The causal effect of another sibling on own fertility –

An estimation of intergenerational fertility correlations by looking at siblings of twins

Martin Kolk[1]

Fertility is associated across generations in contemporary populations. This has been found consistently in developed nations, and the association appears to increase slightly over time (Anderton et al., 1987; Dahlberg, 2013; Murphy, 1999). There is also evidence from developing countries (Murphy, 2012), and there also appear to be an association between child and grandparent fertility, net of that of the intermediate generation (Kolk, forthcoming-a). In other words, there is clear evidence that the number of siblings a person grows up with is correlated with that person's eventual family size.

This association between number of siblings and number of own children are widely documented, but there has been less progress in understanding the source of this association. The current study adds to our knowledge on these mechanisms by looking at the effect of an exogenous change in number of siblings a person grows up with, by means of a twin birth in the family of origin. Using an instrumental variable (IV) approach, I estimate the structural effect of an extra sibling. Thus, it is possible to isolate the role of factors that should be independent of a twin birth (such as parental socioeconomic status, parental fertility preferences, ethnic/religious values), from the effect of growing up with an additional sibling (being related to a different home environment and potentially less parental resources during upbringing).

Previous explanations for intergenerational correlations in fertility

Previous research has suggested several mechanisms to play a role in intergenerational transmission of fertility. The most common explanation of fertility continuities is that parents transmit their values and preferences to their children (Anderton et al., 1987; Johnson & Stokes, 1976). These values could for example be related to the timing of childbearing or ideal number of children, but also be related to factors such as preference for leisure, or a preference towards/against marriage. We also know that socioeconomic status is associated across generations and that different occupational classes have different amounts of children. This could plausibly explain intergenerational transmission of fertility, even though socioeconomic continuities appears to be less important than factors not associated with socioeconomic status (Dahlberg, 2013; Kolk, forthcoming-b). Genetic factors have been shown to be associated with fertility preferences (Kohler et al., 1999; Rodgers et al., 2001).

The number of siblings in the family of upbringing could also in itself influence a person's eventual fertility. The experience of growing up with a sibling could make a person more likely to want to have more children. This could, for example, be because one values growing up in a large family as dependent on the size of the family you grew up in. Having another sibling could also be associated with differences in socioeconomic outcomes (Angrist et al., 2010; Becker & Lewis, 1974), for example due to decreased parental resources, which could affect fertility (Andersson, 2000). However, there appears to be no causal effect of having an additional sibling on socioeconomic status in Sweden, from IV estimates from a twin birth (Åslund & Grönqvist, 2010).

A causal instrumental variable approach, using twins as an exogenous source of fertility variation allows us to distinguish between (a), the effect of transmission of values and preferences (including those with a genetic basis), and parental socioeconomic background, and other parental and/or group specific characteristics, which would be independent of the mother of the index person giving birth to twins, and thus not be causally linked to the chance birth of an additional sibling, and (b), the causal effect of growing up an additional sibling.

Research Design & Method

The effect of a twin sibling on eventual family size is estimated using two-stage least squares (2SLS) regression. In the first stage the effect of a twin birth on parental fertility is estimated. In the second stage, similar to other studies on

[1] Stockholm University Demography Unit - martin.kolk@sociology.su.se

intergenerational transmission of fertility, the fertility of children is regressed on their parents' fertility, but parental fertility is estimated only from the first stage estimate of an exogenous increase in fertility due to a parental twin birth.

The index population consists of Swedish born individuals who can be linked to their parents and siblings. My instrumental variable, an additional sibling in the family due to a twin birth, include index persons whose mother gave birth to twins. I exclude twins themselves from the sample as twins have different fertility patterns, and only examine siblings of twins. I only look at siblings born before a twin birth, and exclude non-twin siblings born after a twin birth. I also examine the effect of a twin birth at different parities. The effect of a twin birth on eventual family size is most likely larger if the twin is born at a later parity, as it then is more of a chance that parents 'overshoot' their desired number of children. A common concern in instrumental variables studies are weak instruments (in this case, the effect of a twin birth on parental family size). A twin birth is a very strong predictor of parental family size, on average a multiple birth in the parental generation, increased the eventual parental family size of 0.73-0.86 children (see table 1). Thus a weak IV, a common concern of IV-designs, is not a problem here.

