LUND UNIVERSITY
School of Economics and Management

Master's Programme in Economics

Suffering From Attention

A Study Using Twins to Isolate a Favourite Child Effect

by

Gustav Lundström

Department of Economics

NEKN01

Master Essay I (15 credits)

June 2022

Supervisor: Ana Rodriguez-Gonzalez

Abstract

I examine a favourite child's effect on educational attainment and mental and physical health and well-being. In the favourite child effect, based on parental preferences in the sex composition of children using a proxy for parental investment, I can assess how parents distinguish inputs based on the sex of their children. Identification utilising data from three longitudinal studies in the United Kingdom and limiting the data to a maximum of three siblings per sibship. I investigate the random assignment of sex in the presence of a twin pair, which allows for causal inference. I am able to analyse the causal effect of being the lone sex child on quality outcomes since parents have no control over the sexes and twin status of their children. The sibling of the favourite child is more likely to develop depression during adolescence than the sibling of a same-sex sibship. I find no indication of a difference between sibships of two sexes and single-sexed sibships in terms of educational attainment and physical health.

Keywords: Sibling sex composition, parental investment, schooling, mental and physical health

Contents

List of Tables

1 Introduction.

It is often said that you do not get to choose your family. This lottery is determined by nature and, by definition, impacts everyone. For most people, the closest ties formed during their lifetime are those with their immediate family. Therefore, the impact these individuals have on the person is entirely deterministic and has long-lasting and sincere effects. Characterising the family as one cohesive unit would be a disservice to how human behaviour and social bonds work and form. Therefore we should consider the family as a network of different relationships interconnected by biological and social bonds. As a result, it is plausible to believe that some ties are stronger than others. The phrase 'Apple of one's Eye' is used to describe someone highly cherished by someone, perhaps by their parent. Are there any truths to this phrase in that there are differences in development for children whom their parents favour on the basis of specific parental sex composition preferences?

There is a long tradition in social science of analysing the effect of parental fertility choices on the family. Becker & Lewis (1973) and later Becker & Tomes (1976) pioneered this topic in economics by developing the quantity/quality trade-off model that formalises how household fertility responds to child quality measures. Becker's theoretical framework supports the idea that the number of resources spent by the parents and the number of children within each family determines a child's quality. An increase in quantity, i.e. an additional child, will lead to a rise in the marginal cost of the quality of children. The literature on parental resources and how children are affected by fertility decisions is not, however, only limited to the economic field, ranging from the 'resource dilution' model[1] of psychology (Zajonc & Markus, 1975; Zajonc, Markus, & Markus, 1979; Zajonc & Mullally, 1997) and the 'confluence' model of sociology[2] (Eijck & De Graaf, 1995; Downey, 1995; Powell & Steelman, 1990). The medical literature also contributes to the relationship between quality measures such as health and amount of siblings within families (Richiardi et al., 2004; Altieri & Hemminki, 2007). In principle, these models investigate the same phenomenon in that parental resources are scarce and

[1]The resource dilution model states that parental resources are limited. As the number of children in a household grows, the resources acquired by each kid must inevitably decrease. Since siblings compete with parents for their time, energy, and financial resources, thus the fewer there are, the better.

[2]The confluence model states that birth order and family size affect educational attainment. Within the framework, intellectual growth within the home setting depends on the cumulative impacts of the intellectual environment, which consists mainly of the intelligence of siblings and parents.

that an increase in quantity dilutes the quality of the children. Hence, the interest of this topic is shared interdisciplinary, but the scientific approach makes the topic diverse.

Expanding on Becker's framework, succeeding economic empirical papers focused on testing Becker's hypothesis by applying different data sets and quality measures. The research focused on estimating measurements for quality by analysing the negative relationship between fertility and labour outcomes for primarily mothers and children (Rosenzweig & Schultz, 1987). Also, Dahl & Moretti (2008) finds in their research that the gender of the first child affects future parental fertility decisions using U.S census data. The authors' evidence points toward U.S parental preferences favouring boys by concluding that the number of children in families with a firstborn girl is significantly higher. This discrepancy between the division of parental resources among sexes highlights that within the quantity/quality trade-off model, both parental fertility choices and parental investment choices exist.

In terms of the underlying mechanisms behind purely correlational findings, both Browning (1992) and Angrist & Evans (1996) argued that research based on analysing the effects of childbearing without controlling for household endogeneity leads to biased estimates. To avoid this bias (Rosenzweig & Wolpin, 1980; Rosenzweig & Schultz, 1987) used multiple births from both census data from India and survey data from Malaysia to create a natural experiment and achieve causal inference. The basic idea behind using twins in the quantity/quality model is that the identification strategy enables the researcher to control for endogenous variation in family size (Angrist, Lavy, & Schlosser, 2010). Some researchers have used instrumental variables proposed by (Angrist, Imbens, & Rubin, 1996) in tandem with 1980 and 1990 census Public Use Micro Data Samples data to exploit preferences in the sex composition of children in the U.S (Angrist & Evans, 1996). Research has also made use of the instrumental variable approach and longitudinal studies such as the National Child Development Study from the U.K (Iacovou, 2001) and the Korean Household Panel Study from South Korea (Lee, 2008). Using the sex composition of firstborn twins and subsequent fertility decisions Angrist & Evans (1996) can capture the effect of the older child having an extra younger child. What the authors found was no evidence of a quantity/quality trade-off. However, the marginal estimates are limited to what order of parity, i.e. sibship size, the study sets out to investigate. Lee (2008) finds that sibling size has adverse effects on per-child investment in education

by extending the instrumental variable approach to include a proxy of parents' monetary investment in children's education as a quality measure for overall parental investment. The approach Lee designs makes marginal estimates of the second child which is enabled by the use of detailed study data.

Parallel to effects coming from the parental channel, there is also a sizeable literature on the impact of the sibling channel. Since siblings are direct reference points to each other, effects may vary depending on the gender composition of the sibship. Some empirical papers disentangle the hidden mechanisms behind sex composition by estimating the relationship between siblings' sex differences and education attainment (Butcher & Case, 1994). Convincing evidence comes from a more recent study by Peter et al. (2015) where the authors employ a twin strategy; they find different effects of sibling sex composition for men and women. Using the extensive Swedish Twin Registry (STR), the authors can compare men and women separately. The resulting estimates show that both men and women benefit from same-sex competition in terms of labour outcomes and family formation.

By combining the effects of parental and sibling channels, Brenøe (2021) analysed how the family environment affects women's conformity to gender roles. Using Danish registry data and an identification strategy centred on firstborn women with either second-born brothers or sisters, Brenøe finds that women with brothers conform more to traditional gender roles measured in labour and marital outcomes. She also introduces an applicable mechanism for how parental investment interacts with different family sex compositions.

This thesis contributes to the field by exploring the relationship between parental investment and quality metrics within the quantity/quality trade-off model. I investigate a causal *favourite child* effect resulting from the child's role as a certain sex composition facilitator. In households with an only-sexed sibling, the unique-sexed kid facilitates either parental preference for one gender over another or parental preference for a mix of sexes over a complete set of same-sexed siblings. The preferred combination of sexes within the sibship is totally dependent on the unique-sexed sibling; consequently, parental involvement should vary. I also account for exogenous variation in family size and concentrate on families with three children and twins regardless of their zygotic status. My primary identification technique removes selection biases in parental fertility decisions since parents cannot choose the sex and twin status of their offspring based on maternal and

paternal ages. In addition, I investigate the trade-offs proposed by the quantity/quality trade-off model using the child quality indicators: educational achievement, psychological and physiological health and well-being.

I hypothesise that since parental resources are scarce and parental preferences determine investment in children, the outcome of the child who gets the most attention from parents will be of the highest quality. Previous research has used the parental-child and the intra-sibling channel to determine how children are affected by differences in sex compositions which enables me to analyse the effect of being a favourite child.

I use data from the 1958 National Child Development Study, NCDS, the 1970 British Cohort Study, BCS and the 2001 Millennium Cohort study, MCS, provided by the Centre for Longitudinal Studies at the Institute of Education, University of London, containing information on twins and sibling composition among cohort members. By developing a strategy that centres on twins, I guarantee that there is a twin in the sibship, which leads to a relatively small sample. Treatment assignment takes one of two forms. In the first analysis, I estimate a direct effect of being the favourite child. Treatment is assigned if a cohort member is the only-sex to the remainder of the sibship, e.g. a family of one boy and two girls. In the second analysis, I estimate the indirect effect of the favourite child. Treatment in the indirect analysis is assigned if the cohort member is of the same sex as one sibling leading to one remaining only-sex sibling, e.g. the cohort member is a boy in a family consisting of two boys and one girl. The control group in both analyses are defined as the group where the whole sibship is of the same sex, e.g. the cohort member is a boy within a family of three boys. I postulate that the two treatment assignments will produce the same effect but in polar opposite directions when the control group is the reference group.

My analysis yields some interesting results. Firstly, the favourite child effect has a significant adverse effect on teenagers' mental health in siblings of only-sex children in three child sibships, measured by the prevalence of depression by the medical examiner. Compared to other studies on how parental preferences in child sex composition affect child development, this study expands on quality metrics by including analyses on educational attainment and psychological and physiological health and well-being. Secondly, no favourite child effect is found for the analyses of educational attainment and physiological health.

Section 2 describes my identification strategy and model selection of interest, while Section 3 outlines and discusses the data and its statistical features. In Section 4, I present my results and potential explanations. I highlight potential mechanisms and sensitivity checks in Section 5 and conclude the thesis in Section 6.

