University of Groningen

The effect of a sibling's gender on earnings and family formation

Péter, Noemi; Lundborg, Petter; Mikkelsen, Sara; Webbink, Dinand

Published in:
Labour Economics

DOI:
[10.1016/j.labeco.2018.06.006](https://doi.org/10.1016/j.labeco.2018.06.006)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2018

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Péter, N., Lundborg, P., Mikkelsen, S., & Webbink, D. (2018). The effect of a sibling's gender on earnings and family formation. *Labour Economics*, 54, 61-78. <https://doi.org/10.1016/j.labeco.2018.06.006>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

The effect of a sibling's gender on earnings and family formation[☆]

Noemi Peter^{a,*}, Petter Lundborg^{b,d}, Sara Mikkelsen^b, Dinand Webbink^{c,d,e}

^a Faculty of Economics and Business, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands

^b Department of Economics, Lund University, P.O. Box 7082, Lund SE-220 07, Sweden

^c Erasmus School of Economics Rotterdam, Burgemeester Oudlaan 50, PA Rotterdam 3062, The Netherlands

^d Institute of Labor Economics (IZA), Germany

^e Tinbergen Institute, The Netherlands

ARTICLE INFO

JEL classification:

J00
J16
J24
J13

Keywords:

Sibling gender
Sex composition
Twins
Income
Fertility

ABSTRACT

We examine how the gender of a sibling affects labor market outcomes and family formation. Identification is complicated by parental preferences: if parents prefer certain sex compositions over others, children's gender affects not only the outcomes of other children but also the existence of potential additional children. We employ two empirical strategies that both address this problem. First, we use a large sample of singletons to estimate whether first-borns are affected by the gender of their second-born sibling. Second, we look at a sample of dizygotic (i.e. non-identical) twins. We find that a same-sex sibling increases men's earnings and family formation outcomes (marriage and fertility), as compared to an opposite-sex sibling. The results for women are similar but the effects are smaller in magnitude and less robust. We argue that the income result for men could be driven by competition between brothers, as we find that men with brothers choose higher paying occupations. For women, we find suggestive evidence that the income premium may come partly from lower unemployment, which could be due to shared job search networks. The effects on family formation might stem from differential parental treatment for men, and from competition between sisters for women.

1. Introduction

Siblings play a significant role in most people's lives. They influence the environment in which children grow up and often remain important figures in adulthood. They provide reference points to each other and also alter interactions with parents (Adams, 1999; Keim et al., 2009; McHale et al., 2003). Given that siblings affect family environment along several dimensions, they may contribute to individuals' motivation, conformity to social norms and access to different types of information. While these observations suggest that siblings have a big potential to have an impact on various long-term outcomes, evidence on causal effects in sibling relationships is still scarce.¹ We try to shed more light on

the role of siblings by studying the impact of one particular factor: we examine how the gender of a sibling affects individuals' labor market outcomes and family formation.

We expect gender to be a relevant factor because having a brother implies a different family environment and peer relationship than having a sister. Research on family conditions shows that parents are more likely to gender-differentiate their parenting in case of opposite-sex children than in case of same-sex children (McHale et al., 2003). Differential parental treatment is also present in adulthood. In particular, parents support the family formation of daughters more than the family formation of sons, for example by providing more informal childcare (Danielsbacka et al., 2011; Pollet et al., 2009). In addition to the parental

[☆] We are grateful to Hessel Oosterbeek, Erik Plug, Thomas Buser, Paul J. Devereux, Zsombor Cseres-Gergely and to seminar participants in Amsterdam, at the 2012 Summer Workshop of the Institute of Economics, Hungarian Academy of Sciences and at the ESPE Annual Conference 2014 for their comments and suggestions. Part of this research was carried out at the University of Amsterdam during the PhD studies of Noemi Peter. Lundborg acknowledges support from the Swedish Research Council for Health, Working Life, and Welfare, FORTE [grant numbers 2010-0745]. Part of the data used in this paper comes from the Swedish Interdisciplinary Panel (SIP), administered by the Centre for Economic Demography, Lund University, Sweden. The authors have no other funding sources or interests to disclose.

* Corresponding author.

E-mail addresses: a.n.peter@rug.nl (N. Peter), petter.lundborg@nek.lu.se (P. Lundborg), sara.mikkelsen@nek.lu.se (S. Mikkelsen), webbink@ese.eur.nl (D. Webbink).

¹ This literature is small but emerging, with studies such as Altonji et al. (2017); Bingley et al. (2017); Black et al. (2017); Dahl et al. (2014); Joensen and Nielsen (2018); Nicoletti and Rabe (2014) and Breining (2014). Another related literature is the one on birth order effects (see e.g. Black et al., 2005; Booth and Kee, 2009; de Haan, 2010; Breining et al., 2017 and Breining and Doyle, 2015).

<https://doi.org/10.1016/j.labeco.2018.06.006>

Received 24 March 2016; Received in revised form 13 June 2018; Accepted 24 June 2018

Available online 7 July 2018

0927-5371/© 2018 Elsevier B.V. All rights reserved.

treatment channel, siblings can have direct effects as well. These effects can differ by the gender composition of the sibship as same-sex sibships may be more competitive (Conley, 2000). Brothers and sisters also provide different reference points to their siblings because of gender differences in general. For example, women typically marry and have children at a younger age, work in different type of jobs and have lower earnings than men on average. In sum, those with a sister are exposed to different influences than those with a brother. We examine whether this has an impact on individuals' labor market outcomes and family formation. We focus on these outcome variables as many gender differences relate to these domains.

We investigate this question with two different empirical strategies. First, we look at singletons who have at least one younger sibling. Specifically, we analyze how the gender of the second-born sibling affects the outcomes of the first-born. We examine men and women separately, thus, we compare first-born men (women) who have a second-born brother to first-born men (women) who have a second-born sister. For this strategy we make use of the Swedish Multigenerational Register which contains family links for the entire population of Sweden. This enables us to create a large sample of first-born singletons (around 1 million observations for men and women altogether). We also have access to background variables which we use to validate our identification assumption empirically, that is, that the sex of the sibling is as good as randomly assigned. In addition, we use high-quality data from various registers to investigate not only the main effects but also provide insights about factors such as occupational choice, unemployment and heterogeneity in the effect of a sibling's gender. In this additional analysis we provide suggestive evidence about potential underlying mechanisms.

Our second empirical strategy is to look at twins and analyze how the gender of a co-twin affects the outcomes of the other twin. We compare men (women) who have a co-twin brother to men (women) who have a co-twin sister. For this strategy we make use of data from the Swedish Twin Registry which contains information on twins' zygosity. This information is important for our purposes because only the sex of dizygotic (i.e. non-identical) co-twins is random; the sex of monozygotic (i.e. identical) twins is always the same as the sex of their co-twin (see Section 2). Therefore, estimates on the sample of all twins could suffer from "zygosity bias". That is, the coefficient of the co-twin's gender could pick up potential differences between dizygotic and monozygotic twins. We avoid this problem by restricting the estimation sample to dizygotic twins.

The literature on the effect of siblings' gender is small but recently expanding. The findings and their interpretations are both mixed.² The pioneering study of Butcher and Case (1994) found that women with any sisters attained lower education than women with only brothers, but their results could not be replicated by others (Kaestner, 1997; Conley, 2000; Hauser and Kuo, 1998).³ Gielen et al. (2016) examine the earnings of twins and closely spaced singleton sibling pairs. They find that those with a same-sex sibling have higher earnings than those with an opposite-sex sibling, except in the case of closely spaced singleton men.⁴ Brenøe (2018) focuses on the impact of a sibling's gender on women's choice of occupation and partner. Specifically, the main outcomes are whether women and their partners work in male dominated occupations, such as those within Science, Technology, Engineering and Mathematics (STEM). She finds that brothers decrease (increase) the likelihood that women (women's partners) work

in such occupations. Consistent with this finding, she also finds that brothers decrease women's earnings. Rao and Chatterjee (2018) and Cools and Patacchini (2018) both look at self-reported wages from the US. The two studies use two different surveys and reach different conclusions. Rao and Chatterjee (2018) finds that the wages of women do not change by sibling sex composition, while the wages of men are increasing in the proportion of siblings who are brothers. They also find that according to a job search questionnaire, same-sex siblings serve more often as contacts to get a job. They argue that this might explain their finding on men's wages. Cools and Patacchini (2018) find that the wages of men do not change by sibling sex composition, but women with brothers earn less. Their sample is too young to measure actual family formation, but self-reported intentions show that women with brothers are more family-centric. The authors conclude that these family-focused intentions could partly explain their findings.

In terms of empirical approach, the tradition is to include both older and younger siblings in the sample.⁵ However, since parents may prefer certain sex compositions over others (Angrist and Evans, 1998), the gender of older children can affect subsequent fertility decisions. This means that the gender of an earlier born child influences the selection of a potential later born child into the sample. As we explain in Section 2 and show empirically in Section 4, this can lead to biased estimates since parental preferences may affect children's outcomes.

Both of our empirical strategies circumvent this problem. The singleton strategy avoids the selection problem since the gender of the second-born child cannot affect the existence of the first-born child retroactively. The twin strategy exploits the fact that twins are born at the same time, so parents cannot make decisions about one twin based on the gender of the other twin.⁶ By executing these two empirical strategies we aim to obtain clear evidence on the effect of a sibling's gender. As we use two different samples for the two strategies, we can get a better understanding of the robustness of the effects.

Our paper provides several other contributions as well. We give a comprehensive picture of the effects as we focus not only on labor market outcomes but also on the family formation of both genders. Our sample of singletons is substantially larger than the samples of the other studies. We have high quality register data not only on important outcome variables like earnings, but also on several additional variables that can be used to learn more about the underlying channels.

In both samples, we find that a same-sex sibling increases men's earnings, their probability of ever getting married and having children, and their number of children. Our results for women are less consistent between the two samples. In the singleton sample, we find that women with a same-sex sibling earn more and are slightly more likely to form a family. However, we do not find these effects in the twin sample, which consists of older cohorts.

We find that our main results are unlikely to be explained by growing up in a larger family or by economies of scale. Instead, we argue that the positive effect on men's income could be largely driven by competition between brothers, as we find that men with brothers choose higher paying occupations. For women, we find no substantial impact on occupational choice. Instead, the positive effect on women's earnings could partly come from lower unemployment. We argue that a plausible explanation for the reduction in unemployment is that women can use a sister's job search network more efficiently than the network of a brother. For family formation, we find empirical patterns that suggest that differential parental treatment could be an important channel for men, while competition between sisters might explain the effects for women.

² We focus here on results from Western countries. For results on other populations, see e.g. Parish and Willis (1993), Morduch (2000), Chen et al. (2017) and Jayachandran and Pande (2017).

³ Similarly, a draft by Pettersson-Lidbom et al. (2008) finds little evidence of an effect on various educational and labor market outcomes.

⁴ Gielen et al. (2016) do not investigate specific social mechanisms behind their results as they focus on the relation between a potential biological factor and wages. We will discuss their approach and its differences from our strategy in Section 4.

⁵ The two new working papers, Cools and Patacchini (2018) and Brenøe (2018), and a draft by Pettersson-Lidbom et al. (2008) do not follow this tradition. Rao and Chatterjee (2018) do a robustness check where they only use the gender of the next younger sibling as treatment.

⁶ Selective abortion and IVF techniques were not available in the time period that we examine.

The paper unfolds as follows. Section 2 explains our empirical approach in more detail and Section 3 describes the data. We present the results in Section 4 and discuss potential explanations in Section 5. Finally, Section 6 concludes.

2. Empirical strategy

Our aim is to estimate the causal effect of a sibling's gender. The traditional approach is to compare those who have a brother to those who have a sister. However, this approach can lead to biased results when the estimates are based on all kinds of siblings. The problem arises because the gender of older children can influence whether parents proceed to have a next child.

