

Data and Methods

The data comes from the 10% public use sample of the 2011 South African Census. The micro data at the individual level has observations for each person within a household from all nine provinces of South Africa. I retained only sons or daughters in each household whose biological parents were alive by the time of the census and matched them with data on their mothers. After identifying the order of birth among two or more siblings using their dates of birth, I selected those observations where the age of the firstborn child is less than or equal to 18. Finally, two groups of samples were used in the final analysis. The next subsection discusses these in detail.

The 2+ and 3+ samples

The 2+ sample consists of only firstborn children with one or more siblings (so the total number of children per mother would be two or more). I selected only firstborns aged 8 to 18.¹ The twins IV (Twins2) in this sample is an indicator for whether the second birth was a multiple birth or not. But I have excluded firstborns who are twins themselves because of the special characteristics of twins that could confound the results. Hence, all firstborns in this sample are singletons. Similarly, the same sex IV (SameSex12) is an indicator for whether the sex of the first and second births are the same or not. In order to look at the effect by gender, I disaggregated this into two indicators: whether the first two births were both boys (Boy12) or both girls (Girl12).²

The first two columns of [Table 1](#) present summary statistics on the variables in the 2+ sample. The sample consists of almost 95,000 firstborns and has a balanced sex composition. The average number of kids is 2.7 and the mean age of the firstborns is 13. About three-quarters of the observations belong to Black African mothers, while coloured and white mothers each account for approximately 11%. The rate of occurrence of twins in the second birth is 1%. On the other hand, about 52% had a sibling of the same gender from second birth. Given that 50% of the sample are males, we have an almost equal proportion of matching first two births for each gender: 26.1% for boys and 25.5% for girls.

The 3+ sample is a subsample of the 2+ sample, where both non-twin first and second-born kids with one or more siblings are included. Here also, I selected children aged 8 to 18 since I can only observe variations in schooling outcomes for kids aged 8 or older. The twins IV in this sample (Twins3) is a dummy for whether the third birth was a twin birth or not. The same sex IV, on the other hand, is an indicator for whether the first three kids were of the same sex. This is termed SameSex123. As in the 2+ sample, I have disaggregated this indicator into two indicators based on the sex of children in the first three births. The Boy123 and Girl123 are indicators for whether the first three children are all boys or all girls, respectively. Adding the two will give SameSex123.

Columns 3-6 of [Table 1](#) presents summary statistics on the 3+ sample. Columns 3 and 4

¹South Africa has a compulsory schooling law where children who have turned 6 by June 30 are required to attend grade 1 (education.gov.za, 2021). To look at 8 year olds and above would allow me to observe variations in educational achievement and grade retention.

²The sum of Boy12 and Girl12 gives the SameSex12 dummy, i.e., $\text{SameSex12} = \text{Boy12} + \text{Girl12}$.

focus only on firstborns while columns 5 and 6 include both first and second-borns. There is no substantial difference across many of the variables between the firstborns in the 3+ sample and their counterparts in the 2+ sample. A similar result follows when looking at the entire 3+ sample (i.e., both first and second-borns; $N = 57,358$). The average family size in this sample is 3.6 and the average age of the children is 13. Just as in the 2+ sample, we have an almost perfectly balanced sex composition. The twin rate at the third birth, 1.3%, is a little higher than in the 2+ sample. About 27% of the observations are in households where the first three births are of the same sex. Again, this is well balanced across the two sexes (about 13.5% each). Black African mothers now account for a larger proportion (82%), while the share of white mothers declined to 6.1%. One can also see that the share of mothers with no schooling is higher in the 3+ sample compared to the 2+ sample (8% vs. 5%).

Model Specification and Outcome Variables

The structural equation that captures the causal relationship of interest is

$$y_i = X_i\beta + \gamma n_i + \epsilon_i \quad (1)$$

where y_i is an outcome variable, X_i is a vector of covariates including characteristics of the children and their mothers, and n_i is the total number of children (including the subject) in the household. Because the number of children n_i is endogenous, OLS estimation of (1) gives a biased and inconsistent estimate of γ . I use both the birth of twins and the sameness of sex in earlier births to instrument for family size. The rest of the current subsection discusses the outcome variables explored in the study. The next subsection has a detailed discussion on the proposed instruments.

The study will look at the effect of sibship size on four outcome variables: educational attainment, whether a kid is lagging behind in grade, whether the kid attends a private school, and the labour force participation of the mother. The first two are final outcome variables whereas the second two are considered as intermediate outcome variables (Cáceres-Delpiano 2006).

Following Rosenzweig and Wolpin (1980), an educational attainment index was constructed by taking the ratio of the highest grade achieved by a child at the time of the census to the mean grade of all kids of the same age, month of birth, and province. The reason month of birth was chosen as a grouping variable is that laws on admission are based on the age of a child at a specific point in time during the admission year.³ Hence, the month of birth is one determining factor for the variation in grade achievement of children of the same age. However, individual provinces have their own rules on registration timelines as well as implementation. From Table 1, the educational attainment index has the expected value of 1 in both the 2+ and 3+ samples. Figure 1 shows a histogram of this index in both samples, which provides a more informative picture of its distribution.

Similarly, a dummy for left behind is constructed by comparing the highest grade of a child to the mode of the highest grade of all kids of the same age, month of birth, and province in

³According to the Admission Policy for Ordinary Public Schools, children aged five turning six by 30 June are required by law to be admitted to grade 1.

the relevant sample. It takes a value of 1 if a child’s highest grade is below the mode and 0 if it is equal to or above the mode. This essentially measures the same thing as the educational attainment index but has some advantages over the latter. Particularly, as it uses mode instead of mean, it is not affected by extreme values. But as a discrete measure it does not take into account the extent of the lagging (or progressing). It can be considered as a useful complement to the educational attainment index. According to [Table 1](#), 38.6 % in the 2+ sample and 37.6 % in the 3+ sample are lagging behind using this measure.

The third outcome variable, “private school” is a dummy for whether the child attends a private school (= 1) or a public school (= 0). Like in many countries, private schools offer better quality education and are preferred to public schools in South Africa. They are also more expensive. Thus enrollment in private school can reflect costly parental investment in the education of children. Attending a private school can also lead to better educational outcomes. The available evidence (?) suggests that educational outcomes are better for kids attending private school, although we should be careful to interpret this as showing causation, not merely correlation (Cáceres-Delpiano [2006](#)). From [Table 1](#), about 9% in the 2+ sample and 6.3% in the 3+ sample attend private school. These might sound low but are not surprising given that private schools are expensive in South Africa.

The last outcome variable, mothers’ labour force participation (Mothers’ LFP) is also a dummy for whether a mother is in the labour force (= 1) or not (= 0). The relationship between a mother’s labour market outcome and family size has been explored in many studies (*cite a few studies*). As an intermediate outcome, a mother being in the labour force has ambiguous effects on child outcomes (Cáceres-Delpiano [2006](#)). On the one hand, working mothers generate income that they can invest on their children, but, on the other hand spend less time caring for their children. Either way, it is an important channel between fertility and the welfare of children.

