of the firstborn children across the number of siblings. The analysis that compares groups of children with different number of siblings is based on propensity score, i.e. on an estimated probit model of the probability of belonging to the group of individuals with a specific number of siblings. The variables included in this model should ideally include all factors that simultaneously affect family size and health outcomes. We include parental characteristics, i.e. mothers year of birth and her marital status as well as education attainment of parents (mothers' or fathers', whichever is highest), we also use information on any infections and diseases at first pregnancy of the mother (urinary tract infections, chronic kidney disease, diabetes mellitus, epilepsy, ulcerative colitis, systemic lupus erythematosus, chronic hypertension) as well as the duration of pregnancy in days. Additionally, we control for the characteristics of our individuals: the year of their birth, gender and health outcomes at birth (Apgar scores at birth, height at birth, BMI at birth)[1]. The results of the estimation of propensity scores are presented in the Table A2 in the Annex. After the propensity for a specific number of siblings is estimated, individuals are matched based on this index. There are a number of matching algorithms that can be used to this end (Caliendo & Kopeinig, 2008), we applied nearest neighbor matching algorithm because of its good balancing properties.

Within the instrumental variable framework, we will adopt the "twin-first approach" proposed by Rosenzweig and Wolpin (1980). The key idea is to use the data on multiple births in order to construct an instrumental variable that is a variable affect family size choice and health outcome only through family size. While decisions on higher party births are non-random, experiencing multiple births may be regarded as an outcome of a random process and not a result of deliberate decisions driven by a calculus considering future child welfare. Thus, information on twin births can be applied to identify the impact of the number of siblings on the child health using an instrumental variable design. A potential threat to internal validity of instrument based on multiple births is related to use

[1] Apgar is a test performed by a doctor, midwife, or nurse on a newborn child after birth. The health care provider examines the breathing effort, heart rate, muscle tone, reflexes and skin color. Each of these categories is scored with 0, 1, or 2, with higher scores indicating better child health outcomes. The scores are summed up to an index that shows whether the infant needs medical assistance in order to adjust to conditions outside the mother's womb(Pediatrics, 2006).

of in-vitro fertilization, implemented in Sweden since 1982, because in-vitro fertilization has been shown to raise the probability of multiple births and affect child health outcomes. We therefore focus our analysis on cohorts born before the introduction of in-vitro fertilization (specifically, we use data for cohorts born 1975-1980).

Applying the "twin-first approach" in the analysis of child health outcomes is complex because children born in multiple births have on average lower weight at birth and also poorer health outcomes later in life. Therefore, in this paper we apply the solution proposed by Black, Devereux, and Salvanes (2005): multiple births are utilized to construct the instrumental variable, and next, children born in multiple births are excluded from the sample. The analytical sample encompasses first born children, who are older siblings of children born in multiple births, and analysis shows the effects of the number of younger siblings on older siblings' health outcomes.

Instrumental variables models are estimated using a two-stage least square (2SLS) estimation procedure. These models consist of two equations. In the first equation the dependent variable is the number of siblings, and the explanatory variables include the instrumental variable as well as characteristics of individuals. Since the risk of multiple birth varies across women's age, we control for mothers age at the birth of an individual. We also include all other variables that were used to estimate the propensity score. The second equation models health outcomes of children as a function of covariates used in the first stage and the predicted values from the first equation instead of the endogenous measure of family size. The procedure adjusts also the covariance matrix of coefficients in the second stage in order to provide correct standard errors of estimates.

