Family Size and Educational Attainment in England and Wales

Tak Wing Chan   Morag Henderson
Department of Social Science
UCL Institute of Education

Rachel Stuchbury
Centre for Longitudinal Study
Information and User Support, UCL

October 29, 2018

Abstract

In this paper, we use linked census data from England and Wales to investigate whether large family size leads to lower educational attainment. Using twin births and the sex composition of the sibling group as instrumental variables, the evidence of a family size effect on educational attainment is rather uncertain. Similar results are obtained when we use occupational attainment as the dependent variable. We also demonstrate the confounding of birth order and family size effects, and show that an adjusted birth order index proposed by Booth and Kee (2009) provides an effective solution to this estimation problem.

1 Background and research question

There is a well established association between family size and educational outcomes: individuals with more siblings tend to do worse in education. This finding has been so widely and so consistently replicated that, for Kuo and Hauser (1997, p. 73), it is 'inarguable that large sibships inhibit educational

1

attainment.' Moreover, the magnitude of this association is quite large. Typically, an extra sibling is associated with one-fifth of a year less schooling which, as Steelman *et al.* (2002, p. 248) point out, 'exceed[s] those of all other familial variables (e.g. parental SEI, farm background, and family "intactness") with the exception of parental education.'

Although this association is based on observational data, it is often interpreted in causal terms. But parents who have many children are likely to be different from those with fewer children. Some of these differences are unobserved and they might affect educational outcomes. In other words, the family size effects are subject to the omitted variable bias. Furthermore, parents are likely to make joint decisions about fertility and children's education. This means that treating family size as exogenously determined is problematic. For these reasons, it is unclear whether large family size really inhibits educational attainment. In this paper, we address this quesion with linked census data from England and Wales, using twin births and the sex composition of the sibling group as instruments for family size.

1.1 Resource dilution and its critics

The most widely accepted interpretation of the family size effects is resource dilution (Blake, 1989). The simple yet appealing kernel of this argument is that parents have a finite amount of resources to invest in children. With more children, parental resources will be spread more thinly, which then adversely affects children's educational attainment. In a similar vein, economists speak of a tradeoff between the quantity and quality of children (Becker and Lewis, 1973; Becker and Tomes, 1976).

Subsequent research has elaborated on how resource dilution works. For example, Steelman and Powell (1991) use data from the High School and Beyond Study to show that parents are less willing and less able to pay for children's higher education as family size increases. Downey (1995) uses data from the 1988 National Education Longitudinal Survey to explore the functional form of the family size effects for different kinds of resources.

The inverse relationship between family size and educational outcomes has long been considered as 'one of the most consistent findings in the status attainment literature' (Downey, 1995, p. 746). But the near unanimity of findings from previous generations of research has been challenged in recent years. First, scholars working with data from other countries have reported more variation in the strength and sometimes even the direction of that association. For example, Park (2008, p. 874) analyses PISA (Programme for International Student Assessment) data and shows that 'countries with stronger public support for childcare, universal child benefits, larger public

expenditures on education and family, show a much less negative effect of growing up in larger families.' Maralani (2008, p. 693) shows that 'in urban areas [in Indonesia], the association between family size and children's schooling was positive for older cohorts but negative for more recent cohorts ...rural areas show no significant association ...for any cohort.'

Clearly, public policies and local contexts matter. Public resources might to some degree substitute for private means. And cultural norms regarding how much schooling children should have, whether parents could draw on the support of the wider kinship network in child-rearing, and the extent to which there is economy of scale in raising multiple children, etc. affect the association between family size and children's schooling. Recognising this, Gibbs *et al.* (2016, p. 731) argue for a conditional resource dilution model in which 'within-family dilution processes are contingent on ...state- and community-level investments ...when the burden of parenting is broadly shared, we expect within-family dilution to be less consequential.'

A second challenge concerns whether family size really has a causal impact on educational attainment. When Guo and VanWey (1999) analyse NLSY (National Longitudinal Survey of Youth) cognitive test scores data with OLS models, they obtain the usual negative associations. But these negative associations disappear when they analyse the data with sibling models or fixed effects models. As sibling models control for unobserved factors that are shared by siblings within the family, and fixed effects models control for unobserved individual traits that are stable over time, they argue that the often reported negative association is due to unobserved confounders and is thus spurious.

These findings are provocative and suggestive. But not all scholars are (fully) persuaded. For example, Phillips (1999) points out that the analytical sample of Guo and VanWey is one with relatively young mothers. So their results might not be generalisable to the US population as a whole. Similarly, Downey *et al.* (1999) argue that Guo and VanWey's sample is one with widely spaced siblings, where the parental resources constraint is less restrictive. We note, however, that Workman (2017) has recently replicated the results of Guo and VanWey with data from the Early Childhood Longitudinal Study–Birth Cohort. With fixed effects models, he shows that '[t]he birth of a sibling was not significantly associated with lower cognitive development, even when the age spacing between the siblings was small' (Workman, 2017, p. 462).

The debate between Guo and VanWey (1999) and their critics also highlights the need to take into account other dimensions of the sibling group, in particular, birth order, birth spacing, and sex composition. We will discuss how we deal with birth spacing and sex composition in our analysis below. But regarding birth order, the findings are quite mixed. In a review essay,

3

Steelman *et al.* (2002, p. 256) note that the birth order effects on educational attainment 'are much more modest in scope than those documented for sibship size. Furthermore, the effects do not consistently veer in the same direction.' In contrast, Black *et al.* (2005, p. 680) use population register data from Norway and show that once birth order is controlled for, 'the family size effects are reduced to close to zero.' Their view is that the 'birth order effects appear to drive the observed negative association between family size and child education' (Black *et al.*, 2005, p. 685).

The difficulty here is partly due to the high correlation between family size and birth order, which makes it hard to obtain precise estimate of either variable (Iacovou, 2007). Because Black *et al.* (2005) draw on population register data, they have enough information to disentangle the two effects. With smaller data sets, this might not be possible. To deal with this problem, Booth and Kee (2009) propose an adjusted birth order index, which deflates the absolute birth order (i.e. firstborn, secondborn, and so on) by the average birth order within the family. That is to say,

$$\text{adjusted birth order} = \frac{\text{absolute birth order}}{\text{average birth order}} = \frac{\text{absolute birth order}}{(\text{family size} + 1)/2}.$$

Thus, the adjusted birth order of the first born in a family with two children is 2/3. But if there are three children, the adjusted birth order of the first born is 1/2. In the English and Welsh census data that we analyse in this paper, family size correlates with absolute birth order at $r = .68$. With the adjusted index, the correlation drops to .09. We will explore the joint impact on educational attainment of family size and birth order, using both the absolute and the adjusted indices for the latter.