Instrumental Variables, Interpretation, and First Stages

The main motivation behind the use of instrumental variables is to uncover a causal relationship between family size and outcomes for children and mothers. Following the potential outcomes framework, what we are trying to estimate using the birth of twins and the sameness of sex of the first two births as natural experiments for the number of children is what Angrist and Imbens ([1995](#)) called the Average Causal Response (ACR). Angrist and Imbens ([1995](#)) have shown that, given mild regularity assumptions, the Two Stages Least Squares (TSLS) estimates using a binary instrument Z to instrument for an endogenous variable that takes a finite set of integer values (say $0, 1, \dots, J$) is the weighted average of per-unit causal effects (that is, the effect of going from 0 to 1, the effect of going from 1 to 2, and so on) along the length of an appropriately defined causal response function. More formally, let N_{0i} and N_{1i} be the potential number of children in the family if a binary instrument Z_i were equal to 0 and 1, respectively. The observed number of children n_i is then, $n_i = N_{0i} + (N_{1i} - N_{0i}) \cdot Z_i$. Similarly let $Y_i(j)$ be the potential outcome as a function of j , where j takes possible values for n_i in the set $\{0, 1, \dots, J\}$. Only one of these potential outcomes is realized and we denote the observed outcome by y_i . In the simplest case with no covariates, the IV estimator using a binary instrument Z gives the

Wald estimator.⁴ Then,

$$\beta_w = \frac{\mathbb{E}(y_i|Z_i = 1) - \mathbb{E}(y_i|Z_i = 0)}{\mathbb{E}(n_i|Z_i = 1) - \mathbb{E}(n_i|Z_i = 0)} = \sum_{j=1}^J \omega_j \cdot \mathbb{E}(Y_i(j) - Y_i(j-1)|N_{1i} \geq j > N_{0i})$$

where

$$\omega_j = \frac{\mathbb{P}(N_{1i} \geq j > N_{0i})}{\sum_{i=1}^J \mathbb{P}(N_{1i} \geq i > N_{0i})}$$

The unit causal response, $\mathbb{E}(Y_i(j) - Y_i(j-1)|N_{1i} \geq j > N_{0i})$, is the average difference in potential outcomes for *compliers* at point j , that is, individuals driven by the instrument from a treatment intensity less than j to at least j .⁵ Thus, the Wald estimator, β_w , is “a weighted ACR for people from families induced by an instrument to go from having fewer than j to at least j children, weighted over j by the probability of crossing this threshold.” (Angrist, Lavy, and Schlosser 2010, p. 787). The numerator in the expression for the weights ω_j above, $\mathbb{P}(N_{1i} \geq j > N_{0i})$, represents the proportion of compliers at point j .⁶ This is normalized by $\sum_{i=1}^J \mathbb{P}(N_{1i} \geq i > N_{0i})$, which can be shown to be equal to the Wald first stage (see Angrist and Pischke 2009, p. 183). That is,

$$\mathbb{E}(n_i|Z_i = 1) - \mathbb{E}(n_i|Z_i = 0) = \sum_{i=1}^J \mathbb{P}(N_{1i} \geq i > N_{0i})$$

There are three assumptions that lie behind the ACR theorem: (i) independence and exclusion; (ii) the existence of a first Stage; and (iii) monotonicity (Angrist and Pischke 2009). The independence assumption requires that the instrument be independent of potential outcomes and potential treatment assignments. This is sometimes alternatively stated by saying that the instrument need to be as good as randomly assigned. Both the twins and same sex instruments were originally proposed as natural experiments for fertility on the ground that both are virtually randomly assigned (Rosenzweig and Wolpin 1980; Angrist and Evans 1998). However, this is no longer obvious because various threats to the randomness of both instruments have been proposed in the literature (see the discussion below). The exclusion assumption, on the other hand, says that the instrument used has no effect on the outcomes other than through its effect on family size. Although this is closely related to the independence assumption, it is distinct from it. The exclusion restriction is a “a claim about a unique channel for causal effects of the instrument” (Angrist and Pischke 2009, p. 153). Again, various authors have pointed out ways in which exclusion restriction might fail to hold for both instruments. I have tried to address some of the important concerns in the upcoming subsections.

The other important assumption needed to estimate the ACR is monotonicity. To understand the idea behind this assumption, we need to introduce the four mutually exclusive groups comprising of our quasi-experimental population: compliers, always takers, never takers, and

⁴The interpretation of the ACR is more elaborate when covariates are added, but the basic idea is preserved. See Angrist and Imbens (1995, p. 437).

⁵We are using the word “compliers” in the context of the treatment-effects framework outlined in Angrist, Imbens, and Rubin (1996). See the discussion below for more explanation.

⁶Individuals with $N_{1i} \geq j > N_{0i}$ for any j in the support of n_i are considered as compliers.

defiers. These were first outlined in the treatment-effects framework of Angrist, Imbens, and Rubin 1996. Always takers always get treated regardless of their assignment, never takers never get treated regardless of their assignment, and compliers get treated when the instrument is switched on and don't get treated otherwise. The defiers, on the other hand, are a very strange group who get treated when the instrument is switched off and don't get treated when it is switched on. The assumption of monotonicity rules out the presence of defiers since they complicate the link between ACR and the reduced form (Angrist and Pischke 2009).⁷ In other words, monotonicity rules out the case where the instrument pushes some people into treatment while pushing others out. In our case, monotonicity requires that the proposed instrument moves fertility in one direction only. In particular, we assume that the potential number of children when the instrument is switched on is at least as large as it would have been when it is switched off; i.e., $N_{1i} \geq N_{0i}$. This is automatically satisfied by the twins instrument as fertility is always increased because of a twin birth for any mother. However, monotonicity need not hold for the same sex instrument as there could be parents who prefer children of the same sex and therefore go on to have another child if the sexes of the first two or three children are different (Huber 2015). In the subsection on the same sex instrument, I discuss a partial check for monotonicity using first stage regressions.

The Twins Instrument

The first stage for the twins instrument is

$$n_i = X_i\theta + \delta \cdot \text{Twins}_i + \xi_i \quad (2)$$

where Twins_i is an indicator for the occurrence of twins in the second (Twins2) or third (Twins3) birth, corresponding to the 2+ and 3+ samples, respectively. There are two rationales for considering twins at a fixed parity and looking at outcomes for older non-twin children. First, it accounts for the fact that twins are more likely in a larger family. Hence, using an indicator for a twin birth at any parity as an IV will be invalid (Öberg 2021). Second, looking at older children before the twin birth avoids the confounding effect from birth order that would occur after a twin birth (Black, Devereux, and Salvanes 2010). For instance, if the second birth is a twin birth, then a younger sibling will be a fourth child. So, the twin birth affects both family size and the birth order for any siblings in later births. The obvious drawback of focusing on children prior to the occurrence of twins is that it doesn't allow to look at the effect of birth order on the children.

Column 1 of Table 2 report the first stage results for the twins instruments. In both the 2+ and 3+ samples, twins lead to a positive and statistically significant increase in the number of children. A twin second birth increases the number of children by 0.83 whereas a twin third birth leads to 0.76 additional children, controlling for the other covariates in the model. These estimates remain virtually the same after controlling for same sex dummies by gender (column 4). The twins instrument is also relevant in both samples as evident from the sufficiently large

⁷Angrist and Pischke 2009 discuss monotonicity in the context of the LATE framework of Imbens and Angrist 1994. But this carries over directly to the ACR since the ACR is just an extension of the LATE (see Angrist and Pischke 2009, p. 181).

F statistics from the Wald test for exclusion restrictions (963 for the 2+ sample; 484 for the 3+ sample).