The key assumptions of an IV design are that the instrument is uncorrelated with the outcome variable. Overall, this is likely to be true as a twin birth is a (mostly) random physiological event, and not associated with other characteristics of the parent. This assumption is overall robust, and has been used in a large number of IV-twin designs (e.g. Åslund & Grönqvist, 2010; Black et al., 2005; Holmlund et al., 2013; Jacobsen et al., 1999; Rosenzweig & Wolpin, 1980; Rosenzweig & Zhang, 2009), but there are some potential biases which could violate the independence assumption. One such possibility is in-vitro fertilization (IVF); twin births are much more common after IVF, as the laboratory often uses several fertilized eggs. IVF could due to economic costs be associated both with twin births and socioeconomic background. However, in vitro-fertilization did not exist when the parental generation in this study completed their childbearing. Another potential bias is the fact that (dizygotic) twinning increases with age (Bulmer, 1959; Hoekstra et al., 2008; Imaizumi, 2003). Thus, both the age profile and the indirectly socioeconomic status (as older mothers often have more resources) of twin mothers might differ slightly from other mothers. As an additional robustness check I do therefore examine both mothers with same sex twins (dizygotic and monozygotic twins) and opposite sex twins (only dizygotic twins), as only dizygotic twinning is associated with age (Hoekstra et al., 2008; Imaizumi, 2003). Thus, we get an indirect estimate on whether this factor may bias the results.

The fertility of twins has been shown to be lower than that of other individuals (Tollebrant, 2001). Therefore I only examine the fertility of siblings of twins, as these would experience a larger family size without themselves being a twin. A twin birth could change the home environment substantially, birth intervals and other factors might differ significantly after the birth of a twin and it is possible that there is a selection in parents who have kids after a twin birth. I therefore also exclude siblings born after a twin birth (cf. Black et al., 2005). One could argue, that growing up with two older twins as siblings could differ from growing up with two older non-twin siblings, and that this could be related to differences in eventual childbearing. However, it appears quite unlikely that this difference would be substantive compared to the exogenous increase of 0.73-0.86 new siblings related to a twin birth.

Data

The data for the study is a collection of administrative registers for the complete Swedish population. Children are connected to their parents through a unique personal identification number. In order to measure intergenerational transmission of family size, one needs to observe the complete reproductive histories of two generations. This is available through government registers, and also sets the limit for the cohorts under observation. The younger generation in the study is the 1940 to the 1965 cohort of Swedish born men and women. This generation is linked to their parents through the Swedish multigenerational register. The quality of the register is high. Information on the mother is available for virtually the entire cohort, missing information on fathers is less than 5%. Both parents and potential siblings had to be alive in Sweden at some point after 1960 in order to be included in the digitized population registers. Fertility is measured for both generations using the number of ever born children. Number of siblings for the younger generation is calculated from the number of full-siblings to the sibling in the registers. Consequently, only information for siblings born before a twin birth is used to analyze fertility correlations. Only information on a twin birth at parity 2, 3 and 4 are analyzed. The percentage twin births of all births at these parities are 1.0%, 1.2% and 1.3%, respectively. To make sure that there are no systematic differences in the models between index-persons with and without siblings who are twins, correlations are only estimated for individuals having at least 2, 3 and 4 siblings, respectively.

Results

The results of a twin birth at parity 2, 3 and 4 can be found in Table 1. The OLS estimate of an additional sibling on own completed fertility was 0.076 which is consistent with previous estimates on intergenerational transmission of fertility in Sweden (Dahlberg, 2013; Kolk, forthcoming-a). I also estimated ordinary OLS regressions regressing fertility of children on fertility of parents for the sample used in the 2SLS models, separately for men and women, born 1st with at least 2 siblings, 1st and 2nd born with at least 3 siblings, and finally 1st, 2nd and 3rd born with at least 3 siblings. As can be seen in Table 1 correlations are lower for men than for women. The effect of number of siblings is stronger for early birth order index persons than later born index persons. The effect varies from 0.070 to 0.044 additional own children for every additional sibling for men, and 0.102 to 0.068 for women. These results are consistent with earlier estimates of intergenerational transmission of fertility, both in Sweden and elsewhere (Dahlberg, 2013; Murphy, 1999; Murphy & Wang, 2001).