For example, in developed countries a typical empirical finding is that parents of two boys or two girls are more likely to have additional children than parents of a boy and a girl (Angrist and Evans, 1998; Hank and Kohler, 2000). Clearly, this indicates parental preferences for children's sex composition. At the same time, the exact nature of these preferences is less obvious. One possibility is that this pattern arises because some families prefer a gender mix. However, another possibility is that this is not about preferring a mix, but some families prefer boys while others prefer girls. According to this explanation, some parents with same-sex children proceed to a next child because their first two children are not from the preferred sex. Families with a boy and a girl do not proceed, since they have a child from the preferred gender for sure. Note that for this explanation to be true, it is not necessary to observe differential progression after the first child. If the number of boy-preferring families equals the number of girl-preferring families, the same number of families will proceed to a second child after a first-born boy as after a first-born girl.⁷

As this example demonstrates, parental preferences can be of various types. Thus, parents who proceed to a next child after a boy may have different preferences than parents who proceed after a girl. If people with different preferences raise their children differently, comparing individuals with an older brother to individuals with an older sister will lead to biased estimates. As we show in Appendix A.1, this intuition can be formalized as a selection bias problem, as in Heckman (1979).

To avoid the selection bias we need a setting where the sibling's gender cannot affect parents' decisions. We look at two such settings in this paper. First, we analyze first-born singletons who have a second-born singleton sibling (and they may also have additional younger siblings). We investigate how the older child is affected by the gender of the second-born sibling. This strategy is free from the selection bias since the gender of the second-born child cannot affect the existence of the first-born child retroactively. Second, we analyze a sample of dizygotic (DZ) twins.⁸ This strategy circumvents the selection problem because twins are born at the same time, and hence parents cannot make decisions about one twin based on the sex of the other twin. By executing these two strategies we aim to obtain strong evidence on the causal effect of a sibling's gender.

We analyze men and women separately both in the singleton and in the twin setting. Thus, altogether we conduct four comparisons: in comparison 1 (2) we compare first-born singleton men (women) who have a second-born brother to first-born singleton men (women) who have

a second-born sister; in comparison 3 (4) we compare men (women) who have a co-twin brother to men (women) who have a co-twin sister. All four comparisons have the same general structure, so our empirical strategy can be summarized by a simple OLS regression that is estimated on four different samples. Specifically, we estimate

$$Y_i = \alpha + \beta G_i + \gamma X_i + U_i \quad (1)$$

where Y_i denotes the outcome of individual i who is a first-born singleton man (woman) in comparison 1 (2), and a twin man (woman) in comparison 3 (4).⁹ G_i denotes the gender of the second-born singleton in comparisons 1 and 2, and the gender of the co-twin in comparisons 3 and 4. X_i denotes our control variables. In the singleton samples we control for birth year fixed effects. In the twin samples, in addition to birth year fixed effects, we control for the presence of older siblings.¹⁰

For the interpretation of the effects, it is useful to let the sibling gender dummy G_i indicate a brother for the male samples and a sister for the female samples. This enables us to look at the estimates as the effect of having a same-sex sibling, for both men and women.

Since the gender of existing children can affect the number of younger siblings, the estimates have to be interpreted as a total impact that includes the family size channel. Controlling for family size would lead to bad control bias (Angrist and Pischke, 2009), so instead we will investigate the family size channel separately in Section 5.2. The analysis in that section indicates that the effects that we find are not driven by this channel.

Our key identification assumption is that sex is random, which is in line with the established view in biology.¹¹ In Section 3 we look at empirical patterns in the singleton data and find that they are consistent with this assumption: the spacing of births and parental background cannot explain variation in G_i .¹² In the twin setting, zygosity could be an additional factor that threatens identification. By zygosity we refer to the fact that twinning can occur in two ways (Phillips, 1993). Dizygotic (DZ) twins arise from two different zygotes, which result from the separate fertilization of two different ova by two different sperms. Their sex is determined the same way as the sex of singletons: the offspring will be male (female) if the sperm bears a Y (X) chromosome. In contrast, monozygotic (MZ) twins arise when one zygote splits into two genetically identical units, so MZ twin pairs are always same-sex. This means that conditional on own sex, the sex of the co-twin correlates with zygosity. Thus, if we could not differentiate twins by zygosity, differences between opposite-sex and same-sex twins could reflect not only the impact of the sibling's gender, but also the impact of having a genetically identical sibling or other potential differences between MZ and DZ twins. To make sure that our estimates are free from this “zygosity bias”, we sought out a dataset which contains zygosity information and base our twin strategy solely on DZ twins.

⁹ In the twin samples, individuals can come from the same family (e.g. we can obtain two observations from a male-male twin pair). Therefore we will estimate standard errors such that we take into account potential clustering at the level of the family of origin.

¹⁰ To avoid sample size problems, we do not restrict the twin sample to those with no older siblings, but control for the presence of older siblings instead. This is controlling for already realized fertility that cannot be affected by the sex composition of the twin pair. In case of the singletons there are a substantial number of births and hence we can restrict the sample to first-borns for an easier interpretation of the effect.

¹¹ See e.g. Reece et al. (2010). Note that medical technologies for sex selection were unavailable in the time period that we examine. As far as natural sex selection is concerned, Wilcox et al. (1995) found that the timing of intercourse in relation to ovulation has no significant effect. Similarly, Gray et al. (1998) found no evidence for the hypothesis that maternal hormones influence sex selection.

¹² We do not have access to such background variables in the twin sample (see Section 3 for details about our data).

⁷ Andersson et al. (2006) investigate births in Nordic countries between 1961 and 1999. They find that differences in parity progression ratios fluctuate over time, so there is no imbalance in second-birth probability that would be stable over the entire period. This pattern is consistent with heterogeneous parental preferences that exhibit changing frequencies in the population. Our own data for Swedish cohorts 1938–1977 shows similar fluctuations: in 1938–1949 and in the 60s, parents are slightly more likely to have a second child after a first-born boy than after a first-born girl, but the pattern is reversed in 1970–1977 and there is no imbalance in the 50s.

⁸ Unlike monozygotic (MZ) twins, dizygotic (DZ) twin pairs can be opposite-sex or same-sex; see further discussion at the end of this section.

3. Data

We use two different datasets in this paper, one for singletons and one for twins. We start this section by explaining how our singleton dataset is compiled, and then proceed to describe the source of the twin data.

Our dataset on singletons comes solely from Swedish administrative registers. We identify first-born singletons born 1938–1977 with a second-born full singleton sibling using family links in the Multigenerational register. We take 1938 as the first cohort because we want to select individuals without older siblings and the Multigenerational register does not contain information on siblings before 1932. The 6 years cutoff is reasonable given that it is atypical to space the birth of siblings further apart.¹³ For the same reason, we restrict the sample to individuals whose second-born sibling was born at most 6 years apart. Note that, in addition to the second-born sibling, the first-born singletons may have other younger siblings.

We add information from several registers in order to create outcome variables. Permanent income is our main labor market outcome but we also consider whether a sibling's gender has an impact on years of schooling since this in turn is bound to affect earnings. We derive years of schooling from highest education registered in the education register (utbildningsregistret, UREG 1990–2011), and when that is not available, we use educational information from the 1970 census. The income data comes from the national register on taxable income (IoT, Inkomst- och taxeringsregistret). We use records from years 1971–2011 (all converted to 2010 prices). The income variable measures total taxable annual labor income: it includes wages, income from self-employment, pensions, sickness benefits and other taxable benefits.¹⁴ From any given year, we use the available income records of those who are aged between 25 and 64 in that year. The main source of income in this age range is typically the labor market. Since we have income records from many years, we have several observations per person. We take their average to get a permanent income measure for each individual. We use the natural logarithm of this measure, $\ln(\text{permanent income})$ as outcome variable. Note that people may have zero income in one year or another; however, almost nobody has zero income in all years so taking the logarithm leads to very few (less than one percent) missing observations. A consequence of this approach is that permanent income will be an average over different number of observations in different cohorts. This implies potential heteroskedasticity, even though cohort fixed effects are included in the regressions. We use robust standard errors to take this into account.

We analyze fertility by looking at whether the individual had any children and the number of children. The registers provide information on children born to the singletons up to 2012; thus, our sample is at least 35 years old at the time of measurement. To analyze timing, we calculate age at first birth from the birth year of children. Furthermore, we define “ever married” as a dummy which takes the value one if the individual has ever been married, and zero otherwise. Cohabitation is not included in “ever married” for singletons because cohabitation is only observed in the registers if the couple has children together.

Our twin analysis is based on a sample of dizygotic twins who were born in Sweden in 1926–1958. The data comes from the SALT project (Screening Across the Lifespan Twin Study) of the Swedish Twin Registry (STR), which surveyed twins irrespective of their sex composition (Lichtenstein et al., 2002). The STR had records of every twin birth from

the national birth register and they contacted all available twins for a phone interview.¹⁵ The process started with a pilot in 1996–1997 and then full scale data collection took place in 1998–2002. The oldest cohorts were surveyed first, followed by younger cohorts.¹⁶ The interviewers tried to get as many responses as possible; e.g. people were called back at least five times if they were unavailable. This led to the remarkably high response rate of 74%.

The STR determined zygosity based on questions about intrapair similarity in childhood. This classification was validated in the pilot: 13 DNA-markers were analyzed in blood samples. Zygosity assignment proved correct in 99% of the cases (see Lichtenstein et al., 2002).

The SALT survey provides information on marital status, children and siblings as well. The questions on children and siblings refer to biological children and siblings (other than the co-twin) who are still alive. We use this information to construct family formation variables for the twin sample.¹⁷ For twins' marital status, the dummy “ever married” includes cohabitation.

For the labor market outcomes of the twins, we are able to use the same administrative registers as for the singleton sample. However, in case of the twin sample our access to the registry data does not go further than 2007, so we construct permanent income using records from 1971 to 2007 and years of schooling using records from 1990 to 2007. In a few cases where education is unobserved in the register, we use information from the survey.

For both samples, we balance across the outcome variables, i.e. we drop individuals whose income measure, education, marital status or number of children is missing.¹⁸ We present summary statistics in Table 1, where column 1 shows statistics for first-born singletons whose second-born sibling is of the opposite sex, while column 2 shows statistics for first-born singletons whose second-born sibling is of the same sex. Similarly, column 3 is for dizygotic twins whose co-twin is of the opposite sex, while column 4 is for dizygotic twins whose co-twin is of the same sex. Recall that the SALT surveyed individuals, not twin pairs, so inclusion in the twin sample is not conditional on the participation of the co-twin. These statistics provide a first impression of the effect of having a brother or a sister. Although the differences are quite small, we observe that male singletons with a brother, compared to those with a sister, are more likely to get married and to have children. They also have more children and earn somewhat more. Although the differences are less pronounced, the same patterns emerge for female singletons with a sister compared to those with a brother. For male and female twins, having a co-twin of the same sex is mostly consistent with what we observe for the singletons.

Since the sex ratio in the general population is slightly imbalanced (about 51% males, 49% females), the share of same-sex siblings is 51% for singleton men and 49% for singleton women. These shares are fifty-fifty for our dizygotic twins, which is in line with previous research on

¹⁵ The SALT also contains twins from cohorts 1886–1925 but there are disproportionately few opposite-sex twins for these cohorts because of differences in the data collecting process. We do not use these early cohorts in our research, only the complete cohorts of 1926–1958.

¹⁶ Note that participants are 41–74 years old at the time of data collection. A population above 40 is desirable for our purposes, since fertility is typically completed by that age.

¹⁷ We imputed data for childless women. Women who never gave birth were not asked questions about children, thus information on children are coded ‘missing’ for these women. We do not have access to the variable on childbirth so we infer zero children for women who did not answer questions in the children section but who did answer questions in the preceding section of the survey. Since data on men is complete, we can check the accuracy of our imputation strategy. Reassuringly, 99.7% of men who answered questions in the preceding section also answered the question on children.

¹⁸ Since age at first birth cannot be calculated for individuals who do not have children, we do not balance the sample with respect to this variable.

¹³ In almost ninety percent of the cases, the first-born and the second-born have an age difference of at most 6 years. Starting the sample from a later cohort does not change much in this statistic.

¹⁴ Many benefits became taxable in 1974 and therefore they are included in the income measure since then. This means that our variable was somewhat more restrictive in the pre-1974 years. Nonetheless, recent work by Björklund et al. (2009) suggests that this is not a major concern, as benefits constitute a very small share of total earnings. In any case, the estimates are similar if we use income records starting from 1974.

Table 1
Summary statistics.