1.2   Instrumental variable approach

Increasingly, scholars have used instrumental variables to determine whether family size has a causal impact on children's educational outcomes. For an instrumental variable (IV) to give consistent estimates, it needs to meet two conditions: (1) that it correlates strongly with the endogenous variable (in the present case, family size); (2) that it does not correlate with the outcome of interest (educational attainment) except through the endogenous variable. Whether these two conditions are met need to be considered carefully in each case. But as scholars in this field have used different instruments, and have applied them to data drawn from a diverse set of countries, it is perhaps unsurprising that they have come to quite different conclusions.

One of the most commonly used IVs in this literature is twin births. The idea is that if twin births are random occurrences, then they would

exogenously change family size. For example, parents planning to have one child end up with two children because of a twin birth. We have already noted that in Norway birth order appears to drive the family size effects. In the same paper, Black *et al.* (2005, p. 669) also show that with 'twin births as an instrument, family size effects become negligible.' In a study that is based on linked census and register data from Israel, Angrist *et al.* (2010) use twin births and sex composition of the sibling group as instruments. Their 'results are remarkably consistent in showing no evidence of a quantity–quality trade-off' (Angrist *et al.*, 2010, p. 773).

Other scholars, also using twin births as IV, have come to quite different conclusions. For example, Rosenzweig and Wolpin (1980) analyse survey data from India and report a tradeoff between child quantity and quality. Li *et al.* (2008, p. 223) analyse data from a 1% sample of the 1990 Chinese Census and report 'a negative effect of family size on children's education . . . the effect of family size is more evident in rural China, where the public education system is poor.' Ponczek and Souza (2012, p. 64) apply the same approach to the 1991 Brazilian Census data, and find 'negative effects [of family size] on educational outcomes for boys and girls.' In an analysis of the 1980 US Census Five-Percent Public Use Micro Sample, Cáceres-Delpiano (2006, p. 738) reports that 'a larger family generated by twins in a later birth reduces the likelihood that older children attend private school . . . [But] the impact of family size on . . . grade retention, is less clear.'

Closer inspection of these results, however, suggests a more ambiguous picture. For example, of the fourteen IV estimates of the family size effects that Li *et al.* report, only eight are negative and statistically significant, two at the 5% level and six at the 10% level (Li *et al.*, 2008, Table 3, columns 3 and 6). Even for the rural China subsample where the family size effects are expected to be stronger, only two of the seven IV estimates are negative and significant, one each at the 5% and 10% levels (Li *et al.*, 2008, Table 4, column 3). In the Brazilian case, of the forty-four IV estimates of the family size effects on educational outcomes (e.g. school attendance, grade progression, literacy, years of schooling), 12 are negative and significant at the 5% level, 4 are negative and significant at the 10% level, and the remaining 28 are not significantly different from zero (Ponczek and Souza, 2012, Tables 7 and 8). As regards Cáceres-Delpiano (2006), the results are also quite mixed: the IV estimate of the family size effects is negative and significant for attending private school in one of the two subsamples, and they are not significant for grade retention for both subsamples (Cáceres-Delpiano, 2006, Table 5).

Next to twin births, the other commonly used instrument for family size is the sex composition of the sibling group. If parents with two children prefer to have a child of each sex, then those with two boys or two girls are more

likely to have a third child than those with a boy and a girl (Angrist and Evans, 1998). As children's sex is randomly assigned (at least in countries where sex-selective abortion is rare), the sex composition of the first two children could be used as an IV for family size. The validity of this instrument depends critically on the assumption that sibling sex composition has no direct impact on children's outcome. The evidence on this is, however, unclear, with Butcher and Case (1994) and Conley (2000) reporting evidence of a direct effect, while Hauser and Kuo (1998) and Kaestner (1997) finding the opposite.

Acknowledging this issue, especially for girls, Conley and Glauber (2006) apply the sex composition IV to boys only. Based on data from the 1990 US census, they report that 'for second-born boys, increased sibship size reduces the likelihood of private school attendance by six percentage points and increases grade retention by almost one percentage point. Sibship size has no effect on first-born boys' (Conley and Glauber, 2006, p. 722).

The preference for a mixed-sex sibling group is likely to be culturally specific. In countries where there is a strong son-preference, the dynamics might be quite different. For example, Lee (2008) argues that in Korea parents are more likely to have a second child if their firstborn is female. Assuming that sex-selective abortion at first birth is rare and also that, despite the son-preference, parents treat sons and daughters equally after they were born, Lee (2008) uses the firstborn's sex as an instrument and applies it to data from the Korean Household Panel Survey. He shows that cross-sectional estimates of the family size effects are biased. But 'even after controlling for unobserved heterogeneity, a greater number of siblings have adverse effects on per-child investment in education' (Lee, 2008, p. 855). Kugler and Kumar (2017) use a similar strategy and report comparable results for India.

Overall, the various IV-based papers have yielded different results, some of which seem quite fragile. Several points are notable here. First, the diverse findings speak to the need to take local contexts seriously. It is quite possible, as Li *et al.* (2008) argue, that the quantity–quality tradeoff works differently in different countries.

Second, the family size effects might also change over time. As noted already, Black *et al.* (2005) use twin births as IV and report no family size effects on educational attainment in Norway. But when they repeat the same analyses with data for Norwegian men of more recent cohorts, they find the opposite, i.e. large family size has negative effects on IQ-score and the probability of high school graduation. They argue that this change might be related to increased female labour force participation in Norway since the 1970s, as 'extra (particularly unexpected) children may be more costly when female labor market participation is higher' (Black *et al.*, 2010, p. 52).

6

Similarly, Marteleto and de Souza (2012) document changes in the family size effects over time in Brazil. Using twin births as IV, they show that in Brazil large family size used to have *positive* effects on childen's educational attainment, especially in the poorer North and Northeast of the country, but these effects have disappeared since the 1990s.

Third, model specification matters. As Black *et al.* (2005, p. 676) point out, not all studies control for birth order, and so the reported 'family size effects could be confounded by the omission of birth order controls.'

Fourth, the nature of the data matters. Studies that are based on a single census are essentially cross-sectional in nature. Scholars using this type of data do not have information about the children's eventual educational attainment. So they use intermediate indicators, e.g. type of school attended or grade retention, as dependent variables instead. Whilst intermediate outcomes provide useful information on the mechanisms of the family size effect, they should not be equated with educational attainment. Furthermore, single-census data probably contain more measurement error in some key variables such as family size, as younger siblings might be born after the census, while some older siblings might have already left the household before the census.