2 Identification Strategy

In this study, I attempt to estimate the causal effect of the *favourite child* based on sibling sex composition in terms of educational attainment and psychological and physiological health and well-being as child quality measures. However, I cannot achieve causal inference by comparing children of different sexes in households with various sibling sex compositions and family sizes. Such comparison leads to biased estimates since parental preferences may influence the decision to have another child and the investment each child receives. According to the quantity/quality trade-off model, this additional child will dilute the parental resources for the whole sibship and, depending on the sex, alter the parental investment distribution among the siblings.

To circumvent the selection bias, I restrict the sample to a sibship size of three siblings and examine the random assignment of sex where a twin pair is present. Since the parents cannot decide the sex and the twin status of the children, I can estimate the causal effects of being the only-sex on quality outcomes. This strategy utilises twins since they are born simultaneously, preventing parents from making decisions based on the sex of either one of the twins. The strategy achieves causal inference using this random sex and twin status assignment, conditional on maternal and paternal age. Therefore, the key identifying assumption is that conditional on third parity sibship size, maternal and paternal age, and families have a twin pair at the first or second birth, the sex of the child is random.

My research design requires limiting the analysis to households with three children since this gives me a majority of a specific sex in the sibship. Although three-child households are not the norm in most countries in the West, it is worthwhile to address the challenging internal validity difficulties inherent to estimating family sex-composition preferences, even if doing so compromises external validity.

The first model specification of the empirical analysis for the direct *favourite child*

effect is:

$$Y_i = \alpha_1 + \rho_1 D_i + X_i'\gamma_1 + \eta_i, \tag{2.1}$$

and the second model specification of the analysis for the indirect *favourite child* effect is:

$$Y_i = \alpha_2 + \rho_2 D_i^* + X_i'\gamma_2 + \eta_i^*. \tag{2.2}$$

In these specifications, Y_i measures quality outcomes of cohort member i, who is a singleton sibling constrained in third parity sibship size in *General Sample* or is a sibling constrained in third parity sibship size where a twin pair is present in *Twin Sample*[3]. The regressor, D_i denotes the treatment status of cohort member i, equals 1 if cohort member is only-sex child in sibship of three and equals 0 if cohort member is of the same sex as siblings. D_i^* is the treatment indicator of cohort members i, equals 1 if cohort member is of the same sex as one of its siblings while the other sibling is of the only-sex and equals 0 if all siblings are of the same sex. ρ_1 and ρ_2 are the estimate of interest in each analysis as it measures effects coming from differences in parental investment between siblings depending on children sexes. The principal controls, X_i, is a vector of fixed control variables for: brith order, twin status, gender, year of birth fixed effects, mother and father's age at cohort member's birth and mother and father's ethnicity. γ_1 and γ_2 are vectors of the control variables associated coefficients. The constants, α_1 and α_2, the regression error terms, η_i and η_i^*, for cohort member i makes out the remainder of the linear regression model.

In the analysis of *favourite child* effects, I will perform a set of regressions to test how the estimates change when controlling for observable determinants. The first and most rudimentary specification includes a model where measures of quality outcomes, Y_i, are regressed on treatment indicator, D_i or D_i^*, depending on whether the estimates are measuring the direct or the indirect effects of *favourite children*. The second specification includes the principal control variables, X_i, separately for the direct and indirect models of interest. The third specification introduces a control for family size endogeneity, in tandem with the principal control variables, by only including families where two siblings are twins, making D_i or D_i^* estimate the causal direct and indirect effect of *favourite children*. Finally, the fourth specification includes an interaction term between gender

[3]For a schematic overview of how I define treatment and control in sibship I refer to Table A1 in the Appendix

and treatment indicator to account for differences in the direct and indirect effect of *favourite children* between boys and girls.

2.1 Inverse Probability Weighting Model

This study also provides an extensive sensitivity, or meta-analysis, of the estimates of the models in Equation 2.1 and Equation 2.2. In this robustness test, I use the computed sample weights provided by the MCS to handle conditional independence and selection bias problems. To give the reader sufficient knowledge about how the sample weights are calculated and used within this analysis, I will follow Verbeek's (2017, p. 276-278) specification of the inverse probability weighting model developed by Rosenbaum & Rubin (1983). The first step to construct the weights is to run a Logit model where the binary outcome, A_i (1 for response and 0 otherwise) are regressed on observed characteristics among cohort members, x_i. The observed characteristics only include measures from previous sweeps to circumvent bad controls[4]. Between two sweeps, denoted $s_i = 0$ and $s_i = 1$, calculating the predicted probability of attrition produces the estimated propensity scores, $\hat{p}(x_i)$. Furthermore, by assuming conditional independence[5] and overlap[6] a consistent estimator can be derived for the causal inference centered on weighting based on estimated propensity scores. The specification of the weighted causal estimator is formally expressed as:

$$\hat{\rho}_{weight} = \sum_{i=1}^{N} (\hat{w}_i(x_i)A_i - [1 - \hat{w}_i(x_i)]A_i) \tag{2.3}$$

where N is the number of observations in the data set and the weights, $\hat{w}_i(x_i)$, are given by:

$$\hat{w}_i(x_i) = \frac{s_i/\hat{p}(x_i)}{\sum_{j=1}^{N} s_j/\hat{p}(x_j)}.$$

I insert the calculated MCS weights into Equation 2.3 along with normalised weights of ones for the NCDS and the BCS to construct the sensitivity check of my analysis.

[4] As described in the User Guide to sample weights of the MCS (University of London, Institute of Education, Centre for Longitudinal Studies, 2020b)

[5] If potential outcomes, Y_{0i}, Y_{1i}, are independent of response conditional on covariates x_i, they are also response of treatment conditional on a propensity score $p(x_i)$. The formal expression of the conditional independence assumption is $(Y_{0i}, Y_{1i}) \perp\!\!\!\perp s_i | p(x_i)$ where $p(x_i) = P(s_i = 1|x_i) = E(s_i = 1|x_i)$

[6] The overlap assumption states that $0 < P(s_i = 1|p(x_i)) < 1$, i.e. for each value of the propensity score there exists both response and not

3 Data

The data used when conducting the analysis consists of three samples taken from the British population in three different periods from 1958 to 2002 in England, Scotland and Wales. The first data source is the 1958 National Child Development Study[7], an ongoing interdisciplinary longitudinal study of around 17,00 newborns (observations) and their respective family members sampled in one week in 1958. The second data source, also of multi-purpose use, is from the longitudinal 1970 British Cohort Study[8], which consists of over 17,000 observations comparable to NCDS but sampled in one week of 1970. The third sample is the Millennium Cohort Study[9], consisting of around 19,00 observations similar to both NCDS and BCS but sampled between 2000-2002. All included studies used in the analysis follow cohorts through childhood, adolescence, and adulthood and are of longitudinal form. The data on cohort members consists of health-related, child development, and cognitive ability assessments. The studies are all available at the Centre for Longitudinal Studies at the Institute of Education, University of London.

In addition, data on cohort members and their associated families contain information on educational attainment and social, economic, and demographic aspects, allowing for numerous analyses to be conducted with the quantity/quality trade-off model. In the quantity/quality trade-off model, one may investigate the quality of children in a variety of ways. Educational or occupational achievement measurements have been employed in correlational and causal studies with varied results to identify the quality of children (Butcher & Case, 1994; Kaestner, 1996; Conley, 2000; Amin, 2009; Peter et al., 2015; Brenøe, 2021; Black, Devereux, & Salvanes, 2005). However, little economic research has focused on the estimation of psychological and physiological health metrics (Bhai, 2015; Baez, 2008). Therefore, I deem it appropriate to elaborate on how prior research has created empirical methodologies to estimate educational outcomes in the quantity/quality trade-off model while expanding on what quality indicators this model may estimate.

I have data points until age 16 in the NCDS and BCS and 17 in the MCS when conducting this study. The age restriction of cohort members is dependent on information

[7]For the NCDS sample, data is received by the Centre for Longitudinal Studies at the Institute of Education, University of London (2020c)

[8]For the MCS sample, data is also received by the Centre for Longitudinal Studies at the Institute of Education, University of London (2013; 2021; 2021)

[9]For the MCS sample, data is also received by the Centre for Longitudinal Studies at the Institute of Education, University of London (2022a, 2022b, 2022c, 2021b, 2021a, 2020a, 2021c)

received from follow-up interviews conducted by each study. The different studies vary in questionnaires and the number of follow-up interviews since the initial sweep. The NCDS and the BCS made four sweeps between birth and age 16, while the MCS, with its more frequent sweep count of six between birth and age 17, differ. Since the data from the MCS's most recent sweep of 2022 is not accessible to the public, the choice of maxima sweep inclusion in my analysis depends on the latest MCS sweep of 2018. I am also limited to this constraint in cohort members' age since my identification strategy depends on families with twins, making the sample size of one study too small for inference testing since twins are rare in the population (Statista, 2020). Therefore, I include studies from the same institution that follow cohort members of the same country of origin to minimise any heterogeneity in country-specific fixed effects.

When it comes to matching cohort members between different sweeps in each study, I use the fact that respondents are given a unique I.D to their family. The unique I.D also makes it possible to match individual health and educational assessments, which provides further information on the cohort members' development throughout life. For example, if there exists a cohort member with a family of five that includes a mother, a father, and two siblings, of which one is a twin sibling with the cohort member, this whole family shares the same I.D. Sorting through other variables enables me to find each family member's role in the cohort member's life. For this study, including twins irrespective of zygotic status allows me to control for endogeneity in family size compared to other cohort members. As twin cohort members are included in pairs, more information is detailed throughout the different sweeps. The twin data is more extensive than the singleton data since the cohort members are the primary focal points in the sweeps.