In addition to being relevant, the twins instrument needs to satisfy the exclusion restriction. A related assumption required by the LATE framework is that it also be random or at least as good as randomly assigned. The randomness of twin births has been challenged in a number of ways. There is a common finding that older mothers are more likely to give birth to twins than younger mothers (Angrist, Lavy, and Schlosser 2010). I also found this to be the case in the 2+ sample but not the 3+ sample. Mothers who give birth to twins in their second birth are about a year older at first birth compared to mothers who don't. The age of the mother is also related to the validity of the instrument in another way. An important concern with twins is that they affect both the number of children and the timing of a third (fourth) child for families with two (three) children. The only families for whom twins increase family size are those who have an unintended extra child because of the birth of twins (Öberg 2021). That is, they have one more child than the desired number of children. For parents who haven't attained their desired number of children, a parity-specific twin birth allows them to have the intended number of children unexpectedly fast. For these families the birth of twins only affects the timing of having children, not family size. Although the desired number of children is unobservable, it is likely to be correlated with the age of the mother. Older mothers would have probably reached their intended number of births compared to younger mothers. So twin births are likely to increase family size for women who experience it later in life, while it would only affect the timing of having children for mothers who experience it early in life (Cáceres-Delpiano 2006).

In order to address this, as well as the observation that twinning is related to the age of the mother, I run separate regressions for several subsamples based on the age of the mother. For ages from 21 to 40, I constructed subsamples by taking firstborns in the 2+ sample whose mothers are at least as old as the given age. For instance, 21+ would have mothers who are at least 21 years old, 22+ mothers who are at least 22, and so on, all the way to 40 (40+). I then run both OLS and IV for each of these subsamples for all outcome variables and plotted the coefficient estimates and confidence intervals in Figure 3. The general picture reveals that the results remain qualitatively the same when we progressively include older mothers, suggesting that systematic biases due to the age of the mother is not a concern. The details of the Figure are discussed in the results section.

Another challenge to the randomness of twins is the observation that mothers who use fertility enhancing treatments are more likely to have multiple births. Since the use of fertility enhancing treatments is usually unobservable, there is a potential bias from using twin births as an instrument (Braakmann and Wildman 2016). Although availability and utilization is low, Assisted Reproductive Technologies (ARTs) such as in vitro fertilizations (IVFs) are becoming more prevalent in South Africa. In fact, South Africa has the second highest number of registered IVF centres in the whole continent in 2019 (Ombelet and Onofre 2019). Studies analysing ART registry data from sub-Saharan Africa have observed high multiple pregnancy rates associated with the utilization of ARTs (Botha, Shamley, and Dyer 2018; Dyer et al. 2019). There are indications of a rise in the number of multiple births over time in our sample also that could be associated with the evolution of ARTs. Using the dates of births of all respondents in the 2011

South African Census, I plotted the number of multiple births (i.e., twins, triplets, quadruplets, etc.) per 1000 live births from 1970 onwards. This is shown in [Figure 2](#). Although there is no historical data on the utilization of ARTs in South Africa since its introduction in the 1980s (ART registries were established in the late 2000s), the rising trend of multiple births across all population groups is consistent with the introduction and (slow) dissemination of fertility treatments.

A related threat to the randomness of twin births is the observation that the occurrence of twins is related to the mother’s health and health related behaviour. Bhalotra and Clarke ([2019](#)) show that twinning is strongly related to morbidity, smoking status, availability of reproductive health services and other health indicators even for mothers not using fertility treatments. They also show that the education of the mother is positively associated with twinning, consistent with the hypothesis that educated women are more likely to engage in health seeking behaviour (as well as to seek fertility enhancing treatments). Although the 2011 South African census does not have data on health indicators, I have been able to replicate the conjecture that twinning is more likely in subpopulations that would seek health treatments. [Figure 2](#) reveals that starting from the mid-1990s the occurrence of multiple births is systematically higher among white South Africans as compared to non-whites. This would make sense as white mothers arguably have both the awareness and the capacity when it comes to utilizing health treatments including ARTs. To check whether this apparent difference poses a challenge, I report the same set of results (first stages, OLS, and 2SLS) using the twins instrument for whites and non-whites. Without going into details (see the results section), the results, reported in [Table 4](#), are robust to division by mother’s population group, suggesting that there is no evidence of a potential bias stemming from systematic differences in the occurrence of multiple births.

Rosenzweig and Zhang ([2009](#)) point out a further challenge to the exogeneity of twin births. They argue that the occurrence of twins changes the behaviour of parents in such a way that they reallocate resources away from twins and towards older singleton children. They point out two aspects of twins that affect inter-child allocations: “(i) the closer spacing of twins, which makes investing in the average quality of the twins more costly compared with investing in non-twins, and (ii) the lower endowments of twins, which will affect the resources allocated to non-twins depending on whether parents reinforce or compensate for endowment differences across children.” (Rosenzweig and Zhang [2009](#), p. 1152). Here the lower endowment of twins refers to the fact that twins usually have lower birth weight, lower survival rate, lower cognitive achievement and so on. If such resource reallocation is taking place, then estimates using twins as an instrument will be biased towards zero.

Following Angrist, Lavy, and Schlosser ([2010](#)), I estimated reduced-form twins effects on outcomes in samples in which twins are unlikely to affect family size (i.e., there is no first stage) to see if resource reallocation is a problem. The rationale for focusing on no-first-stage samples is that because the birth of twins has no effect on family size in these subsamples, effects of confounding factors related to twinning should surface. The two subsamples with a zero first stage are those from families with closely spaced births (< 2 years between the first two births) and those whose mothers were not educated, both from families with at least three children in

the 2+ sample.⁸ Columns 1 and 2 in Panel A of [Table 5](#) shows that the first stage regressions in such families yield insignificant coefficient estimates at the conventional 5% level. In Panel B, I estimate reduced form regressions of all four outcomes on the Twin2 dummy. If what Rosenzweig and Zhang (2009) propose is the case, we ought to see that older singletons have a higher outcome in these no first stage samples. But the opposite is true: in both subsamples, twin births do not significantly affect any of the outcomes for older children. The only exception is the probability of private school attendance in the subsample of uneducated mothers. But even in this case, firstborn children have *lower* not higher likelihood of attending private school, in defiance of the hypothesis that nontwin firstborn children are favoured. Hence, there is no evidence for a reallocation effect that could potentially confound our IV estimates.

The Same Sex Instrument

Instrumenting with the gender mix of siblings in the family was first introduced by Angrist and Evans (1998). They used it as a source of exogenous variation of family size to investigate the labour supply response of mothers to fertility. They explain their rationale for doing so as follows:

This instrument exploits the widely observed phenomenon of parental preferences for a mixed sibling-sex composition. In particular, parents of same-sex siblings are significantly and substantially more likely to go on to have an additional child. Because sex mix is virtually randomly assigned, a dummy for whether the sex of the second child matches the sex of the first child provides a plausible instrument for further childbearing among women with at least two children. (Angrist and Evans (1998, p. 451))

There is plenty of empirical support for the claim that parents prefer to have children of both genders and that this preference drives the decision to have additional children (Norling 2018; Bisbee et al. 2015). The first stage regression on the same sex dummy is also consistent with this result. The results in [Table 2](#) show that in both the 2+ and 3+ samples, the same sex dummies have a positive and statistically significant relationship with the number of children born to a mother. The same relationship holds when it is disaggregated by gender.