PRELIMINARY RESULTS

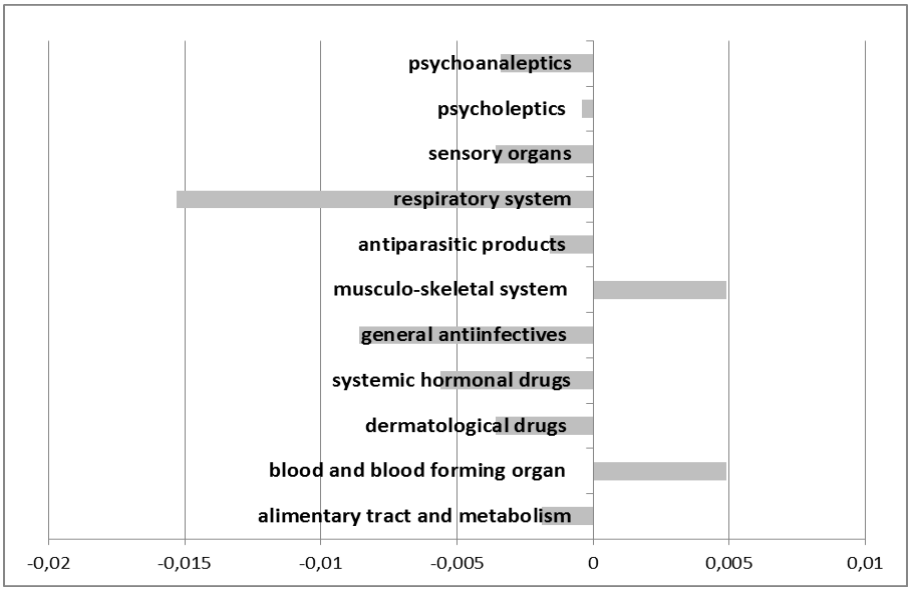
Descriptive statistics

In the first step, we examine correlation between family size and use of drugs according to ATC groups; see Figure 1. In this part of analysis, we do not take into account any confounding factors that may simultaneously affect family size and drug use. In general, the correlation between the number of siblings and drug use is rather small. Contrary to what we could expect based on theoretical models of dilution of parental resources; this correlation appears negative in case of most types of drugs. This suggests that children raised in larger families are at lower risk of suffering from diseases that require medical treatment.

Interestingly, in line with the hygiene hypothesis, we can observe a negative correlation between family size and consumption of drugs related to respiratory system, which indicates fewer problems with allergy and asthma. Individuals who were raised in larger families tend to receive less anti-

infectives, hormonal drugs and dermatological drugs. There are only two categories where we observe a slightly positive effect of family size: drugs for musculo-skeletal system and blood and blood forming organs.

Figure 1. Correlation between family size and drug use according to ATC groups



Source: Swedish Register Data from the Umeå SIMSAM Lab

Results from propensity score matching

In the next step, we present (Table 1) results from the analysis using propensity score matching where we control for a range of factors which potentially jointly determine family size and health outcomes of children. This step is important because children raised in larger families may differ from children with fewer siblings in a way that matters for their health outcomes. After controlling for a range of such characteristics, we still observe a negative correlation between family size and consumption of drug related to respiratory system and anti-infectives. The relationship between family size and drugs for musculo-skeletal system as well as blood and blood forming organ turns out to be statistically insignificant at any standard levels. All in all, we do not find any strong evidence supporting the predictions of theoretical models that emphasize disadvantage related to growing up in a large family.

Table 1. The impact of family size on drug use according to ATC groups – results from PSM

	Effect of transition from 1 to 2 siblings		Effect of transition from 2 to 3 siblings		Effect of transition from 3 to 4 siblings	
Outcome	ATT	Std. Err.	ATT	Std. Err.	ATT	Std. Err.
alimentary tract and metabolism	-0,002	0,002	-0,004	0,003	0.000	0.007
blood and blood forming organ	0,000	0,001	-0,001	0,002	0.003	0.004

dermatological drugs	0,001	0,002	0,000	0,003	0.000	0.007
systemic hormonal drugs	-0,002	0,001	-0,001	0,002	-0.003	0.005
general antiinfectives	-0,011	0,002	-0,011	0,005	0.000	0.010
musculo-skeletal system	-0,003	0,002	-0,002	0,003	0.005	0.007
antiparasitic products	0,000	0,001	-0,001	0,002	-0.003	0.003
respiratory system	-0,010	0,002	-0,012	0,004	-0.008	0.008
sensory organs	-0,001	0,001	-0,005	0,003	0.008	0.006
psycholeptics	-0,003	0,001	-0,002	0,002	0.002	0.005
psychoanaleptics	-0,003	0,001	0,001	0,003	-0.010	0.006

Source: Swedish Register Data from the Umeå SIMSAM Lab. Note: nearest neighbor algorithm has been used to match individuals with larger number of siblings.

