TWIN BIRTH AND MATERNAL CONDITION

Sonia Bhalotra and Damian Clarke*

Abstract—Twin births are often construed as a natural experiment in the social and natural sciences on the premise that the occurrence of twins is quasi-random. We present population-level evidence that challenges this premise. Using individual data for 17 million births in 72 countries, we demonstrate that indicators of mother's health, health-related behaviors, and the prenatal environment are systematically positively associated with twin birth. The associations are sizable, evident in richer and poorer countries—evident even among women who do not use in vitro fertilization—and hold for numerous different measures of health. We discuss potential mechanisms, showing evidence that favors selective miscarriage.

I. Introduction

TWINS have intrigued humankind for more than a century (Thorndike, 1905). In behavioral genetics, demography, and psychology, monozygotic twins are studied to assess the importance of nurture relative to nature (Polderman et al., 2015). In the social sciences, twin births are also used to denote an unexpected increase in family size, which assists causal identification of the impact of fertility on investments in children and on women's labor supply (Rosenzweig & Wolpin, 2000, 1980a; Bronars & Grogger, 1994; Black, Devereux, & Salvanes, 2005). A premise of studies that use twin differences or the twin instrument is that twin births are quasi-random and have no direct impact (except through fertility) on the outcome under study. We present new population-level evidence that challenges this premise. Using 16,962,165 births in 72 countries, of which 462,246 (2.73%) are twins, we show that the likelihood of a twin birth varies systematically with maternal condition. In particular, our estimates establish that mothers of twins are selectively healthy.[1]

We document that the association of twin births and maternal condition is meaningfully large and widespread. We show that is evident in richer and poorer countries and that it holds for sixteen different markers of maternal condition, including health stocks and health conditions prior to pregnancy (height, obesity, diabetes, hypertension, asthma, kidney disease, smoking), exposure to unexpected stress in pregnancy, and measures of the availability of medical professionals and prenatal care.[2] The effects are sizable, with a 1 standard deviation improvement in the indicator tending to increase the likelihood of twinning by 6% to 12%.

Previous research has documented that twins have different endowments from singletons; for example, twins are more likely to have low birthweight and congenital anomalies (Hall, 2003; Rosenzweig & Zhang, 2009). We focus not on differences between twins and singletons but rather on differences between mothers of twins and singletons, which indicate whether occurrence of twin births is quasi-random. It is known that twin births are not strictly random, occurring more frequently among older mothers, at higher parity, and in certain races and ethnicities (Hall, 2003; Bulmer, 1970), but as these variables are typically observable, they can be adjusted for (as in Rosenzweig & Wolpin, 1980a).[3] Similarly, it is well documented that women using artificial reproductive technologies (ART) are more likely to give birth to twins (Vitthala et al., 2009), but ART use is recorded in many birth registries, so it can be controlled for and a conditional randomness assumption upheld (Cáceres-Delpiano, 2006; Angrist, Lavy, & Schlosser, 2010). The reason that our finding is potentially a major challenge is that maternal condition is multidimensional and almost impossible to fully measure and adjust for. To take a few examples, fetal health is potentially a function of whether pregnant women skip breakfast (Mazumder & Seeskin, 2015), they suffer bereavement in pregnancy (Black et al., 2016), or they are exposed to air pollution (Chay & Greenstone, 2003).

Our underlying hypothesis is that twins are more demanding of maternal resources than singletons, and as a result, conditions that challenge maternal health are more likely to result in miscarriage of twins than of singletons. We discuss the role of alternative mechanisms including nonrandom conception and maternal survival selection. We provide evidence in favor of the selective miscarriage mechanism using U.S. Vital Statistics data for 14 to 16 million births. Selective miscarriage is similarly the mechanism behind the stylized fact that weaker maternal condition is associated with a lower probability of male birth (Trivers & Willard, 1973; Almond & Edlund, 2007). We confirm this in our data,

Received for publication June 1, 2017. Revision accepted for publication July 30, 2018. Editor: Rohini Pande.

