ORIGINAL PAPER

Quasi-experimental evidence for the causal link between fertility and subjective well-being

Jan Priebe[1,2]

Abstract

This article presents causal evidence on the impact of fertility on women's subjective well-being using quasi-experimental variation due to preferences for a mixed sibling sex composition (having at least one child of each sex). Based on a large sample of women from 35 developing countries, I find that having children increases mothers' life satisfaction and happiness. I further establish that the positive impact of fertility on subjective well-being can be explained by related increases in mothers' satisfaction with family life, friendship, and treatment by others.

Keywords Causality · Fertility · Happiness · Subjective well-being · Local average treatment effect

JEL classification C21 · D10 · I31 · J17

1 Introduction

The last 100 years have witnessed substantial declines in fertility rates across all high-income countries. These days, several developed countries are below replacement-level fertility, facing populating aging and being unable to reproduce themselves over an extended period of time.[1] The trend of falling fertility levels has not been limited to high-income countries, with the majority of developing and middle-income countries

---

[1]Total fertility rates are below replacement level (2.1) in all developed countries with the USA currently at 1.9 and several other nations such as Germany (1.4), Italy (1.4), Japan (1.4), Poland (1.4), Singapore (1.2), and Spain (1.3) showing rates far away from the replacement threshold. Likewise, many middle-income countries such as Brazil (1.8), Russia (1.8), or Thailand (1.5) exhibit low fertility levels (UN, 2015a).

✉ Jan Priebe
   jan.priebe@giga–hamburg.de

1   University of Göttingen, Göttingen, Niedersachsen, Germany

2   German Institute of Global and Area Studies, Neuer Jungfernstieg 21, 20354 Hamburg, Germany

experiencing rapid movements towards replacement-level fertility (Strulik and Vollmer, 2015; UN, 2015a).

To better understand fertility behavior and the existence of low-fertility regimes, the economic literature has recently turned towards examining the role of parenthood on subjective well-being (SWB).[2] Empirical evidence on this topic—mostly from cross-sectional regressions—has often found an insignificant or even negative effect of having children on SWB, which could help explain the trend towards declining and low fertility levels (Alesina et al. 2004; Blanchflower, 2008; Clark, 2007; Deaton and Stone, 2014; Di Tella et al., 2003, Di Tella and MacCulloch, 2006; Hansen et al. 2009; Dolan et al., 2008; Kohler et al. 2005; Margolis and Myrskylä, 2011; Stanca, 2012; Stutzer and Frey, 2006).

The predominant result that parents are better off without children is surprising given that most of the world is pervaded by strong cultural beliefs that children increase the well-being of parents (Margolis and Myrskylä, 2011). Related research, however, has provided some rationalization for the finding by showing that parents experience higher levels of stress and anxiety (Buddelmeyer et al., 2018; Deaton and Stone, 2014; Evenson and Simon, 2005; Hamermesh and Lee, 2007), increased anger and depression (Nomaguchi and Milkie, 2003), and more worries about sufficient family income (Stanca, 2012) compared with non-parents.[3]

Despite plausible explanations for the absence of a positive correlation between having children and SWB, it should be pointed out that only limited causal evidence exists that examines how fertility affects SWB. Establishing causality in this context is difficult given that fertility decisions are endogenous for multiple reasons. First, concerns about reversed causality need to be addressed given that several studies have pointed out selection effects indicating that happier couples are more likely to have children (Cetre et al., 2016; Moglie et al., 2015; Parr, 2010). Second, econometric results obtained from simple ordinary least squares (OLS), matching on observables, and panel fixed-effect specifications might be biased due to the inability to control for certain (time-varying) confounding variables such as personality, job aspirations, partnership stability, sexual activity, and growing into adulthood (Myrskylä and Margolis, 2014). Third, several of the control variables used in the fertility-SWB literature are simultaneously factors influencing and outcomes of the very same relationship, which therefore requires robustness checks involving different covariate specifications. Since the magnitude, direction, and significance levels of the coefficient of interest are quite sensitive to the choice of covariates (Clark and Oswald, 2002; Herbst and Ifcher, 2016; Margolis and Myrskylä, 2011), the available descriptive evidence is in general difficult to interpret.

---

[2] Please see Sacerdote and Feyrer (2008) for an overview and literature review on factors traditionally associated with fertility choices such as (1) the desire of parents to produce high-human capital children, (2) labor incomes which affect the opportunity cost of men's and women's time, (3) women's labor force participation, (4) child mortality, (5) availability of effective contraception for women, and (6) autonomy and status of women.

[3] Some strand of the literature even asked the question why people in fact have children despite seemingly negative net effects of parenthood on SWB highlighting the role of social norms (Morgan and King, 2001; Vanassche et al. 2013), biased affective forecasting (Gilbert, 2006), and old-age security motives (Herbst and Ifcher, 2016).

My data and empirical setting addresses these difficulties. I leverage data from UNICEF's Multiple Indicator Cluster Surveys (MICSs) using all available datasets in which women's complete birth history and SWB information were collected. The causal identification strategy is borrowed from the labor market and child quantity-quality trade-off literature (Aaronson et al., 2017; Angrist et al., 2010) and exploits quasi-experimental variation in family size due to preferences for a mixed sibling sex composition (Angrist and Evans, 1998).

Employing a local average treatment effect (LATE) framework, I establish several novel facts about the relationship between fertility and SWB. My first finding is that while, similar to other studies, the OLS estimates here indicate a negative relationship between fertility and SWB, the relationship is positive and statistically significant for the causal estimates for the subpopulation of compliers. In fact, instrumental variable (IV) estimates suggest that having a third child increases SWB between 0.45 and 0.58 units. Second, I provide empirical evidence that an increase in certain dimensions of life satisfaction, namely family life, friendship, and treatment by other people, are more closely related to the overall increase in SWB due the birth of a third child.

My study advances the literature on fertility and SWB in three ways. First, I provide causal evidence that addresses concerns regarding the likely endogeneity between fertility and SWB with the previous literature being confined to (i) cross-section and pooled regression models (Alesina et al. 2004; Aassve et al. 2012; Clark, 2007; Deaton and Stone, 2014; Di Tella et al. 2003; Di Tella and MacCulloch, 2006; Herbst and Ifcher, 2016; Margolis and Myrskylä, 2011; Stanca, 2012; Stevenson and Wolfers, 2009; Vanassche et al., 2013), (ii) panel models with fixed effects (Clark and Oswald, 2002; Stutzer and Frey, 2006), and (iii) event studies (Baetschmann et al., 2016; Clark et al., 2008; Clark and Georgellis, 2013; Frijters et al., 2011; Myrskylä and Margolis, 2014; Pedersen and Schmidt, 2014).[4]

Second, bearing in mind the main causal identification strategy, I in particular contribute to the substantially less-developed literature on the effect of fertility on SWB at the intensive margin (an additional child) in contrast to the extensive margin (becoming a parent). Estimates at the intensive margin are less frequently reported, with the majority of studies estimating coefficients on motherhood only, which hides possible differential effects by the intensive and extensive margin. Those studies that provide estimates at the intensive and extensive margin, either by simply controlling for the number of children or by estimating effects separately by birth order, show that both estimates tend to go into the same direction (Herbst and Ifcher, 2016; Margolis and Myrskylä, 2011; Myrskylä and Margolis, 2014; Stanca, 2012).

---

[4] Recent exceptions are Conzo et al. (2017) and Mu and Xie (2017) who implement an instrumental variable approach with the gender of the first child as an instrument as proposed by Lee (2008). The instrumentation strategy in those two studies is, however, problematic. In Conzo et al. (2017), a statistically insignificant coefficient on fertility is obtained based on a sample of 236 women which raises concerns about sufficient statistical power to detect any effect. In Mu and Xie (2017), the instrumentation strategy is even more problematic given the practice of sex-selective abortion that affects first and second births all over China (Chu, 2001) which is likely to violate the exclusion restriction.

Third, my analysis focuses on developing countries, for which only very little evidence on the relationship between fertility and SWB yet exists.[5] Scholars have argued that the underlying mechanism and relative importance of circumstantial factors such as cultural norms, the availability of formal and family child care mechanism, and access to effective contraceptives differs between developing and developed countries, with consequences for the effect of fertility on SWB (Margolis and Myrskylä, 2011). Furthermore, studying fertility behavior, more specifically the fertility-SWB link, in developing countries seems particularly rewarding given that these countries are fundamental to global fertility and population trends (UN, 2015b, 2017), as well as international economic growth and welfare improvements (WB, 2010).

I proceed as follows. In section 2, I describe the data. In section 3, I present the identification strategy and describe the main results. In Section 4, I show additional robustness checks and examine results for different dimensions of life satisfaction as well as heterogeneous treatment effects. Finally, I conclude in section 5.

2 Data

My analysis draws on data from UNICEF's Multiple Indicator Cluster Surveys (MICSs). Over the last two decades, the number of countries covered by MICS has increased while the core questionnaires have undergone several changes.

Following the introduction of the so-called "round 4" type of questionnaires in 2011, MICS included for the first time a module on SWB. Since MICS questionnaires are country-specific, there are notable differences across countries concerning the adoption of the SWB module. First of all, not all countries decided to include the SWB module in the data collection process. Second, some countries only adopted a reduced version of the module which excluded some SWB questions—in particular those related to specific dimensions such as health, friendship, and housing. Third, countries use different age thresholds for respondents of the SWB module. While the default SWB module collected information for women age 15 to 24 years only, several countries increased the age range (e.g., 15 to 49 years).

Starting with the "round 6" type of questionnaires in 2017/8, the implementation of SWB questions is more consistent across countries. SWB questions are asked to all respondents irrespective of age with the SWB module being consolidated to focus exclusively on two SWB outcomes (overall life satisfaction, happiness) only.

In addition to SWB information, the causal identification strategy requires detailed birth information such as each child's birth order, age, gender, and twinning status. This information is routinely collected in MICS's birth history module which is implemented in most but not all countries/surveys. Consequently, MICS rounds that did not implement the birth history module and respondents who did not answer the questions in the birth history module had to be dropped from the sample.

---

[5] The most widely cited study in this field is from Margolis and Myrskylä (2011) which uses data from the World Value Surveys. The authors report a negative correlation between fertility and women's SWB for developing countries. A negative relationship is uncovered too in Conzo et al. (2017) using data from rural Ethiopia while Mu and Xie (2017) find no statistically significant effect in their data from rural China.

Column 5 of Table 10 in the appendix depicts my analytical sample. The compiled dataset comprises 251,057 women with at least 2 children. As described above, some surveys administered the SWB module only to women below a certain age which consequently reduces the sample size. For example, while the MICS 5 dataset for Senegal comprises 820 women with at least two children ("extended sample"), only about 48 of these women ("core sample") were below the age of 25 and therefore eligible for the SWB module. Furthermore, out of these 48 women, about 42 answered the complete SWB module ("reduced sample") including SWB questions on friendship, health, and housing.

In total the core analytical sample comprises 102,798 women with at least two children from 35 countries.

2.1 Outcome definitions

In MICS, SWB information is collected on life satisfaction and happiness. The related questions use an ordinal response scale from 1 to 5 which I keep for the main analysis.[6] I focus mainly on the results obtained for life satisfaction while also presenting robustness checks for happiness. The choice of life satisfaction over happiness as the central indicator in the analysis was made to achieve consistency between the general SWB question and SWB questions focusing on particular domains of life such as family life, friendship, and health with the latter ones being framed as life satisfaction questions exclusively.

2.2 Descriptive statistics

Column 1 in Table 1 displays descriptive statistics at the mother level. On average, mothers are 32.7 years old and had their first birth at the age of 20.5. The majority of women (90%) is married at the time of the survey with most women (61%) residing in rural areas. Furthermore, Table 1 shows that most women state to be very satisfied with their lives with the average score on the different subjective well-being questions ranging from 3.44 to 4.58.[7]

As discussed in more detail in Section 3, the causal identification strategy rests on various LATE assumptions with the mixed-sibling sex composition of the first two born children functioning as instrument (Z). Columns 2 and 3 of Table 1 depict descriptive statistics by the main instrument (sibling sex composition) with column 10 showing p values for a test in differences between columns 2 and 3. With respect to the control

---

[6] The life satisfaction question asks "How satisfied are you with your life overall?". In MICS 4 and MICS 5 rounds, the response categories are very satisfied, somewhat satisfied, neither satisfied nor unsatisfied, somewhat unsatisfied, and very unsatisfied. In MICS 6 rounds, the response categories range from 1 (very unsatisfied) to 10 (very satisfied). In order to achieve consistency between the different MICS rounds, I recoded MICS 6 responses to be between 1 and 5. Across all MICS rounds, the happiness variable is based on the question "Taking all things together, would you say you are very happy, somewhat happy, neither happy nor unhappy, somewhat unhappy or very unhappy?". Comparing OLS and IV results for the MICS 6 sample using the original vs. the transformed scales, I find that the estimates are qualitatively unaffected by the transformation. Results are available from the authors upon request.

[7] Please see Table 11 for details on the construction of variables used in this study. Furthermore, please see Table 12 in the appendix for descriptive statistics regarding alternative parameters that characterize the distribution of each variable.