The results from the same sex IV should also be interpreted as a LATE that applies only to the subpopulation of compliers — i.e., those families whose number of children was increased as a result of having had two (three) first children of the same sex (Bisbee et al. 2015; Öberg 2021). However, just as in the case of the twins instrument, there have been objections to the exclusion restriction with respect to the same sex IV. There is a possibility of gender-specific economies of scale that could reinforce child quality investments when parents have children of the same sex (Clarke 2018). Rosenzweig and Wolpin (2000) point out that cost efficiencies associated with gender mix will confound the effect of an exogenous increase in the number of children with direct child-rearing cost effects on outcomes. To test this, I run reduced form regressions on subsamples with zero first stages, similar to the no-first-stage analysis for the

⁸These are households who are likely to have large families anyway. In such a case, the birth of twins is likely to have little or no effect on family size.

twins instrument. I consider outcomes for firstborn boys in the 2+ sample, so the appropriate instrument to consider would be the Boy12 dummy. The two subsamples include mothers with less than 2 years of spacing in the first two births and those with no schooling. The results, which are reported in [Table 5](#), show that there is no reduced form relation between the Boy12 instrument and any of the outcomes for the firstborn child. That is, a firstborn male child who has a younger sibling of the same gender does not appear to benefit in anyway as suggested by the same-sex cost advantage hypothesis.⁹ Hence, there is no ground to suspect a violation of the exclusion restriction based on household efficiencies related to same sex sibships.

⁹The only statistically significant effect is observed on the mother's labour force participation in the subsample with close spacing (column 3). But we are focusing here on the direct effects on the child's outcomes. Even though there would be indirect effects, we should have seen it in the regressions where child outcomes are the dependent variables.

References

- Angrist, Joshua, Victor Lavy, and Analia Schlosser (Oct. 2010). “Multiple Experiments for the Causal Link between the Quantity and Quality of Children”. In: *Journal of Labor Economics* 28.4. Publisher: The University of Chicago Press, pp. 773–824. ISSN: 0734-306X. DOI: [10.1086/653830](https://doi.org/10.1086/653830). URL: <https://www.journals.uchicago.edu/doi/full/10.1086/653830> (visited on 12/15/2021).
- Angrist, Joshua D. and William N. Evans (1998). “Children and Their Parents’ Labor Supply: Evidence from Exogenous Variation in Family Size”. In: *The American Economic Review* 88.3. Publisher: American Economic Association, pp. 450–477. ISSN: 0002-8282. URL: <https://www.jstor.org/stable/116844> (visited on 12/13/2021).
- Angrist, Joshua D. and Guido W. Imbens (June 1995). “Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity”. In: *Journal of the American Statistical Association* 90.430. Publisher: Taylor & Francis eprint: <https://www.tandfonline.com/doi/pdf/10.1080/01621459.1995.10476535>, pp. 431–442. ISSN: 0162-1459. DOI: [10.1080/01621459.1995.10476535](https://doi.org/10.1080/01621459.1995.10476535). URL: <https://www.tandfonline.com/doi/abs/10.1080/01621459.1995.10476535> (visited on 04/11/2022).
- Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin (1996). “Identification of Causal Effects Using Instrumental Variables”. In: *Journal of the American Statistical Association* 91.434. Publisher: [American Statistical Association, Taylor & Francis, Ltd.], pp. 444–455. ISSN: 0162-1459. DOI: [10.2307/2291629](https://doi.org/10.2307/2291629). URL: <https://www.jstor.org/stable/2291629> (visited on 03/09/2022).
- Angrist, Joshua D. and Jörn-Steffen Pischke (Dec. 2009). *Mostly Harmless Econometrics: An Empiricist’s Companion*. Princeton University Press. ISBN: 0691120358.
- Öberg, Stefan (Sept. 2021). *The casual effect of fertility: The multiple problems with instrumental variables for the number of children in families*. en. preprint. SocArXiv. DOI: [10.31235/osf.io/peuvz](https://doi.org/10.31235/osf.io/peuvz). URL: <https://osf.io/peuvz> (visited on 12/13/2021).
- Bhalotra, Sonia and Damian Clarke (Dec. 2019). “Twin Birth and Maternal Condition”. In: *The Review of Economics and Statistics* 101.5, pp. 853–864. ISSN: 0034-6535. DOI: [10.1162/rest_a_00789](https://doi.org/10.1162/rest_a_00789). URL: https://doi.org/10.1162/rest_a_00789 (visited on 03/09/2022).
- Bisbee, James et al. (Oct. 2015). *Local Instruments, Global Extrapolation: External Validity of the Labor Supply-Fertility Local Average Treatment Effect*. Working Paper 21663. Series: Working Paper Series. National Bureau of Economic Research. DOI: [10.3386/w21663](https://doi.org/10.3386/w21663). URL: <https://www.nber.org/papers/w21663> (visited on 03/09/2022).
- Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes (2010). “Small Family, Smart Family? Family Size and the IQ Scores of Young Men”. In: *The Journal of Human Resources* 45.1. Publisher: [University of Wisconsin Press, Board of Regents of the University of Wisconsin System], pp. 33–58. ISSN: 0022-166X. URL: <https://www.jstor.org/stable/20648936> (visited on 03/09/2022).
- Botha, Barend, Delva Shamley, and Silke Dyer (Apr. 2018). “Availability, effectiveness and safety of ART in sub-Saharan Africa: a systematic review”. In: *Human Reproduction Open* 2018.2, hoy003. ISSN: 2399-3529. DOI: [10.1093/hropen/hoy003](https://doi.org/10.1093/hropen/hoy003). URL: <https://doi.org/10.1093/hropen/hoy003> (visited on 03/12/2022).