*Bhalotra: University of Essex; Clarke: Universidad de Santiago de Chile.

We are grateful to Paul Devereux, James Fenske, Judith Hall, Christian Hansen, Martin Karlsson, Toru Kitagawa, Magne Mogstad, Cheti Nicoletti, Carol Propper, Adam Rosen, Paul Schulz, Margaret Stevens, Atheen Venkataramani, Marcos Vera-Hernandez, Frank Windmeijer, Emilia Del Bono, Climent Quintana-Domeque, Pedro Ródenas, Libertad González, Hanna Mühlrad, Anna Aevarsdottir, Martin Foureaux Koppensteiner, Ryan Brown, Pietro Biroli, Rohini Pande, and three anonymous referees, along with various seminar audiences and discussants for helpful comments and/or sharing data. Any remaining errors are our own. An earlier version of this paper was circulated as Part 1 of "The Twin Instrument," IZA discussion paper 10405.

A supplemental appendix is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/rest_a_00789.

[1]Twins are not as rare as we may think: one in eighty live births and hence one in forty newborns is a twin. In general and, for instance, in the United States, there is a positive trend in twin births.

[2]We also show that a positive association of the chances of having twins with health-related behaviors in pregnancy (healthy diet, smoking, alcohol, drug consumption), although we do not rely on this because behaviors in pregnancy may reflect a response to the mother's knowledge that she is carrying twins.

[3]Other correlates identified in the medical literature but not reflected in social science research include high concentrations of follicle-stimulating hormone in women, season and seasonal light, height, urbanization, and starvation (Hall, 2003), with mixed results (based on small samples) when considering social class (Campbell, Campbell, & MacGillivary, 1974; Campbell, 1998). These results have not been documented in the economics or social science literature. In our discussion of mechanisms, we discuss the difference between monozygotic and dyzygotic twins.

The Review of Economics and Statistics, December 2019, 101(5): 853–864
© 2018 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology.
Published under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.
https://doi.org/10.1162/rest_a_00789

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/rest_a_00789 by guest on 28 March 2024

TABLE 1.—THE QUANTITY–QUALITY AND FERTILITY–FLFP TRADE-OFFS: ESTIMATES USING THE TWIN INSTRUMENT

Authors	Data/Context	Twin Use	Estimates		Maternal Controls
			OLS	IV	
A. Quantity–Quality					
Rosenzweig and Wolpin (1980a)	India, rural survey; outcome is standardized schooling	Twin ratio in OLS[a]	−2.483 (0.740)		None
Black et al. (2005)	Norway, administrative data	IV	−0.060 (0.003) −0.076 (0.004) −0.059 (0.006)	−0.038 (0.047) −0.016 (0.044) −0.024 (0.059)	Age and education
Cáceres-Delpiano (2006)	United States, Census 5% file; outcome is behind educational cohort	IV	0.011 (0.000) 0.017 (0.001)	0.002 (0.003) 0.010 (0.006)	Age, education, and race
Li et al. (2008)	China, census 1% file; outcome is educational enrollment	IV	−0.031 (0.001) −0.038 (0.002)	0.002 (0.009) −0.024 (0.014)	Age and education
Rosenzweig and Zhang (2009)	China, twin survey	RF[b,f]	−0.307 (0.160) −0.225 (0.172)	No birthweight control Birthweight control	Age
Angrist et al. (2010)	Israel, Census 20% file	IV	−0.145 (0.005) −0.143 (0.005)	0.174 (0.166) 0.167 (0.117)	Age, place of birth, race
Black et al. (2010)	Norway, administrative data; outcome is IQ	IV		−0.149 (0.052) −0.170 (0.052) −0.115 (0.080)	Age and education
Åslund and Grönqvist (2010)	Sweden, administrative data	IV	−0.113 (0.004) −0.132 (0.006) −0.100 (0.009)	0.022 (0.048) −0.043 (0.048) −0.042 (0.083)	Age and education
Ponczek and Souza (2012)	Brazil, Census 10% file	IV (girl) IV (boy)	−0.277 (0.015) −0.283 (0.015) −0.233 (0.010) −0.230 (0.010)	−0.372 (0.198) −0.634 (0.194) −0.137 (0.146) −0.060 (0.164)	Age and education
Marteleto and de Souza (2012)	Brazil, household survey	IV	−0.248 (0.003) −0.240 (0.003)	0.064 (0.076) 0.131 (0.055)	Age, education, and family income
Mogstad and Wiswall (2016)	Norway, administrative data	IV[c]		0.053 (0.050) −0.051 (0.053) −0.107 (0.059)	Age and education
B. Fertility and female labor force participation					
Rosenzweig and Wolpin (1980b)	United States, pooled demographic surveys	RF[d]	−0.371 (0.212) 0.142 (0.102)	Short-term estimate Long-term estimate	None
Bronars and Grogger (1994)	United States 1970 and 1980 5% Census	RF[d]	−0.036 (0.036) −0.035 (0.017)	1970 Census 1980 Census	Age at first birth
Angrist and Evans (1998)	United States 1980 5% Census	IV	−0.176 (0.002)	−0.057 (0.011)	Age, age at first birth
Jacobsen et al. (1999)	United States 1970 and 1980 5% Census	IV[e]		−0.021 (0.014) −0.025 (0.008)	Age at first birth cubic
Cáceres-Delpiano (2012)	Pooled demographic surveys, developing countries	IV	−0.014 (0.001) −0.010 (0.001) −0.009 (0.001)	−0.029 (0.012) −0.016 (0.012) −0.022 (0.012)	Age, education, literacy status, country dummies