Results from instrumental variable models

In the final step, we estimate instrumental variable models where we control for both observable factors that affect family size and health outcomes of children as well as factors that have the same confounding effect but cannot be directly captured in our data. Again, we examine the impact of family size separately for use of drugs of specific anatomic therapeutic chemical groups. For most categories of drugs, this effect is weak and statistically non-significant (Table 2). The effect of the number of siblings on using drugs related to musculo-skeletal system is stronger than in the analysis using propensity score matching. In case of the effect of family size on use of general antiinfectives, the effect is close to the one estimated by using propensity score matching, but the estimates are rather imprecise. The coefficient for use of drugs related to the respiratory system is close to zero, suggesting that the positive effect of the number of siblings observed in descriptive statistics and in analysis using propensity score matching is related to some unobserved factors that jointly affect family size and immune system of children. Hence, the mechanism described by the hygiene hypothesis is not confirmed here. We do not find support for hypothesis that siblings per se that affect the hygiene at the household and hence contribute to the development of immune system. Our results suggest instead that parents who decide to have larger families tend to be able to provide their children with conditions that prevent diseases related to the respiratory system.

Table 2. The impact of family size on drug use according to ATC groups – results from IV models

Outcome	Coef.	Std. Err.
alimentary tract and metabolism	-0.005	(0.009)
blood and blood forming organ	-0.000	(0.005)
dermatological drugs	-0.002	(0.009)
systemic hormonal drugs	0.001	(0.007)
general antiinfectives	-0.009	(0.013)
musculo-skeletal system	-0.018	(0.009)
antiparasitic products	-0.002	(0.004)
respiratory system	-0.002	(0.011)
sensory organs	-0.003	(0.007)
psycholeptics	0.005	(0.007)
psychoanaleptics	0.003	(0.008)

Source: Swedish Register Data from the Umeå SIMSAM Lab

The differences in results from analysis that controls solely for observed factors, i.e. propensity score matching, and the analysis designed to eliminate biased from unobservables, suggest that there may be some conditions which encourage parents to have more children and in the same time affect the risk of suffering from problems related to the respiratory and musculo-skeletal system. Precisely because it is not possible to capture these factors directly, even when using very detailed register data, we can only speculate what these factors may be. Still, to give an example, we can indicate the findings from recent demographic and epidemiological research, which suggests that the urban and rural settings provide diverging context for childbearing decisions of parents (Kulu, 2013; Kulu, Vikat, & Andersson, 2007) and in the same time have a different effect on the risk of allergy (Nicolaou, Siddique, & Custovic, 2005) and on the musculo-skeletal system (Vavken & Dorotka, 2011). If we take such unobservable factors into account by adopting a quasi-experimental design, having many siblings turns out to have a positive effect on musculo-skeletal system and to be neutral with respect to the diseases treated with drugs for the respiratory system.

CONCLUSIONS

Theoretical models in demographic and epidemiological literature provide conflicting predictions regarding the impact of the number of siblings on health outcomes. On the one hand, in large families, parental financial support, as well as personal attention, are distributed across a larger number of siblings, and therefore the parental investments per each child may be lower than in small families. On the other hand, it can be argued that growing up in a larger household may raise exposure to infections in early life and hence it may contribute to the development of the immune system and prevent diseases in the adult life.

There are a number of challenges related to testing hypotheses related to the impact of growing up in a large family on health in adulthood. First of all, it is difficult to collect detailed data that provide information on early life conditions, socio-economic background and health outcomes in adulthood. Second, parental decisions to have another child may be driven by the same factors which simultaneously have impact on the offspring's health, and these factors may bias the results from standard regression analysis. We tackle these challenges by using unique record linkage data from Swedish Registers and methods which control for both observed and unobserved confounders.

Overall, our results show that growing up in a large family does not have a detrimental effect on health. If anything, some health outcomes turn out to be better among individuals who had more siblings. However, it is questionable whether it is raised exposure to infections early in life per se that prevents from diseases in adulthood.

ANNEX

Table A1. Description of sample used in analysis.