When looking at the first stage results, the effect of a twin birth on eventual number of siblings, we can see that the effect varies from 0.73 to 0.86 additional children related to a twin birth. Predictably, the effect increases by parity, and a later birth has a stronger effect on eventual family size. Standard scores for the first stage estimates are in all cases above 50 (not shown) showing that there is no concern of weak instruments.

To answer the main research question, to isolate the role of parental factors only indirectly correlated with their number of children (such as values, socioeconomic background, religion etc.), from the structural effect of having an extra sibling, the estimates of an additional sibling is used to estimate the second stage equation. As can be seen in Table 1, estimates from the second stage are in all cases lower than baseline OLS results. Results differ for men and women. For men estimates of intergenerational fertility correlations are weaker when estimated using the second stage equations. For a parity-2 twin birth second stage coefficients are 0.026 compared to 0.070 for the OLS estimates, for a parity-3 twin birth the estimates are 0.044 compared to 0.056, and for a parity-4 twin birth -0.013 compared to 0.044. However, none of the second stage equations show a statistically significant different from the OLS results[2]. When looking at women there are bigger difference between OLS and second stage estimates. Second stage correlations for a parity-2 twin birth are 0.102 compared to -0.056 for the OLS estimates, for a parity-3 twin birth 0.086 compared to -0.005, and for a parity-4 twin birth 0.068 compared to -0.014. These differences are all statistically significant. Interestingly, all of the estimated intergenerational fertility correlations disappear when looking at an exogenous change in the number of siblings compared to the simple correlations in number of children. Thus, the results clearly indicate that, at least for women, the structural role of another sibling is of no importance for population level fertility correlations. Of further notice, is that 1st born siblings followed by a twin birth, have *lower* fertility if experiencing a birth of one additional sibling (3 or more siblings compared to 2 or more siblings). Thus, the experience of having another sibling is negatively correlated with eventual own family size, a finding contrary to all previous population level findings of intergenerational fertility correlations. In the future, I also aim to estimate separate correlations for siblings experiencing two same sex, and opposite sex twins, a pseudo measure for the zygozity of the twin birth.

[2] I performed a z-test of the form $(b_1 - b_2)/(SEb_1^2 + SEb_2^2)^{0.5}$, (Clogg et al., 1995)

Table 1: Effect of family size of origin on number of children, using twin birth as a measure of an exogenous increase in family size of origin.
Men and women born in Sweden between 1940 and 1965.