	First-born singletons (cohorts: 1938–1977)		Dizygotic twins (cohorts: 1926–1958)	
	OS SB sibl.	SS SB sibl.	OS co-twin	SS co-twin
<i>Panel 0: Men</i>				
Ever married ^a	0.670 (0.470)	0.679 (0.467)	0.875 (0.330)	0.889 (0.315)
Any children	0.792 (0.406)	0.797 (0.402)	0.817 (0.387)	0.830 (0.376)
Number of children	1.775 (1.236)	1.792 (1.235)	1.865 (1.239)	1.908 (1.234)
Age at first birth	28.901 (5.431)	28.790 (5.419)	27.882 (5.154)	27.854 (5.229)
Number of siblings ^b	1.752 (1.108)	1.833 (1.113)	2.006 (1.890)	1.992 (1.815)
Years of schooling	12.158 (2.679)	12.168 (2.677)	10.778 (3.120)	10.781 (3.131)
Ln(permanent income)	12.455 (0.566)	12.463 (0.561)	12.434 (0.397)	12.447 (0.385)
Number of obs.	257174	271432	6759	6668
<i>Panel 1: Women</i>				
Ever married ^a	0.745 (0.436)	0.750 (0.433)	0.896 (0.305)	0.896 (0.305)
Any children	0.872 (0.334)	0.874 (0.332)	0.875 (0.330)	0.871 (0.335)
Number of children	1.983 (1.140)	1.985 (1.134)	1.984 (1.170)	1.996 (1.183)
Age at first birth	26.448 (5.154)	26.362 (5.138)	25.064 (4.774)	24.822 (4.725)
Number of siblings ^b	1.756 (1.111)	1.841 (1.134)	2.048 (1.867)	2.031 (1.858)
Years of schooling	12.571 (2.579)	12.563 (2.583)	10.839 (2.976)	10.662 (2.981)
Ln(permanent income)	12.104 (0.579)	12.109 (0.576)	11.935 (0.557)	11.914 (0.603)
Number of obs.	256788	241953	7401	7341

Note: OS means opposite-sex, SS means same-sex, SB means second-born. Standard deviations in parentheses. The singleton sample consists of individuals whose second-born sibling was born at most 6 years later.

^a Marriage includes cohabitation in case of the twin sample, but not in case of the singleton sample.

^b Number of siblings does not include the co-twin.

Table 2
Randomness of the sex of the second-born sibling.

	Second-born is female
<i>Panel 0: Men</i>	
Birth spacing	–0.000 (0.000)
Age 1st birth mother	–0.000 (0.000)
Age 1st birth father	0.000** (0.000)
High parental educ.	0.002 (0.002)
Mixed parental educ.	0.002 (0.002)
Missing parental educ.	0.001 (0.003)
Observations	528,606
R-squared	0.000
<i>Panel 1: Women</i>	
Birth spacing	0.000 (0.000)
Age 1st birth mother	0.000 (0.000)
Age 1st birth father	–0.000 (0.000)
High parental educ.	–0.002 (0.002)
Mixed parental educ.	–0.001 (0.002)
Missing parental educ.	0.003 (0.003)
Observations	498,741
R-squared	0.000

Note: The sample consists of first-born singletons born 1938–1977 whose second-born sibling was born at most 6 years later. Birth spacing means the age difference between the first-born and the second-born child (in months). The regressions control for cohort fixed effects. Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

twins.¹⁹ The pattern in the twin sample is consistent with the notion that male fetuses are more vulnerable and hence a male co-twin can die in utero more easily (Hansen et al., 1999; Mizuno, 2000).²⁰ From Table 1 we can also see that there are somewhat more women in the twin sample. This is probably because women are more likely to respond to surveys (Singer et al., 2000).²¹

For the singleton sample we have richer data available from the registers than for the SALT sample, which we exploit in additional ways. First, we observe educational level not only for the singletons themselves but also for their parents. We utilize this data to get a measure of the socio-economic background of the individuals. We first determine whether the parents have more or less years of schooling than the me-

dian among parents who had their first child in the same year, separately for mothers and fathers. Based on their parents' education, we subsequently create four categories for the singletons: "High parental education" (23%) means that both parents are higher educated than the median; "Mixed parental education" (27%) means that one of the parents has more years of schooling than the median while the other's education is at or below the median; "Low parental education" (38%) means that both parents have at most median years of schooling; "Missing parental education" (12%) denotes that either parent has missing years of schooling. In Section 5 we use this classification of socio-economic background to examine the potential mechanisms behind our results. In addition, we can use the parental educational categories together with additional background variables (age difference between the first-born and second-born sibling in months and the mother's and father's age at the first birth in years) to verify that the gender of the second-born sibling is as good as randomly assigned. More specifically, we run regressions of second-born sex on these background variables, separately for men and women. The results are shown in Table 2. As we can see, the background variables do not explain variation in the gender of the second-born, which supports our identification assumption.

Besides parental background, we also have several additional register-based variables for our singleton sample, such as occupational choice, work industry and unemployment, which we can use to explore potential mechanisms behind the effects of the sibling's gender. The variables used for this additional analysis will be described in more detail in Section 5 (also see Table A.6 in Appendix A.4).

¹⁹ The fifty percent share is so widely accepted that it is even called a "rule", the so-called Weinberg's differential rule (WDR) (Benirschke et al., 2012). Although James (1979) raised questions about this, several studies found that it is supported empirically (Vlietinck et al., 1988; Fellman and Eriksson, 2006; Husby et al., 1991). Thus, Elston et al. (2002) conclude that this is a reasonable approximation and actual deviations are small.

²⁰ If fetal vulnerability is related to later health, brothers in the twin sample might be healthier on average than brothers in the singleton sample, which might lead to observing more positive effects of brothers in the twin sample than in the singleton sample.

²¹ Note that we analyze men and women separately, so this will simply mean that the analysis of the latter will be based on more observations. The response rate is presumably unrelated to the sex of the co-twin since the fifty percent share of sisters is in line with earlier research on twins (see footnote 19).

Table 3
Main results for first-born singletons.

	Ln(permanent income)	Years of schooling	Ever married	Any children	Number of children	Age at first birth
<i>Panel 0: Men</i>						
SS sibling	0.007*** (0.002)	0.006 (0.007)	0.009*** (0.001)	0.006*** (0.001)	0.018*** (0.003)	−0.124*** (0.016)
Observations	528,606	528,606	528,606	528,606	528,606	419,902
Outcome mean	[12.46]	[11.93]	[0.70]	[0.80]	[1.78]	[28.39]
<i>Panel 1: Women</i>						
SS sibling	0.005*** (0.002)	−0.003 (0.007)	0.005*** (0.001)	0.002** (0.001)	0.001 (0.003)	−0.074*** (0.015)
Observations	498,741	498,741	498,741	498,741	498,741	435,395
Outcome mean	[12.09]	[12.27]	[0.76]	[0.88]	[1.91]	[26.23]

Note: The sample consists of first-born singletons born 1938–1977 whose second-born sibling was born at most 6 years later. “SS sibling” is a dummy which takes 1 if the second-born sibling is of the same sex as the first-born and 0 if the second-born sibling is of the opposite sex. “Ever married” is a dummy which takes 1 if the individual has ever been married, 0 otherwise. Permanent income is average income in years 1971–2011 (if subject is aged 25–64 in the given year). All regressions control for cohort fixed effects. Robust standard errors are shown in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4. Results

In this section we show results on the effect of a sibling’s gender. As explained in the empirical strategy section, we analyze men and women separately and define the explanatory variable of interest G_i as having a sibling of the same gender in both the male and female sample. Note that the coefficients capture a total impact which can result from several underlying mechanisms; we will discuss these after the main findings are presented.

The results for first-born singletons are shown in Table 3. Recall that these estimates show the effect of the second-born’s gender on the outcomes of first-born singletons who have at least one younger sibling. For men, we find that those with brothers have a higher permanent income than those with sisters. This effect does not seem to arise via education as the next column shows that the sibling’s gender has no impact on years of schooling. We can also see that men with brothers have a higher probability to ever be married and have children. The estimates also suggest that they have their first child at a younger age.²² For singleton women, the estimated effects of having a same-sex sibling are similar to the results on men, although the magnitudes are smaller. On income, there is a small positive effect of having a sister. Furthermore, the family formation estimates show that women with sisters have a slightly higher likelihood of getting married and having children than women with brothers. The estimates also suggest that these women start having children at a younger age. The results are robust to the inclusion of additional controls, as can be seen from Table A.3 in Appendix A.3.²³

Since we observe income over several years (1971–2011) we can also check whether the income effect differs during the life-cycle. Fig. A.1 in Appendix A.3 shows results from running separate regressions for income at different ages. The outcome is the natural logarithm of a five-year moving income average. Men with brothers have a higher average

income than men with sisters already in their mid 20s and the premium is increasing in age up to the mid 50s. This suggests that men with brothers have a steeper income profile than men with sisters. For women, a same-sex sibling does not seem to alter the slope of the income profile. The small estimated income effect is more or less constant from the mid 30s and onwards.

Recall that our identification strategy circumvents the selection bias. As we argued in Section 2 (and Appendix A.1), this does not hold for alternative strategies such as comparing individuals with older brothers to individuals with older sisters, since these two types of individuals may come from families with different types of gender preferences. In Table A.4 of Appendix A.3 we show estimated effects of having an older sibling of the same sex for second-born singletons. These estimates confirm the notion of selection bias as they are completely different from the results of Table 3 for a number of outcomes. For example, the coefficients are zero for men’s income and marriage, and the estimates on any children and number of children are negative for both men and women. These findings support our argument that one should not include younger siblings in the estimation sample.

The main results for twins are reported in Table 4. For men, the results are in line with what we observe for male singletons. On permanent income we see a positive impact (significant at the 10% level), while there is no impact on education.²⁴ The gender of the co-twin also affects family formation: those with brothers are more likely to get married and have children, and they also have more children. The results for women show a different pattern. There is no significant effect on income. Women with sisters have lower education, which is in line with the early study of Butcher and Case (1994) but not with later studies. For family formation, the estimates are small and insignificant. The exception is the estimated coefficient for age at first birth, which shows that women who get children do so at a younger age if they have a sister compared to a brother.

As we explained in the empirical strategy section, these estimates are free from zygosity bias as they are based solely on DZ twins. Comparing these estimates to results on all twins (including MZs) can be helpful, since zygosity information is often difficult to obtain for researchers. We investigate this issue in Appendix A.2 and find that zygosity bias can indeed distort most of the results, with the severity of the distortion varying by outcome and sample.

²² Age at first birth is conditional on having any children, which decreases the sample size for this outcome and warrants an indicative interpretation of the coefficients.

²³ Also note that using a probit or logit model to estimate the impact on the probability of ever getting married or having children does not alter the results for twins or singletons (untabulated results). As another robustness check, we also looked at pooling $\ln(\text{income})$ from different years, instead of calculating permanent income over the years. In addition to the usual controls, these pooled regressions on $\ln(\text{income})$ also include year fixed effects and use weights such that every person is represented equally. We find that the results on income are qualitatively the same as in the main specifications, in both the twin and singleton samples (untabulated results).

²⁴ Also, Ahrenfeldt et al. (2015) find no effect on test scores for twins.

Table 4
Main results for dizygotic twins.

	Ln(permanent income)	Years of schooling	Ever married	Any children	Number of children	Age at first birth
<i>Panel 0: Men</i>						
SS co-twin	0.012* (0.007)	−0.013 (0.055)	0.013** (0.006)	0.013* (0.007)	0.044** (0.022)	−0.035 (0.101)
Observations	13,427	13,427	13,427	13,427	13,427	10,978
Outcome mean	[12.44]	[10.78]	[0.88]	[0.82]	[1.89]	[27.87]
<i>Panel 1: Women</i>						
SS co-twin	−0.014 (0.010)	−0.122** (0.049)	−0.000 (0.005)	−0.004 (0.006)	0.013 (0.020)	−0.207** (0.086)
Observations	14,742	14,742	14,742	14,742	14,742	12,836
Outcome mean	[11.92]	[10.75]	[0.90]	[0.87]	[1.99]	[24.94]

Note: The sample consists of dizygotic twins born 1926–1958. “SS co-twin” is a dummy which takes 1 if the co-twin is of the same sex as the individual, and 0 if the co-twin is of the opposite sex. “Ever married” is a dummy which takes 1 if the individual has ever been married or cohabited, 0 otherwise. Permanent income is average income in years 1971–2007 (if subject is aged 25–64 in the given year). All regressions control for older sibling, missing sibling information and cohort fixed effects. The standard errors shown in parentheses are robust and clustered by the family of origin. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

In our twin analysis we included dummy variables for the presence of, and missing information on, older siblings.²⁵ With these variables we aim to control for variation in outcomes that arises because of differences in family constellation. However, the impact of the co-twin's gender itself may differ depending on the presence of older siblings. In Appendix A.3, Table A.5, we show results from including interactions between the variables on older siblings and having a same-sex co-twin. The estimated effects for twins who reported zero older siblings are shown in the first row of each panel. They are in line with the earlier results on men's income and family formation. Although the estimated effects are not statistically significant due to the inclusion of the interaction terms, they are of similar magnitude as in Table 4. On the other hand, the negative effect on women's education appears to be driven by those with older siblings.