Finally, the fragility of some of the findings highlights the need to consider carefully the validity of the instruments. Indeed, scholars using different instruments on the same data set have drawn diametrically opposite conclusions (see e.g. De Haan, 2010; Jaeger, 2008). We will discuss the main instrument that we use in this paper, i.e. twin births, after introducing the data.

2   Data

The data that we analyse come from the 'Office for National Statistics Longitudinal Study' (LS in short). LS links census and life events data (i.e. registration of births, cancers and deaths) for approximately one per cent of the population of England and Wales. The original LS sample consists of all individuals enumerated in the 1971 census who were born on one of four selected dates in a calendar year. At each subsequent census (and between censuses) individuals born on those four dates are added to the LS sample. The data available include not only information about LS members on the census form, but also equivalent detail for every member of their household, although only LS members themselves are linked across censuses.

We select LS members aged 4–7 in the 1971, 1981 and 1991 censuses, forming three cohorts (see Figure 1). The data from the 'baseline censuses'
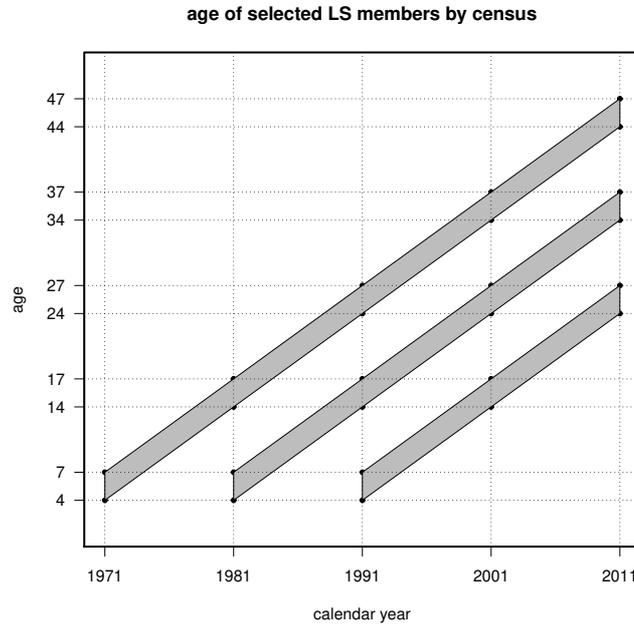
Figure 1: The three cohorts of LS members

are linked respectively to the 1981, 1991, 2001 censuses when LS members were aged 14–17 ('baseline plus 10'), and then further to the 1991, 2001, 2011 censuses when they were aged 24–27 ('baseline plus 20'). For LS members of the 1971 (1981) cohort, we also have information about them up to their mid-40s (mid-30s). At the baseline census, we set the lower age limit at 4 because LS members would be at least 24 year old at 'baseline plus 20'. Most individuals would have completed their education by that age. The upper age limit is set at 7 because we want to minimise the chance of LS members having older siblings who had already left home at the baseline census.

From the baseline censuses, we know the following about LS members: sex, date of birth, region of residence, mother's age when LS members were born, parent's educational attainment and housing tenure. In cases where we have information on the educational attainment of both parents, we take the higher of the two. We also have data on parental social class. But we have not included social class in the analysis because of a relatively high level of missing data. Moreover, adding social class to the models does not change the substantive results. Using data from the baseline census and the baseline-plus-10 census, we are able to determine whether LS members lived with both parents up to their mid-teens (which is our measure of intact family), how many siblings they have, their birth order, and the sex composition of their

Table 1: Distribution of LS members by number of siblings (column %)

	1971	1981	1991	overall
0	6.47	8.00	8.75	7.64
1	34.99	47.41	43.75	41.40
2	29.34	27.27	29.43	28.76
3	15.99	11.04	11.84	13.23
4	7.10	3.70	3.72	5.04
5	3.30	1.44	1.52	2.19
6	1.54	0.62	0.55	0.96
7	0.70	0.29	0.25	0.43
8+	1.25	0.50	0.44	0.78
N	32,063	24,060	25,922	82,045

sibling group.

We use the household roster on the census form to identify the siblings (and other relations) of LS members. This approach is, of course, not error-free. For example, we might miss siblings who are very much older or very much younger than LS members. To gauge the degree of measurement error of the key variable of family size, we compare it with the response to a direct question on *marital fertility* that was put to women in the 1971 census. Of course, the latter is also an imperfect measure, but the two measures correlate very highly with $r = .91$.

Crucially, using the exact date of birth of siblings, we are also able to determine whether LS members are part of a twin, and whether there were any twins among their younger siblings. We use the latter as the main instrumental variable. Finally, from the baseline-plus-20 census, we know whether LS members have obtained a Bachelor degree by their mid-20s, which is our main dependent variable. In a set of supplementary analyses that is restricted to the first two cohorts, we also use LS members' social class position when they are in their mid-30s as the dependent variable.

Table 1 reports the distribution of LS members by family size and cohorts. Reflecting the trend of declining fertility in this period, we see a higher share of LS members have no sibling or just one sibling in the 1981 or 1991 cohort compared to the 1971 cohort. But even for the 1971 cohort, 87% of its members have no more than three siblings. Table 2 provides some descriptive statistics of our data.

Table 2: Descriptive statistics (percentages*)

	1971	1981	1991	overall
degree	32.0	43.3	47.3	39.7
home owner	53.7	62.6	67.3	60.6
social tenant	32.3	30.8	26.5	30.0
private tenant	14.1	6.6	6.3	9.4
parent edu (none)	82.7	80.6	77.3	80.4
parent edu (inter)	12.0	8.7	10.0	10.4
parent edu (degree)	5.3	10.7	12.7	9.3
female	49.1	49.0	49.0	49.0
intact family	63.9	59.7	50.1	58.3
twin births	3.6	2.9	3.2	3.3
LSM is part of twin	1.9	1.7	1.9	1.9
mother's age† (mean)	26.8	26.2	26.9	26.7
mother's age (sd)	5.3	4.8	5.0	5.0
siblings (mean)	2.1	1.6	1.7	1.8
siblings (sd)	1.4	1.2	1.2	1.3
birth order (mean)	2.1	1.9	1.8	2.0
birth order (sd)	1.3	1.0	1.0	1.1

Note: * except for mother's age, number of siblings and birth order;
† mother's age when LS member was born.

3 Twin births as instrumental variable

The main instrumental variable for family size that we use is twin births. For this instrument to give consistent estimates, we need to assume that twin births are unrelated to children's educational attainment except through family size. Given this consideration, we exclude from our analyses those LS members who are themselves part of a twin. This is because twins tend to have lower birth weight and are generally less healthy at birth which, in turn, are related to lower educational attainment and earnings as adults (Case *et al.*, 2005; Black *et al.*, 2007; Rosenzweig and Zhang, 2009).