	OLS	OLS	first stage	second stage	z-test	OLS	first stage	second stage	z-test	OLS	first stage	second stage	z-test
	everyone	1st born men	1st born men	1st born men	OLS vs 2SLS	1-2nd born men	1-2nd born men	1-2nd born men	OLS vs 2SLS	1-3rd born men	1-3rd born men	1-3rd born men	OLS vs 2SLS
Nr of siblings	0.067***	0.070***		0.026		0.056***	0.632	0.044*		0.044***	0.067	-0.013	
(CI 95%)	0.066 - 0.068	0.066 - 0.074		-0.026 - 0.079		0.052 - 0.060		-0.005 - 0.092		0.039 - 0.049		-0.073 - 0.047	
p-value					0.104				0.632				0.067
2nd birth a twin			0.741***										
(CI 95%)			0.710 - 0.772										
3rd birth a twin							0.814***						
(CI 95%)							0.786 - 0.843						
4th birth a twin											0.833***		
(CI 95%)											0.796 - 0.870		
Constant	1.666***	1.615***	2.719***	1.734***		1.668***	3.579***	1.712***		1.717***	4.594***	1.980***	
(CI 95%)	1.663 - 1.669	1.604 - 1.626	2.716 - 2.722	1.591 - 1.878		1.653 - 1.684	3.576 - 3.582	1.539 - 1.885		1.691 - 1.742	4.590 - 4.598	1.703 - 2.256	
Observations	2,582,295	428,879	428,879	428,879		404,511	404,511	404,511		232,583	232,583	232,583	
R-squared	0.007	0.003	0.005	0.002		0.002	0.008	0.002		0.001	0.008	-0.001	
		OLS	first stage	second stage	z-test	OLS	first stage	second stage	z-test	OLS	first stage	second stage	z-test
VARIABLES		1st born women	1st born women	1st born women	OLS vs 2SLS	1-2nd born women	1-2nd born women	1-2nd born women	OLS vs 2SLS	1-3rd born women	1-3rd born women	1-3rd born women	OLS vs 2SLS
Nr of siblings		0.102***		-0.056**		0.086***		-0.005		0.068***		-0.014	
(CI 95%)		0.099 - 0.106		-0.106 - -0.006		0.082 - 0.090		-0.050 - 0.041		0.063 - 0.073		-0.072 - 0.043	
p-value					0.000				0.000				0.005
2nd birth a twin			0.726***										
(CI 95%)			0.694 - 0.757										
3rd birth a twin							0.798***						
(CI 95%)							0.769 - 0.828						
4th birth a twin											0.827***		
(CI 95%)											0.788 - 0.866		
Constant		1.694***	2.716***	2.125***		1.741***	3.584***	2.066***		1.814***	4.594***	2.195***	
(CI 95%)		1.683 - 1.704	2.713 - 2.719	1.988 - 2.262		1.726 - 1.755	3.581 - 3.587	1.903 - 2.228		1.790 - 1.838	4.590 - 4.599	1.931 - 2.459	
Observations		407,736	407,736	407,736		381,491	381,491	381,491		221,330	221,330	221,330	
R-squared		0.008	0.005	-0.011		0.005	0.007	-0.001		0.003	0.008	-0.001	

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Discussion

Overall the use of twin births as an exogenous source of variance in the number of siblings shows clear evidence for weak or no structural effect on an additional sibling in istself. This finding is of major help to understand the causes of intergenerational fertility correlations. The results show that factors such as economic consequences of parental resources and parental investment of an unexpected additional child does not indirectly influence the number of own children an individual has. This is in line with the weak causal effects of family size on socioeconomic status found in previous research (Åslund & Grönqvist, 2010). The results also indicate that the main reason for intergenerational fertility correlations is not the experience of growing up in a larger family per se. For example, researchers have theorized that the experience of a having a larger family of upbringing would cause a later preference for an own large family size or towards a general family orientation.

Instead, the main explanation for fertility correlations in developed countries appears to be related to intergenerational correlations in other factors that parents and their children share, that are not directly associated with the number of siblings as such. Factors like shared ideals about number of children, timing of childbearing, and contraceptive use, could still explain intergenerational correlations in fertility. Similarly, if parents and children share socioeconomic traits this could also be important. Broader shared characteristics such as religious values, ethnic characteristics, and regional patterns could also be the source of these correlations. That fertility correlations are non-causal in the sense that they are not directly related to the size of family of origin does not in any way alter the importance of fertility correlations as a demographic phenomenon. The effect of intergenerational transmission of fertility on population dynamics is caused by the population level correlations in fertility between individuals(e.g. Murphy & Wang, 2003), and sources of these correlations are less important.

The conclusions drawn are limited by the validity of the instrumental variables. In particular, if parents of twins differ in substantive ways in for example socioeconomic status or age from other parents, the results are potentially biased. It is also possible that it is different to grow up with older siblings if they themselves are twin. These assumptions are not testable but further robustness checks will be carried out to estimate the potential influence of these violations. Another limitation is that the numbers of twin births are too low for an exact estimation of second stage fertility correlations, even in a dataset that includes the entire population of Sweden.

References

Andersson, G. (2000). The Impact of Labour-Force Participation on Childbearing Behaviour: Pro-Cyclical Fertility in Sweden During the 1980s and the 1990s. *European Journal of Population/Revue européenne de Démographie, 16*(4), 293-333.