Taking the results from our two estimation strategies (singletons and twins) together, we get strong evidence that a sibling's gender has a positive impact on men's earnings, probability of family formation and number of children. The magnitudes of the effects are small but not trivial. For instance, the estimated earnings premium for men with brothers is about one percent over the whole life cycle. This is similar to previous findings on other dimensions of family environment, such as birth order effects (Black et al., 2005). Moreover, male twins with a co-twin brother have about 0.04 more children. This is equivalent to at least 2% more children.

The effects for women are less clear cut. The effect on the schooling of twin women is not robust to family composition and is not observed in the singleton sample. The effects on women's income and probability of family formation are only observed in the singleton sample and they are smaller than the corresponding effects for men. However, women with sisters have their first child at an earlier age in both samples. In Fig. A.2 in Appendix A.3 we examine the result on timing further by plotting the effect of a same-sex sibling on the probability of being a mother by a certain age. In the singleton sample, we find that women with sisters are more likely to have children already at the end of their teenage years. Nonetheless, most of the effect comes from women's fertility decisions in their 20s, in both the singleton and the twin sample. In the twin sample women with brothers eventually catch up, while in the singleton sample this happens only partly, leading to the small increase in the probability of having children that we observe already in Table 3.

Differences in results for women might come from the fact that the twin sample consists of older cohorts. Women's educational and labor market position improved substantially over the years, and family planning also became easier.²⁶ It is possible that these improvements are important for the beneficial effect of sisters to arise. Indeed, when we look at differences by cohort, we find that the same-sex coefficient on women's income and education is lower in the early cohorts (untabulated results). However, this does not apply for the results on family formation.

Another reason why results between the two samples might differ is spacing. To explore this possibility we look at results from a sample of singletons who are closely spaced. Table 5 shows results on first-born singletons whose younger siblings are born within 24 months. For men, the estimates become larger in magnitude compared to the main results in Table 3. It is also notable that the estimated effect on men's income is very similar to that of twins. For women, the magnitude of the income effect increases while the effects on family formation disappear. Thus, a closer spacing might explain why we did not find effects on the family formation of twin women.

The approach of comparing the effect of a sibling's gender for twins and closely spaced singletons has been used by Gielen et al. (2016) to examine the relation between a biological factor and wages. In particular, their study is based on the Twin Testosterone Transfer (TTT) hypothesis, which is the assumption that testosterone can transfer between twins in utero. Under this assumption, those with a male co-twin would be exposed to higher prenatal testosterone than those with a female co-twin.²⁷ The main result of Gielen et al. (2016) is that same-sex male twins earn significantly more than opposite-sex male twins while closely

²⁶ Swedish female labor force participation was roughly 50% in the mid 1960s but increased rapidly during the following decade and has stayed somewhat above 80% since the 1980s (Statistics for Swedish women aged 25–64 extracted from <http://www.stats.oecd.org/index.aspx>). Social attitudes towards the use of contraceptives and abortion in Sweden became more liberal in the 1960's. In the mid-1970s important legal changes and policy efforts started. A new abortion law was enacted in 1974 making abortion free on request. Contraceptive services were expanded throughout the country in the same years (see Oláh and Bernhardt, 2008; Swedish Institute, 1994).

²⁷ The inspiration for the TTT hypothesis comes from animal studies which found that testosterone can transfer between littermates (e.g. vom Saal and Bronson, 1980; Even et al., 1992). The impetus to the TTT hypothesis was given by Miller (1994), who speculated that testosterone may transfer in case of humans as well and argued that human twins should be studied to uncover potential signs of testosterone transfer. Several papers examined the outcomes of twins in this vein - for reviews, see Cohen-Bendahan et al. (2005a) and

²⁵ Missing sibling information means that the person did not answer the questions on the (birth year of) siblings (about eight percent of the sample). Almost sixty percent of the twins report at least one sibling who is older.

Table 5
Main results for closely spaced singletons (at most 24 months spacing).

	Ln(permanent income)	Years of schooling	Ever married	Any children	Number of children	Age at first birth
<i>Panel 0: Men</i>						
SS sibling	0.010*** (0.003)	0.030** (0.015)	0.013*** (0.002)	0.007*** (0.002)	0.025*** (0.007)	−0.134*** (0.034)
Observations	129,492	129,492	129,492	129,492	129,492	102,205
Outcome mean	[12.42]	[11.94]	[0.68]	[0.79]	[1.82]	[28.53]
<i>Panel 1: Women</i>						
SS sibling	0.011*** (0.003)	0.017 (0.014)	0.001 (0.002)	0.001 (0.002)	−0.007 (0.007)	−0.037 (0.031)
Observations	122,398	122,398	122,398	122,398	122,398	106,912
Outcome mean	[12.07]	[12.33]	[0.75]	[0.87]	[2.04]	[25.89]

Note: The sample consists of first-born singletons born 1938–1977 whose second-born sibling was born at most 24 months later. “SS sibling” is a dummy which takes 1 if the second-born sibling is of the same sex as the first-born and 0 if the second-born sibling is of the opposite sex. “Ever married” is a dummy which takes 1 if the individual has ever been married, 0 otherwise. Permanent income is average income in years 1971–2011 (if subject is aged 25–64 in the given year). All regressions control for cohort fixed effects. Robust standard errors are shown in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

spaced male singletons are unaffected by the gender of their sibling. The authors interpret their finding such that prenatal testosterone increases the wages of men.

Our paper is not specifically focused on the TTT hypothesis, partly because of the lack of clear evidence behind this assumption.²⁸ Nonetheless, given the related context it is noteworthy that in our paper the effect on men’s income is also significant for singletons, which would be difficult to explain with testosterone transfer.²⁹ There are at least two important differences between the empirical strategy of [Gielen et al. \(2016\)](#) and our paper which could explain the difference in findings. First, we focus on first-born singletons whereas [Gielen et al. \(2016\)](#) include singletons irrespective of birth order. As we demonstrated with [Table A.4](#), estimates of the effect of an older sibling’s gender on younger siblings suffer from selection bias, which distorts the results on men’s income towards zero. Second, when we look at closely spaced singletons we use a cutoff of 24 months, which is a common spacing (more than 20% of second-borns are spaced within this range in our data). In contrast, [Gielen et al. \(2016\)](#) choose a cutoff spacing of 12 months in their main specification. This is a very atypical sample: only 1% of second-borns are spaced within this range in our data, and less than 4% of siblings in the data of [Gielen et al. \(2016\)](#). Since preg-

nancies normally last about 9 months, there might be special reasons for a birth spacing of at most 12 months (e.g. religious beliefs, unplanned pregnancies), which could also alter sibling interactions or how they are raised. To examine this issue empirically, we run the estimates on men’s income on the sample of first-borns whose second-born sibling is at most 12 months apart. We found that the coefficient is not only insignificant but actually turns negative: it is −0.012, with a standard error of 0.015. Thus, this small sample of atypically shortly spaced siblings seems to be under the influence of special factors, which obscure the effect that is identified in the main sample of first-borns and in the sample of 24-month-spaced first-borns.³⁰

5. Potential mechanisms

Our results for men show that having a sibling of the same sex increases earnings and the probability of family formation. The pattern is similar for women but less strong and less clear because of differences in results between twins and singletons. In this section we try to shed some light on the potential mechanisms behind these results, using additional register data linked to the singleton sample (see [Table A.6](#) for a description of the additional variables). The analysis is not exhaustive; we focus on channels which we can relate to our data. In [Section 5.1](#), we look at differential parental treatment and direct sibling-to-sibling influences as potential explanations. In [Section 5.2](#), we investigate two mechanisms related to constraints: economies of scale and family size. It should be noted that the analysis of the mechanisms cannot be given a causal interpretation. Instead, this section provides suggestive evidence about the underlying mechanisms by uncovering a number of relevant empirical facts.

5.1. Differential parental treatment and direct sibling-to-sibling influences

A sibling’s gender may affect long-term outcomes indirectly via differential parental treatment and directly via sibling-to-sibling influences. In this subsection we review and empirically explore these two types of mechanisms, first in relation to labor market outcomes and then in relation to family formation.

Differential parental treatment occurs when parents treat boys and girls differently. [McHale et al. \(2003\)](#) reviews research on the family contexts of gender development and finds that such parental be-

[Tapp et al. \(2011\)](#). The reviews find that the results of this literature lack consistency.

²⁸ To the best of our knowledge, the only direct test of this hypothesis is provided by the data of [Abeliovich et al. \(1984\)](#) which do not support the TTT assumption. In particular, [Abeliovich et al. \(1984\)](#) measured testosterone levels in the amniotic fluid of twin fetuses. Their data show that testosterone levels are not elevated for twins with a male co-twin. Other studies that measure testosterone are also not supportive of the assumption that it can transfer from a fetus to others. In particular, several studies failed to find a difference in maternal testosterone levels by the sex of the fetus (see [Glass and Klein, 1981](#); [van de Beek et al., 2004](#); [Cohen-Bendahan et al., 2005b](#); [Rodeck et al., 1985](#); [Toriola et al., 2011](#); only [Meulenberg and Hofman, 1991](#) is an exception) and those that measured testosterone in both amniotic fluid and maternal serum found no correlation between the two measures ([van de Beek et al., 2004](#); [Rodeck et al., 1985](#)).

²⁹ For women, we can rule out the TTT channel in an additional way. Animal studies have shown that females enter puberty later if they develop between male fetuses instead of female fetuses ([Ryan and Vandenberg, 2002](#)). In the human setting this would imply that women with co-twin brothers would enter puberty later than women with co-twin sisters. Thus, we estimate the effect of a co-twin sister on age at menarche (first menstruation) and find a statistically insignificant positive coefficient, which is inconsistent with the predictions of the hypothesis.

³⁰ As a robustness check, [Gielen et al. \(2016\)](#) also look at 24-months spacing, but they do so using both older and younger siblings. As we pointed out above, such estimates can suffer from the selection bias.

Table 6
Exploring labor market mechanisms.

	Share of women in occup.	Ln(occup. income)	Unemp. avg.	Unemp. avg.	Ln(permanent income)	Ln(permanent income)	Ln(permanent income)	Ln(permanent income)
<i>Panel 0: Men</i>								
SS sibling	−0.000 (0.001)	0.005*** (0.001)	−0.001** (0.000)	−0.001 (0.000)	0.007*** (0.001)	0.006*** (0.001)	0.003*** (0.001)	0.002*** (0.001)
Same industry				−0.005*** (0.001)				
Unemp. avg.						−1.409*** (0.005)		−1.175*** (0.005)
Ln(occup. inc.)							0.769*** (0.002)	0.678*** (0.002)
Observations	456,561	456,561	456,561	456,561	456,561	456,561	456,561	456,561
Outcome mean	[0.28]	[12.72]	[0.09]	[0.09]	[12.54]	[12.54]	[12.54]	[12.54]
<i>Panel 1: Women</i>								
SS sibling	0.000 (0.001)	−0.000 (0.001)	−0.002*** (0.001)	−0.001*** (0.001)	0.004*** (0.001)	0.002* (0.001)	0.004*** (0.001)	0.002*** (0.001)
Same industry				−0.003*** (0.001)				
Unemp. avg.						−1.227*** (0.005)		−1.037*** (0.005)
Ln(occup. inc.)							0.725*** (0.002)	0.614*** (0.002)
Observations	444,383	444,383	444,383	444,383	444,383	444,383	444,383	444,383
Outcome mean	[0.69]	[12.51]	[0.12]	[0.12]	[12.19]	[12.19]	[12.19]	[12.19]

Note: The sample consists of first-born singletons born 1938–1977 whose second-born sibling was born at most 6 years later. “SS sibling” is a dummy which takes 1 if the second-born sibling is of the same sex as the first-born and 0 if the second-born sibling is of the opposite sex. Permanent income is average income in years 1971–2011 (if subject is aged 25–64 in the given year). The occupational variables are described in Table A.6. The sample is balanced across outcomes and control variables. All regressions control for cohort fixed effects. Robust standard errors are shown in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

havior is more common in case of opposite-sex children than in case of same-sex children. In addition, Brenøe (2018) provides evidence that parents of opposite-sex children gender-specialize their parenting more than parents of same-sex children. As a result of these differences, individuals with opposite-sex siblings might make more gender-stereotypical choices than those with same-sex siblings. For example, Brenøe (2018) finds that brothers decrease women’s likelihood to work in male dominated occupations, such as those within Science, Technology, Engineering and Mathematics (STEM). This can lead to differences in earnings because of the general wage gap between traditionally male- and female-dominated occupations.