	Mean	Std. Dev.	Min	Max
characteristics of index persons				
year of birth	1977	1.7	1975	1980
men	51.5%	0.5	0.0	1.0
women	48.5%	0.5	0.0	1.0
Apgar scores at birth	8.8	1.3	0.0	10.0
height at birth	50.4	2.3	22.0	60.0
BMI at birth	13.6	1.4	1.4	77.1
duration of pregnancy in days	281.7	12.2	161.0	321.0
characteristics of parents of index persons				
mother's year of birth	1953	4	1931	1965
maternal diseases at first pregnancy				
urinary tract infections	1.3%	0.1	0.0	1.0
chronic kidney disease	0.1%	0.0	0.0	1.0
diabetes mellitus	0.3%	0.1	0.0	1.0
epilepsy	0.1%	0.0	0.0	1.0
ulcerative colitis	0.0%	0.0	0.0	1.0
chronic hypertension	0.0%	0.0	0.0	1.0
maternal marital status				
single	4.5%	0.2	0.0	1.0
married / with sambo	95.3%	0.2	0.0	1.0
divorced / widow	0.2%	0.0	0.0	1.0
maternal education				
tertiary	45.0%	0.5	0.0	1.0
secondary	45.0%	0.5	0.0	1.0
primary	6.0%	0.2	0.0	1.0
less than primary	1.8%	0.1	0.0	1.0
unknown	2.3%	0.1	0.0	1.0

Source: Swedish Register Data from the Umeå SIMSAM Lab

Table A2. The results of the estimation of propensity scores of transitions between family sizes

	from 1 to 2 siblings	from 2 to 3 siblings	from 3 to 4 siblings
characteristics of index persons			
year of birth	-0.029 ***	-0.024 ***	-0.035 ***
	(0.002)	(0.004)	(0.009)
gender: female	-0.007	0.014	0.057 *
	(0.007)	(0.013)	(0.027)
Apgar scores at birth	0.007 **	0.001	0.004
	(0.003)	(0.005)	(0.011)
height at birth	0.009 ***	0.004	-0.001
	(0.002)	(0.003)	(0.007)
BMI at birth	0.008 **	0.005	-0.010
	(0.003)	(0.004)	(0.009)
duration of pregnancy	0.000	-0.001 *	-0.001
	(0.000)	(0.001)	(0.001)
characteristics of parents of index persons			
mother's year of birth	0.059 ***	0.053 ***	0.036 ***
	(0.001)	(0.002)	(0.004)
maternal diseases at first pregnancy			
urinary tract infections	-0.046	0.122 *	0.086
	(0.031)	(0.053)	(0.099)
chronic kidney disease	-0.122	0.009	0.157
	(0.092)	(0.170)	(0.332)
diabetes mellitus	-0.330 ***	-0.142	0.421
	(0.072)	(0.158)	(0.278)
epilepsy	-0.405 **	-0.021	-0.120
	(0.139)	(0.291)	(0.609)
ulcerative colitis	-0.274	-0.258	0.000
	(0.163)	(0.389)	(.)
chronic hypertension	-0.173	0.161	0.007
	(0.184)	(0.344)	(0.629)
maternal marital status (ref: married / with sambo)			
single	0.002	0.032	0.047
	(0.016)	(0.028)	(0.055)
divorced / widow	-0.049	0.241	0.000
	(0.088)	(0.173)	(.)
Parental education (ref: secondary)			
tertiary	0.228 ***	0.058 ***	-0.049
	(0.007)	(0.014)	(0.030)
primary	-0.020	0.108 ***	0.092 *
	(0.015)	(0.025)	(0.047)
less than primary	-0.000	0.258 ***	0.069
	(0.027)	(0.048)	(0.090)
unknown	-0.255 ***	-0.139 *	0.082
	(0.025)	(0.056)	(0.122)
constant	-59.387 ***	-56.910 ***	-0.518
	(3.933)	(7.150)	(15.312)