One advantage of the posed maxima restriction is that attrition and heterogeneity in the sample composition are limited. Therefore, I minimise threats to internal validity induced by a non-random group dropout in later sweeps by using sweeps around 16 and 17. A possible explanation would be that people are selected for treatment by further education and drop out of the study later on because of optimising behaviour, i.e. treated are busier in general and do not bother replying to these extensive questionnaires or vice versa.

Regarding the disadvantages of the maxima restriction, the choice of the dependent variable is perhaps the most important one. As I am interested in studying the quality

of cohort members depending on predetermined characteristics and different parental inputs, not having access to variables such as family formation, fertility, and more detailed educational and labour outcomes weakens my possible analysis. These variables are outcomes that generally show themselves later in individuals' twenties and thirties.

3.1 Data Limitations

Longitudinal surveys such as the NCDS, BCS and MCS are subject to attrition. The first and perhaps most observable problem with missing data is that studies experience a reduction in sample size over time. A reduction in sample size might lead to problems with the statistical inference affecting the precision of the estimated coefficients (Verbeek, 2017, p. 150). The three studies are similar in terms of initial inclusion count but differ in the final sweep. Table 1 highlights productive cases, non-productive cases and ineligibility between all sweeps and across the three studies. In Column 1, the difference between the NCDS ($n = 14,647$) and MCS ($n = 10,625$) is the most notable heterogenous comparison. The difference between the BCS ($n = 11,206$) and the MCS is not as remarkable. There are possible explanations for this discrepancy when considering differences in the balancing of cohort attrition for each study. Firstly, the NCDS supplemented their initial cohort size with comparable migrants three times between 1965 and 1975 (Mostafa et al., 2020). This supplementation was to combat missing data stemming from cohort members moving to another country outside the U.K, making them ineligible.

There is also possible heterogeneity in the definition of eligibility between the MCS and the other studies caused by stricter inclusion criteria. The MCS excludes and does not issue questionnaires to all respondents who have not participated on two consecutive occasions (Mostafa & Ploubidis, 2017). Compared to the MCS, both the BCS and the NCDS try to reach these two-time non-response offenders. Furthermore, if you assume that non-response in longitudinal surveys is inevitable, then it would also be reasonable to assume that the study with the most sweeps would experience the most attrition. With this logic, MCS, with its seven sweeps, has a greater likelihood of experiencing attrition than the other studies with three sweeps.

Table 1: *Matching Cohort Members Through Sweeps*

	Productive		Ineligble		Non-productive	
	Freq.	%	Freq.	%	Freq.	%
	(1)		(2)		(3)	
National Child Development Study:						
Birth Sweep (1958)	17,416	98.7	–	–	223	1.3
1^{st} sweep age 7 (1965)	15,425	87.5	1,296	7.2	1,295	7.7
2^{nd} sweep age 11 (1969)	15,337	86.9	1,541	8.4	1,409	8.4
3^{rd} sweep age 16 (1974)	14,647	83.0	1,672	9.0	2,232	13.2
The original sample: $17,638^a$						
British Cohort Study:						
Birth Sweep (1970)	16,569	95.9	–	–	715	4.1
1^{st} sweep age 5 (1975)	12,939	74.9	565	3.3	3,780	21.9
2^{nd} sweep age 10 (1980)	14,349	83.0	585	3.4	2,350	13.6
3^{rd} sweep age 16 (1986)	11,206	64.8	597	3.4	5,481	31.7
The original sample: 17,284						
Millennium Cohort Study:						
9 months sweep (2001)	18,551	96.4	692	3,6	–	–
1^{st} sweep age 3 (2004)	15,590	81.0	853	4.4	1,870	10.6
2^{nd} sweep age 5 (2006)	15,246	79.2	846	4.4	2,605	14.7
3^{rd} sweep age 7 (2008)	13,857	72.0	3,044	15.8	2,219	14.3
4^{th} sweep age 11 (2012)	13,287	69.0	3,317	17.2	2,201	14.4
5^{th} sweep age 14 (2015)	11,726	60.9	4,301	22.4	3,141	22.0
6^{th} sweep age 17 (2018)	10,625	55.2	4,463	23.2	3,463	24.6
The original sample: 19,243						

Notes: – This table presents the proportion of productive, ineligible and non-productive cases in NCDS, BCD and MCS in all conducted sweeps. The reduction in productive cases over time highlights the need to use weights to combat non-random attrition when evaluating results from these studies. Partly author's own calculations from papers: on the National Child Development Study (Mostafa et al., 2020), the British Cohort Study (Mostafa & Wiggins, 2014) and the Millennium Cohort Study up to sweep 5 (Mostafa & Ploubidis, 2017) and for sweep 6, (Mori, 2020).
[a]The original sample was supplemented by migrants born in 1958 three times: in 1965, 1969 and 1975. These inclusions make the total cohort sample 18,558 observations for the NCDS.

Missing data also create a problem if the loss of observations is not random. Suppose attrition is correlated with any observable or unobservable characteristics. It will result in the loss of a specific kind of respondent, and the sample would no longer be representative of the initial sample (Mostafa & Wiggins, 2014). Even though the initial sampling was conducted adequately, this non-random attrition erodes the random selection when time passes. A researcher chooses three different routes to define said attrition: The dropout is 'completely at random', at 'at random', or 'not at random'. Only the 'not at random' approach takes bias induced by attrition into account, while both the 'completely at random' and 'at random' approaches assume the opposite. However, we can tackle this attrition bias using statistical methods under reasonable assumptions, i.e. that we can observe some characteristics. Mostafa & Wiggins also mentioned the two statistical

approaches to tackle missing data in longitudinal studies.

The two statistical approaches used to balance non-random missing observations in longitudinal studies both use observable characteristics to predict what variables affect attrition. As (Angrist & Pischke, 2008, p. 66) describes, sample weights can be used in several ways depending on how data is structured. The primary purpose of assigning weights is to make the estimated regression close to the initial sample's estimates. Firstly, sample weights re-balance the skewed distribution of specific observable characteristics If there is an overrepresentation of non-natives dropping out of the survey, these sample weights emphasise the remaining observations by assigning a higher weight to non-natives than natives. As Angrist & Pischke furthermore describe, the estimation of sample weights within the inverse probability model is equal to the inverse probability of sampling observation i. These sample weights are found within the MCS data set and will be of great use when conducting a meta-analysis in this study.

Parallel to the inverse probability weighting model, there exists a corresponding weighting scheme connected to group-level data. The grouped data, also called stratification or warding, depending on the aggregate level, makes predictions of dropping out by computing means relative to the frequency of the initial sample. In both these strategies, there is an underlying inherent problem connected to the fact that weights are assigned based on characteristics measured throughout time. In all included studies, measured outcome variables such as parental investment and further individual education are recorded in cohort members' teenage years. In contrast, underlying independent variables such as ethnicity and sex are collected in the initial sample, i.e. at birth. The measured cases constrain the resulting estimates in teenage years and at birth since information is only available for those with data points in both sweeps (Goldstein, 2009), undermining the efficiency of the estimates. However, any possible efficiency gains provided by an adequate weighting strategy should be carefully considered since incorrectly estimated weights can do more harm than good (Angrist & Pischke, 2008, p. 69). I will only use the weighting scheme provided by the MCS as a sensitivity check in Section 5.2 for my analysis[10]. This is partly because the MCS is the only study that provides computed weights and because the MCS is the study I suspect has the most attrition. Additionally, following the recommendations of Jones & Ketende (Jones & Ketende, 2010), I will use

[10]Griffiths et al. (2013) discuss extensively how these estimates are derived using probabilistic methods of selection combined with stratification and clustering)

the sample weights of the most recent sweep, i.e. sixth sweep age 17 (2018). Regarding the implementation of weighting schemes on group-level data, I collapse the stratum and warding classification defined within the MCS sample to one overall group since my twin strategy reduces the sample size in each classification extensively.

3.2 Description of the Sample Analysis

Table 2 depicts the three samples in my analysis, including appropriate variables of interest. The table includes a couple of endogenous variables, several variables on the family composition, a list of control variables, and three dependent variables. The first sample is found in Columns 1 to 3 of Table 2 shows descriptive statistics for the NCDS sample. Column 1 shows statistics for the Total Sample, i.e. every cohort member of the NCDS; Column 2 shows statistics for the *General Sample* which is the group of singleton siblings within a sibship size of three; Column 3 depicts the statistics of the *Twin Sample*, i.e. the group of singleton or twins contained in a sibship size of three where a twin is present. Columns 4 to 6 and 7 to 8 show the same descriptive statistics while corresponding to the different sub-samples in the BCS and MCS. There are some heterogeneous notions regarding the NCDS sample compared to the BCS and the MCS sample. For one, the fathers' ethnicity is not available in the data set, which makes the inclusion of that variable inaccessible. Secondly, there are no data on symptoms of depression as a proxy for mental health status in the NCDS. Thirdly, the number of observations in the Twin treatment, Column 3, is substantially fewer than in the other studies. This discrepancy is because there is no way to isolate if singleton cohort members are siblings to twins since I cannot isolate siblings' age within the NCDS data set. The number of twin cohort members is also smaller, which results in a double effect that extensively reduces the total number of observations in that group.