Instead, following Black *et al.* (2005, 2010), Angrist *et al.* (2010), and other scholars in the field, our strategy is to consider the impact of an exogeneous change in family size due to the birth of younger twin siblings at the nth birth ($n = 2, 3, 4$) on the outcomes of LS members (who were born before the nth birth and are themselves singletons). As Black *et al.* (2010, pp. 37–38) argue, this approach offers several advantages. First, it ensures 'that, on average, preferences over family size are the same in the families with twins at the nth birth and those with singleton births.' Second, it avoids the 'selection problems that arise because families who choose to have another child after a twin birth may differ from families who choose to have another child after a singleton birth.' Third, it bypasses the issue that 'families that have more births are more likely to have at least one twin birth.' Fourth, it sidesteps the difficulty that 'twin birth both increases family size and shifts downward the birth order of children born after the twins.' Finally, we exclude LS members who are themselves twins because their birth order is ambiguous (Black *et al.*, 2005).

Other than changing family size, how might having younger twin siblings affect LS members' educational outcomes? There are plausible *a priori* arguments that give opposite predictions. For example, parents might invest more on the education of the less healthy twins at the expense of their healthier non-twin siblings. Such a compensatory strategy would adversely affect LS members. Alternatively, parents might reinforce the difference in endowments at birth by investing more on the education of the healthier singletons at the expense of the less healthy twins. Under this scenario, LS members with younger twin siblings would, on average, do better (Rosenzweig and Zhang, 2009). The reinforcing strategy does not necessarily imply that parents care less about children with lower endowment. Rather, it suggests that parents could invest in their children in different ways. For example, faced with a child that is less likely to do well in school, parents could save what they would have spent on the child's education, invest the sum in other ways, and pass on those non-educational investment to the child later on

(see e.g. Mulligan, 1997). Whether parents in England and Wales follow the compensatory or the reinforcing strategy is of course a matter for empirical examination, though we note that in Norway the empirical evidence is 'consistent with compensating rather than reinforcing parental investments' (Black *et al.*, 2010, p. 35).

In one respect, however, the impact of twin births on the sibling group is unequivocal. Because there is, by definition, no birth gap between twins, twin births would unambiguously reduce the average birth spacing of the sibling group. And since closer birth spacing probably implies a resource crunch for parents, having younger twin-siblings would, through the mechanism of reduced birth spacing, adversely affect LS members. Thus, although we do not directly control for birth spacing in our analysis, our strategy is biased towards finding negative family size effects (Black *et al.*, 2005, pp. 682–683).

A further issue to consider is the increasing use of assisted reproductive technologies (ART). Twin births are much more likely in pregnancies conceived with ART. The UK's Human Fertilisation and Embrology Authority (2015, p. 2) reports that 'in 2008 ... almost a quarter of births resulting from IVF treatment were of two or more babies.' By comparison, the overall rate of twin births in recent years is about 15 per 1,000 (see top-left panel of Figure 3 in Appendix A). If, for example, a significant proportion of couples receiving fertility treatment are career-minded people who have postponed childbearing to the point when they have difficulty with natural conception; and if we further assume that career-minded couples invest more in their children's education, then twin births would be directly related to educational outcomes. To the extent that this is the case, our use of twin births as an instrument for family size is problematic. Given this concern, we review the trends of twin births and of the use of ART in the UK in Appendix A. The upshot of that discussion is that the increasing use of ART probably does not affect the results reported in this paper.

4 Results

4.1 Probit analysis

We begin by fitting to our data a set of probit models, the dependent variable of which is whether the LS member has a Bachelor degree (or higher) at baseline-plus-20. We want to show that our synthetic data set yields the same results as survey data. We also want to demonstrate the confounding of the family size effects and the birth order effects, and how the adjusted birth order index overcomes this problem.

Because of the large sample size, all but one parameters of Table 3 are statistically significant. Furthermore, the estimates are all in the expected direction. For example, reflecting the expansion of higher education, LS members from the later cohorts are more likely to have a Bachelor degree, as are those born later within any cohort. Also, women are more likely than men to be university graduates. The same is true for those born to older mothers or from intact families. LS members with better educated or home-owning parents are also more likely to have a degree. The parameters of interest are a set of dummies contrasting having 1 to 7+ siblings against the reference cateogory of having no sibling. These sibling parameters are all negative in sign and, with the exceptions of the last two, monotonic in magnitude. This result is consistent with the resource dilution argument.

In model 2, we replace the seven sibling dummies with a single linear measure of number of siblings. As this has almost no impact on other parameters or on model fit, in the interest of parsimony, we will use this linear measure in the models that follow. In model 3 we further control for the birth order of LS member. Again, this has little impact on most other parameters. But the estimate for number of siblings changes from -.075 to -.0009 (and quite clearly not statistically significant). This is consistent with Black *et al.* (2005) who show that in Norway when birth order is controlled for, the substantive magnitude of the family size effect becomes much smaller. In model 4, we follow Booth and Kee (2009) and replace absolute birth order with the adjusted birth order index. Under model 4, the estimate for number of siblings is very similar to that of model 2, suggesting that the adjusted birth order index provides an effective way to consider the joint impact of birth order and family size. We will be using the adjusted birth order index in subsequent analyses.

Figure 2 reports the substantive magnitude of selected covariates under model 4. This shows the predicted probabilities, with 95% confidence intervals, of LS members having a degree for different combinations of family size and parental education (top panel), of family size and gender (middle panel), and of family size and intact family status (bottom panel). Other covariates take on their actual values. It can be seen that, in the range of 0 to 3 siblings, the family size effect is fairly linear. Moreover, parental education is, by far, the most important predictor of educational attainment. The probability of having a Bachelor degree among those with parents who are themselves university graduates is about 40% higher than those whose parents have no qualifications. The family size effect, between those with no sibling and 3 siblings is about 9%, which is larger than the size of the gender gap and the gap between those coming from intact and non-intact families (both about 4%). Overall, these results are consistent with previous studies,