- Braakmann, Nils and John Wildman (2016). “Reconsidering the effect of family size on labour supply: the twin problems of the twin birth instrument”. en. In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 179.4. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/rssa.12160> pp. 1093–1115. ISSN: 1467-985X. DOI: [10.1111/rssa.12160](https://doi.org/10.1111/rssa.12160). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/rssa.12160> (visited on 03/10/2022).
- Cáceres-Delpiano, Julio (2006). “The Impacts of Family Size on Investment in Child Quality”. In: *The Journal of Human Resources* 41.4. Publisher: [University of Wisconsin Press, Board of Regents of the University of Wisconsin System], pp. 738–754. ISSN: 0022-166X. URL: <https://www.jstor.org/stable/40057289> (visited on 02/21/2022).
- Clarke, Damian (2018). “Children and Their Parents: A Review of Fertility and Causality”. en. In: *Journal of Economic Surveys* 32.2. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/joes.12202>, pp. 518–540. ISSN: 1467-6419. DOI: [10.1111/joes.12202](https://doi.org/10.1111/joes.12202). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/joes.12202> (visited on 03/07/2022).
- Dyer, Silke et al. (Feb. 2019). “Assisted reproductive technologies in Africa: first results from the African Network and Registry for Assisted Reproductive Technology, 2013”. In: *Reproductive BioMedicine Online* 38.2. Publisher: Elsevier, pp. 216–224. ISSN: 1472-6483, 1472-6491. DOI: [10.1016/j.rbmo.2018.11.001](https://doi.org/10.1016/j.rbmo.2018.11.001). URL: [https://www.rbmojournal.com/article/S1472-6483\(18\)30571-6/fulltext](https://www.rbmojournal.com/article/S1472-6483(18)30571-6/fulltext) (visited on 03/12/2022).
- Huber, Martin (2015). “Testing the Validity of the Sibling Sex Ratio Instrument”. In: *LABOUR* 29.1. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/labr.12045>, pp. 1–14. ISSN: 1467-9914. DOI: [10.1111/labr.12045](https://doi.org/10.1111/labr.12045). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/labr.12045> (visited on 04/27/2022).
- Imbens, Guido W. and Joshua D. Angrist (1994). “Identification and Estimation of Local Average Treatment Effects”. In: *Econometrica* 62.2. Publisher: [Wiley, Econometric Society], pp. 467–475. ISSN: 0012-9682. DOI: [10.2307/2951620](https://doi.org/10.2307/2951620). URL: <https://www.jstor.org/stable/2951620> (visited on 03/09/2022).
- Norling, Johannes (Nov. 2018). “Measuring heterogeneity in preferences over the sex of children”. en. In: *Journal of Development Economics* 135, pp. 199–221. ISSN: 03043878. DOI: [10.1016/j.jdeveco.2018.07.004](https://doi.org/10.1016/j.jdeveco.2018.07.004). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0304387818305121> (visited on 03/09/2022).
- Ombelet, W and J Onofre (2019). “IVF in Africa: what is it all about?” In: *Facts, Views & Vision in ObGyn* 11.1, pp. 65–76. ISSN: 2032-0418. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6822948/> (visited on 03/10/2022).
- Rosenzweig, Mark R. and Kenneth I. Wolpin (1980). “Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment”. In: *Econometrica* 48.1. Publisher: [Wiley, Econometric Society], pp. 227–240. ISSN: 0012-9682. DOI: [10.2307/1912026](https://doi.org/10.2307/1912026). URL: <https://www.jstor.org/stable/1912026> (visited on 12/16/2021).
- (2000). “Natural ”Natural Experiments” in Economics”. In: *Journal of Economic Literature* 38.4. Publisher: American Economic Association, pp. 827–874. ISSN: 0022-0515. URL: <https://www.jstor.org/stable/2698663> (visited on 12/17/2021).
- Rosenzweig, Mark R. and Junsen Zhang (July 2009). “Do Population Control Policies Induce More Human Capital Investment? Twins, Birth Weight and China’s “One-Child” Policy”.

In: *The Review of Economic Studies* 76.3, pp. 1149–1174. ISSN: 0034-6527. DOI: [10.1111/j.1467-937X.2009.00563.x](https://doi.org/10.1111/j.1467-937X.2009.00563.x). URL: <https://doi.org/10.1111/j.1467-937X.2009.00563.x> (visited on 12/13/2021).

Tables and Figures

Figure 1: Histograms of Educational Attainment Index

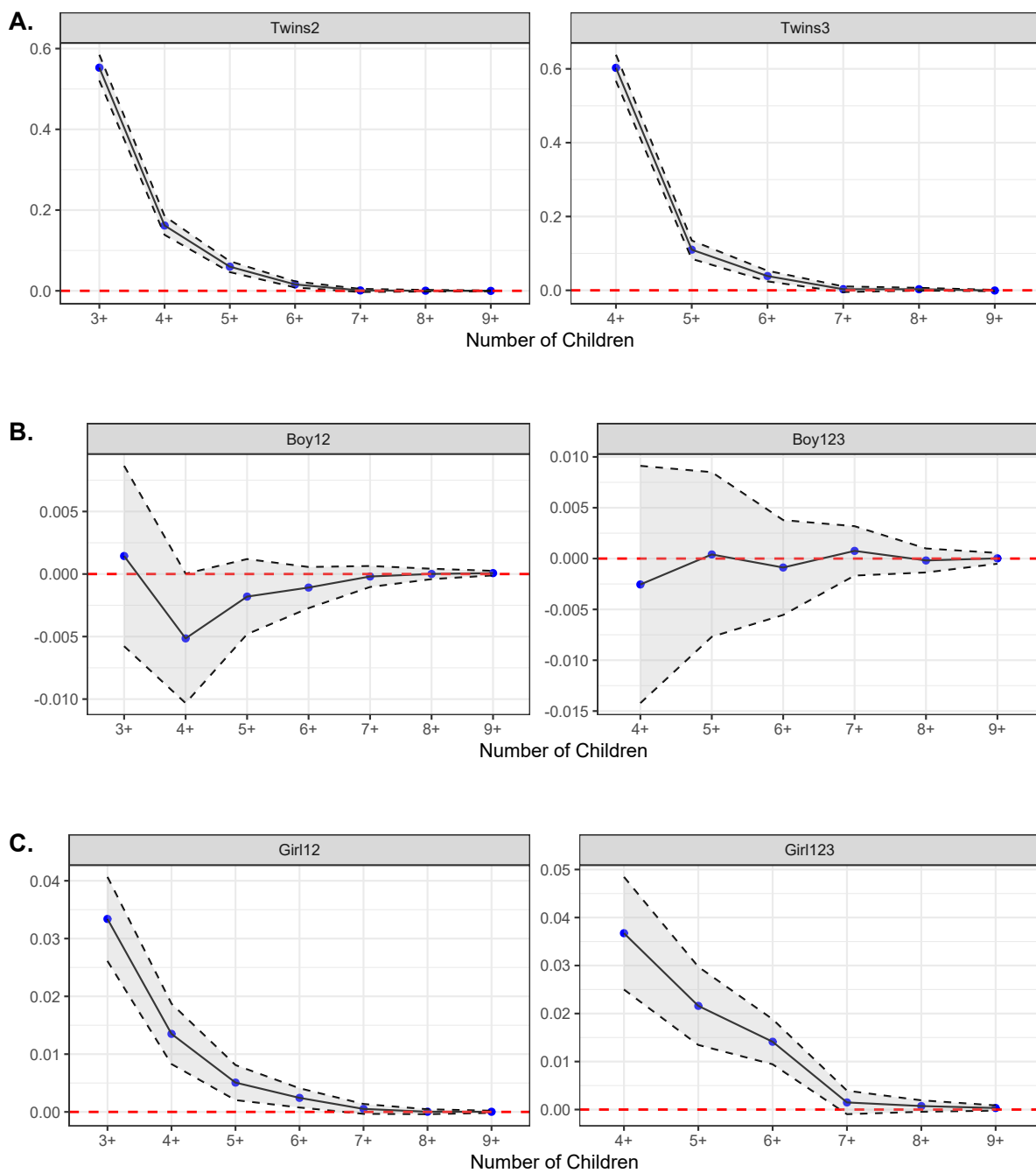


Figure 2: Caption in landscape to a figure in landscape.