As previously mentioned in Section 3.1, approximately a third of the sample is drawn from each study, see Table 2. On average, mothers are in their late twenties when cohort members are born and are of native descent. However, these averages are increasing with time, reflecting the U.K's demographic change historically (Statista, 2021; Government, 2020) which underlines the initial sweeps adequate cohort member inclusion. Generally, mothers of cohort members tend to be younger than their respective fathers. Across the different studies, fathers of cohort members are approximately two years older than

mothers, the most remarkable difference coming from the MCS sample (mothers' age is 28 and fathers' age 32).

Regarding the dependent variables, we see a structural change in the increase of children continuing into further education after the completion of mandatory schooling. Twenty four per cent of cohort members in the NCDS sample report interest in further full-time education, while 56 per cent of the cohort members in the BCS sample report the same interest. The structural shift in further education is further supported by the strong support in the MCS sample, where 66 per cent of cohort members report an interest in further education. Furthermore, the average overweight teenagers, determined by a Body Mass Index over 25, are relatively stable in the NCDS and the BCS samples, being 11 and 12 per cent, respectively. In contrast, the MCS sample reports an overweight average of 27 per cent. Finally, whether any medical examinator reports signs of depression in cohort members, we see sample heterogeneity in that only 2 per cent of the 70s children in the BCS sample are depressed. In comparison, 11 per cent of the 00s children in the MCS sample are depressed which validates a fixed study effect inclusion in the controls, X_i of 2.1 and 2.2. From these descriptive statistics, one could conclude that the children of the 00s are, on average, less healthy in terms of psychological and physiological attributes than the children of the 50s and 70s.

Table 2: *Descriptive Statistics of Analysed Samples*

	National Child Development Study			British Cohort Study			Millennium Cohort Study		
	Total Sample (1)	General Sample (2)	Twin Sample (3)	Total Sample (4)	General Sample (5)	Twin Treatment (6)	Total Sample (7)	General Sample (8)	Twin Sample (9)
Endogenous variable:									
No. siblings	2.463	2	2	1.191	2	2	1.767	2	2
	(1.95)			(1.20)			(1.24)		
Treatment	2,784	2,784	19	3,718	3,718	195	2,733	2,733	152
Family composition:									
First birth (= 1 if CM is firstborn)	35.6	35.3	10.5	57.0	27.5	26.7	36.4	33.5	69.1
	(47.9)	(47.8)	(31.5)	(49.5)	(44.7)	(44.3)	(36.4)	(50.3)	(46.4)
Second birth (= 1 if CM is second-born)	29.4	30.5	89.5	26.3	33.7	73.3	23.0	30.4	30.9
	(45.6)	(46.1)	(31.5)	(44.0)	(47.3)	(44.3)	(42.1)	(46.0)	(46.4)
Gender (= 1 if CM is female)	48.3	48.0	63.1	48.2	45.8	52.8	50.4	53.7	47.4
	(49.9)	(49.9)	(49.6)	(49.9)	(49.8)	(50.0)	(50.0)	(49.8)	(50.1)
Mixed sex twins (= 1 if sibship has mixed sex twin-pair)	0.4	0.3	26.3	0.4	12.7	4.1	0.4	0.2	7.8
	(6.3)	(5.0)	(45.2)	(5.9)	(6.5)	(19.8)	(6.3)	(5.1)	(27.1)
Same sex twins (= 1 if sibship has same sex twin-pair)	0.9	0.8	73.7	0.3	1.3	28.2	.17	1.8	69.7
	(9.5)	(9.1)	(45.2)	(8.5)	(11.2)	(45.1)	(12.9)	(13.1)	(46.1)
Control variables:									
Year of birth (years)	1958	1958	1958	1970	1970	1970	2000.7	2000.7	2000.8
								(46.0)	(46.0)
Mother's age at CM's birth (years)	27.5	27.5	30.44	25.9	25.9	28.3	28.3	28.3	28.8
	(5.72)	(5.54)	(3.88)	(5.53)	(5.33)	(5.11)	(5.86)	(5.86)	(6.58)
Father's age at CM's birth (years)	30.6	30.0	32.3	27.0	27.2	26.08	31.6	31.7	31.4
	(6.45)	(6.18)	(5.34)	(9.79)	(9.09)	(6.05)	(6.48)	(6.45)	(6.26)
Mother's ethnicity									
Non-native (= 1 if mother is non-native)	4.1	2.2	0.0	6.6	4.8	2.3	16.1	15.9	12.5
	(19.8)	(14.8)	(0.0)	(24.7)	(21.3)	(15.0)	(36.7)	(36.6)	(31.8)
Father's ethnicity									
Non-native (= 1 if father is non-native)	–	–	–	10.8	08.7	2.3	14.9	14.3	12.8
				(31.1)	(29.4)	(15.0)	(28.2)	(35.1)	(33.5)
Dependent variables:									
Further education (= 1 if CM is continuing his/her education after age 16)	23.6	22.0	28.6	56.3	56.3	58.0	66.4	71.8	78.9
	(42.5)	(43.4)	(42.8)	(49.1)	(49.6)	(49.5)	(47.2)	(4.50)	(41.1)
Overweight (= 1 if CM has BMI > 25.0 at age 16)	10.7	10.6	13.3	11.6	12.2	13.6	27.5	27.5	27.3
	(30.9)	(30.1)	(35.2)	(32.1)	(32.6)	(13.6)	(44.7)	(44.7)	(44.9)
Depression (= 1 if CM shows signs of depression in medical examination at age 16)	–	–	–	2.3	2.1	1.9	10.3	10.8	9.8
				(15.1)	(14.4)	(13.6)	(30.4)	(30.9)	(30.0)
Observations	18,558	2,784	19	18,593	3,718	195	18,708	2,733	152

Notes – Author's tabulation of data supplied by the Centre for Longitudinal Studies (CLS) at the Institute of Education, University of London. Standard errors are presented in parentheses. The table presents descriptive statistics for the first part of the analysis without controlling for endogeneity in family size. The *Total Sample* consists of all family level observations within each matched study, presented separately for the National Child Development Study, the British Cohort Study and the Millennium Cohort Study. The *Treatment group* fulfils the desired characteristics for being 'treated' by being a three-child family with certain sex composition. The *Twin Sample* assigns treatment similar to the *Full Population Treatment* but also controls for endogeneity in family size by only including twins in *treatment*. The *control variables*, together with a *study fixed effects*, combined *twin* status and *birth order*, are included in X_i, the appropriate *treatment* variable, D_i or D_i^* and the *Dependent variables* in Y_i of Equation 2.1 and Equation 2.2.

Table 3: *Differences in Means for Demographic Variables*

Panel A1: Differences in Means for Parental Characteristics on CMs Gender Status			
	Mean of Male sample	Mean of Female sample	Mean difference
	(1)	(2)	(3)
Mother's age at CM's birth	27.232	27.290	0.058
	(5.789)	(5.785)	(0.055)
Father's age at CM's birth	29.908	30.008	0.100
	(7.760)	(7.788)	(0.076)
Mother is non-native	0.091	0.091	-0.000
	(0.288)	(0.288)	(0.003)
Father is non-native	0.122	0.130	0.008*
	(0.327)	(0.336)	(0.004)
Observations	47,295	8,564	56,946
Panel A2: Balancing Test for Gender Status			
Joint F-static	1.32		
Prob > F	0.26		
Panel B1: Differences in Means for Parental Characteristics on CMs Twin Status			
	Mean of Non-twin sample	Mean of Twin sample	Mean difference
	(1)	(2)	(3)
Mother's age at CM's birth	27.160	28.148	0.989***
	(5.805)	(5.569)	(0.213)
Father's age at CM's birth	29.848	31.211	1.363***
	(7.886)	(7.249)	(0.293)
Mother is non-native	0.092	0.106	0.014
	(0.289)	(0.308)	(0.011)
Father is non-native	0.128	0.132	0.004
	(0.334)	(0.339)	(0.016)
Observations	49,816	782	56,946
Panel B2: Balancing Test for Twin Status			
Joint F-static	4.59		
Prob > F	0.00		
Panel C Differences in Means for Treatment status for the Non-twin and the Twin Population			
	Mean of Non-twin sample	Mean of Twin sample	Mean difference
	(1)	(2)	(3)
General Sample	0.489	0.372	-0.116**
	(0.500)	(0.486)	(0.052)
Twin Sample	0.643	0.508	-0.135***
	(0.479)	(0.502)	(0.044)
Observations	49,816	782	56,946

Notes: – Author's tabulation of data supplied by the Centre for Longitudinal Studies (CLS) at the Institute of Education, University of London. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ based on a two-tail two-sample t-test with test hypothesis that the sample means are equal. Standard errors are presented in parentheses. This table presents differences in means between the genders, Panel A and twin status, Panel B, in observable parental characteristics. Panel C presents mean differences in treatment status between the general and the twin population. The *Mother is non-native*, and *Father is non-native* are binary variables with a value of 0 if parents are of "white-European" descent and a value of 1 if not.