Table 1 Average characteristics and outcomes of always takers, never takers, and compliers

Variable	Overall	Same sex		3rd birth		AT	Comp. 1	Comp. 2	NT	t test
		Yes $Z=1$	No $Z=0$	Yes $D=1$	No $D=0$	$Z=0$ and $D=1$	$Z=0$ and $D=0$	$Z=1$ and $D=1$	$Z=1$ and $D=0$	(2) vs (3) (p value)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Age (years)	32.66	32.73	32.59	35.19	27.41	35.17	27.65	35.21	27.14	0.008
Age at 1st birth (years)	20.50	20.51	20.48	20.44	20.61	20.42	20.60	20.46	20.62	0.233
Married (%)	0.90	0.90	0.90	0.92	0.86	0.92	0.86	0.92	0.86	0.367
Rural (%)	0.61	0.61	0.61	0.63	0.58	0.63	0.58	0.63	0.58	0.557
HH size	6.98	7.00	6.97	7.33	6.26	7.33	6.27	7.34	6.24	0.071
Mothers' education										
Less than completed primary (%)	0.37	0.37	0.38	0.43	0.26	0.43	0.27	0.42	0.25	0.011
Completed primary (%)	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.27	0.26	0.249
Completed junior secondary (%)	0.20	0.20	0.20	0.17	0.26	0.17	0.26	0.17	0.27	0.975
Completed senior secondary or above (%)	0.16	0.16	0.16	0.13	0.22	0.13	0.21	0.14	0.22	0.047
Assets and dwelling										
Asset index	−0.10	−0.11	−0.10	−0.15	−0.01	−0.15	−0.01	−0.14	−0.02	0.292
Floor (main material)	1.37	1.37	1.37	1.37	1.38	1.37	1.38	1.37	1.38	0.880
Roof (main material)	1.76	1.77	1.76	1.79	1.72	1.79	1.72	1.78	1.72	0.702
Exterior wall (main material)	1.92	1.93	1.91	1.88	2.00	1.86	1.99	1.90	2.00	0.263
Cooking fuel (main type)	0.98	0.97	0.98	0.98	0.97	0.99	0.97	0.97	0.98	0.481
Owns TV (%)	0.73	0.73	0.73	0.75	0.69	0.74	0.69	0.75	0.69	0.521

Table 1 (continued)

Variable	Overall	Same sex		3rd birth		AT	Comp. 1	Comp. 2	NT	t test
		Yes $Z=1$	No $Z=0$	Yes $D=1$	No $D=0$	$Z=0$ and $D=1$	$Z=0$ and $D=0$	$Z=1$ and $D=1$	$Z=1$ and $D=0$	(2) vs (3) (p value)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Owns cell phone (%)	0.92	0.93	0.92	0.92	0.93	0.91	0.92	0.92	0.93	0.004
Subjective well-being										
Satisfaction with life overall (SWB)	3.44	3.45	3.43	3.31	3.71	3.30	3.69	3.33	3.74	0.011
Satisfaction with appearance	4.58	4.59	4.57	4.62	4.54	4.62	4.53	4.62	4.55	0.192
Satisfaction with treatment by others	4.38	4.39	4.37	4.48	4.29	4.47	4.28	4.48	4.30	0.010
Satisfaction with residence	4.33	4.34	4.32	4.43	4.22	4.44	4.21	4.43	4.24	0.106
Satisfaction with own health	4.40	4.40	4.39	4.40	4.39	4.40	4.38	4.40	4.40	0.354
Satisfaction with friendships	4.41	4.42	4.39	4.49	4.33	4.48	4.32	4.50	4.33	0.013
Satisfaction with family life	4.47	4.48	4.45	4.53	4.41	4.52	4.39	4.54	4.42	0.009
Happiness	4.10	4.10	4.09	4.04	4.20	4.04	4.19	4.05	4.22	0.010

AT, always takers; *Comp. 1*, never takers and untreated compliers; *Comp. 2*, always takers and untreated compliers; *NT*, never takers

Table 2 Composition of births

Variable	%	% with 3rd child	p value	Obs.
	(1)	(2)	(3)	(4)
First boy	51.51			102,798
First girl	48.49			102,798
Two boys	26.61	68.18	0.00	102,798
Two girls	23.50	70.53	0.00	102,798
Same sex	50.11	69.28	0.00	102,798
Mixed sex	49.89	65.65		102,798
3rd birth	67.47			102,798
Twins at 2nd birth	1.39			102,798

Table reports descriptive statistics for the "core" sample. Column 3 shows results from a test for differences in means (t test) in the treatment variable (3rd birth) against the "mixed sex" category

variables, I mostly find no statistically significant differences comparing women with $Z = 0$ and $Z = 1$ which provides some evidence for the absence of selection effects with respect of sibling sex composition in my core sample. Concerning the SWB outcome variables, I observe statistically significant differences with SWB values being slightly higher in the $Z = 1$ compared with the $Z = 0$ sample. In Section 4, which concerns instrument validity and selection effects, I discuss this latter finding in more detail.

Table 2 reports statistics on the variables used to construct instrumental variables. The gender of the first two children (2 boys or 2 girls) is the same for about half of all women (50.1% vs. 49.9%). A preference for continuing to have a third child among women whose first two children have the same gender is indicated in columns 2 and 3. On average, women who have 2 boys or 2 girls as their first children are about 3.6 percentage points more likely to have a third child.[8]

3 Main results

In this section, I first discuss the principal causal identification strategy. Then I show the main results.

3.1 Econometric approach

I am interested in the average effect of a binary treatment $D \in \{0, 1\}$ (having a third child) on the outcome Y (SWB). Under the plausible assumption of endogeneity, the

---

[8] Using census data from 101 countries, Aaronson et al. (2017) report first stage effects of "same sex" on having a 3rd child ranging from 0.092 to − 0.036 with an average effect across countries of 0.029. Cruces and Galiani (2007) report effects ranging from 0.0336 to 0.413 for Argentina and Mexico while Priebe (2011) reports an effect of 0.031 for Indonesia. Furthermore, please see Table 13 in the appendix for demographic statistics by country.

effect of D is confounded with unobserved factors that affect both the treatment (D) and the outcome (Y). Similar to studies in the context of female labor supply and child quantity-quality trade-offs (Aaronson et al. 2017; Angrist and Evans, 1998; Angrist et al., 2010), I use the same-sibling sex composition of the two first born children as a binary instrument $Z \in \{0, 1\}$ for D assuming that Z is correlated with D but not with Y.

Adopting a potential outcome notation I denote by $D(z)$, the potential treatment state for instrument $Z = z$. For each subject, only one of the two potential outcomes and treatment states is observed. As discussed in Angrist et al. (1996), the population can be characterized into four types (denoted by $T \in \{a, c, d, n\}$) depending on how the treatment state changes with the instrument. The compliers (c: $D(1) = 1$, $D(0) = 0$) react on the instrument in the intended way by having a third child when $Z = 1$ and abstaining from it when $Z = 0$. The always takers (a: $D(1) = 1$, $D(0) = 1$) always have a third child irrespective of the instrument state, the never takers (n: $D(1) = 0$, $D(0) = 0$) never have a third child, while the defiers (d: $D(0) = 1$, $D(0) = 1$) only have a third child when $Z = 0$.

The four types cannot be directly identified from the data. As shown in Imbens and Angrist (1994) under the further assumptions of IV validity (Eqs. 1 and 2), monotonicity (Eq. 3), and relevance (Eq. 4), the local average treatment effect (LATE) on the compliers is point identified.[9]

3.1.1 IV validity

$$Y \perp (D(1), D(0), Y(1, 1), Y(1, 0), Y(0, 1), Y(0, 0)) \tag{1}$$

$$Y(1, d) = Y(0, d) = Y(d) \; for \; d \in \{0, 1\} \tag{2}$$

3.1.2 Monotonicity

$$\Pr(D(1) \geq D(0)) = 1 \tag{3}$$

Equation 1 states that the instrument Z is as good as random and unrelated with factors affecting the treatment (having a third child) and/or the outcome (SWB) which implies that the error terms in the 2SLS model are independent from Z. Equation 2 stipulates that Z must not have a direct effect on Y other than through D, i.e., satisfy the exclusion restriction. In addition, Eq. 3 requires that the potential treatment state of any individual does not decrease in Z. Equation 3 rules out the existence of defiers (type $T = d$) because for the latter group $D(1) < D(0)$.

---

[9] For simplicity, I present in this section a discussion for the case in which the LATE of compliers is obtained without covariates. In the case that covariates are included, the required assumptions change accordingly. In general, the required assumptions in the case that covariates are included are less strict since only conditional but not full independence needs to hold. Please see Huber (2015) and Huber and Wüthrich (2019) for a more detailed discussion on this topic.

3.1.3 Relevance

$$E\left(D|Z=1-E(D|Z=0)\neq 0\right) \tag{4}$$

While the above assumptions on IV validity (Eqs. 1 and 2) and monotonicity (Eq. 3) imply the existence of compliers, this is in the empirical setting only satisfied if the first-stage effect of the instrument (same-sibling sex composition) is positive, statistically significant, and sufficiently large to shift the treatment decision (having a third child) at least for a subpopulation when switching from $Z=0$ to $Z=1$.

The LATE parameter of interest for compliers can be consistently estimated by 2SLS. Since the instrument is quasi-randomly assigned, the parameters of interest could in principle be estimated by 2SLS in a model without covariates. While I present such estimates, I believe that conditioning 2SLS estimates on a minimal set of covariates is most likely to fulfill the conditional independence assumption from Eqs. 1 and 2. The principal 2SLS specification is depicted in Eqs. 5 and 6 below

$$D_{c,i} = \gamma Z_{c,i} + X'_{c,i}\varphi + \alpha_c + \sigma_t + \varepsilon_{c,i} \tag{5}$$

$$SWB_{c,i} = \beta \widehat{D}_{c,i} + X'_{c,i}\lambda + \pi_c + \tau_t + \mu_{c,i} \tag{6}$$

where $D_{c,\,i}$ refers to the treatment variable (binary indicator of having a 3rd child) for woman i in country c. Furthermore, $X_{c,\,i}$ is a vector of controls including mother's age and a rural/urban indicator while $\propto_c$ and λ_c refer to country fixed effects, σ_t and τ_t are survey year fixed effects, and $\varepsilon_{c,\,i}$ and $\mu_{c,\,i}$ are the error terms. The first-stage effect of the instrument Z is captured by the parameter γ with Z being binary and taking the value 1 if the first two born children are either 2 boys or 2 girls. The main coefficient of interest is β which represents the LATE estimand on compliers.

3.2 First-stage estimates and instrument validity

3.2.1 Relevance

As discussed above for the credible identification of β, it is important to show that the instrument has a meaningful effect on fertility outcomes in order to trigger exogenous fertility increases.

OLS and IV first-stage effects are shown in the main results table (Table 3).[10] Women whose first two born children are of the same gender are about 3.5 percentage points more likely to have a third child. This finding is robust to whether (columns 4 to 6) or not (columns 1 to 3) I include control variables X. When splitting up the same-

---

[10] Throughout this paper, I present regression results with standard errors clustered at the community level. For rural areas, communities are equivalent to villages while in urban areas "community" refers to municipalities.

Table 3 OLS and IV (2SLS) estimates on life satisfaction

Parameter	Without covariates			With covariates		
	OLS	IV	IV	OLS	IV	IV
		Z = same sex	Z = 2 boys, 2 girls		Z = same sex	Z = 2 boys, 2 girls
	(1)	(2)	(3)	(4)	(5)	(6)
Mother had a 3rd child	− 0.017*	0.570**	0.142	− 0.009	0.579**	0.435*
	(0.009)	(0.241)	(0.209)	(0.011)	(0.261)	(0.234)
Observations	102,798	102,798	102,798	102,798	102,798	102,798
r2	0.2020	0.1721	0.1998	0.2043	0.1817	0.1914
First-stage results						
Same sex	0.0360*** (0.0030)	0.0360*** (0.0030)		0.0330*** (0.0020)	0.0330*** (0.0020)	
2 boys			0.0240*** (0.0030)			0.0190*** (0.0030)
2 girls			0.0500*** (0.0030)			0.0490*** (0.0030)
F-stat (1st stage)		172.0240	108.2720		189.7190	112.8520
Hansen's J (p value)			0.0000			0.1570

All regressions include year and country fixed-effects. Covariates include mother's age (dummy variables for each year), mother's age at first birth (dummy variables for 5-year intervals), gender of the 1st child, and location (rural vs. urban). Standard errors are depicted in parentheses and clustered at the community level

*/**/***Significance levels at 10/5/1% respectively

sibling sex composition instrument into 2 separate instruments (2 boys, 2 girls), the results (columns 3, 6) suggest that the fertility response is somewhat stronger for having 2 girls as first two births in comparison with having 2 boys.

3.2.2 Validity

A possible concern in any IV study is correlation between the instruments and potential outcomes, either because of confounding or violations of the exclusion restriction. In the following, I discuss and examine to what extent key assumptions of LATE identification are likely to hold.