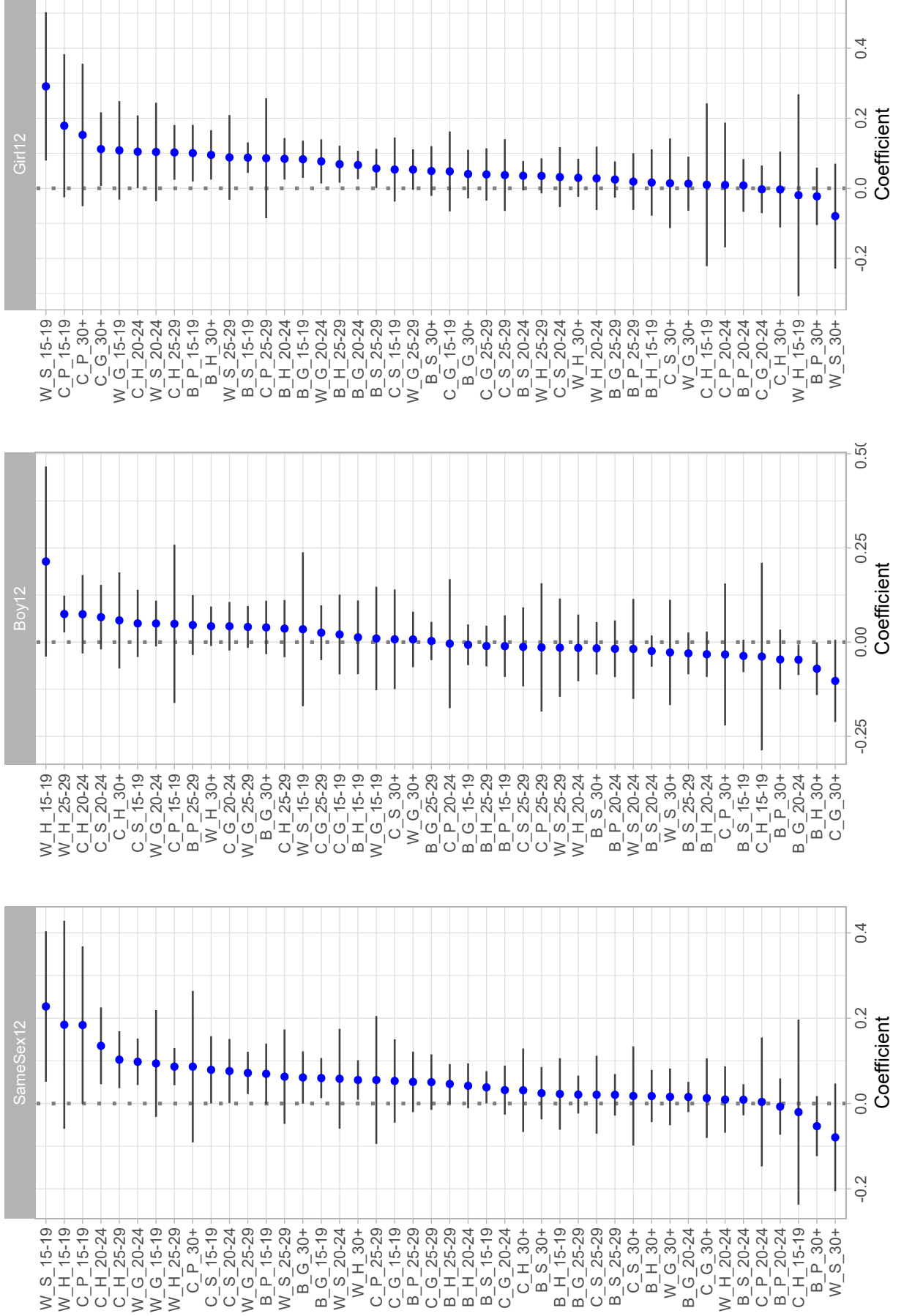


Figure 3: Multiple Births Per 1000 Live Births by Mother's Population Group

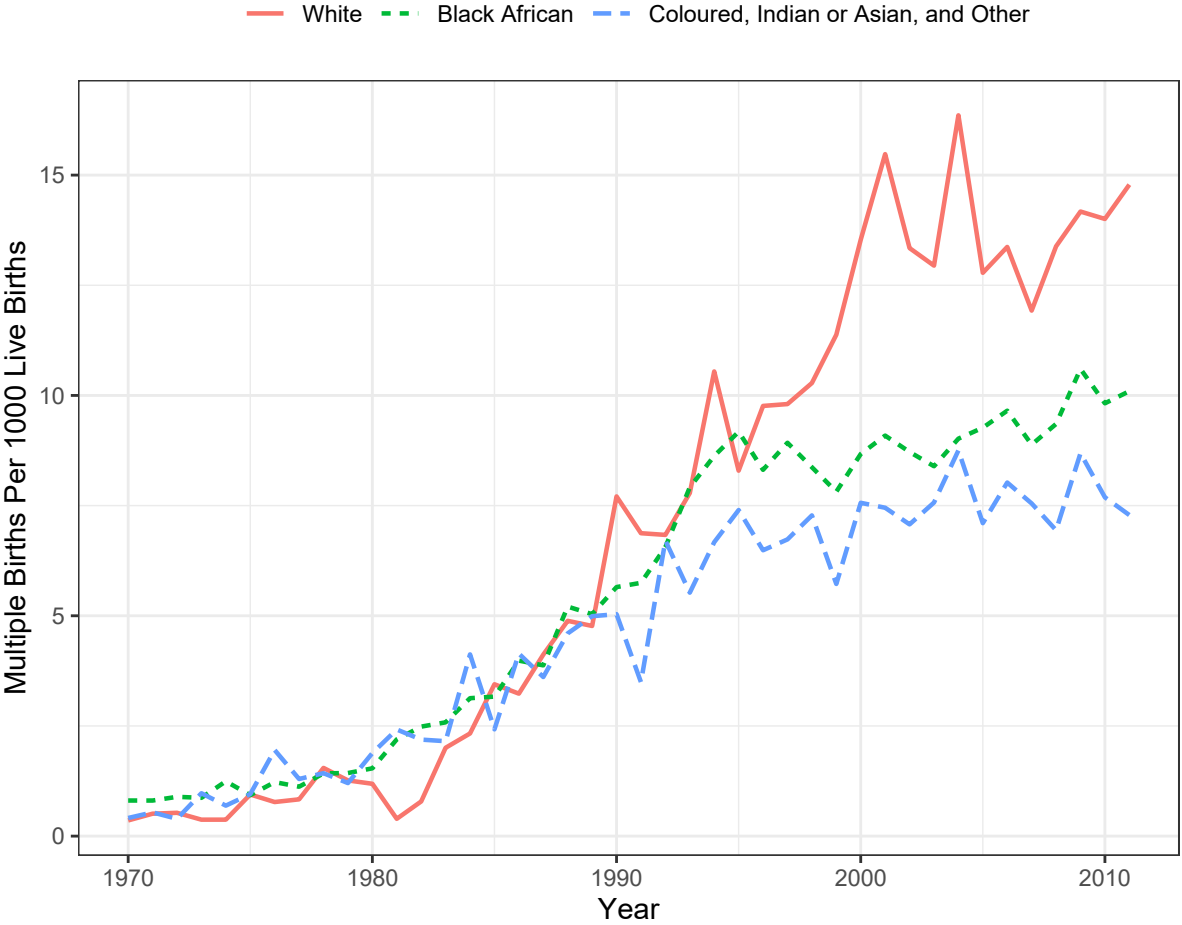


Figure 4: Plot of Coefficient Estimates and 95% Confidence Interval (2+ Sample)