13

Table 3: Probit regressions of family size on educational attainment ($N = 69,456$)

	model 1		model 2		model 3		model 4	
	β	s.e.	β	s.e.	β	s.e.	β	s.e.
1981 cohort	.225	.012	.226	.012	.231	.012	.236	.012
1991 cohort	.261	.012	.262	.012	.252	.012	.255	.012
age within coh	$-.014$	.004	$-.014$	.004	$-.013$	.004	$-.013$	.004
female	.133	.010	.133	.010	.135	.010	.135	.010
mother's age[†]	.016	.001	.016	.001	.030	.001	.031	.001
intact family	.567	.016	.567	.016	.551	.016	.550	.016
intermediate	1.187	.020	1.186	.020	1.152	.020	1.150	.020
degree	.137	.010	.138	.010	.129	.010	.129	.010
social tenant	$-.566$	.012	$-.567$	.012	$-.545$	.012	$-.545$	.012
private tenant	$-.253$	.018	$-.253$	.018	$-.246$	.018	$-.246$	.018
1 sibling	$-.062$	.019						
2 siblings	$-.176$	.020						
3 siblings	$-.259$	.023						
4 siblings	$-.352$	.030						
5 siblings	$-.389$	.042						
6 siblings	$-.636$	.066						
7+ siblings	$-.461$	.069						
# sibling			$-.089$	.004	$-.000*$	.006	$-.086$	.004
birth order					$-.160$	.007		
adj birth order							$-.377$	.016
constant	$-.854$	.042	$-.750$	.039	-1.045	.042	$-.769$	.039
R^2	.138		.138		.142		.143	
Log likelihood	-40272.43		-40279.72		-40057.24		-40011.94	

Note: region dummies included in the models but not shown; * *not* statistically significant at the conventional 5% level, all other parameters are statistically significant; [†] mother's age when LS member was born.
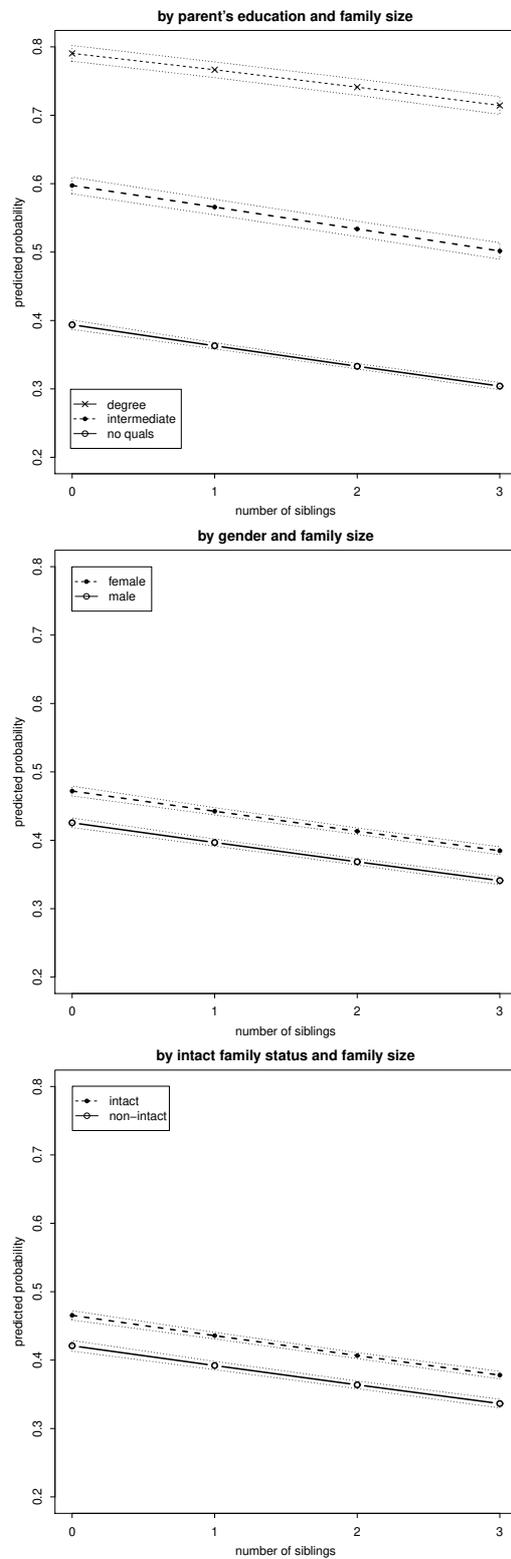
Figure 2: Predicted probability of LS members obtaining a degree under model 4

confirming the family size effects that are found with conventional analyses. We have repeated the analyses of this section with linear probability models. The results are qualitatively very similar to the probit models of Table 3.

4.2   IV-probit analysis

Our main instrument for family size is twin births. Following Black *et al.* (2005, 2010), we apply this IV to three subsamples of the LS data. In the first subsample, LS members are first-born children in families with at least two births. The twin birth, if there was one, happened at the second birth. In the second subsample, LS members are the first or second-born children in families with at least three births, and the twin birth, if it happened, was the third birth. In the third subsample, LS members are the first, second or third-born children in families with at least four births, and the twin birth, if there was one, was the fourth birth. The size of these subsamples are substantial, with $N = 23{,}175$, 18,145, and 8,582 respectively. But the number of twin births are still relatively small, being 102, 75 and 24 respectively.

In column 1 of Table 4, we first show that for all three subsamples a probit analysis yields negative and statistically significant estimate for number of siblings, and for adjusted birth order (included in the second and third subsamples only). These results are consistent with those reported in Table 3 and with the resource dilution argument.

We then fit IV-probit models to the data. For all three subsamples, twin birth turns out to be a strong predictor of family size in the stage 1 regression. So there is no concern for weak instrument. At stage 2, our IV estimates for sibship size are very far from statistically significant in all three subsamples, while adjusted birth order remains negative and significant for the second and third subsamples. Thus, our result is consistent with Black *et al.* (2005) and Angrist *et al.* (2010) in showing no family size effect. That is to say, LS members who have more siblings than their parents had expected because of the birth of younger twin siblings are *not* less likely to obtain a Bachelor degree.