Confounding variables If instruments are virtually randomly assigned, then IV estimates should be valid even without conditioning on covariates. Covariates, however, might be included because the conditional independence assumption and the exclusion restrictions are more likely to be valid after conditioning.[11] In my main specifications, I include covariates similar to Angrist and Evans (1998) and Angrist et al. (2010). More specifically, the main control variables relate to mother's age, the gender of the first child, and location (rural vs. urban areas).

Monotonicity and exclusion restrictions As discussed above, LATE identification requires assumptions 1 (conditional independence); Eqs. 1 and 2) and 2 (monotonicity; Eq. 3) to hold. In particular, it has been argued that Eq. 2 (assumption 1) might not hold and that the sibling sex composition of the first two children has a direct effect on the outcome variable. For instance, Rosenzweig and Wolpin (2000) argue that the sibling sex composition of the first two children can introduce investment and expenditure effects due to economies of scale in household expenditures through clothes-sharing that might be more likely among children of the same sex. Empirical evidence for this hypothesis is mixed (Bütikofer 2011; Priebe (2011) which suggests that such expenditure effects can exist but appear to be very country and context-specific. Furthermore, it might be that the sibling sex composition of the first two children directly influences SWB irrespective of the outlined expenditure channel.

While I cannot proof that assumptions 1 and assumptions 2 ultimately hold in my setting, I provide supporting empirical evidence from two different approaches.

Approach 1: Statistical tests

Assumptions 1 and 2 above provide testable implications of the identifying assumptions as shown in Eq. 7 below. Namely, $f(y, D = 1 | Z = 1) - f(y, D = 1 | Z = 0) = f(y(1), T = c)$ and $f(y, D = 0 | Z = 0) - f(y, D = 0 | Z = 1) = f(y(0), T = c)$ imply for y in the support of Y that:

$$f(y, D = 1 | Z = 1) \geq f(y, D = 1 | Z = 0), \quad f(y, D = 0 | Z = 0) \geq f(y, D = 0 | Z = 1) \quad (7)$$

If one or both of the inequalities depicted in Eq. 7 are violated, at least one of the three assumptions (IV validity (Eqs. 1 and 2) and monotonicity (Eq. 3)) is violated. To formally test

---

[11] Moreover, 2SLS estimates might be more precise if conditioning on covariates reduces some of the variability in the dependent variable (Angrist and Pischke, 2009).

for Eq. 7, Kitagawa (2015) proposes a test on resampling a variance-weighted two sample Kolmogorov-Smirnov-type statistic. An alternative testing approach is presented in Mourifie and Wan (2017) who show that a modified version of Eq. 7 fits the intersection bounds framework of Chernozhukov et al. (2013).[12] Both proposed tests apply to unconditional outcomes but can be adopted to test (8) conditional on observed covariates, if the latter are binned into subsets (Huber and Wüthrich, 2019). In contrast to Kitagawa (2015), the test of Mourifie and Wan (2017) can in addition be applied to the full covariate specification. In both tests, the "null hypothesis" shall not be rejected in order for Eq. 7 to hold.

Table 14 in the appendix shows the results from testing Eq. 7 implementing the proposed tests of Kitagawa (2015) and Mourifie and Wan (2017).[13] The first row shows results for the full sample while rows 2 to 22 depict test results for particular subsets of the data. From Table 14, I conclude that the LATE identification assumption appears to hold.

The statistical tests discussed above relate to the case in which the number of endogenous variables (D) equals the number of instruments (Z). With respect to testing IV validity (exogeneity) of the instruments, more traditional overidentification tests are available for the case that the number of instruments exceeds the number of endogenous variables. With respect to the sibling sex composition instrument, I can split the default instrument into two mutually exclusive instruments (2 boys and 2 girls). Estimating Eqs. 5 and 6 with these two instruments allows us therefore to report in addition results from conventional overidentification tests. Consequently, I report for all main regression results additional specifications (2 boys and 2 girls) and show the respective overidentification test statistics. As can been seen from the main regression tables below, I find that the instruments pass tests for exogeneity which I believe provides support that key LATE identification assumptions hold.

Approach 2: Testing plausible channels

With respect to the sibling sex composition instrument, it has been argued that the sibling sex composition may affect outcomes due to economies of scale in household expenditures through clothes-sharing that might become more likely among children of the same sex. The MICSs do not collect expenditure information that would allow us to directly test whether the sibling sex composition of the first two born children affects household expenditure patterns and levels. The surveys, however, gather information on various dwelling characteristics and asset possession. If economies of scale indeed exist and are of a meaningful economic size, one would expect that women whose first two children are of the same gender tend to be (a) better off than parents with a mixed-sibling sex composition or (b) able to invest more into the quality of its children.

In order to test for potential welfare effects stemming from the sibling sex composition, I estimate regressions similar to Eq. 5 above but with various types of dwelling characteristics and two asset possession indicators as dependent variables.

---

[12] Alternative testing constraints are derived in Huber and Mellace (2015).

[13] To implement the test of Kitagawa (2015), I rely on the R-package (version February 2019) kindly provided to us by Toru Kitagawa. To implement Mourifie and Wan (2017), I follow the authors' instructions and utilize the STATA package discussed in Chernozhukov et al. (2015). Implementing the "clrtest" command in STATA does not provide p values associated with testing the "null hypothesis" but "only" information on whether the "null hypothesis" was rejected or not. In Table I, report the results for testing the "null hypothesis" against a 90% significance level.

Table 15 (columns 1 to 5) in the appendix depict estimates for various dwelling characteristics. In general, I do not find that the sibling sex composition of the first two children is leading to improvements in dwelling characteristics. Perhaps one could argue that economies of scale do not necessarily show up in improvements of dwelling characteristics which involve substantial costs to households in developing countries but rather in relatively less expensive assets. As shown in Table 1 TVs (cell phones) are owned by 73 (92) percent of women in my sample which suggests that these are comparatively affordable items. Columns 6 to 7 report estimates on tv and cell phone possession. Again, there is no evidence for sibling sex composition being related to differences in asset possession.

Possibly, economies of scale in household expenditures due to sibling sex composition do not manifest themselves in asset possession and dwelling characteristics but in investments in children. To examine this channel, I run regressions similar to those presented in Table 15. Results from this exercise are displayed in Tables 16 and 17 in the appendix. Again, I find that sibling sex composition does not seem to be related to differences in outcomes—in this case, investments into 1st and 2nd born children.

I am aware that statistical and empirical tests on whether the exclusion restrictions for the sibling sex composition hold have its limitations. However, the analyses conducted above seem to suggest that there are no obvious violations of the LATE identification assumptions which are consistent with a causal interpretation of the IV estimates.

3.3 OLS and IV estimates: main results

Table 3 presents the main findings. Columns 1 and 4 show OLS results of the effect of the treatment variable (having a third child) on life satisfaction. I find that having a third child decreases life satisfaction between 0.009 and 0.017 units. In the model without covariates, the coefficient of interest is statistically significant at the 10% level while in the model with covariates it becomes statistically insignificant. The OLS results are largely in line with estimates of fertility on SWB from cross-sectional OLS regressions as discussed in the Introduction. Columns 2, 3, 5, and 6 depict the corresponding IV estimates for the sibling sex composition instrument. The IV estimates are larger and become positive and statistically significant once I condition on a minimal set of covariates. According to the IV estimates, having a third child increases life satisfaction by about 0.57 units[14].[15]

---

[14] It is noteworthy that conditioning on a minimal set of covariates not only seems to increase precision in my case but also makes it more likely that the conditional independence assumption holds. Comparing Hansen's J statistics between columns 3 and 6 suggests that the conditional independence assumption is only not rejected in the latter case.

[15] Lee (2018) points out that in the case of estimating LATE with multiple instruments (e.g., 2 boys and 2 girls) conventional standard errors need additional corrections as do statistics derived for overidentification. To assess whether the results are robust to adjusting standard errors as outlined in Lee (2018), I estimate regressions using the STATA package "mlr2sls" provided by Seojeong Lee. I find that both, coefficients and standard errors, are very close to the ones provided by standard STATA packages in my context. For instance, for the specification used in column 6, a coefficient of 0.457 with a standard error of 0.241 was obtained. Since the conventional overidentification test indicates that instruments easily pass and since the "mlr2sls" package does not allow to report first-stage and overidentification results, the remaining analysis continues using conventional STATA packages.

4 Robustness checks and extensions

4.1 Alternatives to 2SLS estimation

Since the dependent variable (life satisfaction) is ordinal with covariates being included, 2SLS might not give the best approximation of the conditional expectation function (CEF). In this subsection, I discuss results when using a semi-parametric (Abadie, 2003) and a non-parametric approximation (Frölich, 2007) for the CEF.[16] Table 18 in the appendix shows that the results are largely unaffected from changes in the estimation method.

4.2 Alternative covariate specifications

To examine whether results are robust to the inclusion of specific control variables, I re-run the main regression specifications by including additional covariates related to mothers' marital status, wealth quintile, and education level. Table 19 in the appendix illustrates that the main effects remain similar in terms of sign, magnitude, and statistically significance irrespective of the tested covariate specifications. Similar as before, coefficients are smaller in magnitude and statistically less significant when using 2 boys, 2 girls as instruments compared with the single instrument case. However, even in the specification with 2 instruments (2 boys, 2 girls), coefficients are positive and statistically significant at the 10% level.

4.3 Sensitivity analysis assuming exogeneity

Despite the results from Section 3.2 on instrument validity, I cannot ultimately prove that all LATE identification assumptions are fulfilled. With respect to assumptions regarding instrument exogeneity, I therefore provide bounds following Conley et al. (2012) to assess how sensitive the results are to violations of the exclusion restriction. The basic idea presented in Conley et al. (2012) can best be discussed in re-writing Eqs. 5 and 6 with the second stage including the additional term $\theta Z_{c,\,i}$.

$$D_{c,i} = \gamma Z_{c,i} + X'_{c,i}\varphi + \alpha_c + \sigma_t + \varepsilon_{c,i} \tag{8}$$

$$SWB_{c,i} = \beta \widehat{D}_{c,i} + \theta Z_{c,i} + X'_{c,i}\lambda + \pi_c + \tau_t + \mu_{c,i} \tag{9}$$

Previously, it was assumed that $\theta = 0$ resulting in point estimates for β. One way to loosen the IV assumptions is to remove the assumption that θ is precisely equal to zero.

---

[16] To implement the estimators of Abadie (2003) and Frölich (2007) I rely on the software packages provided by the two authors. In both cases estimates are only available when specifying a single instrument. Therefore, results are only obtain for IV specifications concerning the standard sex composition instrument but not for the case of splitting up the instrument into 2 separate dummy variables (2 boys, 2 girls). Furthermore, the results remain very similar too if I run IV ordered probit regressions. Results are available from the authors upon request.

In the framework of Conley et al. (2012), researchers can select priors for θ in a range of flexible ways.[17] Table 20 in the appendix provides bound estimates at the 95% significance level for β for various assumptions regarding the value of θ. For values of $\theta \le 0.0025$, β remains positive and statistically significant while for values of $\theta \ge 0.0025$, β loses statistical significance at the specified significance level.

To put the selected ranges of θ into better perspective, I relate to discussions and simulations presented in Conley et al. (2012) and Clarke and Matta (2018). The main regressions provided point estimates of β in the magnitude of 0.579. In this context, a value for θ of 0.0025 assumes a rather small direct effect of Z on Y (about 1/300 of the effect of β) with the simulations in Conley et al. (2012) (Clarke and Matta (2018)) assuming ratios of 1/10 (1/30) respectively. Therefore, while the main result of a positive and statistically significant effect of having a third child on SWB is robust to mild violations of the exclusion restriction (small values of θ), it is overall rather sensitive to assumptions about θ. Given that the first-stage effect of Z on D is usually rather small for the sibling sex composition instrument, the sensitivity of IV results to possible violations of the exclusion restriction is a common result though (Conley et al., 2012).

4.4 Alternative instruments: twinning

Comparing results obtained from the sibling sex composition instrument with alternative instruments provides a specification check since the omitted variable bias associated with each type of instrument should act differently with different instruments generating different average causal effects. One reason behind this is that the strength of the link between first-stage effects and the subpopulations affected by each underlying experiment differs as does the range of fertility outcomes induced by each instrument.

In this sub-section, I follow closely the empirical strategy outlined in Angrist et al. (2010) and provide alternative estimates for the effect of fertility on SWB using quasi-experimental variation in fertility due to twin births (Rosenzweig and Wolpin, 1980). Similar to Angrist et al. (2010), I estimate by 2SLS IV models in which the "twin" instrument substitutes for the sibling sex composition instrument in the first-stage and models in which the "twin" instrument and the sibling sex composition instrument enter jointly in the first stage.[18]

4.4.1 Twinning at birth order 2

Besides its function as specification check for omitted variable bias, the use of the "twin" instrument sheds further light on the external validity of the previous results.[19] Estimates generated by any particular IV strategy only captures effects on individuals

---

[17] To implement the approach of Conley et al. (2012), I utilize the STATA package "plausexog" and follow the instructions discussed in Clarke and Matta (2018).