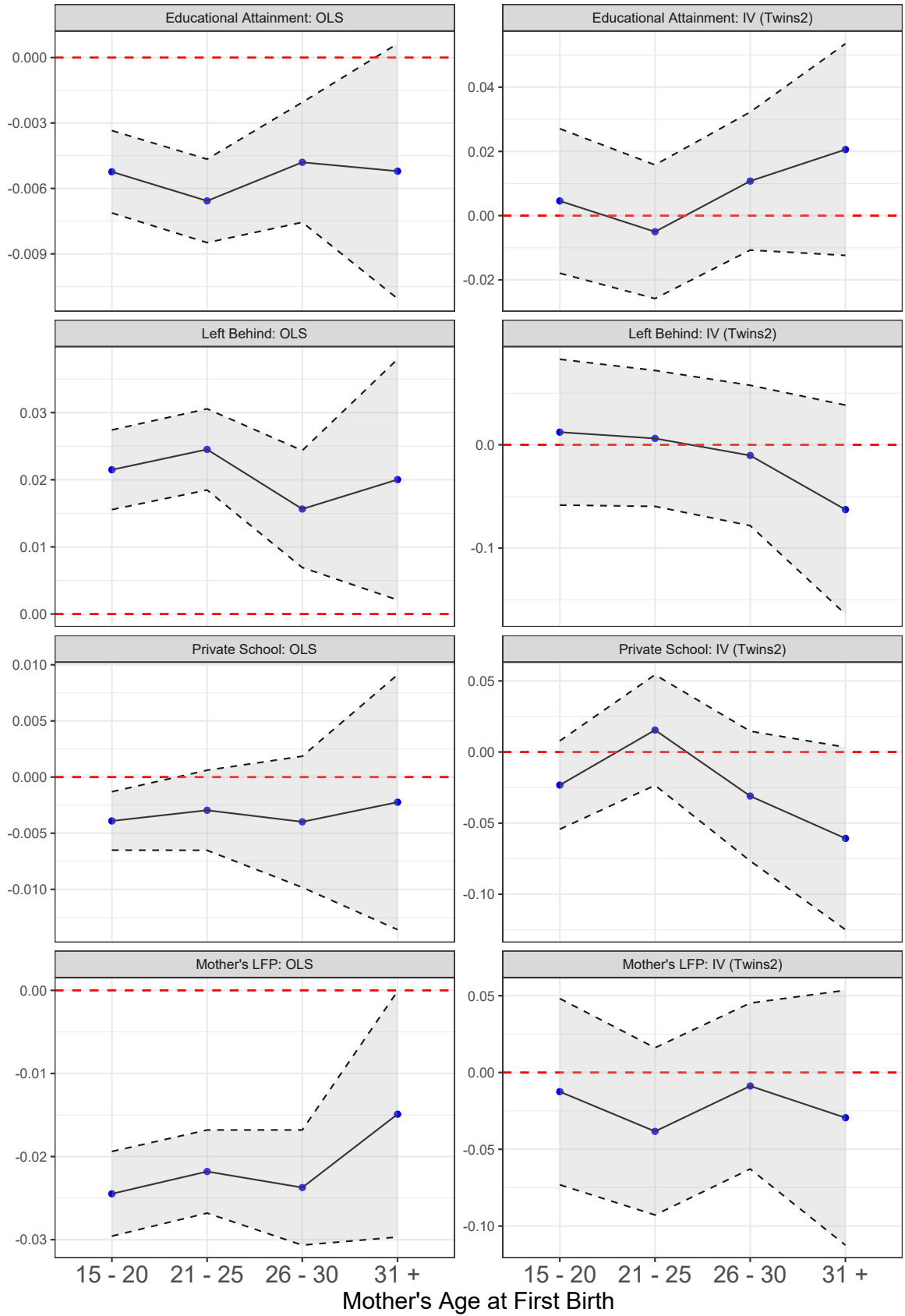


Table 1: Summary Statistics

| | 2+ Sample | | 3+ Sample | | | |
|-----------------------------|------------|-------|------------|-------|------------------------|-------|
| | Firstborns | | Firstborns | | First and Second Borns | |
| | Mean | sd | Mean | sd | Mean | sd |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Number of Children | 2.665 | 0.904 | 3.592 | 0.87 | 3.592 | 0.87 |
| Male | 0.499 | | 0.495 | | 0.499 | |
| Age (Years) | 13.14 | 3.061 | 14.88 | 2.271 | 13.13 | 2.912 |
| Age (Months) | 163.1 | 36.8 | 184 | 27.18 | 162.8 | 34.96 |
| Twins2 | 0.01 | | | | | |
| SameSex12 | 0.517 | | | | | |
| Boy12 | 0.261 | | | | | |
| Girl12 | 0.255 | | | | | |
| Twins3 | | | 0.013 | | 0.013 | |
| SameSex123 | | | 0.274 | | 0.274 | |
| Boy123 | | | 0.138 | | 0.138 | |
| Girl123 | | | 0.136 | | 0.136 | |
| Grade Achieved | 6.471 | 2.921 | 7.865 | 2.328 | 6.386 | 2.761 |
| Educational Attainment | 1 | 0.155 | 1 | 0.139 | 1.002 | 0.159 |
| Left Behind | 0.386 | | 0.401 | | 0.376 | |
| Private School | 0.089 | | 0.065 | | 0.063 | |
| Mother's Age (Years) | 36.57 | 5.561 | 37.43 | 4.808 | 37.43 | 4.808 |
| Age at First Birth | 22.94 | 4.86 | 22.05 | 4.417 | 22.05 | 4.417 |
| Mother in the Labour Force | 0.763 | | 0.713 | | 0.713 | |
| Father in Household | 0.608 | | 0.606 | | 0.608 | |
| Mother's Population Group | | | | | | |
| Black African | 0.752 | | 0.819 | | 0.819 | |
| Coloured | 0.109 | | 0.094 | | 0.094 | |
| White | 0.107 | | 0.061 | | 0.061 | |
| Indian or Asian | 0.03 | | 0.023 | | 0.023 | |
| Other | 0.003 | | 0.003 | | 0.003 | |
| Mother's Level of Education | | | | | | |
| No schooling | 0.05 | | 0.079 | | 0.079 | |
| Some Primary | 0.089 | | 0.129 | | 0.129 | |
| Completed Primary | 0.046 | | 0.061 | | 0.061 | |
| Some Secondary | 0.362 | | 0.391 | | 0.391 | |
| Grade 12/Std 10 | 0.285 | | 0.224 | | 0.224 | |
| Higher | 0.168 | | 0.117 | | 0.117 | |
| Mother's Marital Status | | | | | | |
| Never Married | 0.153 | | 0.125 | | 0.125 | |
| Married | 0.635 | | 0.662 | | 0.662 | |
| Living Together | 0.18 | | 0.187 | | 0.187 | |
| Separated | 0.011 | | 0.012 | | 0.012 | |
| Divorced | 0.018 | | 0.012 | | 0.012 | |
| Widower | 0.003 | | 0.002 | | 0.002 | |
| <i>N</i> | 94,977 | | 28,679 | | 57,358 | |

The standard deviations for proportions is not presented.