To offer evidence for the identifying assumption that sibling sex is random, Column 3 of Panel A1 in Table 3 evaluates whether parental characteristics vary based on cohort members' sex. By only analysing observable control variables, there is only a significant difference, at the ten per cent level, between the two sexes in father's ethnicity. Panel A2 displays the results of a joint F-test, or a balancing test, that examines the control variables' ability to predict the sex of the cohort member. The F-test fails to reject the joint significance, which supports the identifying assumption of interest. In Column 3 of

Panel B1 in Table 3, I perform the same mean difference test of parental characteristics but for twin status. There are significant differences, at the 1 per cent level, between twins and non-twins in both mother's and father's age at cohort members' birth. These findings replicate the results of John A. H Waterhouse (1950) that twins are more likely for older women. Panel B2 presents the joint F-test of the control variables' ability to predict the twin status of the cohort member. The F-test reject the null hypothesis of no joint significance of these variables. Lastly, In Panel C of Table 3, I perform the same balance test but for the actual treatment variable: i.e. being the only-sex child in the *General Sample* vs the *Twin Sample* and having a sibling that is of the only sex in the *General Sample* and the *Twin Sample*. There is a significant difference, at the 5 per cent level, between the two populations when it comes to treatment status. A potential explanation for this discrepancy is that twins tend to be of the same sex in this sample which makes the possible combinations of being the only sex and having a sibling that is only-sex less likely.

4 Results

I present the results of being the *favourite child*, based on sibling sex composition, in terms of educational attainment and psychological and physiological health and well-being as quality measures in this section. I provide the main results of this thesis in subsection 4.1 by analysing the direct effect of being of the only sex in a sibship compared to being of the same sex as your siblings in a sibship on three measures. The three measures are: whether the child will continue into further education after the age of eighteen, the prevalence of overweight or obesity measured in Body Mass Index over 25 in cohort members and whether the child shows signs of depression when examined by a medical professional. Next, subsection 4.2 outlines these same effects but instead on children that are of the same sex as one other sibling and have one additional opposite-sexed sibling, which I denote as the indirect effect.

4.1 The Direct Effects of the Favourite Child

Table 4 reports the main results for the effect of being the *favourite child*. There is an overall tendency of a statistical non-significant relationship in educational attainment

and psychological and physiological health and well-being between children of the only sex compared to siblings of a same-sex sibling composition.

Panel A, Column 1 depicts the most rudimental model regressing whether cohort members will continue to further education on treatment status. Estimates show no statistical difference between a child of the opposite sex as the remaining sibship, continuing into further education from the average of 45.9 per cent. Column 2 includes an expanded functional form by estimating the same effect but also controls for predetermined characteristics such as; maternal and paternal age and ethnicity, sex of cohort members, birth order, twin status and a study fixed effects. Likewise, these estimates show no difference in the only-sex children pursuing further education from the same average. Column 3 expands on this controlling strategy by including control for family size endogeneity which is the causal estimate of interest. It also shows no difference in only-sex children from the average of 63.6 per cent. However, the more determinants I include, the less precise the estimates become reflected in the standard errors of 22.7 per cent compared to those of Column 1, which is 1.9 per cent. Finally, Column 4, with its included interaction-term between sex and treatment, reports no statistically difference in the likelihood of further education between the two sexes when comparing the sibships with two sexes and one sex coming from the only-sex child channel.

Panel B reports that only-sexed children have a lower risk of being overweight or obese measured in Body Mass Index when including time-invariant determinants such as parental background information, birth order and twin status. Column 1, using the only-sex children as a determinant, shows a statistically insignificant 0.7 per cent decrease in the prevalence of overweight or obesity compared to a full sibship of the same sex. In contrast, Column 2 reports a statistically significant lower risk of these health measures as only-sexed children experience on average 3.6 per cent lower risk of being overweight or obese than their same-sex sibship counterparts. The estimates of being only-sex siblings on the health outcomes also become less precise the more determinants I include, which is reflected in the increase in standard errors. This increase is likely driven by the sample size reduction when restricting only sibships where twins are present. The causal effects specification of Column 3 is not significantly different from zero between the prevalence of overweight and obesity and only-sexed siblings compared to sibships of the same sex. These findings point toward a downward bias in the estimates not controlling for sample

size endogeneity and the prevalence of overweight and obesity. Turning our attention to the last specification in Column 4, when the analysis accounts for differences in the only-sex effect between boys and girls, we also observe this non-significance neutral effect of being an only-sexed sibling.

Panel C reports a causal relationship between only-sexed children and having the same risk of medical examiner reporting depression in teenagers as children of same-sex sibships. However, the magnitude of the reported standard errors becomes more salient when I account for additional determinants, probably due to the estimation excluding the NCDS data from the prevalence of depression estimates. The NCDS exhibit the fewest treatment observations, which may cause problems when estimating the effects of being an only-sexed child on quality outcomes. Therefore, the following estimates are robust when analysing the relationship between sex composition in families and depression. From the causal estimates of Column 3, we see a statistically insignificant difference of experiencing depression through teenage years if you are the only-sexed sibling in a family of the siblings compared to a family of the same sex and sibship parity. These results show that the purely correlational estimates not accounting for sample size endogeneity do not suffer from any bias. When analysing these effects separately for boys and girls in Column 4, I do not find any statistically significant differences between the sexes of being the only-sexed sibling compared to siblings of a same-sex sibship.

Turning our attention to the significance levels of the quality measures, there is an overall tendency of non-significance, even at the 10 per cent level, in the effect of being an only-sex sibling compared to the sibship of same-sex siblings. The neutral result may reflect an absence of an effect in this analysis. It may also be the result of imprecision induced by the sample having slight variation or being too small (Verbeek, 2017, p. 24). With an ocular inspection of observations in Columns 1 to 3 in Panel B, where the difference is the most evident, we see major differences due to restricting the sample to sibships with a twin pair. The first specification, Column 1, reports 2,978 observations, and the second specification, Column 2, is relatively stable with its 1,499 observations. In contrast, Columns 3 and 4, with their respective specifications, exhibit only 64 observations. This is a substantial drop in the number of observations, affecting the inference testing. Verbeek (2017, p. 30-31) states that type II errors are increasingly unlikely in large samples. The toppled effect of reduced sample size might cause these estimates to

become non-significant.

Table 4: *Effects of being Only-Sex Sibling on Child Development*

	General Sample		Twin Sample	
	(1)	(2)	(3)	(4)
Panel A: Further Education				
Treatment	0.5	2.9	-5.3	-15.1
	(1.9)	(2.6)	(22.7)	(26.3)
Interaction term	–	–	–	23.2
				(31.6)
Observations	2,870	1,365	60	60
Average	45.9	45.9	63.6	63.6
Panel B: Overweight				
Treatment	-0.7	-3.6*	-20.6	-22.8
	(1.4)	(2.1)	(21.5)	(23.1)
Interaction term	–	–	–	21.4
				(23.3)
Observations	2,978	1,499	64	64
Average	17.6	17.6	13.1	13.1
Panel C: Depressed				
Treatment	-0.6	-0.2	1.2	2.5
	(1.2)	(1.3)	(16.0)	(17.0)
Interaction term	–	–	–	-2.0
				(15.2)
Observations	1,819	1,543	68	68
Average	7.4	7.4	5.6	5.6
No controls	✓			
Principal controls		✓	✓	✓
Twin treatment			✓	✓
Interaction term				✓

Notes: – Standard errors are in the parentheses.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The author multiplied the estimates by 100 to get the percentages. Each column in this table is one separate regression indicated by the checkmarks. Column 1 runs a regression of the outcome variables on thetreatment indicator of being an only-sex sibling compared to a sibling of a same-sex sibship. Column 2 shows the same regressions while controlling for maternal and parental age and ethnicity, sex of CM and a study fixed effects. Column 3 expands on this by controlling for family size endogeneity and is the causal estimate of being an only-sex sibling. Column 4 also estimates the causal estimate of being only-sex while including an interaction term between sex (= 1 if girl, = 0 if boy) and being only-sex, i.e. controlling for differences between the boys and girls.

4.2 The Indirect Effects of the Favourite Child

Table 5 reports the regressional results of the indirect effect of the *favourite child*. Compared to the direct results of being a *favourite child* explored in the previous subsection, there seems to be a greater difference in the quality measures between the same-sex sibling with an only-sex sibling and the sibship with the same sex in all siblings. This discrepancy highlights that the differentiating made by parental inputs based on preferences in

child sex composition is most evident in the non-unique sex child outcomes.

Panel A Column 1 reports a neutral educational attainment result for the individual who shares their sex with one of their siblings compared to the individual who shares their sex with two siblings in a sibship of size three. When I add additional determinants such as predetermined maternal and paternal age and ethnicity, sex of cohort members, birth order, twin status and a study fixed effects in Column 2, results remain stable in that there is no statistically significant difference between treatment and control. Furthermore, in Column 3, when the model specification is expanded to include a control for sample size endogeneity, the magnitude of the estimates increases. At the same time, it remains not significantly different from the control group. This relationship remains when I consider differences between boys and girls in the *favourite child* effect, Column 4.

Panel B Column 1 suggests that being of the same sex as one of your siblings while the remaining sibling is of the opposite sex compared to a sibship of the same sex leads to a reduction of 2.7 per cent at the 5 per cent significance level. In Column 2, where I include additional determinants of family characteristics, this negative relationship increases to 3.3 per cent, at the one per cent significance level. Including other regressors that account for sample size endogeneity in the specification increases the standard errors that the indirect impact of *favourite children* is no longer statistically distinct from zero. The statistically significant estimates of Columns 3 and 4 may result from a sample size reduction. Still, they can also shed light on the true causal estimates of the indirect effects of *favourite children* and their siblings. The purely correlational finding of Columns 1 to 2 suffers from downward bias.