[18] Alternative instruments used in the literature refer to infertility shocks (Aguero and Marks, 2008; Bratti et al., 2020). Since the MICS data does not allow to capture infertility shocks, I could not apply this instrument as robustness check.

[19] It should be noted though that with respect to both, the sex composition and twin instrument, reasonable external validity to various cultural contexts can be expected once context-level heterogeneity is taken into account (Aaronson et al., 2017; Bisbee et al., 2017; Dehejia et al., forthcoming).

Table 4 IV (2SLS) estimates on SWB with alternative instruments

Parameter	Z = same sex	Z = 2 boys, 2 girls	Z = twin	Z = same sex, twin	Z = 2 boys, 2 girls, twin
	(1)	(2)	(3)	(4)	(5)
Mother had a 3rd child	0.579** (0.261)	0.435* (0.234)	0.282* (0.159)	0.359*** (0.134)	0.327** (0.130)
Observations	102,798	102,798	102,798	102,798	102,798
r2	0.1817	0.1914	0.1988	0.1954	0.1969
First-stage results					
Same sex	0.0330*** (0.0020)			0.0340*** (0.0020)	
2 boys		0.0240*** (0.0030)			0.0240*** (0.0030)
2 girls		0.0500*** (0.0030)			0.0500*** (0.0030)
Twin			0.2410*** (0.0080)	0.2410*** (0.0080)	0.2970*** (0.0060)
F-stat (1st stage)	189.7190	112.8520	961.6910	584.2130	399.9840
Hansen's J (p value)		0.1570		0.3320	0.3240

All regressions include year and country fixed-effects. Covariates include mother's age (dummy variables for each year), mother's age at first birth (dummy variables for 5-year intervals), gender of the 1st child, and location (rural vs. urban). Standard errors are depicted in parentheses and clustered at the community level

*/**/***Significance levels at 10/5/1% respectively

affected by the instrument (Imbens and Angrist, 1994) which leads to concerns about the external validity of IV estimates (Moffit, 2005). As discussed in more detail in Angrist et al. (2010), twin estimates generate the average causal effect of treatment on the non-treated where treatment is defined as a dummy for having another child.[20] In contrast, sibling sex composition instruments identify the local average treatment for a different population of compliers in which the complier population, however, is less complete given that not all non-treated are affected by sibling sex composition.[21]

As shown in Table 2 among mothers who have at least 2 children, the twinning rate is about 1% at 2nd birth.[22] Furthermore, as depicted in Table 4, the twin instrument has a strong first-stage effect on fertility. In fact, a multiple second birth increases the likelihood of a mother to have a third child by about 25–30 percentage points.[23]

While Eq. 3 (monotonicity) is fulfilled by design with the twin instrument, it should be noted that nowadays several concerns about the validity of the exclusion restrictions (Eqs. 1 and 2) exist. For instance, it has been argued that parents might allocate resources away from twins towards older singleton-birth children (Rosenzweig and Zhang, 2009). If consequently the allocation pattern of resources across children influences women's subjective well-being than the twin instrument, it would potentially violate the exclusion restriction. Furthermore, it has been argued that selection into twinning is not random even after controlling for various demographic and household characteristics (Bhalotra and Clarke, 2019). Likewise, concerns about twinning having a direct effect on SWB might exist.

Table 4 presents results with respect to twin instrument and combinations of twin and sibling sex composition instrument.[24] Results across all specifications show that the effect of twins on SWB is positive and statistically significant (at the 10% level) too. The obtained LATE estimand is, however, smaller in magnitude compared with one obtained from the sibling sex instrument.

4.4.2 Twinning at different birth orders

While the twin instrument has faced some criticism regarding its validity (Bhalotra and Clarke, 2019; Rosenzweig and Wolpin, 2000; Rosenzweig and Zhang, 2009), it

---

[20] With twins there are no never-takers so the non-treated consist only of compliers with the twin instrument switched off. Because twinning is as good as randomly assigned, causal effects for the latter population are the same as causal effects on all compliers.

[21] One could plausibly argue that having an additional child is more likely to be a surprise in the twinning setting compared to the 'more deliberate' decision to have a third child in the sex composition setting. While twinning, then leads, ceteris paribus, to higher average fertility levels, one could argue that the share of 'unwanted' children is larger in the case of twinning than in the case of the sibling sex composition instrument. Therefore, the causal effect of fertility on SWB could be hypothesized to be more positive in the latter case.

[22] Reported twin rates at 2nd birth are 0.95% in Aaronson et al. (2017), 0.96% in Rosenzweig and Zhang (2009) for China, and 0.85–1.2% in Angrist and Evans (1998).

[23] Aaronson et al. (2017) find an average effect of 0.418, Bhalotra and Clarke (2019) report an effect of 0.789 in their sample of 72 countries while Ponczek and Souza (2012) report an effect of 0.852 for their Brazilian data. In comparison to these studies the effect obtained in my sample is relatively small. I believe there are two principal reasons behind this difference. First, the MICS are more likely to be conducted in high-fertility developing countries compared to the DHS. Second, the sample includes a relatively larger share of women with completed fertility.

[24] As discussed in Angrist et al. (2010) and Bhalotra and Clarke (2019), the selected control variables for twin instrument are largely similar to the ones chosen for the sex composition instrument.

Table 5 OLS and IV (2SLS) estimates on happiness

Parameter	Without covariates		With covariates	
	OLS	IV	OLS	IV
		Z = same sex		Z = same sex
	(1)	(2)	(3)	(4)
Mother had a 3rd child	− 0.030*** (0.007)	0.409** (0.160)	0.007 (0.008)	0.441** (0.173)
Observations	102,798	102,798	102,798	102,798
r2	0.0773	0.0344	0.0827	0.0509
First-stage results				
Same sex	0.0360*** (0.0030)	0.0360*** (0.0030)	0.0330*** (0.0020)	0.0330*** (0.0020)
F-stat (1st stage)		172.0240		189.7190

All regressions include year and country fixed-effects. Covariates include mother's age (dummy variables for each year), mother's age at first birth (dummy variables for 5-year intervals), gender of the 1st child, and location (rural vs. urban). Standard errors are depicted in parentheses and clustered at the community level

*/**/***Significance levels at 10/5/1% respectively

provides the advantage, conditional on identification assumptions hold, that researchers can explore the external validity of the previously obtained IV estimates. More precisely, the twin instrument can be applied to other samples than my main analytical sample (women with at least 2 children—"2+ sample"). In Table 21 in the appendix, I present alternative IV results from twinning at 1st (3rd) birth in samples of women with at least one child (three children—"3+ sample"). While I find that IV coefficients of having another child are positive across all specifications, only the specifications for higher birth orders are statistically significant. Hence, the results suggest that concerns about the external validity of my 2+ sample results might be justified.

4.5 Results on happiness

Life satisfaction and happiness are arguably mostly hedonic measures of SWB based on pleasure. While questions on life satisfaction are considered to be linked closer to cognitive aspects of well-being, i.e., judgements one can make about one's life, happiness measures are more closely linked to pure emotional hedonic pleasure (Clark and Senik, 2011). Despite its conceptual differences, both measures are usually highly correlated.[25] In the core sample, the correlation between the two measures is 0.52. Since life satisfaction and happiness capture different aspects of SWB and given that the correlation between the two measures is not perfect, I re-estimate the main results with happiness as dependent variable. As shown in Table 5, results of the effect of having a third child on happiness are very similar to those in Table 3. As before, OLS estimates tend to be very small and marginally statistically significant or statistically

---

[25] For instance, Clark and Senik (2011) using data for 21 European countries report a correlation coefficient of 0.61 between life satisfaction and happiness.

Table 6 IV (2SLS) estimates on different SWB dimensions

Panel A: Z = Same sex

Parameter	Overall	Family	Friendship	Health	Residence	People	Appearance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mother had a 3rd child	0.579** (0.261)	0.812* (0.433)	0.850* (0.453)	0.295 (0.427)	0.614 (0.470)	0.847* (0.451)	0.386 (0.368)
Observations	102,798	30,073	30,073	30,073	30,073	30,073	30,073
First-stage results							
Same sex	0.0330*** (0.0020)	0.0240*** (0.0050)	0.0240*** (0.0050)	0.0240*** (0.0050)	0.0240*** (0.0050)	0.0240*** (0.0050)	0.0240*** (0.0050)
F-stat (1st stage)	189.719	26.6830	26.6830	26.6830	26.6830	26.6830	26.6830

All regressions include year and country fixed-effects. Covariates include mother's age (dummy variables for each year), mother's age at first birth (dummy variables for 5-year intervals), gender of the 1st child, and location (rural vs. urban). Standard errors are depicted in parentheses and clustered at the community level

*/**/***Significance levels at 10/5/1% respectively

Table 7 IV (2SLS) estimates for heterogeneous treatment effects

Panel A: Wealth level

Parameter	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	(1)	(2)	(3)	(4)	(5)
Mother had a 3rd child	1.041 (0.670)	0.419 (0.499)	0.998* (0.544)	− 0.401 (0.610)	0.889 (0.556)
Observations	29,921	23,216	20,242	17,113	12,306
r2	0.2357	0.2208	0.1333	0.1788	0.0696
First-stage results					
Same sex	0.0240*** (0.0040)	0.0350*** (0.0050)	0.0360*** (0.0050)	0.0350*** (0.0060)	0.0430*** (0.0070)
F-stat (1st stage)	29.5800	51.3010	44.1380	34.4950	34.9400

Panel B: Education

Parameter	Less than primary	Primary	More than primary
	(1)	(2)	(3)
Mother had a 3rd child	0.220 (0.467)	0.979* (0.568)	0.510 (0.336)
Observations	38,344	27,010	37,444
r2	0.1731	0.1878	0.1549
First-stage results			
Same sex	0.0300*** (0.0040)	0.0290*** (0.0050)	0.0400*** (0.0040)
F-stat (1st stage)	74.2970	40.2630	91.7830

All regressions include year and country fixed-effects. Covariates include mother's age (dummy variables for each year), mother's age at first birth (dummy variables for 5-year intervals), gender of the 1st child, and location (rural vs. urban). Standard errors are depicted in parentheses and clustered at the community level

*/**/***Significance levels at 10/5/1% respectively

Table 8 OLS and IV (2SLS) estimates on life satisfaction by age

Parameter	Age < 30 years			Age > 29 years		
	OLS	IV	IV	OLS	IV	IV
		Z = same sex	Z = 2 boys, 2 girls		Z = same sex	Z = 2 boys, 2 girls
	(1)	(2)	(3)	(4)	(5)	(6)
Mother had a 3rd child	−0.030** (0.014)	0.398 (0.390)	0.463 (0.365)	−0.002 (0.017)	0.651* (0.334)	0.391 (0.299)
Observations	40,678	40,678	40,678	62,120	62,120	62,120
r2	0.2602	0.2430	0.2374	0.1532	0.1323	0.1456
First-stage results						
Same sex	0.0320*** (0.0040)	0.0320*** (0.0040)		0.0350*** (0.0030)	0.0350*** (0.0030)	
2 boys			0.0190*** (0.0060)			0.0210*** (0.0040)
2 girls			0.0450*** (0.0060)			0.0510*** (0.0040)
F-stat (1st stage)		57.0130	32.6370		161.9090	96.4210
Hansen's J (p value)			0.6440			0.0440

All regressions include year and country fixed-effects. Covariates include mother's age (dummy variables for each year), mother's age at first birth (dummy variables for 5-year intervals), gender of the 1st child, and location (rural vs. urban). Standard errors are depicted in parentheses and clustered at the community level

*/**/***Significance levels at 10/5/1% respectively

insignificant. In contrast, IV estimates are all positive and statistically significant. If there is a difference, then the estimates suggest that the effect of having a third child on SWB is slightly smaller in magnitude for happiness than for life satisfaction.[26]

4.6 Results on other dimensions of life satisfaction

As there are different aspects in life, having a third child might affect certain dimensions of life satisfaction but not necessarily others (Adler et al. 2017; Van Praag et al., 2003). In this subsection, I present 2SLS estimates for six different dimensions of life satisfaction—more specifically satisfaction with family life, friendship, health, current residence, treatment by other people, and appearance. Since questions on the above six dimensions were only included in some of the MICS 4 and MICS 5 round of questionnaires, the number of observations drops markedly to about 30,000. Table 6 displays the respective results.

While I find that all coefficients are positive, only the coefficients for friendship, family life, and treatment by other people are statistically significant.

It has often been emphasized in the literature that having children can be rewarding and burdensome at the same time and that positive and negative effects of having children could potentially offset each other which could explain the absence of a positive correlation between fertility and SWB in many developed countries. In this context, sociological and psychological studies stress that a positive impact of children on SWB often operates through an increase in social connectedness after having children (Gallagher and Gerstel, 2001; Umberson and Gove, 1989; Nomaguchi and Milkie, 2003). The results in Table 6 are compatible with and supportive of this line of reasoning.