Anderton, D. L., Tsuya, N. O., Bean, L. L., & Mineau, G. P. (1987). Intergenerational Transmission of Relative Fertility and Life Course Patterns. *Demography, 24*(4), 467-480.

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), 773-824.

Åslund, O., & Grönqvist, H. (2010). Family Size and Child Outcomes: Is There Really No Trade-Off? *Labour Economics, 17*(1), 130-139.

Becker, G. S., & Lewis, H. G. (1974). Interaction between Quantity and Quality of Children *Economics of the Family: Marriage, Children, and Human Capital* (pp. 81-90): UMI.

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.

Bulmer, M. G. (1959). The Effect of Parental Age, Parity and Duration of Marriage on the Twinning Rate. *Annals of human genetics, 23*(4), 454-458.

Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical Methods for Comparing Regression Coefficients between Models. *American Journal of Sociology*, 1261-1293.

Dahlberg, J. (2013). Family Influence in Fertility: A Longitudinal Analysis of Sibling Correlations in First Birth Risk and Completed Fertility among Swedish Men and Women. *Demographic Research, 29*(9), 233-246.

Hoekstra, C., Zhao, Z. Z., Lambalk, C. B., Willemsen, G., Martin, N. G., Boomsma, D. I. (2008). Dizygotic Twinning. *Human Reproduction Update, 14*(1), 37-47.

Holmlund, H., Rainer, H., & Siedler, T. (2013). Meet the Parents? Family Size and the Geographic Proximity between Adult Children and Older Mothers in Sweden. *Demography*, 1-29.

Imaizumi, Y. (2003). A Comparative Study of Zygotic Twinning and Triplet Rates in Eight Countries, 1972-1999. *Journal of biosocial science, 35*(2), 287-302.

Jacobsen, J. P., Pearce III, J. W., & Rosenbloom, J. L. (1999). The Effects of Childbearing on Married Women's Labor Supply and Earnings: Using Twin Births as a Natural Experiment. *Journal of Human Resources*, 449-474.

Johnson, N. E., & Stokes, C. S. (1976). Family Size in Successive Generations: The Effects of Birth Order, Intergenerational Change in Lifestyle, and Familial Satisfaction. *Demography, 13*(2), 175-187.

Kohler, H.-P., Rodgers, J. L., & Christensen, K. (1999). Is Fertility Behavior in Our Genes? Findings from a Danish Twin Study. *Population and Development Review, 25*(2), 253-288.

Kolk, M. (forthcoming-a). Multigenerational Transmission of Family Size in Contemporary Sweden. *Population Studies*.

Kolk, M. (forthcoming-b). "Understanding Transmission of Fertility across Multiple Generations – Socialization or Socioeconomics?". *Research in Social Stratification and Mobility*.

Murphy, M. (1999). Is the Relationship between Fertility of Parents and Children Really Weak? *Social Biology, 46*(1-2), 122-145.

Murphy, M., & Wang, D. L. (2001). Family-Level Continuities in Childbearing in Low-Fertility Societies. *European Journal of Population, 17*(1), 75-96.

Murphy, M., & Wang, D. (2003). The Impact of Intergenerationally-Transmitted Fertility and Nuptuality on Population Dynamics in Contemporary Populations. In J. L. Rodgers & H.-P. Kohler (Eds.), *Biodemography of Human Reproduction and Fertility* (pp. 209-228). Boston: Kluwer Academic Publishers.

Murphy, M. (2012). Intergenerational Fertility Correlations in Contemporary Developing Countries. *American Journal of Human Biology, 24*(5), 696-704.

Rodgers, J. L., Hughes, K., Kohler, H. P., Christensen, K., Doughty, D., Rowe, D. C. (2001). Genetic Influence Helps Explain Variation in Human Fertility: Evidence from Recent Behavioral and Molecular Genetic Studies. *Current Directions in Psychological Science, 10*(5), 184-188.

Rosenzweig, M. R., & Wolpin, K. I. (1980). Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment. *Econometrica: journal of the Econometric Society, 48*(1), 227.

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.

Tollebrant, J. (2001). Tvillingars Livsmönster. *Välfärd, 4*, 8-9.