Table 2: First Stage Regressions

| | <i>Dependent variable:</i> | | | |
|---------------------------|----------------------------|---------------------|---------------------|---------------------|
| | Number of Children | | | |
| | (1) | (2) | (3) | (4) |
| Panel A: 2+ Sample | | | | |
| Twins2 | 0.830*** (0.027) | | | 0.830*** (0.027) |
| SameSex12 | | 0.037*** (0.005) | | |
| Boy12 | | | 0.016** (0.008) | 0.015** (0.008) |
| Girl12 | | | 0.057*** (0.008) | 0.057*** (0.007) |
| F | 962.7 | 47.38 | 31.02 | 342. |
| Observations | 94,977 | 94,977 | 94,977 | 94,977 |
| R ² | 0.184 | 0.177 | 0.177 | 0.185 |
| Panel B: 3+ Sample | | | | |
| Twins3 | 0.761*** (0.035) | | | 0.762*** (0.035) |
| SameSex123 | | 0.051*** (0.011) | | |
| Boy123 | | | 0.018 (0.014) | 0.019 (0.014) |
| Girl123 | | | 0.083*** (0.015) | 0.084*** (0.015) |
| F | 484.1 | 22.19 | 15.95 | 172.3 |
| Observations | 57,358 | 57,358 | 57,358 | 57,358 |
| R ² | 0.152 | 0.143 | 0.143 | 0.153 |

Notes: Covariates for all regressions include age (in months) and sex of child, mother's characteristics, dummies for districts, and a dummy for whether the father resides in the household. The regressions for the 3+ sample in addition include dummies for birth order. F statistics are from the Wald test for exclusion restrictions of the relevant instrument(s) from the regression. Robust standard errors (for the 2+ sample) and cluster robust standard errors (for the 3+ sample; clustered by mother's ID) are in parenthesis. *** Significant at 1%, ** Significant at 5%, * Significant at 10%.

Table 3: OLS and 2SLS Estimates of The Effect of The Number of Children

| | 2+ Sample | | | | 3+ Sample | | | |
|------------------------|----------------------|-------------------|-------------------|-------------------------|----------------------|-------------------|-------------------|---------------------------|
| | OLS | | 2SLS | | OLS | | 2SLS | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Educational Attainment | -0.006*** (0.001) | 0.005 (0.006) | -0.013 (0.023) | 0.004 (0.006) | -0.008*** (0.001) | 0.012 (0.009) | 0.052* (0.028) | 0.016* (0.008) |
| Left Behind | 0.023*** (0.002) | -0.004 (0.019) | 0.084 (0.074) | 0.001 (0.018) | 0.025*** (0.003) | -0.025 (0.025) | -0.025 (0.078) | -0.025 (0.024) |
| Private School | -0.003*** (0.001) | -0.017 (0.011) | -0.021 (0.041) | -0.017* (0.010) | -0.003** (0.001) | 0.010 (0.016) | -0.025 (0.047) | 0.007 (0.015) |
| Mothers LFP | -0.019*** (0.002) | -0.021 (0.016) | 0.027 (0.061) | -0.019 (0.015) | -0.023*** (0.003) | 0.038 (0.030) | 0.112 (0.093) | 0.045 (0.028) |
| IV used | - | Twins2 | Boy12 Girl12 | Twins2, Boy12 Girl12 | - | Twins3 | Boy123 Girl123 | Twins3, Boy123 Girl123 |

Notes: Covariates for all models include age (in months) and sex of child, mother's education, dummies for mother's population group and income range, dummies for districts, and a dummy for whether the father resides in the household. The regressions for the 3+ sample are clustered by mother's ID. The regressions using the SameSex12 and SameSex123 instruments were excluded for reasons of space. The results are qualitatively similar to the ones in columns 3 and 7, respectively, and are available upon request. Robust standard errors are in parentheses. *** Significant at 1%, ** Significant at 5%, * Significant at 10%.

Table 4: Heterogeneity by Mother's Population Group (Whites vs. Non-Whites)

| | | 2+ Sample | | | 3+ Sample | | | | |
|-----------------------------------|--|--------------------------------|-------------------|--------------------------------|-------------------|--------------------------------|-------------------|--------------------------------|-------------------|
| | | Whites | | Non-Whites | | Whites | | Non-Whites | |
| | | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Panel A: First Stage [†] | | | | | | | | | |
| Twins2/ Twins3 | | 0.942*** (0.044) [447.7] | | 0.809*** (0.030) [732.5] | | 0.861*** (0.068) [162.9] | | 0.756*** (0.037) [422.2] | |
| F | | | | | | | | | |
| Panel B: OLS & 2SLS [‡] | | | | | | | | | |
| | | OLS | IV | OLS | IV | OLS | IV | OLS | IV |
| Educational Attainment | | -0.001 (0.002) | 0.019 (0.012) | -0.005*** (0.001) | 0.003 (0.007) | -0.009 (0.005) | 0.026 (0.018) | -0.007*** (0.001) | 0.011 (0.010) |
| Left Behind | | 0.018* (0.009) | -0.034 (0.042) | 0.020*** (0.002) | -0.001 (0.021) | 0.023 (0.020) | -0.009 (0.082) | 0.024*** (0.003) | -0.025 (0.027) |
| Private School | | 0.021** (0.007) | -0.010 (0.035) | -0.004*** (0.001) | -0.017 (0.011) | 0.047* (0.019) | -0.094 (0.081) | -0.004*** (0.001) | 0.015 (0.015) |
| Mothers LFP | | -0.016** (0.006) | 0.008 (0.027) | -0.020*** (0.002) | -0.024 (0.018) | -0.009 (0.019) | 0.044 (0.080) | -0.023*** (0.003) | 0.040 (0.032) |
| N | | 10,121 | | 84,856 | | 3,494 | | 53,864 | |

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%.

[†] Columns 1-4 have the coefficients for the Twins2 instrument (in the 2+ sample) and in columns 5-8 are the coefficients for the Twins3 instrument (3+ sample). The regressions were run using the same set of controls as in Table 3. Robust standard errors are in parentheses and the numbers in square brackets below the s.e. are F stats for the exclusion of the Twins2 or the Twins 3 variable from the model.

[‡] The regressions control for the same set of covariates as in Table 3 (see notes there). The regressions for the 3+ sample are clustered by mother's ID.

Table 5: Reduced Form Effects in No-First-Stage Samples

| | Twins2 [†] | | Boy12 [‡] | |
|--|-----------------------|------------------------|-----------------------|------------------------|
| | Spacing < 2 (1) | No Schooling (2) | Spacing < 2 (3) | No Schooling (4) |
| Panel A: First Stage | | | | |
| No. of children | 0.266* (0.141) | 0.101 (0.134) | −0.0004 (0.037) | 0.041 (0.046) |
| Panel B: Reduced Form[§] | | | | |
| Educational Attainment | 0.031 (0.025) | 0.006 (0.021) | 0.001 (0.006) | 0.007 (0.007) |
| Left Behind | −0.131* (0.070) | −0.067 (0.067) | 0.013 (0.018) | −0.026 (0.020) |
| Private School | 0.027 (0.053) | −0.011*** (0.004) | 0.005 (0.011) | 0.001 (0.006) |
| Mothers LFP | 0.033 (0.065) | −0.052 (0.067) | 0.034** (0.015) | 0.005 (0.020) |
| <i>N</i> | 3,489 | 3,022 | 2,869 | 2,381 |

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. The same set of covariates as in [Table 3](#) were included in all regressions. Robust standard errors in parentheses.

[†] The subsamples are restricted to firstborns from families with at least three children.

[‡] Subsamples includes only firstborn boys.

[§] In addition to the covariates used in [Table 3](#), the reduced form regressions control for the number of children in the family.