Source: Swedish Register Data from the Umeå SIMSAM Lab

Table A3. Results from the first stage of 2SLS regression

	Family size	
	Coef.	se
Instrumental variable:	0.78	(0.02)
multiple birth at 2nd birth		
characteristics of index persons		
year of birth	-0.02	(0.00)
gender: female	0.00	(0.00)
Apgar scores at birth	0.00	(0.00)
height at birth	0.00	(0.00)
BMI at birth	0.00	(0.00)
duration of pregnancy in days	-0.00	(0.00)
characteristics of parents of index persons		
mother's year of birth	0.04	(0.00)
maternal diseases at first pregnancy		
urinary tract infections	0.03	(0.02)
chronic kidney disease	-0.05	(0.05)
diabetes mellitus	-0.15	(0.03)
epilepsy	-0.18	(0.07)
ulcerative colitis	-0.16	(0.08)
chronic hypertension	0.05	(0.09)
maternal marital status (married / with sambo)		
single	0.03	(0.01)
divorced / widow	0.01	(0.04)
parental education (ref secondary)		
tertiary	0.10	(0.00)
primary	0.10	(0.01)
less than primary	0.05	(0.01)
unknown	-0.13	(0.01)
constant	-40.64	(2.06)
N	169147	

Source: Swedish Register Data from the Umeå SIMSAM Lab

REFERENCES

Altieri, Andrea, Castro, Felipe, Bermejo, Justo Lorenzo, & Hemminki, Kari. (2006). Number of siblings and the risk of lymphoma, leukemia, and myeloma by histopathology. *Cancer Epidemiology Biomarkers & Prevention, 15*(7), 1281-1286.

Arpino, Bruno, & Aassve, Arnstein. (2013). Estimating the causal effect of fertility on economic wellbeing: data requirements, identifying assumptions and estimation methods. *Empirical Economics, 44*(1), 355-385.

Baez, Javier E. (2008). Does More Mean Better? Sibling Sex Composition and the Link between Family Size and Children's Quality.

Becker, Gary S, & Lewis, H Gregg. (1973). On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy, 81*(2), S279-288.

Becker, Gary S, & Tomes, Nigel. (1976). Child Endowments and the Quantity and Quality of Children. *Journal of Political Economy*, S143-S162.

Black, Sandra E, Devereux, Paul J, & Salvanes, Kjell G. (2005). The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics*, 669-700.

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

Burström, Bo. (2002). Increasing inequalities in health care utilisation across income groups in Sweden during the 1990s? *Health policy, 62*(2), 117-129.

Caliendo, Marco, & Kopeinig, Sabine. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys, 22*(1), 31-72.

Desai, Sonalde. (1995). When are children from large families disadvantaged? Evidence from cross-national analyses. *Population Studies, 49*(2), 195-210.

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

Glick, Peter J, Marini, Alessandra, & Sahn, David 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.

Grundy, Emily, & Sloggett, Andy. (2003). Health inequalities in the older population: the role of personal capital, social resources and socio-economic circumstances. *Social science & medicine, 56*(5), 935-947.

Hart, CL, & Smith, G Davey. (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.

Hatton, Timothy J, & Martin, Richard M. (2010). The effects on stature of poverty, family size, and birth order: British children in the 1930s. *Oxford Economic Papers, 62*(1), 157-184.

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.

Kemppainen, L, Mäkikyrö, T, Jokelainen, J, Nieminen, P, Järvelin, M-R, & Isohanni, M. (2000). Is grand multiparity associated with offsprings' hospital-treated mental disorders? A 28-year follow-up of the North Finland 1966 birth cohort. *Social psychiatry and psychiatric epidemiology, 35*(3), 104-108.

Kramers, Pieter GN. (2003). The ECHI project: health indicators for the European Community. *The European Journal of Public Health, 13*(suppl 3), 101-106.

Kulu, Hill. (2013). Why do fertility levels vary between urban and rural areas? *Regional Studies, 47*(6), 895-912.

Kulu, Hill, Vikat, Andres, & Andersson, Gunnar. (2007). Settlement size and fertility in the Nordic countries. *Population Studies, 61*(3), 265-285.

Millimet, Daniel L, & Wang, Le. (2011). Is the quantity-quality trade-off a trade-off for all, none, or some? *Economic Development and Cultural Change, 60*(1), 155-195.