But it should be noted that, as is common with IVs (see e.g. Black *et al.*, 2005; Angrist *et al.*, 2010), the standard error of the IV-probit estimates are much larger than the probit estimates (see Sargan, 1958, for a discussion of the inefficiency of IV estimates). For example, in the first subsample, the IV-probit standard error is .087 compared to that for the probit estimate of .011. Thus, the statistical insignificance of the IV-probit estimate might in part be attributed to its inefficiency which, in turn, is due to the relatively small number of twin births, even for the LS data. Indeed, using a Wald test of exogeneity, we cannot reject the hypothesis of no endogeneity (sample

16

Table 4: Probit and IV-probit regressions of family size on educational attainment in three sub-samples

| | probit | | iv-probit | | | | |
| | | | first stage | | second stage | | |
	β	s.e.	β	s.e.	β	s.e.	N
Subsample 1: first born child in families with 2+ births							
# sibling	$-.039{**}$	.011			$-.085$	.087	$23,175$
IV: twin at 2nd birth			$1.637{***}$	.080			
Subsample 2: first or second born children in families with 3+ births							
# sibling	$-.120{***}$	.015			$-.033$	.117	$18,145$
adj birth order	$-.502{***}$	.046			$-.418{**}$	.121	
IV: twin at 3rd birth			$1.567{***}$	.085			
Subsample 3: first, second or third born children in families with 4+ births							
# sibling	$-.094{***}$	.022			.152	.179	$8,582$
adj birth order	$-.441{***}$	.058			$-.281{*}$	.132	
IV: twin at 4th birth			$1.697{***}$	.155			

Note: * $p < .05$, ** $p < .01$, ** $p < .001$; all other covariates of Table 3 are included in the models but not reported here.

1: $\chi^2(1) = .29, p = .59$; sample 2: $\chi^2(1) = .56, p = .45$; sample 3: $\chi^2(1) = 1.86, p = .17$). So while there is no evidence supporting the resource dilution argument, we need to be cautious in our interpretation of the results.

Table 5 reports the results of several further tests. Since the resource dilution argument is more likely to hold for families with less resource, we restrict our analyses to subsamples of LS members whose parents have no educational qualification. Panel A of Table 5 shows that with these subsamples we obtain broadly the same results as before. That is to say, probit analyses suggest that having more siblings is associated with a lower probability of obtaining a Bachelor degree. But with twin births as instrument, sibship size no longer predicts whether LS members are university graduates.

In Panel B, we consider whether LS members were in professional or managerial occupations (i.e. whether they are in the salariat social class) when they reached mid-30s. Since LS members of the third cohort were in their mid-20s when they were last enumerated in the 2011 census, it can be argued that they have not yet reached occupational maturity (Bukodi and Goldthorpe, 2011). For this reason, we drop the youngest cohort from this set of analysis. But all LS members from the first two cohorts are included irrespective of parental educational level. The results of Panel B are again comparable to those reported earlier. Probit analyses suggest that

Table 5: Further probit and IV-probit regressions of family size. Panel A: restricted to parents with no qualifications. Panel B: using occupational attainment as dependent variable. Panel C: using sex composition of sibling group as instrument.

| | probit | | iv-probit | | | | |
| | | | first stage | | second stage | | |
	β	s.e.	β	s.e.	β	s.e.	N
Panel A: LS members with parents having no qualifications							
Subsample 1: first child in families with 2+ births							
# sibling	−.042**	.012			−.039	.096	18,034
IV: twin at 2nd birth			1.622***	.090			
Subsample 2: first and second children in families with 3+ births							
# sibling	−.119***	.016			−.107	.126	14,687
adj birth order	−.485***	.051			−.472**	.138	
IV: twin at 3rd birth			1.575***	.095			
Subsample 3: first, second and third children in families with 4+ births							
# sibling	−.093***	.023			.149	.183	7,324
adj birth order	−.412***	.063			−.244	.144	
IV: twin at 4th birth			1.695***	.162			
Panel B: LS members of the first and second cohorts							
Membership in the salariat in the mid-30s as dependent variable							
Subsample 1: first child in families with 2+ births							
# sibling	−.073***	.014			.010	.102	13,828
IV: twin at 2nd birth			1.653***	.099			
Subsample 2: first and second children in families with 3+ births							
# sibling	−.109***	.019			−.116	.144	10,756
adj birth order	−.339***	.058			−.346*	.147	
IV: twin at 3rd birth			1.642***	.112			
Subsample 3: first, second and third children in families with 4+ births							
# sibling	−.106***	.027			.268	.259	5,184
adj birth order	−.198**	.072			−.047	.186	
IV: twin at 4th birth			1.406***	.212			
Panel C: LS members are first or second child in families with 3+ births							
Sex composition of sibling group as instrument							
# sibling	−.115***	.015			.061	1.046	18,164
adj birth order	−.501***	.046			−.328	1.045	
IV: sex composition			.019	.010			

Note: * $p < .05$, ** $p < .01$, ** $p < .001$; all other covariates of Table 3 are included in the models but not reported here.

LS members with more siblings are less likely to be members of the salariat. But when we use twin births as instrument, family size no longer predicts occupational attainment.

In Panel C, we use sex composition of the sibling group as IV for family size. Here we consider those LS members who are either first or second born children in families with at least two births. The dependent variable is whether they obtain a Bachelor degree by their mid-20s. The results are again consistent with those reported above. In the probit model, number of siblings is a highly significant predictor of educational attainment. With IV, LS members in single-sex sibling group are marginally more likely to have a third sibling (the sex composition parameter in the first stage regression is significant at $p = .056$). In the second-stage regression, family size, instrumented by the sex composition of the sibling group, does not predict educational attainment.

5 Summary and discussion

In this paper, we use linked census data from England and Wales to test whether family size has a causal impact on children's educational attainment. To the best of our knowledge, this is the first paper which uses LS data to address this question. We show that there is indeed a negative association between family size and educational outcome. Individuals with more siblings are less likely to obtain a Bachelor degree. Consistent with Black *et al.* (2005), we show that this association disappears when birth order is controlled for. This is, however, due to the relatively high correlation between sibship size and birth order. When the adjusted birth order index (Booth and Kee, 2009) is used, we again obtain the negative association between family size and education.

The main finding of this paper is that, when we use twin births as IV, the evidence for the family size effect is rather uncertain. The IV estimates are not significantly different from zero. But this is related to their larger standard errors. This is true whether we analyse the whole sample or restrict the analyses to LS members from the least advantaged background (i.e. those with parents with no educational qualification). We obtain the same result when the dependent variable is salariat class membership (i.e. having a professional or managerial occupation). Likewise, the evidence for the family size effect is ambiguous when we use the sex composition of the sibling group as IV although, in this case, the instrument is rather weak.

Given the uncertain nature of the evidence for a family size effect, we should consider the possibility that parents who have more children are dif-

ferent from those with fewer children, and it is these unobserved differences that might account for the association between family size and children outcomes. It is beyond the scope of this paper to determine what those unobserved differences are. But, for example, it is possible that a higher share of those parents with a large number of children let themselves drift into parenthood (Sawhill, 2014). Using data from the third National Survey of Sexual Attitudes and Lifestyles, Wellings *et al.* (2013) report that 16.2% of pregnancies in the UK are unplanned, a further 29% are 'ambivalent', and only 54.8% of pregnancies are planned. They also show that unplanned pregnancies are associated with, among other things, drug use, depression and low educational attainment. As many of these factors are not (fully) controlled for in research on family size effects, it seems plausible that it is factors such as these which might contribute to the observed association between family size and educational attainment in England and Wales.