Panel C reports the estimates for psychological health and well-being comparing the indirect *favourite child* effects and the sibship of same-sex. Children of the same sex as only one of their siblings compared to a child who shares their sex with two of their siblings appear to have an increased risk of reported depression history, examined by a medical professional. In Columns 1 and 2, estimates are not distinctly different from zero when I do not account for determinants controls for sample size endogeneity (i.e. including sibships where all siblings are singleton children). Column 3 shows that the indirect *favourite child* effect leads to a 12.4 per cent increase in the risk of depression in the teenage years accounting for the downward bias estimates of the specifications of Columns 1 to 2. When I consider differences between the two sexes in having an only-sex

sibling, boys are on average 14.6 per cent more likely to experience depression in their teenage years than a sibling of a same-sex sibship. However, girls do not experience a similar increase in the risk of depression when examined by a medical professional. These results show that there seem to be differences between boys and girls when it comes to being the sibling of an only-sex sibling.

This thesis focus on testing the hypothesis that parents favour one of their children based on sex composition preference. The idea is that given this favouring behaviour, additional resources will be spent on the child, and the subsequent outcomes of the favourite child will be of higher quality. This thesis's quality measures are educational attainment and psychological and physiological health, and well-being. From the result presented in this section, we observe some stylised facts regarding these quality measures when comparing children of only sex and children of the same sex as the whole sibship, conditional on sibship parity. On average, unique-sexed children do not experience an increased risk of depression than children from sibships of the same sex. I do, however, observe differences in how siblings of a unique-sex child experience a higher likelihood of depression than the same-sex sibship. These findings point out that each individual in same-sex sibship experiences a lower risk of reporting depression in their teenage years than children in sibships of two sexes. However, the findings expand traditional quality measures in the quantity/quality trade-off model by inserting psychological health outcomes. While previous research has investigated the relationship between fertility and mental health on similar data sets (Lawson & Mace, 2009, 2010), none has used the same empirical strategy as this study to estimate how parental preferences in sex composition affect a child's mental well-being. The following section will provide evidence that the effect of being a favourite child runs from the parental channel since parents differentiate investments based on preferences in children's sex composition.

Table 5: *Effects of Sibling Being Only-Sex Sibling on Child Development*

	General Sample		Twin Sample	
	(1)	(2)	(3)	(4)
Panel A: Further Education				
Treatment	1.5	-10.7	13.0	-15.1
	(2.2)	(2.2)	(14.6)	(16.7)
Interaction term	–	–	–	13.9
				(25.9)
Observations	4,185	1,976	90	90
Average	46.9	46.9	59.3	59.3
Panel B: Overweight				
Treatment	-2.7**	-3.3*	-11.2	-9.2
	(1.2)	(1.8)	(10.6)	(12.5)
Interaction term	–	–	–	-6.2
				(20.4)
Observations	4,187	2,050	100	100
Average	16.2	16.2	15.9	15.9
Panel C: Depressed				
Treatment	-0.1	-0.6	12.4**	14.6**
	(1.1)	(1.1)	(4.9)	(5.8)
Interaction term	–	–	–	-6.4
				(9.1)
Observations	2,538	2,153	104	104
Average	7.3	7.3	3.2	3.2
No controls	✓			
Principal controls		✓	✓	✓
Twin treatment			✓	✓
Interaction term				✓

Notes: – Standard errors are in the parentheses.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The author multiplied the estimates by 100 to get the percentages. Each column in this table is one separate regression indicated by the checkmarks. Column 1 is a regression where the independent variables are run on the treatment indicator of being a same-sex sibling with an only-sex sibling compared to a sibling of a same-sex sibship. Column 2 shows the same regressions while controlling for maternal and parental age and ethnicity, sex of CM and a study fixed effects. Column 3 expands on this by controlling for family size endogeneity and is the causal estimate of being an only-sex sibling compared to a sibling of a same-sex sibship. Column 4 also estimates the causal estimate of having an only-sex sibling while including an interaction term between sex (= 1 if girl, = 0 if boy) and being only-sex, i.e. controlling for differences in treatment between the sexes.

5 Mechanisms and Sensitivity

In this section, I will outline a potential mechanism behind parental investments as a mediator of preferences in their children's sex composition. Then continue by including a sensitivity check section where I explore the attrition problem and the stability of the results in Section 4.

5.1 Mechanisms Behind Parental Investments

I utilise that within the NCDS, the BCS and the MCS, detailed information exists on how invested the parents are in the cohort members' homework from school to study whether sibling sex composition influences child-parent relationships. For the 1958 children of the NCDS, this parental investment proxy is measured at the age of 7 when both parents are asked whether parents provide substantial help at home. In the BCS, the dataset provides a similar proxy for parental investment where the 1970s teenagers at the age of 16 are asked whether their parents helped with school homework throughout childhood. Lastly, in the MCS, the proxy for parental investment is derived at the age of 11 when asked how often anyone at home help with a child's homework. As I am not worried about the different nuances in parental investment, I normalise the variables for parental investments in all studies to substantial help (= 1) or not (= 0) in the outcome variable of the regressions. I include the same specifications as in the Result section to highlight what direction the bias comes from.

Columns 1 to 4 in Table 6 report the main regressional results of parental investment for both families where the cohort member is only-sexed in the sibship and when the cohort member's sibling is of only-sex. Panel A describes the *General Sample*, i.e. restricted to sibships of three children, while Panel B control for sibship endogeneity by only including families where a twin pair is present. Parental investment in children whose sibling is of the only sex is 4.2 per cent less than investment in children of sibship where all siblings are of the same sex. The effect of being a same-sex sibling with one only-sex sibling compared to a sibship of the same sexes increases when I consider differences between boys and girls. In this specification, parental investment is 7.7 per cent less for boys than the siblings of the same sex. There are, however, no statistically significant differences in parental investment for girls in this specification. In contrast,

there is no significant difference between the only-sex children in sibships in the families of three singleton children compared to sibships where all members are of the same sex. Column 2 *Interaction term* shows that the effect of being a girl with two boy siblings leads to an 8.9 per cent increase in parental investment, at the 10 per cent significance level, compared to the sibship of same-sex siblings. The results from Column 2 in both Panel A and Panel B show the same relationship as Brenøe (2021) where parents differentiate their investments in children based on sex. The differentiation in investment based on sex is, however, biased downwards as the *Twin Sample* reports a higher relationship between sex-based differentiation than the *General Sample*.

The findings of Panel B, where I control for sample size endogeneity, suggest causal differences in parental investments between sibships that consist of two sexes and three siblings and sibships that consist of one gender. These findings contrast with the *General Sample* as it unveils differences in how parents invest differently, in line with my initial hypothesis. Column 1 reports no significant difference, while Column 2, where I consider differences between boys and girls in parental investments, shows no statistical difference for boys when it comes to substantial help in homework from parents. The sex difference between girls that are only-sex in their respective sibship is that girls, on average, get 43.4 more attention from parents in terms of homework aid compared to the average of 51.6 per cent. These estimates are statistically significant at the ten per cent level. Compared to Dahl & Moretti (2008), my findings indicate favouring in the U.K, at least in parental investment, of girls compared to boys in terms of substantial help in school homework.

Table 6: *Effect of Sibling Sex Composition on Parental Time Investment*

	CM is Only-sex		Sibling is Only-sex	
	(1)	(2)	(3)	(4)
Panel A: General Sample				
Treatment	0.1	-4.7	-4.2**	-7.7**
	(2.3)	(3.5)	(2.0)	(3.1)
Interaction term		8.9*		6.3
		(4.7)		(4.1)
Observations	1,851	1,851	2,577	2,577
Average	52.4	52.4	50.5	50.5
Panel B: Twin Sample				
Treatment	-0.2	23.4	14.7	11.2
	(24.7)	(27.9)	(13.1)	(15.9)
Interaction term		43.4*		8.9
		(25.2)		(22.4)
Observations	78	78	118	118
Average	51.6	51.6	55.2	55.2

Notes: – Standard errors are in the parentheses.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The author multiplied the estimates by 100 to get the percentages. Each column represents causal estimates for parental investment in cohort members. Columns 1 and 2 estimate the percentage change in parental investments for only-sex children compared to siblings of a same-sex sibship. Columns 3 and 4 estimate the percentage change in parental investment for children with only-sex siblings compared to siblings of same-sex sibships. The included *Interaction term* is provided to analyse differences in boys and girls (= 1 if girl, = 1 if boy) being only-sex children in a sibship.

5.2 Sensitivity Checks of the Results

To further test the robustness of my two analyses and to tackle the attrition in the studies, I provide a sensitivity check regarding the missing data problem posed in Section 3.1. The Inverse Probability Model outlined in Section 2.1 is a method to adjust for bias due to a 'not at random' dropout in the different longitudinal studies. I will apply the weighting scheme supplied by the MCS in this sensitivity check. This approach limits the sample size slightly since observations from the MCS with negative weights are discarded to provide inference testing with sample weights. Normalised sample weights (= 1) of NCDS and BCS are also included in the study making the total amount of observations relatively similar to the principal analysis.

Table 7 presents the principal sensitivity test of the direct effect of being the favourite child using the inverse probability weighted models and the exact specifications as in the analysis. Compared to the primary analysis results, the estimates of educational attainment for the only-sexed sibling differ from the sibships of the same sex. Panel A, Columns 1 to 2, excluding sample size endogeneity controls, produces a downward bias

in the estimates since they underestimate the true effect. Column 3 reports a statistically significant 28.5 per cent causal increase in the likelihood of continuing into further education for 16-year-old's. In the main analysis, there seems to be some non-random attrition in individuals not planning to continue into further education, which is picked up in the robustness test. When estimating differences between the boys and girls who are only-sex in their respective sibships, no statistically significant difference is observed replicating the findings of the main results.