Estimates and standard errors reported in columns OLS and IV refer to main estimates from each paper. Estimates are included from published articles using large samples of microdata. A comprehensive review is provided in Clarke (2018). Where multiple estimates are reported, unless otherwise indicated, the first line refers to the impact of twins at birth 2, the second line the impact of twins at birth 3, and the third line the impact of twins at birth 4 (if available). In panel A, the estimates refer to the outcome variable "years of education" unless specified in column 2. In panel B, all outcomes are the mother's labor market participation.

[a] Twin ratio is the number of twin births divided by the number of pregnancies.

[b] Coefficients reported are impact of second birth twins on nontwin first births.

[c] Nonlinear estimates are reported in paper. Here linear estimates are presented for comparison with other results.

[d] Reduced form (RF) uses twins at first birth as independent variable.

[e] First line reports estimates from 1970 census, second line reports 1980 census.

[f] Standard errors are calculated from reported *t*-statistics.

showing that twin births are more likely to be girls. Our findings add a novel twist to recent literature documenting that a mother's health and her environmental exposure to nutritional or other stresses during pregnancy influence birth outcomes, with many studies documenting lower birthweight (Currie & Moretti, 2007; Bernstein et al., 2005; Quintana-Domeque & Ródenas-Serrano, 2017). If birthweight is the intensive margin, we may think of miscarriage as an extensive margin response or the limiting case of low birthweight.

Our findings have implications for research that has exploited the assumed randomness of twin births. Studies using twins to isolate exogenous variation in fertility will tend to underestimate the impact of fertility on parental investments in children and on women's labor supply if selectively healthy mothers invest more in children after-birth, and are more likely to participate in the labor market (as discussed in Bloom, Kuhn, & Prettner, 2015). In table 1 we summarize studies using twin births to instrument fertility, documenting the mother-level controls in each study. In some cases, the validity of the conditional randomness assumption is directly probed—for instance, with respect to mother's education (Black et al., 2005; Li, Zhang, & Zhu, 2008; Rosenzweig

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/rest_a_00789 by guest on 28 March 2024