With respective to factors that explain potentially negative effects of children on women's SWB in developed countries, it has been pointed that having children can lead to reductions in spousal affection (Grossbard and Mukhopadhyay, 2013), decreased marital satisfaction (Keizer et al., 2010), decreased sexual activity (Gettler and Oka, 2016), decreased time for work and leisure (Connelly and Kimmel, 2015; Hansen, 2012), and increasing financial pressure (Stanca, 2012; Pollmann-Schult, 2014). Unfortunately, the MICS data does not allow us to examine these channels closer.[27]

4.7 Heterogeneous treatment effects by wealth level and mothers' education

There are several reasons why the effect of having an additional child on women's SWB is likely to depend on a woman's education and wealth level. First, fertility outcomes and preferences in developing countries show a strong education and wealth

---

[26] Please see Table 22 in the appendix for IV estimates based on the twin instrument. IV estimates based on the twin instrument indicate a positive and statistically significant relationship between fertility and happiness. Similar to the previous results on life satisfaction, the twin estimates are, however, smaller in magnitude compared with specifications that use mixed sibling sex composition as instrument.

[27] In line with the reasoning that having children creates adverse effects on the financial situation of parents, Stanca (2012) finds a stronger negative effect of fertility and financial satisfaction than of fertility on general life satisfaction. The SWB module in the MICS asks a question on satisfaction with current income which, however, only refers to a mother's own income and is only asked to those women holding a job (about 60% of women in the core sample). For these reasons and given that selection into jobs is endogenous to fertility choices and levels, I decided to not analyze this questionnaire item.

gradient with poorer and less educated women tending to have more children (actual fertility) and wanting more children (wanted and desired number of children) (Bongarts, 2003; Bongarts and Casterline, 2013). Beside many other factors, traditional social norms that encourage and reward having a third child are more common among women from poorer socio-economic and educational backgrounds (Canning et al., 2013) which could result in stronger (more positive) effects on SWB for poorer and less educated women (Balbo and Arpino, 2016).

Table 9 Testing for selection and global external validity: *p* values

	δ_{DZ}	δ_D	δ_Z
	(1)	(2)	(3)
Age at 1st birth (years)	0.067	0.000	0.011
Married	0.435	0.000	0.579
HH size	0.558	0.000	0.949
Mothers' education			
Less than completed primary (%)	0.518	0.000	0.282
Completed primary (%)	0.107	0.000	0.246
Completed junior secondary (%)	0.115	0.032	0.265
Completed senior secondary or above (%)	0.551	0.000	0.250
Assets and dwelling			
Asset index	0.656	0.000	0.625
Floor (main material)	0.866	0.000	0.704
Roof (main material)	0.464	0.000	0.484
Exterior wall (main material)	0.960	0.842	0.664
Cooking fuel (main type)	0.353	0.019	0.920
Owns TV (%)	0.652	0.000	0.793
Owns cell phone (%)	0.179	0.000	0.017
Health			
# vaccinations (1st born)	0.096	0.396	0.308
# vaccinations (2nd born)	0.748	0.000	0.910
Subjective well-being			
Satisfaction with life overall (SWB)	0.677	0.314	0.284
Satisfaction with appearance	0.740	0.737	0.354
Satisfaction with treatment by others	0.860	0.893	0.123
Satisfaction with residence	0.092	0.049	0.053
Satisfaction with own health	0.627	0.995	0.364
Satisfaction with friendships	0.262	0.326	0.534
Satisfaction with family life	0.902	0.584	0.149
Happiness	0.826	0.564	0.195

All regressions include year and country fixed-effects. Covariates include mother's age (dummy variables for each year), mother's age at first birth (dummy variables for 5-year intervals), gender of the 1st child, and location (rural vs. urban). Standard errors are clustered at the community level

Second, it has been pointed out that access to institutional child care arrangements (formal and informal) can affect the relationship between fertility and SWB (Aassve et al., 2005; Bertrand, 2013; Morgan and King, 2001) with better access being correlated with higher SWB (Glass et al., 2016; Margolis and Myrskylä, 2011). While better off and better educated women are more likely to have access to formal child care arrangements such as kindergartens and nannies, these arrangements play overall a less important role in developing countries in which the vast majority of the population relies on informal means of child care provision via other family members, relatives, friends, and neighbors. To what extent access to child-care support institutions differs along education and wealth gradients is less obvious in a developing country context (Roby, 2011; ODI, 2016).

Ultimately, it is an empirical question to what extent and how the effect of having a third child on SWB differs by wealth and education levels. Table 7 presents results from the LATE framework for different samples. By presenting split-sample estimates, I lose considerable statistical power to detect any statistical significant effects. Despite this limitation, I believe that studying changes in the sign and magnitude of coefficients can still be interpreted albeit with greater care.

Results presented in Table 7 show no obvious pattern among the split sample estimates—neither along the education nor the wealth gradient. With the exception of the sample on wealth quintile 4 (the 2nd richest wealth group), all estimates show positive signs and are of sizeable magnitude.

4.8 Heterogeneous treatment effects by mother's age

The relevant literature on fertility and mother's subjective well-being discusses the role of mothers' age from three different angles. The first strand follows the literature that studies the effect of important life events and shocks on subjective well-being (Clark et al., 2008, 2016). Often, this literature stresses the importance of adaptation processes and therefore distinguishes between short-and long-term effects of a particular life event. In this context, several papers examine the so-called "baseline-hypothesis" that stipulates that life events only have a temporary effect on subjective well-being. Following the rational of the "baseline hypothesis," the effect of fertility on SWB should be smaller in magnitude for older women given that their children are on average already older too (Baetschmann et al., 2016).

The second strand of the literature emphasizes that raising children can be particularly stressful to parents in first years of life and that parents are able to enjoy the benefits of having children in the long-run (Buddelmeyer et al., 2018; Herbst and Ifcher, 2016; Myrskylä and Margolis, 2014). According to this literature, the effect of fertility on SWB should become more positive for older woman given that their children are on average already older too.

A third strand of the literature argues that the timing of having children is reflective of social norms and individual preferences. For instance, having children early in life might be more reflective of following traditional social norms. In contrast, having children later in life might more closely link to reflecting individual preferences regarding fertility. For instance, Cetre et al. (2016) argue that women who have children later in life are, ceteris paribus, more happy than younger mothers since for older women the decision to have another child is rather reflective of their own choice and

preferences. Borrowing from Cetre et al. (2016), I would expect that the effect of fertility on SWB might be more positive for older woman conditional of having children of the same age.

Ultimately, the relationship between fertility and SWB is an empirical question. Table 8 depicts OLS and IV split-sample estimates for women below (columns 1–3) and equal or above (columns 4–6) 30 years of age. The results show that there are no major differences in the fertility-SWB relationship between younger and older women in my sample. All the obtained estimates seem to support the view that older women are deriving higher subjective well-being compared with younger ones.

4.9 Selection and treatment effect heterogeneity

Section 3 showed that OLS and IV estimates of having a third child on SWB are quite different. Naturally in this context OLS and IV estimates are difficult to compare since the former is for the whole population while the IV estimate refers to the complier subpopulation. The circumstance, however, that the LATE estimate differs strongly from OLS points to the circumstance that OLS is likely to overestimate the negative effect of fertility on SWB.

To more formally explore whether the obtained LATE for compliers provides additional information on the causal effect of fertility on SWB for the overall population, I employ simple tests derived from the marginal treatment effect literature (Björklund and Moffitt, 1987; Brinch et al., 2017; Heckman and Vytlacil, 1999, 2005, 2007) and more specifically the work of Kowalski (2016a, b, 2019).

Based on the results presented above, I believe that the LATE of compliers is internally valid. If the LATE is internally valid, then selection into treatment (having a 3rd child) is random among compliers, but selection need not be random in the experiment as a whole. For instance, always takers (never takers) select into (out of) treatment regardless of the random assignment. Moreover, while the LATE for compliers does not depend on the treated outcome of always takers or the untreated outcome of never takers, these latter outcomes can be informative about selection and treatment effect heterogeneity. A difference in the average untreated outcomes of compliers and never takers provides evidence of selection while a difference in the average treated outcomes of compliers and always takers provides evidence of selection, treatment effect heterogeneity, or both (Kowalski, 2016a, b).

Assuming weak monotonicity and linearity in untreated and treated outcomes from always takers to compliers to never takers, Kowalski (2016a, b) proposes a simple difference-in-difference test as shown in Eq. 11 to test for selection effects and treatment effect heterogeneity.

$$Y_{c,i} = \alpha_0 + \delta_{DZ}DxZ + \delta_D D + \delta_Z Z + X'_{c,i}\lambda + \pi_c + \tau_t + \mu_{c,i} \qquad (10)$$

As before, D refers to the treatment (having a 3rd child), Z refers to the child sex composition (same sex), X represents covariates, π are country fixed effects, and μ represents the error term. Y in Eq. 11 stands, depending on the estimated regression, for background and outcome variables. As discussed in more detail in Kowalski (2016a, b,

2019), the δ coefficients provide evidence for selection effects ($\delta_Z \neq 0$), treatment effect heterogeneity ($\delta_{DZ} \neq 0$ in the case of Y representing outcome variables), and different relationships between baseline and intervention take-up ($\delta_{DZ} \neq 0$ in the case of Y representing background variables).

Table 9 shows p values for the δ coefficients from estimating the described diff-in-diff framework using OLS. p values in column 1 for δ_{DZ} are quite large for all relevant outcome variables (subjective well-being and happiness outcomes) which suggests that treatment effect heterogeneity is not necessarily responsible for differences between the OLS and IV estimates. Results concerning p values of δ_Z from regressions on outcome variables, however, are relatively smaller and indicate that selection effects might be present which would imply that the underlying complier population appears to be fundamentally different to the always taker and never taker subpopulations. Therefore, I conclude that the obtained LATE result cannot necessarily be extrapolate from the complier to the overall population.

5 Conclusion

I study the causal link between fertility and mothers' subjective well-being at the intensive margin. More specifically, I examine how women's SWB responds to the birth of a 3rd child using a unique sample of all suitable UNICEF MICS datasets available. Following the seminal work of Angrist and Evans (1998), my causal identification strategy exploits variation in fertility at 3rd birth due preferences for a mixed-sibling sex composition.

Causal LATE estimates for the complier population indicate that having a 3rd child affects SWB positively and in a meaningful magnitude. Furthermore, my analysis shows that similar effects can be found for other dimensions of well-being such as satisfaction with family life, friendship, and treatment by other people which are in line with findings from sociology and psychology that emphasize that having children contribute to social connectedness.

Taking into account that my pooled dataset spans 35 countries with very diverse social, cultural, and economic contexts, I believe that the results provide considerable evidence for the external validity of the estimates for the subpopulation of compliers.

The causal estimates are derived from standard instrumental variable strategies. As it is common in this context, it is impossible to rule out all possible concerns regarding the violation of identifying assumptions. In this paper, I tried to address the concerns relying on various statistical and econometric tests. While the applied tests and analyses seem to suggest that the relevant identifying assumptions by and large hold, I find that the results are sensitive to possible misspecifications and sometimes fulfill necessary identification assumptions only at the margin.

Furthermore, there are two noteworthy limitations of my study that I would like to point out. First of all, I examine the causal relationship at the intensive margin. While this relationship is important and relevant, it does not necessarily shed light on the causal effect of having children or not (extensive margin). Second, the study is data-constraint and cannot rigorously investigate all possible channels that drive the difference between OLS and IV estimates. Clearly, OLS and LATE identify effects for different populations with LATE taking in addition possible endogeneity problems into

account. While I find evidence that the complier population differs from the overall population, I believe that the obtained results are compatible with various sociological, economic, and psychological explanations of why children can provide joy and pleasure to their parents. Nonetheless, more future work needs to be conducted in understanding causal estimates of fertility on subjective well-being.

Compliance with ethical standards

Conflict of interest   The author declares no conflict of interest.

Disclaimer   The views expressed in this paper are those of the authors alone and do not present the views of GIGA.