To explore the differential parental treatment channel we examine whether individuals with an opposite-sex sibling choose more gender-stereotypical occupations than individuals with a same-sex sibling. We use the Swedish version of ISCO occupational codes (4-digit level) from the registers, observed between 2001 and 2011, and assign each person their mode occupation. To measure how gendered each occupation is, we calculate the population share of women employed within that occupation. Column 1 of Table 6 shows that the gender of the sibling has no effect on this dimension of occupational choice. As compared to those with an opposite-sex sibling, those with a same-sex sibling do not choose less stereotypical occupations.³¹

In addition to the indirect influence via parents, siblings might affect each other directly. Such a direct way is that siblings provide reference points to each other (Adams, 1999; Keim et al., 2009; Kuegler,

2009). The gender of the sibling is relevant in this context for two reasons. First, same-sex siblings have more salient reference points which might make these sibships more competitive (Conley, 2000). For example, Joensen and Nielsen (2018) show that students are more likely to choose advanced math-science courses if their older sibling was exogenously induced to do so. Their result is led by same-sex siblings (brothers in particular) and the most likely mechanism is found to be sibling competition. A second reason for gender to matter is that the earnings of men are higher on average, and hence brothers will represent a higher reference point than sisters. Notice that for women, this effect goes in the opposite direction as the saliency effect. Hence, the prediction of this mechanism is not so clear for women’s earnings. In contrast, both the level and the saliency of the reference point imply that we can expect brothers to increase men’s earnings.

One way in which siblings can compete is by choosing more highly paid occupations. To examine this possibility, we look at the general earnings level within each person’s occupation. More specifically, we calculate the natural logarithm of the population average income within their mode occupation. Column 2 of Table 6 shows that men with brothers indeed choose higher paying occupations than men with sisters, while there is no such effect for women. These results are consistent with the predictions of the competition mechanism (and inconsistent with the predictions of the differential parental treatment channel). Thus, sibling competition seems like a plausible explanation, although we have to keep in mind that there could also be other forces at play which we are unable to uncover.

Another way in which siblings can have a direct impact is via the use of each other’s job search networks. The sex of the sibling is relevant in this case because men and women tend to work in different industries. For example, construction and manufacturing of machinery are male dominated industries, whereas health and social work is mainly done by females. In line with the gendered patterns, we find that same-sex siblings are more likely to work in the same industry than opposite-sex siblings (untabulated results). This suggests that they can utilize each other’s networks better. The findings of Rao and Chatterjee (2018) also point to this. In particular, they analyze answers to a job search ques-

³¹ We get similar results when the outcome is a binary variable of gender stereotypical occupation (based on whether the share of women is below or above the median). There are also no effects on entrepreneurship (self-employment). The only specification in which we find a small but significant effect (only for women) is when we use a binary variable for STEM occupations: women are less likely to work in such occupations if they have a brother. This result confirms the finding of Brenøe (2018) that sibling sex-composition can help to understand the choice of STEM fields in particular. However, our data does not support the hypothesis that the sex of the sibling is relevant for choosing gendered occupations more generally.

tionnaire and find that men find a job via their brother about twice as often as via their sister, whereas the reverse holds for women.

To investigate the plausibility of the job search channel, we examine whether the sex of the sibling affects individuals' average unemployment rate (over ages 25–64 observed between 1986 and 2011). The results show that same-sex siblings decrease the incidence of unemployment (see column 3 of Table 6). The estimated effect is somewhat stronger for women. While we cannot determine whether these results are led by gendered job search networks, we can at least check how the coefficients change when we control for a dummy for working in the same industry as the sibling. As we can see in column 4, the coefficient of the same-sex dummy loses significance for men and decreases for women. All in all, the results seem to be consistent with the idea that same-sex siblings benefit from each other's job search networks more than opposite-sex siblings.

In the last four columns of Table 6, we conduct a “horse race” between earnings level within occupation and unemployment.³² Similarly to column 4, this analysis controls for endogenous variables so the results need to be taken with a grain of salt. Nonetheless, the patterns suggest that a large part of the income premium for men could come from working in a higher paid occupation. For women, the coefficient of same-sex does not respond to the inclusion of occupational income, as expected from the results of column 2. However, the coefficient drops to half of its size when unemployment is included among the controls, which could potentially be due to the job search channel.

So far we discussed mechanisms only in relation to labor market outcomes. However, differential parental treatment and social interactions between siblings can also influence family formation. With respect to differential parental treatment, Pollet et al. (2009) and Danielsbacka et al. (2011) find that parents support the family formation of daughters more than the family formation of sons. For example, they provide more informal childcare and are more likely to provide essentials, gifts and extras for the baby. This implies that there is a substitution effect from sisters: as compared to those with brothers, those with sisters get less support from their parents for their own family formation. We can expect differential parental treatment to matter more if the individual and the parents live in the same municipality (e.g. because babysitting requires spatial closeness). Thus, we make use of register data on locational information to descriptively explore this possibility. In particular, in Table 7 we interact the same-sex dummy with an indicator of whether the individual and his/her mother lives in the same municipality when the individual is thirty years old.³³ As expected, the interaction is positive and significant for men's fertility. While this is only suggestive evidence, it indicates that differential parental treatment could explain the findings on men's family formation.³⁴ For women, neither the main results, nor the results in Table 7 indicate the importance of this channel.

As far as direct sibling-to-sibling influences are concerned, recall the arguments that siblings represent reference points to each other and that same-sex siblings compete more than opposite-sex siblings. In terms of family formation sisters represent a “higher” reference point because women typically marry and have kids at an earlier age than men. Thus, the prediction for women is that sisters have a positive effect on their family formation. For men, the prediction is less clear because a brother

Table 7

Exploring differential parental treatment in family formation.

	Ever married	Any children	Number of children	Age at first birth
<i>Panel 0: Men</i>				
SS sibling	0.009*** (0.002)	0.004** (0.002)	0.012** (0.005)	−0.092*** (0.024)
Same muni.*SS	0.003 (0.002)	0.004* (0.002)	0.014** (0.007)	−0.054 (0.033)
Observations	499,427	499,427	499,427	399,676
Outcome mean	[0.68]	[0.80]	[1.80]	[28.85]
<i>Panel 1: Women</i>				
SS sibling	0.002 (0.002)	0.001 (0.001)	−0.001 (0.005)	−0.067*** (0.021)
Same muni.*SS	0.005* (0.002)	0.002 (0.002)	0.001 (0.007)	0.000 (0.030)
Observations	470,214	470,214	470,214	413,492
Outcome mean	[0.75]	[0.88]	[2.00]	[26.41]

Note: The sample consists of first-born singletons born 1938–1977 whose second-born sibling was born at most 6 years later. “SS sibling” is a dummy which takes 1 if the second-born sibling is of the same sex as the first-born and 0 if the second-born sibling is of the opposite sex. The regressions include a dummy for the first-borns who, at age 30, live in the same municipality as their mother (Same muni.). The sample includes those with non-missing information on municipality of residence for themselves and their mother. “Ever married” is a dummy which takes 1 if the individual has ever been married, 0 otherwise. All regressions include cohort fixed effects. Robust standard errors are shown in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

exposes them to a lower reference point but at the same time induces more competition.

To examine whether siblings react to each other's family formation events we again take an explorative approach. In particular, we look at how the individual's year of marriage relates to the sibling's year of marriage. This analysis is of course far from being causal. The aim is simply to examine whether the patterns in the data are consistent with the notion of reference points and competition.

Fig. 1 shows distributions of the difference between the year the first-born got married and the year the second-born got married. The gray connected dots show the actual number of sibling pairs. Observations to the left of zero are those sibling pairs where the first-born got married before the second-born while observations to the right represent the more atypical cases where the reverse is true. Although this is a selective sample, the irregularities at zero clearly stand out. It seems that siblings have a tendency to marry the same year. This pattern is consistent with the argument that individuals do not want to lag behind their siblings in getting married. However, there are many other possible explanations as well; for example, information sharing and mimicking could also lead to such spikes.

To get a sense of the magnitude of the excess mass at zero we use an estimation strategy in the spirit of Chetty et al. (2011). Specifically, we treat the counts as the dependent variable and the differences as the independent variable and fit a polynomial to the counts, including a dummy for the irregularity at zero. The predicted values from this model, when we exclude the contribution of the dummy, constitute the counterfactual distributions shown in black in the figures. We use the difference between the actual and counterfactual count, relative to the counterfactual count, as a crude measurement of the excess mass at zero. For women with sisters, we find that the actual number of observations at zero is 42% higher than the counterfactual. For women with brothers, the excess mass is only 24%. For men, the excess mass is 26% for those with brothers and 36% for those with sisters. The fact that the excess mass is highest for women with sisters is consistent with the prediction that women react more strongly to the family formation events of their sisters. Recall that the predictions were unclear for men, and indeed we

³² Note that for women, the baseline coefficient for income in column 5 is slightly lower than in our main results. This is because the sample is reduced due to balancing over the variables considered in Table 6.

³³ Municipality of residence is observed between 1968 and 2011 so the number of missing observations is minimized at this age for our sample. See also Table A.6 for detailed variable descriptions.

³⁴ Note that living in the same municipality as the mother is endogenous. While the sex of the sibling has no significant effect on this variable for women, we find that men with brothers are more likely to live in the same municipality as their mother than men with sisters.

Fig. 1. Timing of marriage. *Note:* These figures plot the distributions of the following difference: calendar year the first-born (FB) got married - calendar year the second-born (SB) got married. The sample includes sibling pairs where the first-born is born between 1938 and 1977 and where both siblings got married between 1969 and 2012. The gray dots are the actual counts. The black lines are the counterfactual distributions estimated by a 15-degree polynomial, taking out the effect from a dummy specified for the difference zero. The first figure is for first-born women with sisters (96,236 obs.), the second figure is for first-born women with brothers (92,798 obs.), the third figure is for first-born men with brothers (99,453 obs.) and the fourth figure is for first-born men with sisters (98,847 obs.).

do not see that the excess mass would be higher for men with brothers than for men with sisters.

5.2. Mechanisms related to constraints

Those who grow up with a same-sex sibling may be exposed to different kinds of constraints compared to those who grow up with an opposite-sex sibling. In this section, we focus on two potential mechanisms that are related to constraints: economies of scale and family size.

Rosenzweig and Wolpin (2000) argue that families economically benefit from having same-sex children because of hand-me-down savings, for example by reduction in child-expenditures on goods such as clothing. In other words, there might be economies of scale from having same-sex children. Therefore, a potential explanation of why same-sex siblings have better labor market outcomes is that they grew up in relatively better-off families than opposite-sex siblings.³⁵ The cost reductions implied by the economies of scale mechanism is expected to be particularly important for households of low socio-economic status. Thus, we explore this mechanism by examining heterogeneous effects with regard to parental education. Table 8 shows results for our main outcomes when we include interactions between same-sex and parental

education categories, where having parents with low education is the reference category. Overall, the patterns do not seem to support the economies of scale explanation. The estimated effects of having a same-sex sibling do not seem to systematically differ by the socio-economic status of the parents.

The size of the family in which siblings grow up is also related to constraints. Theoretical models suggest that parents face a “quantity-quality trade-off” when deciding to have additional children. Because of both financial constraints (Becker and Lewis, 1973) and time constraints (Lundholm and Ohlsson, 2002), investment per child decreases with the number of children. If same-sex siblings grow up in larger families than opposite-sex siblings, the prediction from theory would be that they have worse labor market outcomes. On a more speculative note, family size could also affect the family formation of the children. On the one hand, they might be inclined towards having more children as they grew up in a bigger family themselves. On the other hand, they might see the constraints imposed by larger families better and decide to have smaller families themselves.