Moffitt, Robert. (2005). Remarks on the analysis of causal relationships in population research. *Demography, 42*(1), 91-108.

Mossialos, Elias, Dixon, Anna, Figueras, Joseph, & Kutzin, Joe. (2002). Funding health care: options for Europe: Open University Press Maidenhead.

Mucci, Lorelei A, Hsieh, Chung-cheng, Williams, Paige L, Dickman, Paul W, Björkman, Lars, & Pedersen, Nancy 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), 499-506.

Nicolaou, N, Siddique, N, & Custovic, A. (2005). Allergic disease in urban and rural populations: increasing prevalence with increasing urbanization. *Allergy, 60*(11), 1357-1360.

Nordin, Martin, Dackehag, Margareta, & Gerdtham, Ulf-G. (2013). Socioeconomic inequalities in drug utilization for Sweden: evidence from linked survey and register data. *Social science & medicine, 77*, 106-117.

Oakes, J Michael, & Johnson, Pamela Jo. (2006). Propensity score matching for social epidemiology. *Methods in social epidemiology, 1*, 370-393.

Pediatrics, American Academy of. (2006). The APGAR score. *Advances in Neonatal Care, 6*(4), 220-223.

Qian, Nancy. (2009). Quantity-quality and the one child policy: The only-child disadvantage in school enrollment in rural China: National Bureau of Economic Research.

Rassen, Jeremy A, Brookhart, M Alan, Glynn, Robert J, Mittleman, Murray A, & Schneeweiss, Sebastian. (2009). Instrumental variables I: instrumental variables exploit natural variation in nonexperimental data to estimate causal relationships. *Journal of clinical epidemiology, 62*(12), 1226-1232.

Rosenzweig, Mark R, & Wolpin, Kenneth I. (1980). Testing the quantity-quality fertility model: The use of twins as a natural experiment. *Econometrica: journal of the Econometric Society*, 227-240.

Rosenzweig, Mark R, & Zhang, Junsen. (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.

Rubin, Donald B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology, 66*(5), 688.

Sandin, Bengt. (1995). Split Visions, Changing Childhoods and the Welfare state in Sweden: Reflections on the Century of the Child.

Sandin, Bengt, Sjoberg, Johanna, & Sparrman, Anna. (2012). *Situating child consumption: Rethinking values and notions of children, childhood and consumption*: Nordic Academic Press.

Smith, Ken R, Mineau, Geraldine P, Garibotti, Gilda, & Kerber, Richard. (2009). Effects of childhood and middle-adulthood family conditions on later-life mortality: evidence from the Utah Population Database, 1850–2002. *Social science & medicine, 68*(9), 1649-1658.

Steelman, Lala Carr, Powell, Brian, Werum, Regina, & Carter, Scott. (2002). Reconsidering the effects of sibling configuration: Recent advances and challenges. *Annual Review of Sociology*, 243-269.

Strachan, David P. (2000). Family size, infection and atopy: the first decade of the'hygiene hypothesis'. *Thorax, 55*(Suppl 1), S2.

Van den Berg, Gerard J, Doblhammer, Gabriele, & Christensen, Kaare. (2009). Exogenous determinants of early-life conditions, and mortality later in life. *Social Science & Medicine, 68*(9), 1591-1598.

Vavken, Patrick, & Dorotka, Ronald. (2011). Burden of musculoskeletal disease and its determination by urbanicity, socioeconomic status, age, and sex: results from 14,507 subjects. *Arthritis care & research, 63*(11), 1558-1564.

Weitoft, G Ringbäck, Rosen, M, Ericsson, Ö, & Ljung, R. (2008). Education and drug use in Sweden—a nationwide register-based study. *Pharmacoepidemiology and drug safety, 17*(10), 1020-1028.

Wettermark, Björn, Hammar, Niklas, MichaelFored, C, Leimanis, Andrejs, Otterblad Olausson, Petra, Bergman, Ulf, . . . Rosén, Måns. (2007). The new Swedish Prescribed Drug Register—opportunities for pharmacoepidemiological research and experience from the first six months. *Pharmacoepidemiology and drug safety, 16*(7), 726-735.