Appendix: Background tables

Table 10   Overview on MICS datasets and analytical sample

| Number | Dataset | Region | Year | Round | # women with 2 children | | | SWB |
| | | | | | Core | Reduced | Extended | Core |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	Lebanon	Asia	2011	MICS 4	118	103	1923	4.34
2	Tunisia	Africa	2011	MICS 4	104	86	3753	4.50
3	Madagascar	Asia	2012	MICS 4	1887	1858	1887	3.82
4	Moldova	Europe	2012	MICS 4	86	86	2624	4.53
5	Ukraine	Europe	2012	MICS 4	135	135	2470	4.02
6	Kenya	Africa	2013	MICS 5	250	244	2037	4.27
7	Kosovo	Europe	2013	MICS 5	203	202	3426	4.53
8	Malawi	Africa	2013	MICS 5	2447	2376	15,610	4.50

Table 10 (continued)

Number	Dataset (1)	Region (2)	Year (3)	Round (4)	# women with 2 children			SWB
					Core (5)	Reduced (6)	Extended (7)	Core (8)
9	Nepal	Asia	2013	MICS 5	644	640	7895	4.06
10	Benin	Africa	2014	MICS 5	1099	1089	9512	3.93
11	Cameroon	Africa	2014	MICS 5	916	887	5584	4.05
12	Dominican Republic	LAC	2014	MICS 5	2522	2444	16,605	4.57
13	El Salvador	LAC	2014	MICS 5	6722	6703	6728	4.64
14	Guinea Bissau	Africa	2014	MICS 5	857	857	5879	4.69
15	Guyana	LAC	2014	MICS 5	396	384	2822	4.56
16	Kyrgyzstan	Asia	2014	MICS 6	418	417	4133	4.67
17	Sao Tome and Principe	LAC	2014	MICS 5	242	232	1801	4.24
18	Swaziland	Africa	2014	MICS 5	504	498	4820	4.36
19	Congo	Africa	2014	MICS 5	491	484	3192	3.96
20	Mali	Africa	2015	MICS 5	1975	1968	11,234	4.13
21	Congo	Africa	2015	MICS 5	549	541	3991	4.08
22	Mauritania	Africa	2015	MICS 5	1051	1044	7636	4.69
23	Senegal	Africa	2015	MICS 5	218	204	3068	4.35
24	Nigeria	Africa	2016	MICS 5	2470	2457	18,843	4.48
25	Paraguay	LAC	2016	MICS 5	405	399	3628	4.56
26	Cote d'Ivoire	Africa	2016	MICS 5	1037	1025	6949	4.19
27	Mongolia	Asia	2016	MICS 5	79	79	1766	4.44
28	Guinea	Africa	2016	MICS 5	892	890	5962	4.31
29	Senegal	Africa	2016	MICS 5	48	42	820	4.63
30	Nigeria	Africa	2017	MICS 5	208	208	1481	4.52
31	Sierra Leone	Africa	2017	MICS 6	9380	0	9920	2.60
32	Pakistan	Asia	2017	MICS 6	3671	0	3977	2.96
33	Iraq	Asia	2018	MICS 6	14,810	0	16,101	2.87
34	Gambia	Africa	2018	MICS 6	7170	0	7551	2.64
35	Mongolia	Asia	2018	MICS 6	6536	0	7048	3.32
36	Suriname	LAC	2018	MICS 6	3179	0	3571	3.75
37	Pakistan	Asia	2018	MICS 6	28,784	0	32,378	3.29
				Total:	102,798	30,073	251,057	3.44

Table 10 reports number of observations from the pooled MICS data that are used in the analyses. "Core" refers to the core sample which uses all observations for which SWB information (overall life satisfaction and happiness) is available. "Reduced" refers to the reduced sample which only considers individuals/surveys that answered the full SWB module (additional SWB questions). "Extended" refers to the extended sample which in addition to the "core" sample includes observations for which SWB information was not collected. Mostly, the difference in sample size between the "core"/"reduced" samples vs. "extended" sample arises due to skipping patterns in the country-specific MICS questionnaire (SWB questions are often asked to younger women only). Column 8 reports mean values on satisfaction with life overall (main outcome indicator)

Table 11 Variable construction from MICS data

Variable (1)	Description (2)
Age	Dummy variables are constructed for each age; e.g., d14 = 1 if woman is 14 years old and 0 otherwise
Age at 1st birth	Dummy variables are constructed for 5-year age intervals; intervals are < 15 years, 15–19, 20–24, 25–29, 30–34,35–39, > 39
Married	Woman is married; binary (1 = yes, 0 = no)
Rural	Household resides in rural area; binary (1 = rural, 0 = urban)
Primary (woman)	Highest degree is primary schooling; binary (1 = yes, 0 = no)
Junior secondary (woman)	Highest degree is junior secondary schooling; binary (1 = yes, 0 = no)
Senior secondary (woman)	Highest degree is secondary schooling or above; binary (1 = yes, 0 = no)
Asset index	Provided by MICS Unicef; larger positive values indicate high welfare
Wealth quintiles	Dummy variable for each wealth quintile; quintiles are country-specific and based on asset index
Floor	Quality of floor material; ordinal (0 = poor, 1 = medium, 2 = good)
Roof	Quality of roof material; ordinal (0 = poor, 1 = medium, 2 = good)
Wall	Quality of wall material; ordinal (0 = poor, 1 = medium, 2 = good)
Cooking	Quality of cooking device; ordinal (0 = poor, 1 = medium, 2 = good)
TV	Owns at least 1 TV; binary (1 = yes, 0 = no)
Cellphone	Owns at least 1 cellphone; binary (1 = yes, 0 = no)
# vaccinations	Additive index constructed as sum of the following vaccinations: 1 BCG, 3× polio, 3× DPT, 1× measles
SWB	MICS 4 and 5 rounds: original scale is simply reversed to facilitate interpretation
	scale 1 (very unsatisfied), 2 (somewhat unsatisfied), 3 (neither satisfied/unsatisfied),
	4 (somewhat satisfied), 5 (very satisfied)
	MICS 6 rounds: original scale is from 0 (very unsatisfied) -10 (very satisfied).
	Conversion to MICS 4/5 scale: 0,1,2 are 1; 3,4 are 2; 5,6 are 3; 7,8 are 4; 9,10 are 5
Happiness	Original scale is simply reversed to facilitate interpretation
	scale: 1(very unhappy), 2 (somewhat unhappy), 3 (neither happy/unhappy),
	4 (somewhat happy), 5 (very happy)
	All MICS rounds used the same classification

Construction of variables follows MICS codebooks as provided by UNICEF

Table 12 Summary statistics

Variable	Mean	Median	SD	Min.	Max.	Obs.
	(1)	(2)	(3)	(4)	(5)	(6)
Age (years)	32.66	32.00	8.57	15.00	49.00	102,798
Age at 1st birth (years)	20.50	20.00	4.40	11.00	45.00	102,798
Married (%)	0.90	1.00	0.30	0.00	1.00	102,798
Rural (%)	0.61	1.00	0.50	0.00	3.00	102,798
HH size	6.98	6.00	3.29	1.00	15.00	102,798
Mothers' education						
Less than completed primary (%)	0.37	0.00	0.48	0.00	1.00	102,798
Completed primary (%)	0.26	0.00	0.44	0.00	1.00	102,798
Completed junior secondary (%)	0.20	0.00	0.40	0.00	1.00	102,798
Completed senior secondary or above (%)	0.16	0.00	0.37	0.00	1.00	102,798
Assets and dwelling						
Asset index	− 0.10	− 0.19	0.94	− 8.89	5.38	102,798
Floor (main material)	1.37	2.00	0.91	0.00	2.00	99,715
Roof (main material)	1.76	2.00	0.62	0.00	2.00	99,717
Exterior wall (main material)	1.60	2.00	0.72	0.00	2.00	99,717
Cooking fuel (main type)	0.92	1.00	0.95	0.00	2.00	102,798
Owns TV (%)	0.73	1.00	0.45	0.00	1.00	89,404
Owns cell phone (%)	0.87	1.00	0.33	0.00	1.00	102,798
Health						
# vaccinations (1st born)	7.48	8.00	1.44	0.00	8.00	2932
# vaccinations (2nd born)	6.72	8.00	2.21	0.00	8.00	11,834
Subjective well-being						
Satisfaction with life overall (SWB)	3.44	4.00	1.50	1.00	5.00	102,798
Satisfaction with appearance	4.58	5.00	0.77	1.00	5.00	30,403
Satisfaction with treatment by others	4.38	5.00	0.90	1.00	5.00	30,437
Satisfaction with residence	4.33	5.00	1.00	1.00	5.00	30,397
Satisfaction with own health	4.40	5.00	0.93	1.00	5.00	30,441
Satisfaction with friendships	4.41	5.00	0.87	1.00	5.00	30,301
Satisfaction with family life	4.47	5.00	0.84	1.00	5.00	30,416
Happiness	4.10	4.00	0.94	1.00	5.00	102,798

Summary statistics are based on the core sample (women with at least 2 children and information on SWB). Variables on subjective well-being are ordinal from 1 (very unsatisfied) to 5(very satisfied). Note that subjective well-being values in this paper are in reverse order from the original MICS data to facilitate interpretation. Please see Table 11 for the coding of variables

Table 13 Composition of births

#	Country	First birth		Second birth				% with 3rd birth given:				% twins at:		Obs
		% Boy	% Girl	% 2 Boys	% 2 Girls	% SS	% MS	2 Boys	2 Girls	SS	MS	2nd	3rd	
(1)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1	Lebanon	49.15	50.85	27.12	26.27	53.39	46.61	25.00	45.16	34.92	32.73	0.85	3.39	118
2	Tunisia	48.08	51.92	27.88	24.04	51.92	48.08	68.97	76.00	72.22	48.00	1.92	0.96	104
3	Madagascar	48.75	51.25	24.06	26.02	50.08	49.92	81.06	80.04	80.53	75.48	1.11	1.27	1887
4	Moldova	50.00	50.00	26.74	26.74	53.49	46.51	30.43	39.13	34.78	20.00	0.00	1.16	86
5	Ukraine	57.78	42.22	31.11	19.26	50.37	49.63	23.81	23.08	23.53	13.43	0.74	2.22	135
6	Kenya	48.80	51.20	22.40	26.00	48.40	51.60	44.64	38.46	41.32	38.76	1.60	2.80	250
7	Kosovo	51.72	48.28	26.11	24.14	50.25	49.75	43.40	55.10	49.02	42.57	2.46	1.97	203
8	Malawi	50.96	49.04	25.66	23.25	48.92	51.08	35.99	41.83	38.76	34.00	1.59	2.25	2447
9	Nepal	46.58	53.42	25.00	22.36	47.36	52.64	27.33	47.22	36.72	33.92	0.16	0.93	644
10	Benin	51.05	48.95	25.02	23.84	48.86	51.14	46.18	40.84	43.58	45.37	1.82	3.37	1099
11	Cameroon	51.86	48.14	28.71	22.93	51.64	48.36	46.77	46.67	46.72	40.18	1.64	1.86	916
12	Dominican Republic	50.28	49.72	26.72	24.78	51.51	48.49	34.27	32.32	33.33	30.34	0.67	1.59	2522
13	El Salvador	50.37	49.63	25.87	24.62	50.49	49.51	55.84	58.85	57.31	55.59	0.79	0.89	6722
14	Guinea Bissau	50.06	49.94	25.67	25.09	50.76	49.24	45.00	45.12	45.06	37.91	1.05	1.98	857
15	Guyana	51.77	48.23	26.77	25.00	51.77	48.23	45.28	39.39	42.44	38.22	1.26	1.26	396
16	Kyrgyzstan	47.37	52.63	23.44	24.64	48.09	51.91	28.57	36.89	32.84	26.73	0.24	0.00	418
17	Sao Tome and Principe	48.35	51.65	27.27	26.03	53.31	46.69	42.42	41.27	41.86	30.09	0.83	0.83	242
18	Swaziland	52.18	47.82	25.40	21.43	46.83	53.17	32.03	46.30	38.56	24.63	1.79	1.39	504
19	Congo	53.56	46.44	27.70	22.00	49.69	50.31	41.18	45.37	43.03	42.11	1.43	3.05	491
20	Mali	52.96	47.04	28.20	22.43	50.63	49.37	52.60	47.40	50.30	46.97	1.06	2.18	1975

Table 13 (continued)

#	Country	First birth		Second birth				% with 3rd birth given:				% twins at:		Obs
		% Boy	% Girl	% 2 Boys	% 2 Girls	% SS	% MS	2 Boys	2 Girls	SS	MS	2nd	3rd	
(1)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
21	Congo	50.46	49.54	26.23	26.78	53.01	46.99	44.44	46.26	45.36	37.98	1.09	2.00	549
22	Mauritania	50.14	49.86	28.26	24.64	52.90	47.10	52.86	46.33	49.82	52.32	1.62	2.00	1051
23	Senegal	54.13	45.87	25.69	18.81	44.50	55.50	41.07	56.10	47.42	41.32	0.92	1.38	218
24	Nigeria	52.67	47.33	26.36	23.28	49.64	50.36	52.07	50.78	51.47	51.53	1.26	1.66	2470
25	Paraguay	48.64	51.36	25.93	21.98	47.90	52.10	23.81	28.09	25.77	28.91	0.99	1.48	405
26	Cote d'Ivoire	53.33	46.67	30.09	22.08	52.17	47.83	41.35	38.86	40.30	36.69	1.45	1.54	1037
27	Mongolia	45.57	54.43	20.25	25.32	45.57	54.43	37.50	30.00	33.33	37.21	3.80	5.06	79
28	Guinea	50.11	49.89	24.33	23.43	47.76	52.24	36.87	40.19	38.50	44.64	1.35	1.79	892
29	Senegal	62.50	37.50	35.42	10.42	45.83	54.17	47.06	0.00	36.36	38.46	0.00	4.17	48
30	Nigeria	50.00	50.00	26.44	19.71	46.15	53.85	63.64	48.78	57.29	45.54	2.40	0.48	208
31	Sierra Leone	51.12	48.88	25.93	22.95	48.88	51.12	73.85	74.08	73.96	70.89	2.08	2.15	9380
32	Pakistan	51.43	48.57	26.75	22.94	49.69	50.31	79.43	83.61	81.36	77.37	1.36	1.39	3671
33	Iraq	52.91	47.09	28.16	22.57	50.73	49.27	79.24	82.68	80.77	77.92	1.66	1.84	14,810
34	Gambia	51.58	48.42	26.88	23.17	50.04	49.96	81.79	82.48	82.11	77.95	2.05	1.95	7170
35	Mongolia	51.81	48.19	26.61	23.91	50.52	49.48	60.21	63.85	61.93	55.66	1.03	1.27	6536
36	Suriname	50.24	49.76	25.61	24.69	50.30	49.70	62.41	64.84	63.60	59.24	1.07	1.16	3179
37	Pakistan	51.65	48.35	26.38	23.71	50.10	49.90	78.60	82.99	80.68	76.48	1.27	1.46	28,784