& Zhang, 2009). However, as is acknowledged in each case, any such tests are at best partial evidence in support of instrumental validity. Importantly, no previous study has attempted to control for maternal health conditions or behaviors. This is pertinent, as it could resolve the ambiguity of the available evidence on the impacts of fertility. In particular, recent studies using the twin instrument challenge a long-standing theoretical prior of Becker and Lewis (1973) in rejecting the presence of a quantity–quality (QQ) fertility trade-off in developed countries (Black et al., 2005; Angrist et al., 2010), but our estimates suggest that this rejection could in principle arise from ignoring the positive selection of women into twin birth. Similarly, research using the twin instrument tends to find that additional children have relatively little influence on female labor force participation (FLFP; see Lundborg, Plug, & Rasmussen, 2017). But, again, these estimates are likely to be downward biased. The results of studies in economics, psychology, education, and biology that instead exploit the genetic similarity of twins will not be biased but will tend to have more restricted external validity than previously assumed.[4]

II. Methodology

In this section, we discuss two distinct approaches to testing our hypothesis that twins are selectively born to healthier mothers. We identify variation in the mother's health before she gives birth to twins and before she knows she will give birth to twins. In the first approach, we use information on her health condition (morbidities, height, weight), health-related behaviors, access to health care, and environmental health stressors. In our second approach, we use as a marker of maternal health the fetal or infant survival rate of her births prior to the birth at which she has twins (with parity-matched counterfactuals). We discuss below the methods used to investigate potential mechanisms.

We conduct three robustness checks. First, we restrict the sample to non-ART births. It is important to demonstrate that our hypothesis holds independent of ART use because there is a positive association of ART with the likelihood of twin births (Vitthala et al., 2009), and ART users are typically more educated and wealthy (Lundborg et al., 2017). Another potential concern is that we are capturing genetic traits that, for instance, are associated with the woman's height or weight and also correlated with her predisposition toward twin birth. This would appear to be a second-order concern since we do not only rely on woman-specific measures of health but also show a positive association of twinning with environmental stressors, health facilities, and health-related behaviors. We nevertheless investigate this concern in two different ways. First, we test whether we can identify a positive association of the probability that a birth is a twin with woman-specific time-varying health indicators conditional on woman fixed effects that sweep out genetic influences. Second, we leverage biomedical research showing that monozygotic (MZ) twins are randomly allocated across mothers, although genetic predispositions may influence the chances of having dizygotic (DZ) twins (Meulemans et al., 1996). Ideally, we would restrict the sample to MZ twins, but MZ versus DZ are not identified in the data. Instead, on the premise that MZ twins are necessarily same sex and about half of all DZ twins are same sex, we investigate our hypothesis restricting the sample to include only same-sex twins. If our results were driven by genetic predispositions, then we should find weaker associations in the same-sex sample. The methods and data used to conduct the robustness checks are discussed with the results. The rest of this section elaborates the specification used in the two main approaches to testing for twin randomness.

A. Across Mothers

To test the null that twin births are "as good as random," we estimate conditional regressions of the form

$$twin_{bjy} = \gamma_0 + \gamma_1 Health_{bjy} + \mu_b + \lambda_y + \varepsilon_{bjy}. \quad (1)$$

Here, *twin* is an indicator of whether a birth of order b born to woman j at age y is a twin. We control for fixed effects for mother's age and parity, as these are known to influence the probability of twin birth. Where births are observed over multiple years, races, or geographic areas, we include the relevant fixed effects. Under the null, the coefficients on maternal health variables $Health_{bjy}$ should not be statistically distinguishable from 0. This is equivalent to a test of (conditional) balance of characteristics of treated (with twins) and control (without twins) mothers. Standard errors are clustered at the level of the mother.