References

Angrist, J. D. and Evans, W. N. (1998). Children and their parents' labor supply: evidence from exogenous variation in family size. *American Economic Review*, **88**(3), 450–477.

Angrist, J. D., Lavy, V., and Schlosser, A. (2010). Multiple experiements for the causal link between quantity and quality of children. *Journal of Labor Economics*, **28**(4), 773–823.

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

Becker, G. S. and Tomes, N. (1976). Child endowment and the quantity and quality of children. *Journal of Political Economy*, **84**, S143–S162.

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

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2007). From the cradle to the labor market: the effect of birth weight on adult outcomes. *Quarterly Journal of Economics*, **122**(1), 409–439.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2010). Small family, smart family: family size and the IQ scores of young men. *Journal of Human Resources*, **45**(1), 33–58.

Blake, J. (1989). *Family Size and Achievement*. University of California Press, Berkeley, CA.

Booth, A. L. and Kee, H. J. (2009). Birth order matters: the effect of family size and birth order on educational attainment. *Journal of Population Economics*, **22**(2), 367–397.

Braakmann, N. and Wildman, J. (2016). Reconsidering the effect of family size on labour supply: the twin problems of the twin birth instrument. *Journal of the Royal Statistical Society, Series A*.

Bukodi, E. and Goldthorpe, J. H. (2011). Class origins, education and occupational attainment in Britain. *European Societies*, **13**(3), 347–375.

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

Cáceres-Delpiano, J. (2006). The impacts of family size on investment in child quality. *Journal of Human Resources*, **14**(4), 738–754.

Case, A., Fertig, A., and Paxson, C. (2005). The lasting impact of childhood health and circumstance. *Journal of Health Economics*, **24**, 365–389.

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

Conley, D. and Glauber, R. (2006). Parental educational investment and children's academic risk: estimates of the impact of sibship size and birth order from exogenous variation in fertilty. *Journal of Human Resources*, **14**(4), 722–737.

De Haan, M. (2010). Birth order, family size and educational attainment. *Economics of Education Review*, **29**, 576–588.

Downey, D. B. (1995). When bigger is not better: family size, parental resources, and children's educational performance. *American Journal of Sociology*, **60**, 746–761.

Downey, D. B., Powell, B., Steelman, L. C., and Pribesh, S. (1999). Much ado about siblings: change models, sibship size, and intellectual development: comment on Guo and VanWey. *American Sociological Review*, **64**(2), 193–198.

21

Gibbs, B. G., Workman, J., and Downey, D. B. (2016). The (conditional) resource dilution model: state- and community-level modificiations. *Demography*, **53**, 723–748.

Guo, G. and VanWey, L. K. (1999). Sibship size and intellectual development: is the relationship causal? *American Sociological Review*, **64**(2), 169–187.

Hauser, R. M. and Kuo, H.-H. D. (1998). Does the gender composition of sibships affect women's educational attainment? *Journal of Human Resources*, **33**(3), 644–657.

Human Fertilisation and Embrology Authority (2015). Improving outcomes for fertility patients: multiple births. Technical report, Human Fertilisation and Embryology Authority.

Iacovou, M. (2007). Family size, birth order, and educational attainment. *Marriage and Family Review*, **42**(3), 35–57.

Jaeger, M. M. (2008). Do large sibships really lead to lower educational attainment? New evidence from quasi-experimental variation in couples' reproductive capacity. *Acta Sociologica*, **51**(3), 217–235.

Kaestner, R. (1997). Are brothers really better? Sibling sex composition and educational achievement revisited. *Journal of Human Resources*, **32**(2), 250–284.

Kugler, A. D. and Kumar, S. (2017). Preference for boys, family size, and educational attainment in India. *Demography*, **54**, 835–859.

Kuo, H.-H. D. and Hauser, R. M. (1997). How does size of sibship matter? family configuration and family effects on educational attainment. *Social Science Research*, **26**(1), 69–94.

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

Li, H., Zhang, J., and Zhu, Y. (2008). The quantity–quality tradeoff of children in a developing country: identification using Chinese twins. *Demography*, **45**(1), 223–243.

Maralani, V. (2008). The changing relationship between family size and educational attainment over the course of economic development: evidence from Indonesia. *Demography*, **45**(3), 693–717.

Marteleto, L. J. and de Souza, L. R. (2012). The changing impact of family size on adolescents' schooling: assessing the exogenous variation in fertility using twins in Brazil. *Demography*, **49**, 1453–1477.

Martin, J. A., Hamilton, B. E., and Osterman, M. J. K. (2012). Three decades of twin births in the United States. NCHS Data Brief 80, U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.

Mulligan, C. B. (1997). *Parental Priorities and Economic Inequality*. University of, Chicago.

Park, H. (2008). Public policy and the effect of sibship size on educational attainment: a comparative study of 20 countries. *Social Science Research*, **37**, 874–887.

Phillips, M. (1999). Sibship size and academic achievement: what we now know and what we still need to know: comment on Guo and VanWey. *American Sociological Review*, **64**(2), 188–192.

Ponczek, V. and Souza, A. P. (2012). New evidence of the causal effect of family size on child quality in a developing country. *Journal of Human Resources*, **47**(1), 64–106.

Rogers, E. M. (2003). *Diffusion of innovations*. Free Press, New York, fifth edition.

Rosenzweig, M. R. and Wolpin, K. I. (1980). Testing the quantity–quality tradeoff fertility model: the use of twins as a natural experiment. *Econometrica*, **48**(1), 227–240.

Rosenzweig, M. R. and Zhang, J. (2009). Do population control policies induce more human capital investment? Twins, birth weight and China's "one-child" policy. *Review of Economic Studies*, **76**, 1149–1174.

Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica*, **26**(3), 393–415.

Sawhill, I. V. (2014). *Generation Unbound: Drifting Into Sex and Parenthood*. Brookings Institute Press, Washington DC.

Smith, L. K., Manktelow, B. N., Draper, E. S., Boyle, E. M., Johnson, S. J., and Field, D. J. (2014). Trends in the incidence and mortality of multiple births by socioeconomic deprivation and maternal age in England: population-based cohort study. *BMJ Open*, **4**, e004514.

Steelman, L. C. and Powell, B. (1991). Sponsoring the next generation: parental willingness to pay for higher education. *American Journal of Sociology*, **96**(6), 1505–1529.