Table reports descriptive statistics for the "core" sample. SS, same sex; MS, mixed sex sibling composition

Table 14 Testing assumptions underlying LATE identification

#	Sample	Kitagawa (2015)	Mourifie and Wan (2017) Without covariates	With covariates
	(1)	(2)	(3)	(4)
1	Full sample	.90	Not rejected	Not rejected
2	Age (< 26)	.99	Not rejected	Not rejected
3	Age (> 25 and < 36)	.77	Not rejected	Not rejected
4	Age (> 35)	.86	Not rejected	Not rejected
5	Age at 1st birth (< 21)	.76	Not rejected	Not rejected
6	Age at 1st birth (> 20 and < 26)	.99	Not rejected	Not rejected
7	Age at 1st birth (> 26)	.61	Not rejected	Not rejected
8	Not married	.93	Not rejected	Not rejected
9	Married	.93	Not rejected	Not rejected
10	Rural	.64	Not rejected	Not rejected
11	Urban	.99	Not rejected	Not rejected
12	Girl (1st child)	.99	Not rejected	Not rejected
13	Boy (1st child)	.53	Not rejected	Not rejected
14	No primary	.99	Not rejected	Not rejected
15	Primary	.51	Not rejected	Not rejected
16	Junior secondary	.20	Not rejected	Not rejected
17	Senior secondary	.99	Not rejected	Not rejected
18	Wealth quintile 1	.93	Not rejected	Not rejected
19	Wealth quintile 2	.99	Not rejected	Not rejected
20	Wealth quintile 3	.44	Not rejected	Not rejected
21	Wealth quintile 4	.89	Not rejected	Not rejected
22	Wealth quintile 5	.99	Not rejected	Not rejected

Column 2 reports p values from testing the null hypothesis as formulated in Kitagawa (2015). Columns 3 and 4 report results from STATA's clrtest command testing the null hypothesis as formulated in Mourifie and Wan (2017). Tests results shown in columns 3 and 4 are at the 90% significance level. Covariates included in column 4 include mother's age, mother's age at first birth, gender of the 1st child, rural/urban status, year fixed-effects, and country fixed-effects

Table 15 OLS estimates on the relationship between sex composition and asset/dwelling characteristics

Panel A: Same sex

Parameter	AI	Floor	Roof	Wall	Cooking	TV	Cellphone
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Same sex	−0.011 (0.008)	−0.004 (0.007)	−0.000 (0.004)	0.006 (0.006)	−0.002 (0.006)	0.003 (0.003)	0.003 (0.003)
Observations	152,279	146,190	146,198	146,198	152,182	131,617	152,279
r2	0.0198	0.2002	0.2743	0.1773	0.4289	0.3071	0.2319

Panel B: 2 boys, 2 girls

Parameter	AI	Floor	Roof	Wall	Cooking	TV	Cellphone
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
2 boys	−0.008 (0.009)	−0.000 (0.008)	−0.004 (0.005)	0.009 (0.007)	−0.002 (0.007)	0.002 (0.004)	0.004 (0.003)
2 girls	−0.014 (0.010)	−0.007 (0.009)	0.004 (0.006)	0.003 (0.007)	−0.003 (0.008)	0.004 (0.004)	0.001 (0.003)
Observations	152,279	146,190	146,198	146,198	152,182	131,617	152,279
r2	0.0198	0.2002	0.2743	0.1773	0.4289	0.3071	0.2319

Asset variables related to floor, roof, wall material, and type of cooking fuel are coded to be between 0 (low quality) and 2 (high quality) following MICS guidelines. TV and cellphone possession are coded as binary variables. All regressions include year and country fixed-effects. Covariates include mother's age (dummy variables for each year), mother's age at first birth (dummy variables for 5-year intervals), gender of the 1st child, and location (rural vs. urban). Standard errors are depicted in parentheses and clustered at the community level

*/**/***Significance levels at 10/5/1% respectively

Table 16 OLS estimates on the relationship between sex composition and children's education outcomes

Parameter	No primary (1)	Primary (2)	Junior secondary (3)	Senior secondary (4)
Panel A: 1st child				
Same sex	0.002 (0.009)	0.002 (0.008)	0.004 (0.007)	−0.009 (0.007)
Observations	9586	9586	9586	9586
r2	0.1557	0.0809	0.2302	0.0578
Panel B: 2nd child				
Same sex	0.001 (0.009)	0.004 (0.008)	0.005 (0.007)	−0.009 (0.007)
Observations	9542	9542	9542	9542
r2	0.1560	0.0760	0.2300	0.0571
Panel C: 1st and 2nd child				
Same sex	0.001 (0.009)	0.003 (0.008)	0.005 (0.007)	−0.009 (0.007)
Observations	19,128	19,128	19,128	19,128
r2	0.1557	0.0780	0.2300	0.0575

Panels A (1st born child), B (2nd born child), and C (1st and 2nd born children) refer to children born to women in our main sample (women with at least 2 children). All regressions include year and country fixed-effects. Covariates include mother's age (dummy variables for each year), mother's age at first birth (dummy variables for 5-year intervals), gender of the 1st child, and location (rural vs. urban). Standard errors are depicted in parentheses and clustered at the community level

*/**/***Significance levels at 10/5/1% respectively

Table 17 OLS estimates on the relationship between sex composition and children's vaccination outcomes

Parameter	Total # (1)	BCG (2)	Polio 1 (3)	Polio 2 (4)	Polio 3 (5)	DPT 1 (6)	DPT 2 (7)	DPT 3 (8)
Panel A: 1st child								
Same sex	−0.063 (0.114)	0.005 (0.014)	0.011 (0.013)	0.021 (0.020)	−0.030 (0.026)	0.010 (0.024)	0.021 (0.023)	−0.024 (0.025)
Observations	723	743	741	739	733	740	736	731
r2	0.3727	0.1067	0.0650	0.0839	0.1160	0.3323	0.5778	0.5301
Panel B: 2nd child								
Same sex	−0.063 (0.114)	0.005 (0.014)	0.011 (0.013)	0.021 (0.020)	−0.030 (0.026)	0.010 (0.024)	0.021 (0.023)	−0.024 (0.025)
Observations	723	743	741	739	733	740	736	731
r2	0.3727	0.1067	0.0650	0.0839	0.1160	0.3323	0.5778	0.5301
Panel C: 1st and 2nd child								
Same sex	−0.063 (0.114)	0.005 (0.014)	0.011 (0.013)	0.021 (0.020)	−0.030 (0.026)	0.010 (0.024)	0.021 (0.023)	−0.024 (0.025)
Observations	723	743	741	739	733	740	736	731
r2	0.3727	0.1067	0.0650	0.0839	0.1160	0.3323	0.5778	0.5301

Panels A (1st born child), B (2nd born child), and C (1st and 2nd born children) refer to children born to women in our main sample (women with at least 2 children). Number of vaccinations is an additive index constructed as sum of the following vaccinations: 1 BCG, 3× polio, and 3× DPT. All regressions include year and country fixed-effects. Covariates include mother's age (dummy variables for each year), mother's age at first birth (dummy variables for 5-year intervals), gender of the 1st child, and location (rural vs. urban). Standard errors are depicted in parentheses and clustered at the community level

*/**/***Significance levels at 10/5/1% respectively

Table 18 Alternative IV estimates on SWB

Parameter	2SLS	Abadie03	Frölich07
	(1)	(2)	(3)
Parameter	0.579** (0.261)	0.572** (0.270)	0.574** (0.279)

All regressions include year and country fixed-effects. Covariates include mother's age (dummy variables for each year), mother's age at first birth (dummy variables for 5-year intervals), gender of the 1st child, and location (rural vs. urban). Standard errors are depicted in parentheses and clustered at the community level

*/**/***Significance levels at 10/5/1% respectively

Table 19 IV (2SLS) estimates on life satisfaction—additional controls

	Z = same sex				Z = 2 boys, 2 girls			
	A	B	C	ABC	A	B	C	ABC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mother had a 3rd child	0.560** (0.262)	0.620** (0.259)	0.516** (0.254)	0.558** (0.256)	0.414* (0.235)	0.452* (0.231)	0.390* (0.229)	0.403* (0.230)
Observations	102,798	102,798	102,798	102,798	102,798	102,798	102,798	102,798
r2	0.1863	0.1990	0.1975	0.2094	0.1961	0.2098	0.2045	0.2183
First-stage results								
Same sex	0.0330*** (0.0020)	0.0330*** (0.0020)	0.0340*** (0.0020)	0.0330*** (0.0020)				
2 boys					0.0190*** (0.0030)	0.0190*** (0.0030)	0.0200*** (0.0030)	0.0190*** (0.0030)
2 girls					0.0480*** (0.0030)	0.0490*** (0.0030)	0.0490*** (0.0030)	0.0480*** (0.0030)
F-stat (1st stage)	188.4780	187.8730	197.0390	192.0600	111.8390	112.7560	116.0300	113.9320
Hansen's J (p value)					0.1510	0.1040	0.1970	0.1230
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Basic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Marital status	Yes	No	No	Yes	Yes	No	No	Yes
Age at 1st birth	No	Yes	No	Yes	No	Yes	No	Yes
Mother's education	No	No	Yes	Yes	No	No	Yes	Yes

All regressions include year and country fixed-effects. Covariates include mother's age (dummy variables for each year), mother's age at first birth (dummy variables for 5-year intervals), gender of the 1st child, and location (rural vs. urban). Standard errors are depicted in parentheses and clustered at the community level

*/**/***Significance levels at 10/5/1% respectively

Table 20 Bound estimates for β assuming different values of θ

	Bounds	
θ	Lower	Upper
	(1)	(2)
0.0005	0.067	1.075
0.0010	0.053	1.075
0.0015	0.038	1.075
0.0020	0.024	1.075
0.0025	0.009	1.075
0.0030	− 0.006	1.075
0.0035	− 0.020	1.075
0.0040	− 0.035	1.075
0.0045	− 0.050	1.075
0.0050	− 0.064	1.075
0.0055	− 0.079	1.075
0.0060	− 0.094	1.075
0.0065	− 0.108	1.075
0.0070	− 0.123	1.075
0.0075	− 0.138	1.075

Columns 2 and 3 report bounds on "having a 3rd child". Bounds are calculated for a 95% confidence interval for our parameter of interest. Results are obtained using STATA's "plausexog" package which implements the procedures outlined in Conley et al. (2012)

Table 21 IV (2SLS) estimates on life satisfaction: twin instrument by birth order

Parameter	B. Order 1	B. Order 2	B. Order 3
	(1)	(2)	(3)
Mother had a 2nd child	0.105 (0.120)		
Mother had a 3rd child		0.282* (0.159)	
Mother had a 4th child			0.410* (0.215)
Observations	144,910	102,798	69,356
r2	0.2380	0.1988	0.1573
First-stage results			
Twins	0.2900*** (0.0080)	0.2410*** (0.0080)	0.1890*** (0.0100)
F-stat (1st stage)	1389.9660	961.6910	390.6320

All regressions include year and country fixed-effects. Covariates include mother's age (dummy variables for each year), mother's age at first birth (dummy variables for 5-year intervals), gender of the 1st child, and location (rural vs. urban). Standard errors are depicted in parentheses and clustered at the community level

*/**/***Significance levels at 10/5/1% respectively

Table 22 IV (2SLS) estimates on happiness (twin instrument)

	Without covariates (1)	With covariates (2)
Mother had a 3rd child	0.178** (0.084)	0.249** (0.103)
Observations	102,798	102,798
r2	0.0677	0.0729
First-stage results		
Twin	0.2970*** (0.0060)	0.2410*** (0.0080)
F-stat (1st stage)	2734.2620	961.6910

All regressions include year and country fixed-effects. Covariates include mother's age (dummy variables for each year), mother's age at first birth (dummy variables for 5-year intervals), gender of the 1st child, and location (rural vs. urban). Standard errors are depicted in parentheses and clustered at the community level

*/**/***Significance levels at 10/5/1% respectively

References

Aaronson D, Dehejia R, Jordan A et al (2017) "The effect of fertility on mothers' labor supply over the last two centuries". NBER working paper no. 32717

Aassve A, Massuco S, Mencarini L (2005) Childbearing and well-being: a comparative analysis of the European Community. J Eur Soc Policy 15(4):283–299

Aassve A, Goisis A, Sironi M (2012) Happiness and childbearing across Europe. Soc Indic Res 108(1):65–86