For ease of exposition, we maintain the subscript y for the woman's age at birth, but most of the health indicators are measured before pregnancy to avoid the potential concern of reverse causality—that twin births cause greater depletion of the mother's health than singleton births or encourage women to adopt different behaviors. These include prepregnancy measures of smoking, diabetes, hypertension, obesity, height, kidney disease, and asthma. Measures of prenatal or medical care are constructed as community-level measures of availability. In a specific case we discuss below, we use an exogenous measure of environmental stress in pregnancy. We also show results for some variables measured in

[4]The twin instrument has been criticized for other reasons. A recent critique of the use of twins to identify the QQ trade-off has argued that parental behaviors may respond to the endowment of twins and not only to the fact that twin births represent a fertility shock. Rosenzweig & Zhang (2009) highlight that twins have lower birth endowments. They argue that if parents reinforce endowments, then they may reallocate resources toward the better-endowed children born before the twins, obscuring any underlying QQ trade-off; this is examined in Angrist et al. (2010) and Fitzsimons & Malde (2014). We remain agnostic on this. Our critique is in principle orthogonal to this critique, providing a different reason that an underlying QQ trade-off may be obscured, relating to endowments and behaviors of mothers. This critique has not been previously considered.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/rest_a_00789 by guest on 28 March 2024

pregnancy—smoking, alcohol, drugs, diet—and for one measure (BMI in developing country data) measured after birth. We flag these variables so that their coefficients can be interpreted with this caveat in mind.[5] Importantly, if we dropped all of the flagged variables, we would still have a fairly compelling breadth of evidence. We add controls for education and, where available, wealth to allow for the fact that education may motivate and wealth may facilitate health-seeking behaviors (Kenkel, 1991; Lleras-Muney & Cutler, 2010). This will confirm that the indicators in *Health* are not simply proxying for socioeconomic status. As discussed above, we will present additional specifications including woman fixed effects in the model and restricting to same-sex twins.

B. Pre-Twin Balance

We perform an alternative test that exploits predetermined birth outcomes within mothers. This essentially involves testing whether women who produce twins had, on average, healthier births before the twin birth, as this would be a measure of predetermined maternal health. For each $n = \{2, 3, 4\}$, we estimate

$$PriorDeath_{b<n,jy} = \alpha_0 + \alpha_1 Twin_{b=n,jy} + \lambda_y + \nu_{jby}, \quad (2)$$

where we restrict the sample to prior birth outcomes of mother j who was fully exposed to the risk of death before birth order $b < n$. Thus, for $n = 2$, the independent variable *Twin* takes the value of 1 if the mother gives birth to twins on her second birth and 0 if she gives birth to a singleton on her second birth. We generalize this to higher birth orders. *PriorDeath* refers to the proportion of pre-twin births of a mother who survived and, for instance, for $n = 2$, this is the survival status of the first birth. When we use the U.S. data, this refers to fetal survival, and when we use the Demographic and Health Survey (DHS) data, this refers to survival from birth through to 12 months of age. However, we also show results for size at birth, a less extreme measure of child health than mortality.[6] If women who give birth to twins are selectively healthy, we will observe $\alpha_1 < 0$. Maternal age fixed effects are included. In appendix B.1, we discuss issues relating to the measurement of maternal health and miscarriage data.

[5] The concern is that these variables may respond to a woman's knowledge that she is carrying twins. If the response is to accentuate the relationships of interest—for instance, if she smokes more—then failing to account for this would lead us to underestimate the relationships of interest. However, if instead she increases her attendance of antenatal care and this more than offsets the resource stress of carrying twins, it is possible that we overestimate the relationship. BMI in one data set is available only after birth. If twin births deplete the mother more, then twin mothers will record lower BMI, and accounting for this would only strengthen our contention.

[6] Infant mortality is widely used as a marker of health, and it has the advantage that it is largely predetermined with respect to the following birth (given gestation is about nine months); to ensure this, we remove children born less than a year after their older sibling. Similarly, miscarriage rates have been shown to respond to maternal condition and are high, even in developed country settings.

III. Data

Not all birth records contain indices of maternal health or health-related behaviors. To estimate equation (1), we sought data that did and that were representative and, given the relative rarity of twins, large. Data sets fulfilling these criteria include administrative birth data from the United States, Spain, and Sweden and household survey data from Chile, the United Kingdom, and 68 developing countries (the DHS) for different sets of years. (Details of temporal and geographic coverage, and summary statistics for each data set are provided in online data appendix B.2.) Together, these data sets include 17 million births from 1972 to 2013. We consistently restrict the sample to women aged 18 to 49 years old and exclude triplets and higher-order multiple births. We take advantage of U.S. Vital Statistics data from 2009 to 2013 that identify ART use by birth, removing the approximately 1.6% of births that were ART assisted.[7] For the developing country sample, on the premise that ART was not available prior to 1990, we split the birth data into pre- and post-1990 samples.