The family size channel is relevant if two conditions are met: 1) the sex composition of siblings affects family size and 2) family size affects the outcomes of children. We can empirically examine both conditions. The first condition is fulfilled for our sample of singletons: having a second-born sibling of the same sex significantly increases the number of siblings, by about 0.08 (untabulated result). This is in line with findings from previous literature (e.g. the estimate is very similar to that of Angrist and Evans, 1998). For twins, we do not find a significant impact

³⁵ The effect might be more pronounced for men as recent research shows that they benefit more from additional resources such as higher school quality (Autor et al., 2016).

Table 8
Exploring the economies of scale mechanism.

	Ln(permanent income)	Years of schooling	Ever married	Any children	Number of children	Age at first birth
<i>Panel 0: Men</i>						
SS sibling	0.007*** (0.002)	0.015 (0.010)	0.009*** (0.002)	0.006*** (0.002)	0.022*** (0.005)	−0.123*** (0.026)
Mixed par. educ.*SS	0.001 (0.004)	−0.000 (0.016)	0.003 (0.003)	0.001 (0.003)	−0.002 (0.008)	0.050 (0.040)
High par. educ.*SS	−0.001 (0.004)	−0.018 (0.018)	−0.004 (0.003)	−0.004 (0.003)	−0.011 (0.009)	−0.038 (0.042)
Missing par. educ.*SS	0.010* (0.005)	−0.016 (0.027)	−0.000 (0.004)	−0.001 (0.004)	−0.006 (0.012)	−0.016 (0.055)
Observations	528,606	528,606	528,606	528,606	528,606	419,902
Outcome mean	[12.46]	[11.93]	[0.70]	[0.80]	[1.78]	[28.39]
<i>Panel 1: Women</i>						
SS sibling	0.007*** (0.002)	0.002 (0.010)	0.003 (0.002)	−0.000 (0.001)	−0.004 (0.005)	−0.093*** (0.023)
Mixed par. educ.*SS	−0.004 (0.004)	−0.012 (0.016)	0.002 (0.003)	0.004 (0.002)	0.005 (0.008)	0.066* (0.037)
High par. educ.*SS	−0.001 (0.005)	0.005 (0.017)	0.003 (0.003)	0.002 (0.003)	0.006 (0.008)	0.043 (0.038)
Missing par. educ.*SS	−0.001 (0.006)	−0.006 (0.026)	0.004 (0.003)	0.007** (0.003)	0.021* (0.011)	−0.047 (0.049)
Observations	498,741	498,741	498,741	498,741	498,741	435,395
Outcome mean	[12.09]	[12.27]	[0.76]	[0.88]	[1.91]	[26.23]

Note: The sample consists of first-born singletons born 1938–1977 whose second-born sibling was born at most 6 years later. “SS sibling” is a dummy which takes 1 if the second-born sibling is of the same sex as the first-born and 0 if the second-born sibling is of the opposite sex. “Ever married” is a dummy which takes 1 if the individual has ever been married, 0 otherwise. Permanent income is average income in years 1971–2011 (if subject is aged 25–64 in the given year). All regressions control for parental education dummies (Low par. edu. is the reference category) and cohort fixed effects. Robust standard errors are shown in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 9
Twins at second parity as instrument for number of siblings.

	Ln(permanent income)	Years of schooling	Ever married	Any children	Number of children	Age at first birth
<i>Panel 0: Men</i>						
Nr. of siblings	0.016 (0.010)	0.054 (0.051)	0.008 (0.008)	0.003 (0.008)	−0.009 (0.023)	0.132 (0.115)
Observations	533,734	533,734	533,734	533,734	533,734	423,998
Outcome mean	[12.46]	[12.16]	[0.67]	[0.79]	[1.78]	[28.84]
<i>Panel 1: Women</i>						
Nr. of siblings	0.013 (0.012)	0.054 (0.048)	0.008 (0.008)	−0.015** (0.007)	−0.069*** (0.022)	0.190* (0.105)
Observations	503,540	503,540	503,540	503,540	503,540	439,537
Outcome mean	[12.11]	[12.57]	[0.75]	[0.87]	[1.98]	[26.41]

Note: This table shows estimated effects of family size on our main outcomes. The sample consists of first-born singletons born 1938–1977 whose second-born siblings, who are either singletons or twins, were born at most 6 years later. The specification is 2SLS, where number of siblings in the family is instrumented by an indicator for twins at second birth. Number of siblings include all full siblings in the first-born’s family. “Ever married” is a dummy which takes 1 if the individual has ever been married, 0 otherwise. Permanent income is average income in years 1971–2011 (if subject is aged 25–64 in the given year). All regressions control for cohort fixed effects. Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

of the sex of the co-twin on the number of siblings. This is also in line with our expectations: since twins already increase family size substantially, the sex composition of the pair is less likely to increase family size further. Recall from Section 4 that our main results are similar for male twins and singletons. This already gives an indication that family size does not drive the results for men.

A commonly used empirical strategy to examine the second condition is to use twin births as instruments for family size. Overall, these studies rarely find significant family size effects on child outcomes. For example, Black et al. (2005) find that education, earnings and teen births are all unaffected by family size in Norway. Åslund and Grön-

qvist (2010) reach similar conclusions about years of schooling and labor market outcomes in Sweden.

We can use the richness of the register data to employ the twin birth IV approach for the particular outcomes considered in our paper. To be consistent with our previous analysis, we focus on the impact of family size on first-born singletons in families with at least two births. Thus, we use twinning at second parity as an instrument for number of siblings. Table 9 shows the results. For men, we find no impact of family size on any of our outcomes. It is also unlikely that family size is driving our main results for women. If anything, the impact of family size for women imply that our estimated effects of having a sister are biased towards zero for family formation outcomes. Recall that singleton women

with sisters have on average 0.08 more siblings than those with brothers. Thus, the estimated coefficients need to be multiplied with this to calculate the bias, which will therefore become very small.

6. Conclusions

In this paper we examined the impact of a sibling's gender on labor market outcomes and family formation. We studied both singletons and twins by using two well-founded empirical strategies and two different data sets. With the help of additional data on singletons, we also tried to shed some light on underlying mechanisms by examining extended dimensions of both the labor market and family formation decisions.

We found that a same-sex sibling increases men's earnings, likelihood to form a family and number of children. These results are consistent between the singleton and twin samples. Same-sex siblings were found to increase the earnings and family formation of women as well, but these effects are smaller and appear only in the singleton sample. We showed that family size and economies of scale are unlikely to drive the observed effects. Instead, we found empirical patterns which are consistent with the importance of direct sibling-to-sibling influences and differential parental treatment.

For men, we found that a large part of the income premium could come from occupational choice: men with brothers tend to choose higher paying occupations. We argue that this may be interpreted as a sign of competition between brothers. When it comes to the positive impact of brothers on men's family formation, our findings suggest that this is more prominent for men who live close to their mother. This pattern is consistent with the idea that differences in parental treatment can be the underlying force. Specifically, we argue that the effect could arise because sisters decrease parental support for family formation, as previous research found that parents support the family formation of daughters rather than the family formation of sons (Pollet et al., 2009; Danielsbacka et al., 2011).

For women, our analysis of the singleton data revealed no substantial impact on occupational choice. Instead, we found suggestive evidence that women benefit from having sisters because of shared job search networks. If this is the main driver of the income result for singleton women, it is not a surprise that we do not find an income effect in the twin sample. The twins belong to an older generation of women with a weaker attachment to the labor market and thus more limited opportunities to help each other via job search networks. With respect to women's family formation, we argue that the small positive effect in the singleton sample might arise due to competition between sisters, since marriage timing patterns seem to be consistent with this mechanism.

Finally, we want to highlight two caveats. First, our findings on the mechanisms need to be interpreted with caution. Some of these results are based on variation that is not exogenous and therefore should not be seen as causal evidence. In addition, there might be other potential explanations that we could not explore with our data. Thus, there is scope for further research on the forces that gave rise to the main effects. Second, it is important to be aware of the specific setting from which our data originates. Although we used two different datasets in our analysis, both of them are tied to the Swedish context. Sweden is a country that has often been described in terms of high gender equality. Moreover, Sweden has a longstanding tradition of family friendly policies. This specific cultural and institutional context may be a determinant of the strength of the mechanisms at play.

Appendix A

A1. The selection bias problem in a formal framework

In Section 2 we provided a non-technical account of how selection bias can arise when one tries to estimate the effect of an older sibling's gender on younger siblings. To show this formally, we need to consider a

set of equations. We will analyze men and women separately so parameters can be different by gender. Nonetheless, for the ease of exposition we omit subscripts for own gender. We start with the equation for the latent variable Y_i^* :

$$Y_i^* = \alpha + \beta^o G_i^o + \gamma X_i + U_i \quad (2)$$

where Y_i^* is the outcome of individual i , G_i^o is the gender of the older sibling and X_i denotes observable exogenous covariates. U_i consists of other relevant variables with $E[U_i] = 0$. Parental preferences are denoted by P_i and they are included in U_i . They are relevant because parents with different preferences may raise their children differently. They are unobservable, so they have to be in U_i instead of X_i . Nonetheless, this does not lead to bias in the estimation of this equation because G_i^o and P_i are independent, due to the random assignment of gender.

The problem is that we cannot estimate the above equation because Y_i^* is a latent variable. Instead of observing Y_i^* , we observe

$$Y_i = \begin{cases} Y_i^* & \text{if } S_i = 1 \\ \text{missing} & \text{if } S_i = 0 \end{cases} \quad (3)$$

That is, we observe outcomes for a selected sample: only for those people who were actually born ($S_i = 1$). This selection depends on both parental preferences (P_i) and the gender of the older child:

$$S_i = f(P_i, G_i^o) \quad (4)$$

If we try to estimate the parameters using the selected sample, we need to consider the expected value of Y_i conditional on $S_i = 1$:

$$E[Y_i | S_i = 1, G_i^o, X_i] = \alpha + \beta^o G_i^o + \gamma X_i + E[U_i | S_i = 1, G_i^o, X_i] = \alpha + \beta^o G_i^o + \gamma X_i + E[U_i | f(P_i, G_i^o) = 1, G_i^o, X_i] \quad (5)$$

The last term makes the selection bias visible. We condition on both $S_i = 1$ and G_i^o . Since selection depends on both P_i and G_i^o , this implicitly defines P_i : only certain parents will proceed to a next child after they have a girl (boy). Recall that U_i includes P_i because parental preferences may affect how children are raised. This implies that $E[U_i | f(P_i, G_i^o) = 1, G_i^o = 1, X_i] \neq E[U_i | f(P_i, G_i^o) = 1, G_i^o = 0, X_i]$, so the estimates suffer from selection bias.

A2. What can researchers expect if they lack zygosity information?

As we explained in Section 2, failing to distinguish twins by zygosity threatens identification. This is because such estimates would pick up not only the effect of the co-twin's gender, but also the impact of having a genetically identical sibling and other potential differences between MZ and DZ twins. Given the complex nature of this zygosity bias, it would be difficult to predict a priori how it would distort estimates.³⁶ However, using our data we can give an indication on what researchers can expect when they do not have access to zygosity information.

Table A.1 reports estimates on all twins (including MZs) in SALT. In case of women, the results are very different from the unbiased estimates of Table 4. The coefficients are insignificant and close to zero for all outcomes, including education and age at first birth. For men, the results on income are similar to Table 4, while the estimates on family formation are somewhat smaller and the probability of having any children is not significant anymore.

One should be cautious when drawing conclusions from these results, as the distortion in a particular sample depends on multiple factors. For example, in our sample the share of DZ twins is almost three quarters. As MZ and DZ twins arise in different ways, their relative frequency can vary a lot from population to population or over time, which will affect the eventual magnitude of the distortion. To simulate this, we examined what happens when we weight observations by zygosity status.

³⁶ For example, how could we predict how a man or a woman would respond to having a genetically identical sibling?