Abadie A (2003) Semiparametric instrumental variable estimation of treatment response model. J Econ 113(2):231–263

Adler MD, Dolan P, Kavetsos G (2017) Would you choose to be happy? Tradeoffs between happiness and the other dimensions of life in a large population survey. J Econ Behav Organ 139(July):60–73

Aguero JM, Marks MS (2008) Motherhood and female labor force participation: evidence from infertility shocks. Am Econ Rev 98(2):500–504

Alesina A, di Tella R, McCulloch R (2004) Inequality and happiness: are European and Africans different? J Public Econ 88(9–10):2009–2042

Angrist JD, Evans WN (1998) Children and their parents labor supply: evidence from exogenous variation in family size. Am Econ Rev 88(3):450–477

Angrist JD, Pischke J (2009) Mostly harmless econometrics: an empiricist's companion. Princeton University Press, Oxfordshire

Angrist JD, Imbens GW, Rubin D (1996) Identification of causal effects using instrumental variables. J Am Stat Assoc 91(434):444–455

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

Baetschmann G, Staub KE, Studer R (2016) Does the stork deliver happiness? Parenthood and life satisfaction. J Econ Behav Organ 130(October):242–260

Balbo N, Arpino B (2016) The role of family orientation in shaping the effect of fertility on subjective well-being: a propensity score matching approach. Demography 53(4):955–978

Bertrand M (2013) Career, family, and the well-being of college-educated women. Am Econ Rev 103(3):244–250

Bhalotra S, Clarke D (2019) Twin births and maternal condition. Rev Econ Stat 101(5):853–864

Bisbee J, Dehejia R, Pop-Eleches C, Samii C (2017) Local instruments, global extrapolation: external validity of the labor supply-fertility local average treatment effect. J Labor Econ 35(S1):99–147

Björklund A, Moffitt R (1987) The estimation of wage gains and welfare gains in self-selection models. Rev Econ Stat 69(1):42–49

Blanchflower, D. (2008). "International evidence on well-being". NBER working paper no. 14318

Bongarts J (2003) Completing the fertility transition in the developing world: the role of educational differences and fertility preferences. Popul Stud 57(3):321–335

Bongarts J, Casterline J (2013) Fertility transition: Is sub-Saharan Africa different? Popul Dev Rev 38(supplement):153–168

Bratti M, Fiore S, Mendola M (2020) The impact of family size and sibling structure on the great Mexico-USA migration. J Popul Econ 33:483–529

Brinch CN, Mogstad M, Wiswall MJ (2017) Beyond LATE with a discrete instrument. J Polit Econ 125(4), 985–1039

Buddelmeyer H, Hamermesh DS, Wooden M (2018) The stress cost of children on moms and dads. Eur Econ Rev 109(October):148–161

Bütikofer, A. (2011) "Sibling sex composition and the cost of children". Working Paper

Canning D, Günther I, Linnemayr S, Bloom D (2013) Fertility choice, mortality expectations and interdependent preferences. An empirical analysis. Eur Econ Rev 63(C):273–289

Cetre S, Clark AE, Senik C (2016) Happy people have children: choice and self-selection into parenthood. Eur J Popul 32(3):445–473

Chernozhukov V, Lee S, Rosen AM (2013) Intersection bounds: estimation and inference. Econometrica 81(2):667–737

Chernozhukov V, Kim W, Lee S, Rosen AM (2015) Implementing intersection bounds in Stata. Stata J 15(1): 21–44

Chu J (2001) Prenatal sex determination and sex-selective abortion in rural Central China. Popul Dev Rev 27(2):259–281

Clark AE (2007) "Born to be mild? Cohort effects don't explain why well-being is U-shaped in age". IZA working paper no. 3170

Clark AE, Georgellis Y (2013) Back to baseline in Britain – adaptation in the BHPS. Economica 80(319): 496–512

Clark AE, Oswald AJ (2002) "Well-being in panels". DELTA. mimeo

Clark AE, Senik C (2011) Is happiness different from flourishing? Cross-country evidence from the ESS. Rev Econ Polit 121(1):17–34

Clark AE, Diener E, Georgellis Y (2008) Lags and leads in life satisfaction: a test of the baseline hypothesis. Econ J 118(529):222–243

Clark AE, D'Ambrosio C, Ghislandi S (2016) Adaptation to poverty in long-run panel data. Rev Econ Stat 98(3):591–600

Clarke D, Matta B (2018) Practical considerations for questionable IVs. Stata J 18(3):663–691

Conley TG, Hansen CB, Rossi PE (2012) Plausibly exogenous. Rev Econ Stat 94(1):260–272

Connelly R, Kimmel J (2015) If you're happy and you know it, clap your hands: how do mothers and fathers really feel about child caregiving. Fem Econ 21(1):1–34

Conzo P, Fuochi G, Mencarini L (2017) Fertility and life satisfaction in rural Ethiopia. Demography 54(4): 1331–1351

Cruces G, Galiani S (2007) Fertility and female labor supply in Latin America: new causal evidence. Labour Econ 14(3):565–573

Deaton A, Stone AA (2014) Evaluative and hedonic well-being among those with and without children at home. Proc Natl Acad Sci 111(4):1328–1333

Di Tella R, MacCulloch R (2006) Some uses of happiness data in economics. J Econ Perspect 20:25–46

Di Tella R, MacCulloch R, Oswald A (2003) The macroeconomics of happiness. Rev Econ Stat 85(4):809–827

Dolan P, Peasgood T, White M (2008) Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. J Econ Psychol 29(1):94–122

Evenson RJ, Simon R (2005) Clarifying the relationship between parenthood and depression. J Health Soc Behav 46(4):341–358

Frijters P, Johnston DW, Shields MA (2011) Life satisfaction dynamics with quarterly life event data. Scand J Econ 113(1):190–211

Frölich M (2007) Nonparametric IV estimation of local average treatment effects with covariates. J Econ 139(1):35–75

Gallagher SK, Gerstel N (2001) Connections and constraints: the effects of children on caregiving. J Marriage Fam 63(1):265–275

Gettler LT, Oka CA (2016) Are testosterone levels and depression risk linked based on partnering and parenting? Evidence from a large population-representative study of U.S. men and women. Soc Sci Med 163(August):157–167

Gilbert DT (2006) Stumbling on happiness. Random House, New York

Glass J, Simon RW, Andersson MA (2016) Parenthood and happiness: effects of work-family reconciliation policies in 22 OECD countries. Am J Sociol 122(3):886–929

Grossbard S, Mukhopadhyay S (2013) Children, spousal love, and happiness: An economic analysis. Rev Econ Househ 11:447–467

Hamermesh D, Lee J (2007) Stressed out on four continents: time crunch or yuppie kvetch? Rev Econ Stat 89(2):374–383

Hansen T (2012) Parenthood and happiness: a review of folk theories versus empirical evidence. Soc Indic Res 108:26–64

Hansen T, Slagsvold B, Moum T (2009) Childlessness and psychological well-being in midlife and old-age: an examination of parental status effects across a range of outcomes. Soc Indic Res 94(2):343–362

Heckman JJ, Vytlacil E (1999) Local instrumental variables and latent variable models for identifying and bounding treatment effects. Proc Natl Acad Sci 96(8):4730–4734

Heckman JJ, Vytlacil E (2005) Structural equations, treatment effects and econometric policy evaluation. Econometrica 73(3):669–738

Heckman JJ, Vytlacil E (2007) Econometric evaluation of social programs, part ii: Using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments. Handb Econ 6:4875–5143

Herbst CM, Ifcher J (2016) The increasing happiness of US parents. Rev Econ Househ 14(3):529–551

Huber M (2015) Testing the validity of the sibling sex ratio instrument. Labour 29(1):1–14

Huber M, Mellace G (2015) Testing instrument validity for LATE identification based on inequality moment constraints. Rev Econ Stat 97(2):398–411

Huber M, Wüthrich K (2019) Local average and quantile treatment effects under endogeneity: a review. Journal of Econometric Methods 8(1):1–27

Imbens GW, Angrist J (1994) Identification and estimation of local average treatment effects. Econometrica 62(2):467–475

Keizer R, Dykstra PA, Poortman AR (2010) Life outcomes of childless men and fathers. Eur Sociol Rev 26(8):1–15

Kitagawa T (2015) A test for instrument validity. Econometrica 83(5):2043–2063

Kohler H-P, Behrman J, Skytthe A (2005) Partner ? Children = happiness? The effects of partnerships and fertility on well-being. Popul Dev Rev 31(3):407–445

Kowalski A (2016a) "Doing more when you're running LATE: applying marginal treatment effect methods to examine treatment effect heterogeneity in experiments". NBER working paper no. 22363

Kowalski A (2016b) "Doing more when you're running LATE: applying marginal treatment effect methods to examine treatment effect heterogeneity in experiments". NBER working paper no. 22363

Kowalski A (2019) "Reconciling seemingly contradictory results from the Oregon health insurance experiment and the Massachussets health reform". NBER working paper no. 24647

Lee J (2008) Sibling size and investment in children's education: an Asian instrument. J Popul Econ 21(4): 855–875

Lee S (2018) A consistent variance estimator for 2SLS when instruments identify different LATEs. J Bus Econ Stat 36(3):400–410

Margolis R, Myrskylä M (2011) A global perspective on happiness and fertility. Popul Dev Rev 37(1):29–56

Moffit R (2005) Remarks on the analysis of causal relationships in population research. Demography 42(1): 91–108

Moglie ML, Mencarini L, Rapallini C (2015) Is it just a matter of personality? On the role of subjective well-being in childbearing behavior. J Econ Behav Organ 117(September):453–475

Morgan SP, King RB (2001) Why have children in the 21st century? Biological predisposition, social coercion, rational choice. Eur J Popul 17(1):3–20

Mourifie I, Wan Y (2017) Testing local average treatment effect assumptions. Rev Econ Stat 99(2):305–313

Mu Z, Xie Y (2017) 'Motherhood penalty' and 'fatherhood premium'? Fertility effects on parents in China. Demogr Res 35(47):1373–1410

Myrskylä M, Margolis R (2014) Happiness: before and after the kids. Demography 51(5):1843–1866

Nomaguchi K, Milkie M (2003) Costs and rewards of children: the effect of becoming a parent on adult's lives. J Marriage Fam 65(2):356–374

ODI (2016) "Women's work: mothers, children and the global childcare crisis". Overseas Development Institute

Parr N (2010) Satisfaction with life as an antecedent of fertility: partner + happiness = children? Demogr Res 21:635–661

Pedersen P, Schmidt T (2014) "Life events and subjective well-being: The case of having children". IZA Discussion Paper No. 8207

Pollmann-Schult M (2014) Parenthood and life satisfaction: Why don't children make people happy? J Marriage Fam 76(2):319–336

Ponczek V, Souza AP (2012) New evidence of the causal effect of family size on child quality in a developing country. J Hum Resour 47(1):64–106

Priebe J (2011) "Child costs and the causal effect of fertility on female labor supply: an investigation for Indonesia 1993–2008". Mimeo

Roby JL (2011) "Children in informal alternative care". UNICEF discussion paper

Rosenzweig MR, Wolpin K (1980) Testing the quantity-quality fertility model: the use of twins as a natural experiment. Econometrica 48(1):227–240

Rosenzweig MR, Wolpin K (2000) Natural experiments in economics. J Econ Lit 38(4):827–874

Rosenzweig MR, Zhang J (2009) Do population control policies induce more human capital investment? Twins, birthweight, and China's 'one child' policy. Rev Econ Stud 73(3):1149–1174

Sacerdote B, Feyrer J (2008) Will the stork return to Europe and Japan? Understanding fertility within developed nations. J Econ Perspect 22(3):3–22

Stanca L (2012) Suffer the little children: measuring the effect of parenthood on well-being worldwide. J Econ Behav Organ 81(3):742–750

Stevenson B, Wolfers J (2009) The paradox of declining female happiness. Am Econ J Econ Pol 1(2):190–225

Strulik H, Vollmer S (2015) The fertility transition around the world. J Popul Econ 28(1):31–44

Stutzer A, Frey BS (2006) Does marriage make people happy, or do happy people get married? J Socio-Econ 35(2):326–347

Umberson D, Gove W (1989) Parenthood and psychological well-being: theory, measurement, and stage in the family life course. J Fam Issues 10(4):440–462

UN (2015a) World fertility report 2015. Department of Economic and Social Affairs. Population Division. United Nations, New York

UN (2015b) World fertility patterns 2015. Department of Economic and Social Affairs. Population Division. United Nations, New York

UN (2017) World population prospects: key findings and advance tables. Department of Economic and Social Affairs. Population Division. United Nations, New York

Van Praag B, Frijters P, Ferrer-i-Carbonell A (2003) The anatomy of subjective well-being. J Econ Behav Organ 51(1):29–49

Vanassche S, Swicegood G, Matthijs K (2013) Marriage and children as a key to happiness? Cross-national differences in the effects of marital status and children on well-being. J Happiness Stud 14(2):501–524

WB (2010) Determinants and consequences of high fertility: a synopsis of the evidence. World Bank, Washington