Equation (2) is estimated using only the DHS and the U.S. Vital Statistics files. The DHS has the complete fertility history, including the survival status and birthweight of all children preceding each twin or singleton birth, and the U.S. birth certificate data allow us to infer earlier miscarriages for every mother as the difference between total reported births and live births. The miscarriage data are discussed further in section IVB.

IV. Results

A. Twin Births and Maternal Condition

In table 2 we present estimates of equation (1) for several countries using multiple indicators of maternal health. We find broadly consistent results across indicators and across samples. In online appendix C, we provide additional discussion of the stability of the general result across countries and levels of economic development. All independent variables in table 2 are standardized as z-scores so that the estimates can be cast as the effects of increasing by 1 standard deviation (SD) the independent variable of interest. Unstandardized results are presented in appendix table A1.

We find that the probability of twin birth is significantly positively influenced by the following indicators of maternal health included independently: not underweight, tall,[8] more

[7] The data since 2009 also include a range of new measures of maternal morbidity and behaviors.

[8] Height is the indicator of health most widely measured in birth and demographic data, and several studies show that it responds to infection and nutritional scarcity in the growing years; for example, individuals exposed to famine and war have been shown to have lower stature in adulthood, (Silventoinen, 2003; Bozzoli, Deaton, & Quintana-Domeque, 2009; Akresh et al., 2012). Previous research has shown widespread associations of short stature among mothers with the risk of low birthweight and infant mortality among their children (Bhalotra & Rawlings, 2013).

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/rest_a_00789 by guest on 28 March 2024

TABLE 2.—EFFECTS OF MATERNAL HEALTH ON TWIN BIRTHS

Health Behaviors/Access			Health Stocks and Conditions		
Variable	Estimate	[95% CI]	Variable	Estimate	[95% CI]
A. United States (N = 13,646,236, % twin = 2.84)					
Smoked before pregnancy	−0.108***	[−0.116, −0.100]	Height	0.612***	[0.604, 0.620]
Smoked trimester 1[a]	−-0.195***	[−0.203, −0.187]	Underweight	−0.156***	[−0.164, −0.148]
Smoked trimester 2[a]	−0.232***	[−0.240, −0.224]	Obese	0.042***	[0.032, 0.052]
Smoked trimester 3[a]	−0.238***	[−0.246, −0.230]	Diabetes	−0.286***	[−0.296, −0.276]
Education	0.800***	[0.790, 0.810]	Hypertension	−0.223***	[−0.233, −0.213]
B. Sweden (N = 1,240,621, % twin = 2.55)					
Smoked (12 weeks)[a]	−0.266***	[−0.301, −0.231]	Height	0.617***	[0.592, 0.642]
Smoked (30–32 weeks)[a]	−0.285***	[−0.312, −0.258]	Underweight	−0.140***	[−0.173, −0.107]
			Obese	−0.113***	[−0.137, −0.089]
			Asthma	−0.015*	[−0.033, 0.003]
			Diabetes	−0.253***	[−0.278, −0.228]
			Kidney disease	−0.079***	[−0.101, −0.057]
			Hypertension	−0.099***	[−0.121, −0.077]
C. United Kingdom (Avon) (N = 10,463, % twin = 2.37)					
Healthy foods[a]	0.538***	[0.256, 0.820]	Height	0.399***	[0.115, 0.683]
Fresh fruit[a]	0.019	[−0.281, 0.319]	Underweight	−0.161	[−0.439, 0.117]
Alcohol (infrequently)[a]	−0.099	[−0.373, 0.175]	Obese	−0.046	[−0.322, 0.230]
Alcohol (frequently)[a]	−0.358**	[−0.630, −0.086]	Diabetes	−0.056	[−0.328, 0.216]
Passive smoke[a]	0.047	[−0.243, 0.337]	Hypertension	−0.480***	[−0.752, −0.208]
Smoked during pregnancy[a]	−0.162	[−0.448, 0.124]			
Education	0.416*	[−0.002, 0.834]			
D. Chile (N = 14,050, % twin = 2.55)					
Smoked during pregnancy[a]	−0.327***	[−0.572, −0.082]	Underweight	−0.183*	[−0.399, 0.033]
Drugs (infrequently)[a]	0.002	[−0.253, 0.257]	Obese	−0.258***	[−0.446, −0.070]
Drugs (frequently)[a]	−0.161***	[−0.196, −0.126]			
Alcohol (infrequently)[a]	−0.072	[−0.362, 0.218]			
Alcohol (frequently)[a]	−0.172***	[−0.213, −0.131]			
Education	0.529***	[0.200, 0.858]			
E. Developing countries (N = 2,050,795, % twin = 2.07)					
Doctor availability	0.092***	[0.059, 0.125]	Height	0.276***	[0.245, 0.307]
Nurse availability	0.060***	[0.029, 0.091]	Underweight	−0.090***	[−0.115, −0.065]
Prenatal care availability	0.103***	[0.076, 0.130]	Obese	0.059***	[0.028, 0.090]
Education	0.141***	[0.110, 0.172]			