The sensitivity check of the estimates of the prevalence of overweight and obesity among only-sex siblings compared to the sibships of the same sex are relatively robust across the specifications. However, there is an increase in the precision of the estimates as the figures in Column 2 in Panel B report a higher degree of statistical significance than the corresponding figures in Table 4. The sensitivity check of the causal relationship underlines the aforementioned downward bias of estimates excluding controls for sample size endogeneity.

Table 7: *Robustness Test for Being Only-Sex Sibling on Child Development*

	General Sample		Twin Sample	
	(1)	(2)	(3)	(4)
Panel A: Further Education				
Treatment	1.3	2.8	28.5*	16.5
	(1.9)	(2.7)	(28.2)	(33.0)
Interaction term	–	–	–	24.6
				(36.9)
Observations	2,601	1,150	47	47
Average	44.5	44.5	54.6	54.6
Panel B: Overweight				
Treatment	-2.1	-7.9**	-15.3	-23.4
	(1.4)	(2.3)	(17.6)	(22.2)
Interaction term	–	–	–	28.1
				(28.6)
Observations	2,557	1,155	48	48
Average	14.8	14.8	8.6	8.6
Panel C: Depressed				
Treatment	2.9*	1.1**	1.1	0.2
	(1.5)	(18.2)	(18.2)	(19.8)
Interaction term	–	–	–	-0.2
				(24.7)
Observations	1,376	1,167	51	51
Average	8.5	8.5	4.4	4.4
No controls	✓			
Principal controls		✓	✓	✓
Twin treatment			✓	✓
Interaction term				✓

Notes: – Standard errors are in the parentheses.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Estimates are multiplied by 100 to get the percentages. Each column in this table is one separate regression indicated by the checkmarks. Columns 1 to 4 are estimates using the inverse probability weighting model to assign weight to the MCS sample. In Column 1, the outcome variables are regressed on the treatment indication of being an only-sex sibling vs a sibling of a same-sex sibship. Column 2 displays the same regressions adjusted for maternal and paternal age and ethnicity, CM's gender, and study fixed factors. Column 3 extends on this by adjusting for variation in family size and provides a causal estimate of being an only-sex sibling compared to a sibling of a same-sex sibship. Column 4 additionally calculates the causal estimate of being only-sex by incorporating an interaction term between sex (= 1 if girl, = 0 if boy) and being only-sex, i.e. controlling for differences between the boys and girls.

Panel C provides the robustness test of the estimates of psychological health between only-sexed siblings and sibships of same-sex. When accounting for non-random attrition, the purely correlational estimates of Columns 1 to 2 show differences between the only-sexed siblings and the sibships of the same-sex compared to the primary analysis results. The robustness test of the causal estimates in Column 3 proves insignificant in line with the main analysis. The main causal estimates of psychological health were not significantly different from zero. When I account for non-random attrition, there seems to be no difference between the only-sexed sibling of a sibship of three and the same-sex sibling of a sibship of three with the same sex throughout the whole sibship.

Turning our attention to the sensitivity check of the indirect effect of favourite children, Table 8 reports the estimates of the inverse probability weighted models using the exact specifications as in the analysis. Panels A and B show similar results as in the main, unweighted analysis for the indirect effect of favourite children. There is a slight difference in the statistical precision in the estimates, probably induced by the discarding of negative weights in the MCS sample depicted in the respective observations. Panel A, there seems to be some 'non-random' attrition in the individuals who do not choose to continue further education since the correlational findings of Column 2 are significant in the robustness test and not in the main analysis. The robustness test provides evidence of downwards bias in the correlational estimates of the prevalence of overweight and obesity, in Panel B, between siblings being only-sex and if the sibship is of the same sex. The causal estimates of psychological health in Columns 3 to 4 in Panel C provide no estimates in the inverse probability weighted model. The discarding of negative weighted individuals reduces the sample size so that no inference testing can be made.

Table 8: *Robustness Test for Sibling Being Only-Sex Sibling on Child Development*

	General Sample		Twin Sample	
	(1)	(2)	(3)	(4)
Panel A: Further Education				
Treatment	-1.2	-4.5*	-12.2	-15.4
	(1.7)	(2.5)	(18.4)	(20.1)
Interaction term	–	–	–	11.2
				(33.7)
Observations	3,711	1,587	77	77
Average	43.1	43.1	52.0	52.0
Panel B: Overweight				
Treatment	-2.9**	-4.8**	-11.9	-14.1
	(1.2)	(2.0)	(13.2)	(14.2)
Interaction term	–	–	–	10.2
				(101.5)
Observations	3,612	1,564	80	80
Average	14.0	14.0	13.9	13.9
Panel C: Depressed				
Treatment	-1.9*	-1.4	–	–
	(1.1)	(1.1)	–	–
Interaction term	–	–	–	–
				–
Observations	1,895	1,605	80	80
Average	5.7	5.7	0.0	0.0
No controls	✓			
Principal controls		✓	✓	✓
Twin treatment			✓	✓
Interaction term				✓

Notes: – Standard errors are in the parentheses.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Estimations are multiplied by 100 to obtain percentages. Each table column represents a separate regression, as indicated by the checkmarks. Estimates utilising the inverse probability weighting model to allocate weight to the MCS sample are displayed in columns 1 through 4. In Column 1, the outcome variables are regressed on the treatment indication of being the only-sex sibling's brother or sister versus being a sibling of a same-sex sibship. Column 2 presents the same regressions corrected for the age and ethnicity of the mother and father, the gender of the CM, and study fixed variables. Column 3 extends this by controlling for family size variance and providing a causal estimate of being an only-sex siblings brother or sister vs a sibling of a same-sex sibship. In addition, column 4 estimates the causal estimate of being only-sex by integrating an interaction term between sex (= 1 if girl, = 0 if boy) and being only-sex.

6 Conclusion

In this study, I analyse a favourite child effect based on parental preferences in the sex composition of children using a proxy for parental investment; I can determine how parents distinguish inputs based on the sex of their children. Using the three longitudinal studies NCDS, BCS and MCS of the U.K, I find that the favourite child's sibling is more likely to experience depression during adolescence than children of a same-sex sibship. However, these findings introduce some caveats: since my identification strategy is centred around sibships of size three, my results are only applicable for families of three children. This restriction induces problems with the external validity since a household of three children may not be the norm in the U.K. However, the results have potential implications for both families and policymakers in understanding how innate characteristics, such as sex, drive outcome variables, such as education and mental and physical health and well-being, depending on family sex composition. It also underlines how parental investment in children affects the psychological health of children who are only-sex in three sibling sibships. I strongly encourage further research to expand on quality measures and proxies for parental investments. Specifically, analysis using the Millennium Cohort Study of later sweeps that will be available shortly could provide evidence of how only-sexed children's marital status and family formation is affected by their parents' preferences in sex composition. Also advised to expand the approach of this study to other countries perhaps to see intra-country differences in parental investment and child outcomes

References

Altieri, A., & Hemminki, K. (2007). Number of siblings and the risk of solid tumours: a nation-wide study. *British journal of cancer*, *96*(11), 1755–1759.

Amin, V. (2009). Sibling sex composition and educational outcomes: A review of theory and evidence for the uk. *Labour*, *23*(1), 67–96.

Angrist, J., & Evans, W. N. (1996). *Children and their parents' labor supply: Evidence from exogenous variation in family size.* National bureau of economic research Cambridge, Mass., USA.

Angrist, J., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, *91*(434), 444–455.

Angrist, J., Lavy, V., & Schlosser, A. (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics*, *28*(4), 773–824.

Angrist, J., & Pischke, J.-S. (2008). Mostly harmless econometrics. In *Mostly harmless econometrics.* Princeton university press.

Baez, J. (2008). Does more mean better? sibling sex composition and the link between family size and children's quality.

Becker, G. S., & Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of political Economy*, *81*(2, Part 2), S279–S288.

Becker, G. S., & Tomes, N. (1976). Child endowments and the quantity and quality of children. *Journal of political Economy*, *84*(4, Part 2), S143–S162.

Bhai, M. (2015). Brothers are better: The effect of sibling sex composition on women's schooling, health, earnings, and labor supply. *Health, Earnings, and Labor Supply (February 23, 2015)*.

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

Brenøe, A. A. (2021). Brothers increase women's gender conformity. *Journal of Population Economics*, 1–38.

Browning, M. (1992). Children and household economic behavior. *Journal of Economic Literature*, *30*(3), 1434–1475.

Butcher, K. F., & Case, A. (1994). The effect of sibling sex composition on women's education and earnings. *The Quarterly Journal of Economics*, *109*(3), 531–563.

Chamberlain, G., University of London, Institute of Education, Centre for Longitudinal Studies, Chamberlain, R. (2013). *1970 british cohort study: Birth and 22-month subsample, 1970-1972.* [data collection]. 3rd Edition. UK Data Service. SN: 2666, DOI: 10.5255/UKDA-SN-2666-2.

Conley, D. (2000). Sibship sex composition: Effects on educational attainment. *Social Science Research*, *29*(3), 441–457.

Dahl, G. B., & Moretti, E. (2008). The demand for sons. *The review of economic studies*, *75*(4), 1085–1120.