We found that as the relative weight of DZ twins decreases, a positive impact emerges on men's education while the results on their family formation gradually lose significance. The result on men's marriage loses significance when the share of DZ twins decreases to 55% and the result on their number of children becomes insignificant when the DZ share decreases to 35%. This is illustrated by Table A.2, which shows estimation results from a weighted regression that simulates what would happen if only 35% of twins were dizygotic. As we can see, in such a case researchers would erroneously conclude that a co-twin's gender has no significant effect on any of the family formation variables.

In sum, zygosity bias distorts most of the results, but to different degrees and sometimes even in different directions. When researchers have no zygosity information, they should be most cautious when drawing conclusion about women's age at first birth and education, and men's probability of having children.

A3. Additional tables and figures

A4. Variable definitions for analyzing potential mechanisms

Table A1
Main results on all twins (including monozygotics).

	Ln(permanent income)	Years of schooling	Ever married	Any children	Number of children	Age at first birth
<i>Panel 0: Men</i>						
SS co-twin	0.014** (0.006)	0.056 (0.049)	0.010** (0.005)	0.009 (0.006)	0.040** (0.020)	−0.033 (0.090)
Observations	17,978	17,978	17,978	17,978	17,978	14,680
Outcome mean	[12.44]	[10.83]	[0.88]	[0.82]	[1.89]	[27.87]
<i>Panel 1: Women</i>						
SS co-twin	−0.004 (0.008)	−0.029 (0.043)	−0.003 (0.005)	−0.006 (0.005)	0.005 (0.018)	−0.068 (0.077)
Observations	20,099	20,099	20,099	20,099	20,099	17,453
Outcome mean	[11.93]	[10.82]	[0.89]	[0.87]	[1.98]	[25.02]

Note: The sample consists of all SALT twins (including monozygotics) born 1926–1958. “SS co-twin” is a dummy which takes 1 if the co-twin is of the same sex as the individual, and 0 if the co-twin is of the opposite sex. “Ever married” is a dummy which takes 1 if the individual has ever been married or cohabited and 0 otherwise. All regressions control for older sibling, missing sibling information and cohort fixed effects. Permanent income is average income in years 1971–2007 if subject is aged 25–64 in the given year). The standard errors shown in parentheses are robust and clustered by the family of origin. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A2
How would results look like if only 35% of twins were dizygotic?

	Ln(permanent income)	Years of schooling	Ever married	Any children	Number of children	Age at first birth
<i>Panel 0: Men</i>						
SS co-twin	0.016** (0.007)	0.119** (0.057)	0.008 (0.006)	0.006 (0.007)	0.036 (0.022)	−0.016 (0.101)
Observations	17,978	17,978	17,978	17,978	17,978	14,680
<i>Panel 1: Women</i>						
SS co-twin	0.006 (0.009)	0.053 (0.049)	−0.005 (0.005)	−0.008 (0.006)	0.000 (0.020)	0.046 (0.087)
Observations	20,099	20,099	20,099	20,099	20,099	17,453

Note: The sample consists of all SALT twins (including monozygotics) born 1926–1958. In all regressions, weights are assigned to observations based on zygosity status, to simulate a situation when only 35% of twins are dizygotic. “SS co-twin” is a dummy which takes 1 if the co-twin is of the same sex as the individual, and 0 if the co-twin is of the opposite sex. “Ever married” is a dummy which takes 1 if the individual has ever been married or cohabited and 0 otherwise. All regressions control for older sibling, missing sibling information and cohort fixed effects. Permanent income is average income in years 1971–2007 (if subject is aged 25–64 in the given year). The standard errors shown in parentheses are robust and clustered by the family of origin. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A3
Robustness check: Main results for first-born singletons.

	Ln(permanent income)	Years of schooling	Ever married	Any children	Number of children	Age at first birth
<i>Panel 0: Men</i>						
SS sibling	0.007*** (0.002)	0.008 (0.007)	0.009*** (0.001)	0.005*** (0.001)	0.018*** (0.003)	−0.119*** (0.016)
Observations	528,606	528,606	528,606	528,606	528,606	419,902
Outcome mean	[12.46]	[11.93]	[0.70]	[0.80]	[1.78]	[28.39]
<i>Panel 1: Women</i>						
SS sibling	0.005*** (0.002)	0.000 (0.006)	0.005*** (0.001)	0.002** (0.001)	0.001 (0.003)	−0.071*** (0.014)
Observations	498,741	498,741	498,741	498,741	498,741	435,395
Outcome mean	[12.09]	[12.27]	[0.76]	[0.88]	[1.91]	[26.23]

Note: The sample consists of first-born singletons born 1938–1977 whose second-born sibling was born at most 6 years later. “SS sibling” is a dummy which takes 1 if the second-born sibling is of the same sex as the first-born and 0 if the second-born sibling is of the opposite sex. “Ever married” is a dummy which takes 1 if the individual has ever been married, 0 otherwise. Permanent income is average income in years 1971–2011 (if subject is aged 25–64 in the given years). All regressions control for cohort fixed effects, age difference in months, dummy variables for the mother’s and father’s age at first birth and parental education. Robust standard errors are shown in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A4
Exploring selection bias: Results for second-born singletons.

	Ln(permanent income)	Years of schooling	Ever married	Any children	Number of children	Age at first birth
<i>Panel 0: Men</i>						
SS sibling	0.000 (0.002)	0.001 (0.007)	0.000 (0.001)	−0.004*** (0.001)	−0.010*** (0.003)	−0.088*** (0.017)
Observations	486,225	486,225	486,225	486,225	486,225	382,219
Outcome mean	[12.44]	[12.00]	[0.65]	[0.79]	[1.75]	[28.89]
<i>Panel 1: Women</i>						
SS sibling	0.005*** (0.002)	0.008 (0.007)	−0.005*** (0.001)	−0.001 (0.001)	−0.018*** (0.003)	0.015 (0.016)
Observations	458,206	458,206	458,206	458,206	458,206	396,531
Outcome mean	[12.10]	[12.43]	[0.71]	[0.87]	[1.92]	[26.55]

Note: The sample consists of second-born singletons born 1938–1977 whose first-born sibling was born at most 6 years earlier. “SS sibling” is a dummy which takes 1 if the first-born sibling is of the same sex as the second-born and 0 if the first-born sibling is of the opposite sex. “Ever married” is a dummy which takes 1 if the individual has ever been married, 0 otherwise. Permanent income is average income in years 1971–2011 (if subject is aged 25–64 in the given years). All regressions control for cohort fixed effects. Robust standard errors are shown in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Fig. A1. Results on income at different ages, for first-born singletons. *Note:* The figures show results from separate regressions by age where the dependent variable is the natural logarithm of a five-year moving average of income. For example, at age 37 we use the average income over ages 35–39. The black lines show the estimated coefficients of the same-sex sibling dummy and the dotted lines show 95% confidence intervals. All regressions include cohort fixed effects. The first figure shows results for first-born singleton men and the second figure shows results for first-born singleton women.

Table A5
Main results for dizygotic twins by family structure.

	Ln(permanent income)	Years of schooling	Ever married	Any children	Number of children	Age at first birth
<i>Panel 0: Men</i>						
SS co-twin	0.015 (0.013)	−0.018 (0.102)	0.009 (0.010)	0.016 (0.012)	0.052 (0.037)	0.068 (0.177)
Older sib.*SS	−0.004 (0.016)	0.021 (0.124)	0.003 (0.012)	−0.005 (0.015)	−0.010 (0.046)	−0.179 (0.219)
Missing sib.*SS	−0.011 (0.025)	−0.089 (0.186)	0.027 (0.024)	−0.003 (0.026)	−0.023 (0.094)	0.053 (0.408)
Observations	13,427	13,427	13,427	13,427	13,427	10,978
Outcome mean	[12.44]	[10.78]	[0.88]	[0.82]	[1.89]	[27.87]
<i>Panel 1: Women</i>						
SS co-twin	−0.014 (0.017)	−0.039 (0.090)	−0.006 (0.009)	−0.008 (0.010)	0.000 (0.035)	−0.178 (0.152)
Older sib.*SS	0.004 (0.021)	−0.160 (0.109)	0.010 (0.011)	0.006 (0.012)	0.025 (0.043)	−0.107 (0.187)
Missing sib.*SS	−0.055 (0.047)	0.223 (0.193)	0.003 (0.024)	−0.000 (0.023)	−0.035 (0.087)	0.571 (0.366)
Observations	14,742	14,742	14,742	14,742	14,742	12,836
Outcome mean	[11.92]	[10.75]	[0.90]	[0.87]	[1.99]	[24.94]

Note: The sample consists of dizygotic twins born 1926–1958. “SS co-twin” is a dummy which takes 1 if the co-twin is of the same sex as the individual, and 0 if the co-twin is of the opposite sex. “Ever married” is a dummy which takes 1 if the individual has ever been married or cohabited, 0 otherwise. Permanent income is average income in years 1971–2007 (if subject is aged 25–64 in the given year). All regressions control for older sibling, missing sibling information and cohort fixed effects. The standard errors shown in parentheses are robust and clustered by the family of origin. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Fig. A2. Results on the probability of being a mother by a certain age. *Note:* The figures show results from separate regressions by age where the dependent variable is an indicator for having any children. The black lines show estimated coefficients of the same-sex sibling/same-sex co-twin dummy and the dotted lines show 90% confidence intervals. The first figure shows results for first-born singleton women and the second figure shows results for twin women. All regressions include cohort fixed effects and the regressions on the twin sample also control for older sibling and missing sibling information.

Table A6
Variable definitions related to the analysis of the potential mechanisms.

Variable	Data and definition
Occupation	Occupation is observed in terms of 4-digit occupational codes SSYK (Swedish version of ISCO) between years 2001 and 2011. Each individual's occupation in the singleton sample is the mode of observed occupations between 2001 and 2011.
Occupational income	We calculate the average income within each occupation using the entire Swedish population aged 25–64 between 2001 and 2011, then each individual in the singleton sample is assigned the natural log of the average income within their mode occupation.
Share of women in occupation	We calculate the share of women (between 0 and 1) employed in each occupation using the entire Swedish population observed between 2001 and 2011, then each individual in the singleton sample is assigned the share of women employed in their mode occupation.
Same industry	Industry is observed in terms of 2-digit industry codes (SNI) between year 1990 and 2010, there are 59 categories observed during this period (where the three largest industries, in terms of number of employed, are health and social work, education and other business activities). Each individual's industry in the singleton sample is the mode of observed industry between 1990 and 2010. Same industry is an indicator variable which equals 1 if the first-born and second-born singleton have the same mode industry, 0 otherwise.
Average unemployment	Unemployment status is observed between years 1986 and 2011, average unemployment is the fraction of years with unemployment status between ages 25 and 64. We observe unemployment at different age ranges for different cohorts, however we always include birth year fixed effects in all regressions to account for such differences.
Same municipality	Municipality of residence is observed between 1968 and 2011. We define an indicator variable which equals 1 if the individual and the mother lives in the same municipality when the individual is thirty years old, 0 otherwise.
Parental education	For each mother and father we determine whether their education is above the median years of schooling within the group of mothers/fathers who had their first child in the same year. Then we define four parental education categories: High = Both parents above median, Mixed = One parent above and the other parent equal/below median, Low = Both parents equal/below median, Missing = One or both parents' education are unobserved.