Steelman, L. C., Powell, B., Werum, R., and Carter, S. (2002). Reconsidering the effects of sibling configuration: recent advances and challenges. *Annual Review of Sociology*, **28**, 243–269.

Warnock, M. (1984). Report of the Committee of Inquiry into Human Fertilisation and Embryology. Technical report, Department of Health and Social Security, London.

Wellings, K., Jones, K. G., Mercer, C. H., Tanton, C., Clifton, S., Datta, J., Copas, A. J., Erens, B., Gibson, L. J., Macdowall, W., Sonnenberg, P., Phelps, A., and Johnson, A. M. (2013). The prevalence of unplanned pregnancy and associated factors in Britain: findings from the third National Survey of Sexual Attitudes and Lifestyles (Natsal-3). *The Lancet*, **382**, 1807–1816.

Workman, J. (2017). Sibling additions, resource dilution, and cognitive development during early childhood. *Journal of Marriage and Family*, **79**(2), 462–474.

A Assisted reproductive technologies in the UK and their implications for the family size effect

The top-left panel of Figure 3 shows that the rate of twin births fell between the early 1950s and the mid-1970s, followed by a sharp rebound. The overall shape of this trend echoes the trend of maternal age over the same period (see the top-right panel). As maternal age is a key predictor of twin births (Smith *et al.*, 2014), this suggests that the fluctuation in the twin births rate is driven to a significant degree by changing maternal age over time. This is borne out by the bottom-left panel in which we plot the trends of twin birth rates for mothers of different age groups. Three points are notable here. First, by and large, older mothers are more likely to have twins, which is consistent with the literature (Martin *et al.*, 2012). Second, at least for mothers younger than 30, twinning rates have stayed relatively stable over time. Third, there is a sharp rise in the rate of twin births since the mid-1980s, especially for mothers aged 30 or over. This is probably due to increasing use of ART

24

among older mothers. A similar pattern is observed for the US, where about one third of the rise of twin births since the 1980s can be attributed to rising maternal age, and the rest is due to ART (Martin *et al.*, 2012).

Regarding the use of ART in the UK, we have a time series and two further data points. The first data point refers simply to the fact that the world's first 'test-tube baby', Louise Brown, was born in Oldham in Northern England in 1978. The time series come from the Human Fertilisation and Embrology Authority (2015) which reports that between 1992 and 2010 the number of IVF/ICSI patients more than tripled, rising from 14,061 to 48,147. For the period between 1979 and 1991, however, we know of just one data point from the 1984 Warnock Report, which states that '[i]n 1983 there were 967 laparoscopies performed for 579 women' (Warnock, 1984, p.34). These data are plotted in the bottom-right panel of Figure 3. Although there is very little information for the 1980s, if the adoption of ART follows the typical S-shaped curve of the diffusion of innovation (Rogers, 2003), it seems likely that relatively few people received fertility treatment for much of this decade. It seems that there are two S-shaped curves in the bottom-right panel of Figure 3. The first of these levelled off in the late 1990s; and we are still in the steep rising part of the second curve at the time of writing. Given that our LS members were born in 1964–67, 1974–77 and 1984–87 respectively, the implications of ART for our use of twin births as IV seem quite small, at least for the first two cohorts. It is also relevant to note that Braakmann and Wildman (2016) use data from the Millennium Cohort Study to assess the implications of ART for using twin births as an IV for family size. Their conclusion is that ART do introduce some bias for children born in the UK in 2000–01, but the magnitude of this bias is small.
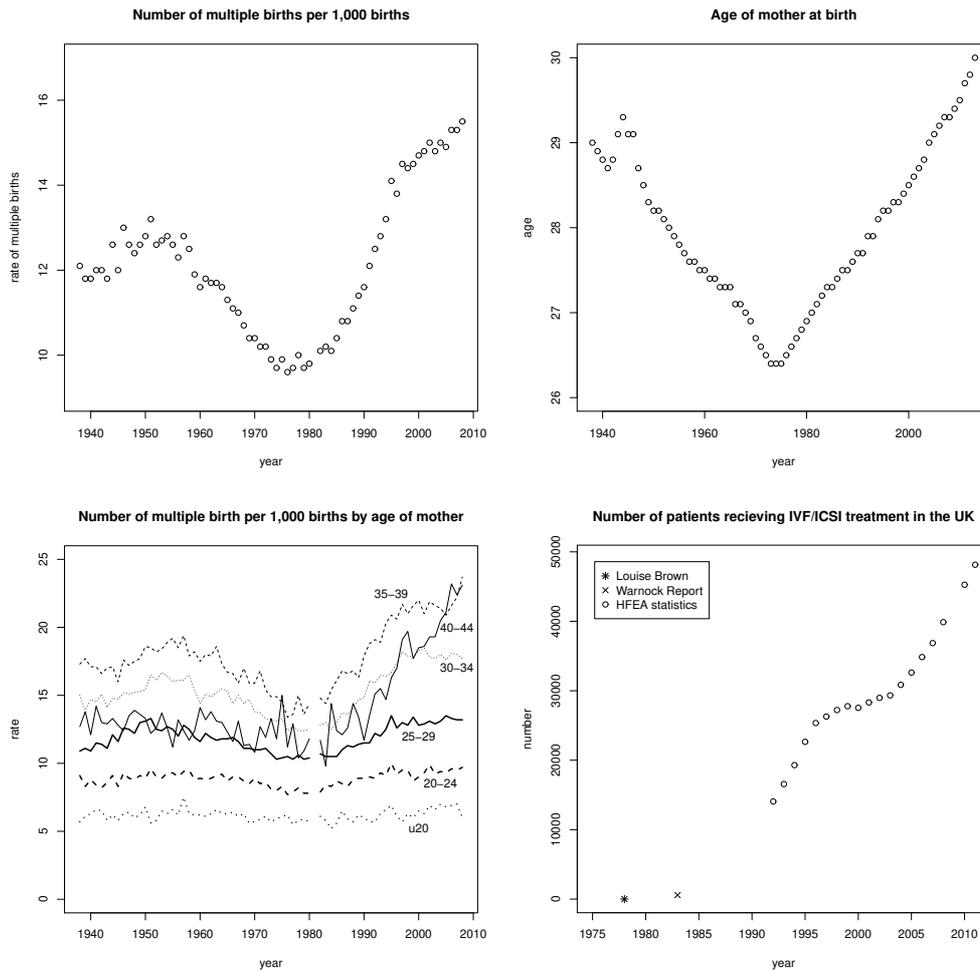
Source: ONS Birth Statistics, Series FM1, various years; Human Fertilisation and Embrology Authority, statistics, various years.

Figure 3: Trends of multiple births and assisted reproductive technologies in the UK