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

Eijck, K. v., & De Graaf, P. M. (1995). The effects of family structure on the educational attainment of siblings in hungary. *European Sociological Review*, *11*(3), 273–292.

Goldstein, H. (2009). Handling attrition and non-response in longitudinal data. *Longitudinal and Life Course Studies*, *1*(1).

Government, U. (2020). *Population of england and wales.* Retrieved 2022-05-16, from https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/national-and-regional-populations/population-of-england-and-wales/latest

Griffiths, L., Rich, C., Geraci, M., Sera, F., Cortina-Borja, M., Pouliou, T., ... Dezateux, C. (2013). Technical report on the enhancement of millennium cohort study data with accelerometer-derived measures of physical activity and sedentary behaviour in seven year olds. *London, UK*.

Iacovou, M. (2001). *Fertility and female labour supply* (Tech. Rep.). ISER Working Paper Series Number 2001-19.

Jones, E., & Ketende, S. (2010). User guide to analysing mcs data using spss. *London: University of London*.

Kaestner, R. (1996). *Are brothers really better? sibling sex composition and educational*

achievement revisited. National Bureau of Economic Research Cambridge, Mass., USA.

Lawson, D. W., & Mace, R. (2009). Trade-offs in modern parenting: a longitudinal study of sibling competition for parental care. *Evolution and Human Behavior*, *30*(3), 170–183.

Lawson, D. W., & Mace, R. (2010). Siblings and childhood mental health: evidence for a later-born advantage. *Social Science & Medicine*, *70*(12), 2061–2069.

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

Mori, I. (2020). *The centre for longitudinal studies (cls) institute of education. millennium cohort study seventh sweep - technical report millennium cohort study seventh sweep (mcs7)* (Tech. Rep.). Retrieved 2022-05-11, from https://cls.ucl.ac.uk/cls-studies/millennium-cohort-study/mcs-age-17-sweep/

Mostafa, T., Narayanan, M., Pongiglione, B., Dodgeon, B., Goodman, A., Silverwood, R., & Ploubidis, G. (2020). Improving the plausibility of the missing at random assumption in the 1958 british birth cohort: A pragramatic data driven approach.

Mostafa, T., & Ploubidis, G. (2017). *Millennium cohort study* (Tech. Rep.). Sixth Survey 2015-2016: Technical report on response (Age 14).

Mostafa, T., & Wiggins, D. (2014). Handling attrition and non-response in the 1970 british cohort study.

Peter, N., Lundborg, P., & Webbink, H. D. (2015). The effect of a sibling's gender on earnings, education and family formation.

Powell, B., & Steelman, L. C. (1990). Beyond sibship size: Sibling density, sex composition, and educational outcomes. *Social Forces*, *69*(1), 181–206.

Richiardi, L., Akre, O., Lambe, M., Granath, F., Montgomery, S. M., & Ekbom, A. (2004). Birth order, sibship size, and risk for germ-cell testicular cancer. *Epidemiology*, 323–329.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.

Rosenzweig, M. R., & Schultz, T. P. (1987). Fertility and investments in human capital: Estimates of the consequence of imperfect fertility control in malaysia. *Journal of Econometrics*, *36*(1-2), 163–184.

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

Statista. (2020). *Rate for multiple births in england and wales from 1990 to 2020.* Retrieved 2022-05-25, from https://www.statista.com/statistics/971604/england-and-wales-multiple-birth-rate/

Statista. (2021). *Average age of mothers at childbirth in england and wales from 1938 to 2020, by child number.* Retrieved 2022-05-16, from https://www.statista.com/statistics/294594/mother-average-age-at-childbirth-england-and-wales-by-child-number/#:~:text=In%202020%20the%20average%20age,%2C%20third%2C%20and%20fourth%20child.

University of London, Institute of Education, Centre for Longitudinal Studies. (2020a). *Millennium cohort study: Sixth survey, 2015.* [data collection]. 7th Edition. UK Data Service. SN: 8156, DOI: 10.5255/UKDA-SN-8156-7.

University of London, Institute of Education, Centre for Longitudinal Studies. (2020b). *Millennium cohort study user guide (surveys 1 – 5).* August 2020. 9th Edition. UK Data Service. http://doc.ukdataservice.ac.uk/doc/5795/mrdoc/pdf/mcs1-5$_u ser_g uide_e d9_2 020 - 08 - 07.pdf$.

University of London, Institute of Education, Centre for Longitudinal Studies. (2020c). *National child development study: Childhood data from birth to age 16, sweeps 0-3, 1958-1974.* [data collection]. 3rd Edition. National Children's Bureau, National Birthday Trust Fund, [original data producer(s)]. National Children's Bureau. SN: 5565, DOI: 10.5255/UKDA-SN-5565-2.

University of London, Institute of Education, Centre for Longitudinal Studies. (2021a). *Millennium cohort study: Fifth survey, 2012.* [data collection]. 5th Edition. UK Data Service. SN: 7464, DOI: 10.5255/UKDA-SN-7464-5.

University of London, Institute of Education, Centre for Longitudinal Studies. (2021b). *Millennium cohort study: Fourth survey, 2008.* [data collection]. 8th Edition. UK Data Service. SN: 6411, DOI: 10.5255/UKDA-SN-6411-8.

University of London, Institute of Education, Centre for Longitudinal Studies. (2021c). *Millennium cohort study: Seventh survey, 2018.* [data collection]. 2nd Edition. UK Data Service. SN: 8682, DOI: 10.5255/UKDA-SN-8682-2.

University of London, Institute of Education, Centre for Longitudinal Studies. (2022a). *Millennium cohort study: First survey, 2001-2003.* [data collection]. 14th Edition. UK Data Service. SN: 4683, DOI: 10.5255/UKDA-SN-4683-6.

University of London, Institute of Education, Centre for Longitudinal Studies. (2022b). *Millennium cohort study: Second survey, 2003-2005.* [data collection]. 11th Edition. UK Data Service. SN: 5350, DOI: 10.5255/UKDA-SN-5350-6.

University of London, Institute of Education, Centre for Longitudinal Studies. (2022c). *Millennium cohort study: Third survey, 2006.* [data collection]. 9th Edition. UK Data Service. SN: 5795, DOI: 10.5255/UKDA-SN-5795-6.

University of London, Institute of Education, Centre for Longitudinal Studies, Butler, N., Bynner, J. (2021). *1970 british cohort study: Age 10, sweep 3, 1980.* [data collection]. 6th Edition. UK Data Service. SN: 3723, DOI: 10.5255/UKDA-SN-3723-7.

University of London, Institute of Education, Centre for Longitudinal Studies, Osborn, A., Dowling, S., Butler, N. (2021). *1970 british cohort study: Age 5, sweep 2 1975.* [data collection]. 5th Edition. UK Data Service. SN: 2699, DOI: 10.5255/UKDA-SN-2699-4.

Verbeek, M. (2017). *A guide to modern econometrics.* John Wiley & Sons.

Waterhouse, J. A. (1950). Twinning in twin pedigrees. *British journal of social medicine*, *4*(4), 197.

Zajonc, R. B., & Markus, G. B. (1975). Birth order and intellectual development. *Psychological review*, *82*(1), 74.

Zajonc, R. B., Markus, H., & Markus, G. B. (1979). The birth order puzzle. *Journal of personality and social psychology*, *37*(8), 1325.

Zajonc, R. B., & Mullally, P. R. (1997). Birth order: Reconciling conflicting effects. *American Psychologist*, *52*(7), 685.

A Appendix

Table A1: *Definitions of Treatment and Control for the General and Twin Samples*

General Sample Treatment			Twin Sample Treatment		
CM is only-sex (1)	Sibling is only-sex (2)	Control (3)	CM is only-sex (4)	Sibling is only-sex (5)	Control (6)
{B*, G, G}	{B*, B, B}	{B*, B, B}	{B*, GG}	{B*, BG}	{B*, BB}
{G, B*, G}	{G, B*, B}	{B, B*, B}	{GG, B*}	{B*, GB}	{B, B*B}
{G, G, B*}	{B, G, B*}	{B, B, B*}	{B*G, G}	{BG, B*}	{B, BB*}
{G*, B, B}	{B*, G, B}	{G*, G, G}	{GB*, G}	{GB, B*}	{B*B, B}
{B, G*, B}	{B, B*, G}	{G, G*, G}	{G, B*G}	{B*B, G}	{BB*, B}
{B, B, G*}	{G, B, B*}	{G, G, G*}	{G, GB*}	{BB*, G}	{BB, B*}
	{G*, G, B}		{G*, BB}	{G, B*B}	{G*, GG}
	{B, G*, G}		{BB, G*}	{G, BB*}	{G, G*G}
	{G, B, G*}		{G*B, B}	{G*, BG}	{G, GG*}
	{G*, B, G}		{BG*, B}	{G*, GB}	{G*G, G}
	{G, G*, B}		{B, G*B}	{BG, G*}	{GG*, G}
	{B, G, G*}		{B, BG*}	{GB, G*}	{GG, G*}
				{G*G, B}	
				{GG*, B}	
				{B, G*G}	
				{B, GG*}	

Notes: – *Describes the order where cohort member B, boy sibling, or cohort member G, girl sibling, is in set S_i, sibship. The *General Sample* describes families with three singleton siblings while the *Twin Sample* describes families with one singleton and two twin siblings. There are six possible combinations for *CM is only-sex* and *Control* while there are twelve possible combinations in set *Sibling is only-sex* in the *General Sample*. Comparing the same possible combinations in the *Twin Sample* these figures change to twelve and sixteen, respectively.