References

- Abeliovich, D., Leiberman, J.R., Teuerstein, I., Levy, J., 1984. Prenatal sex diagnosis: testosterone and FSH levels in mid-trimester amniotic fluids. *Prenatal Diagnosis* 4, 347–353.
- Adams, B.N., 1999. Cross-cultural and U.S. kinship. In: Sussman, M.B., Steinmetz, S.K., Peterson, G.W. (Eds.), *Handbook of Marriage and the Family*. Springer US, Boston, MA, pp. 77–91.
- Ahrenfeldt, L., Petersen, I., Johnson, W., Christensen, K., 2015. Academic performance of opposite-sex and same-sex twins in adolescence: a Danish national cohort study. *Hormones Behav.* 69, 123–131.
- Altonji, J.G., Cattani, S., Ware, I., 2017. Identifying sibling influence on teenage substance use. *J. Hum. Resour.* 52 (1), 1–47.
- Andersson, G., Hank, K., Rønsen, M., Vikat, A., 2006. Gendering family composition: sex preferences for children and childbearing behavior in the Nordic countries. *Demography* 43 (2), 255–267.
- Angrist, J.D., Evans, W.N., 1998. Children and their parents' labor supply: evidence from exogenous variation in family size. *Am. Econ. Rev.* 88, 450–477.
- Angrist, J.D., Pischke, J.-S., 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, NJ.
- Åslund, O., Grönqvist, H., 2010. Family size and child outcomes: is there really no trade-off? *Labour Econ.* 17, 130–139.
- Autor, D., Figlio, D., Karbownik, K., Roth, J., Wasserman, M., 2016. School quality and the gender gap in educational achievement. *Am. Econ. Rev.* 106 (5), 289–295.
- Becker, G.S., Lewis, H.G., 1973. On the interaction between the quantity and quality of children. *J. Polit. Econ.* 81 (2, Part 2), S279–S288.
- van de Beek, C., Thijssen, J.H., Cohen-Kettenis, P.T., van Goozen, S.H., Buitelaar, J.K., 2004. Relationships between sex hormones assessed in amniotic fluid, and maternal and umbilical cord serum: what is the best source of information to investigate the effects of fetal hormonal exposure? *Hormones Behav.* 46, 663–669.
- Benirschke, K., Burton, G.J., Baergen, R.N., 2012. *Multiple pregnancies. Pathology of the Human Placenta*. Springer.
- Bingley, P., Lundborg, P., Lyk-Jensen, S., 2017. Brothers in Arms: Spillovers from a Draft Lottery. Technical Report. IZA. Discussion Paper 10483.
- Björklund, A., Jäntti, M., Lindquist, M.J., 2009. Family background and income during the rise of the welfare state: brother correlations in income for Swedish men born 1932–1968. *J. Public Econ.* 93 (5), 671–680.
- Black, S.E., Breining, S., Figlio, D.N., Guryan, J., Karbownik, K., Nielsen, H.S., Roth, J., Simonsen, M., 2017. Sibling Spillovers. Working Paper 23062. National Bureau of Economic Research.
- 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. *Q. J. Econ.* 120 (2), 669–700.
- Booth, A.L., Kee, H.J., 2009. Birth order matters: the effect of family size and birth order on educational attainment. *J. Popul. Econ.* 22, 367–397.
- Breining, S., Doyle, J., 2015. Birth Order and Parental Investment for Boys vs. Girls: evidence from the Genetic Lottery. PhD thesis. Aarhus University. Chapter 1.
- Breining, S.N., 2014. The presence of ADHD: spillovers between siblings. *Econ. Lett.* 124 (3), 469–473.
- Breining, S.N., Joseph J. Doyle, J., Figlio, D.N., Karbownik, K., Roth, J., 2017. Birth Order and Delinquency: Evidence from Denmark and Florida. Working Paper 23038. National Bureau of Economic Research.
- Brenøe, A.A., 2018. Origins of Gender Norms: Sibling Gender Composition and Women's Choice of Occupation and Partner. Working Paper.
- Butcher, K.F., Case, A., 1994. The effect of sibling sex composition on women's education and earnings. *Q. J. Econ.* 109 (3), 531–563.
- Chen, S.H., Chen, Y.-C., Liu, J.-T., 2017. The impact of family composition on educational achievement. *J. Hum. Resour. forthcoming*. doi:10.3368/jhr.54.1.0915.7401R1.
- Chetty, R., Friedman, J.N., Olsen, T., Pistaferri, L., 2011. Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: evidence from Danish tax records. *Q. J. Econ.* 126 (2), 749–804.
- Cohen-Bendahan, C.C., van de Beek, C., Berenbaum, S.A., 2005. Prenatal sex hormone effects on child and adult sex-typed behavior: methods and findings. *Neurosci. Biobehav. Rev.* 29, 353–384.
- Cohen-Bendahan, C.C., van Goozen, S.H.M., Buitelaar, J.K., Cohen-Kettenis, P.T., 2005. Maternal serum steroid levels are unrelated to fetal sex: a study in twin pregnancies. *Twin Res. Hum. Genet.* 8 (2), 173–177.
- Conley, D., 2000. Sibship sex composition: effects on educational attainment. *Social Sci. Res.* 29 (3), 441–457.
- Cools, A., Patacchini, E., 2018. The Brother Earnings Penalty. Working Paper.
- Dahl, G.B., Løken, K.V., Mogstad, M., 2014. Peer effects in program participation. *Am. Econ. Rev.* 104, 2049–2074.
- Danielsbacka, M., Tanskanen, A.O., Jokela, M., Rotkirch, A., 2011. Grandparental child care in Europe: evidence for preferential investment in more certain kin. *Evol. Psychol.* 9 (1), 3–24.
- Elston, R.C., Olson, J.M., Palmer, L., 2002. *Biostatistical Genetics and Genetic Epidemiology*. John Wiley and Sons.
- Even, M.D., Dhar, M.G., vom Saal, F.S., 1992. Transport of steroids between fetuses via amniotic fluid in relation to the intrauterine position phenomenon in rats. *J. Reprod. Fert.* 96, 709–716.
- Fellman, J., Eriksson, A.W., 2006. Weinberg's differential rule reconsidered. *Hum. Biol.* 78, 253–275.
- Gielen, A.C., Holmes, J., Myers, C., 2016. Prenatal testosterone and the earnings of men and women. *J. Hum. Resour.* 51 (1), 30–61.
- Glass, A., Klein, T., 1981. Changes in maternal serum total and free androgen levels in early pregnancy: lack of correlation with fetal sex. *Am. J. Obstetrics Gynecol.* 140, 656–660.
- Gray, R.H., Simpson, J.L., Bitto, A.C., Queenan, J.T., Li, C., Kambic, R.T., Perez, A., Mena, P., Barbato, M., Stevenson, W., Jennings, V., 1998. Sex ratio associated with timing of insemination and length of the follicular phase in planned and unplanned pregnancies during the use of natural family planning. *Hum. Reprod.* 13 (5), 1397–1400.
- de Haan, M., 2010. Birth order, family size and educational attainment. *Econ. Educ. Rev.* 29 (4), 576–588.
- Hank, K., Kohler, H.-P., 2000. Gender preferences for children in Europe: empirical results from 17 FFS countries. *Demograph. Res.* 2.
- Hansen, D., Möller, H., Olsen, J., 1999. Severe periconceptional life events and the sex ratio in offspring: follow up study based on five national registers. *Br. Med. J.* 319 (7209), 548–549. doi:10.1136/bmj.319.7209.548.
- Hauser, R.M., Kuo, H.-H.D., 1998. Does the gender composition of sibships affect women's educational attainment? *J. Hum. Resour.* 33 (3), 644–657.
- Heckman, J.J., 1979. Sample selection bias as a specification error. *Econometrica* 47, 153–161.
- Husby, H., Holm, N., Gernow, A., Thomsen, S., K. K.K., Gürtler, H., 1991. Zygosity, placental membranes and Weinberg's rule in a Danish consecutive twin series. *Acta Geneticae Medicae et Gemellologiae* 40, 147–152.
- James, W.H., 1979. Is Weinberg's differential rule valid? *Acta Geneticae Medicae et Gemellologiae* 28 (1), 69–71.
- Jayachandran, S., Pande, R., 2017. Why are Indian children so short? The role of birth order and son preference. *Am. Econ. Rev.* 107 (9), 2600–2629.
- Joensen, J.S., Nielsen, H.S., 2018. Spillovers in education choice. *J. Public Econ.* 157, 158–183.
- Kaestner, R., 1997. Are brothers really better? Sibling sex composition and educational achievement revisited. *J. Hum. Resour.* 32 (2), 250–284.
- Keim, S., Klärner, A., Bernardi, L., 2009. Who is Relevant? Exploring Fertility Relevant Social Networks. MPIDR Working paper WP 2009-001.
- Kuegler, A., 2009. A Curse of Comparison? Evidence on Reference Groups for Relative Income Concerns. Policy Research Working Paper 4820.
- Lichtenstein, P., Faire, U.D., Floderus, B., Svartengren, M., Svedberg, P., Pedersen, N.L., 2002. The Swedish twin registry: a unique resource for clinical, epidemiological and genetic studies. *J. Internal Med.* 252, 184–205.
- Lundholm, M., Ohlsson, H., 2002. Who takes care of the children? The quantity-quality model revisited. *J. Popul. Econ.* 15 (3), 455–461.
- McHale, S.M., Crouter, A.C., Whiteman, S.D., 2003. The family contexts of gender development in childhood and adolescence. *Social Dev.* 12 (1), 125–148.
- Meulenberg, P.M.M., Hofman, J.A., 1991. Maternal testosterone and fetal sex. *J. Steroid Biochem. Mol. Biol.* 39 (1), 51–54.
- Miller, E.M., 1994. Prenatal sex hormone transfer: a reason to study opposite-sex twins. *Personality Individual Differences* 17, 511–529.
- Mizuno, R., 2000. The male/female ratio of fetal deaths and births in Japan. *Lancet* 356 (9231), 738–739.
- Morduch, J., 2000. Sibling rivalry in Africa. *Am. Econ. Rev.* 90 (2), 405–409.
- Nicoletti, C., Rabe, B., 2014. Sibling Spillover Effects in School Achievement. IZA Discussion Paper. IZA DP No. 8615.
- Oláh, L. Sz., Bernhardt, E.M., 2008. Sweden: combining childbearing and gender equality. *Demograph. Res.* 19, 1105–1144.
- Parish, W.L., Willis, R.J., 1993. Daughters, education, and family budgets: Taiwan experiences. *J. Hum. Resour.* 28 (4), 863–898.
- Petersson-Lidbom, P., Thoursie, P.S., Vlachos, J., 2008. Does child gender affect sibling and family outcomes? Draft paper.
- Phillips, D.I.W., 1993. Twin studies in medical research: can they tell us whether diseases are genetically determined? *Lancet* 341, 1008–1009.
- Pollet, T.V., Nelissen, M., Nettle, D., 2009. Lineage based differences in grandparental investment: evidence from a large British cohort study. *J. Biosocial Sci.* 41, 355–379.
- Rao, N., Chatterjee, T., 2018. Sibling gender and wage differences. *Appl. Econ.* 50 (15), 1725–1745.
- Reece, J.B., Urry, L.A., Cain, M.L., Wasserman, S.A., Minorsky, P.V., Jackson, R.B., 2010. *Biology - Campbell, ninth ed.* Benjamin Cummings.
- Rodeck, C.H., Gill, D., Rosenberg, D.A., Collins, W.P., 1985. Testosterone levels in midtrimester maternal and fetal plasma and amniotic fluid. *Prenatal Diagnosis* 5, 175–181.
- Rosenzweig, M.R., Wolpin, K.I., 2000. Natural "natural experiments" in economics. *J. Econ. Lit.* 38 (4), 827–874.
- Ryan, B., Vandenberg, J., 2002. Intrauterine position effects. *Neurosci. Biobehav. Rev.* 26, 665–678.
- vom Saal, F.S., Bronson, F.H., 1980. Sexual characteristics of adult female mice are correlated with their blood testosterone levels during prenatal development. *Science* 208, 597–599.
- Singer, E., van Hoewyk, J., Maher, M.P., 2000. Experiments with incentives in telephone surveys. *Public Opin. Q.* 64, 171–188.
- Swedish Institute, 1994. *Family Planning in Sweden. Fact Sheets on Sweden*. Technical Report. ISSN-1101-6124.
- Tapp, A.L., Maybery, M.T., Whitehouse, A.J., 2011. Evaluating the twin testosterone transfer hypothesis: a review of the empirical evidence. *Hormones Behav.* 60, 713–722.
- Toriola, A., Väärasmäki, M., Lehtinen, M., A. A.Z.-J., Lundin, E., Rodgers, K., Lakso, H., Chen, T., Schock, H., Hallmans, G., Pukkala, E., Toniolo, P., Grankvist, K., Surcel, H., Lukanova, A., 2011. Determinants of maternal sex steroids during the first half of pregnancy. *Obstetrics Gynecol.* 118 (5).
- Vlietinck, R., C. D., R. D., den Bergh, H. V., M. T., 1988. The validity of Weinberg's rule in the East Flanders prospective twin survey (EFPTS). *Acta Geneticae Medicae et Gemellologiae* 37, 137–141.
- Wilcox, A.J., Weinberg, C.R., Baird, D.D., 1995. Timing of sexual intercourse in relation to ovulation. *N. Engl. J. Med.* 333 (23), 1517–1521.