[a]Conditions that are measured during pregnancy, and so may be behavioral responses to twins.

Each coefficient represents a separate regression of child's birth type (twin or singleton) on the mother's health behaviors and conditions. In each sample, all mothers aged 18 to 49 are included. The dependent variable (Twins) is mutliplied by 100, and the independent variables are standardized as z-scores, so coefficients are interpreted as the percentage point change in twin births associated with a 1 standard deviation increase in the variable of interest. All models include fixed effects for age and birth order and, where possible, for wealth (panels A and D) and for gestation of the birth in weeks (panels A and B). Unstandardised and conditional results are included in online appendix tables A1 and A2. Results are robust to the inclusion of education as a quadratic term (appendix table A3). Standard errors are clustered by mother. * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

educated, having greater access to medical or antenatal care, not having smoked before pregnancy, not having any of a range of morbidities prior to conception (obesity, diabetes, hypertension, asthma, kidney disease), and averting risky behaviors in pregnancy (smoking, alcohol, drugs, unhealthy diet). The effects are sizable, with a 1 SD improvement in the indicator tending to increase the likelihood of twinning by 6% to 12% in most cases, relative to a mean of about 2.7% in the (global) sample. There are smaller effects from fresh fruit consumption and larger effects from height. We shall see when we present the pre-twin survival test results below that these effect sizes are broadly comparable to the difference in U.S. data of about 7% in rates of miscarriage of first births between mothers who go on to have twins at second birth and mothers who do not. This similarity of orders of magnitude contributes plausibility to our argument that miscarriage is a mechanism. We directly test this mechanism in section IVB.

Using all available measures of health for each country, we also calculated a factor index of maternal health (as in Biroli, 2016; see appendix D). Mothers of twins consistently have a higher score than mothers of singletons, but as the variables available for each country are different, the scores are not comparable across countries. The statistical significance of these health indicators is robust to running regressions, which condition on all available indicators of the mother's health and, importantly, education (appendix tables A2–A3). Our results all hold after correcting test statistics for large sample sizes that increase the likelihood of rejecting a null, following Deaton, 1997; see appendix table A4). First, we elaborate our findings by country. Then we present results from alternative

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/rest_a_00789 by guest on 28 March 2024

approaches and the robustness checks concerned with the role of genetic traits.

Estimates for the United States. We pool all non-ART births in the United States from 2009 to 2013. We estimate that a 1 SD increase in rates of smoking before pregnancy is associated with a 0.11 percentage point (pp) lower chance of a twin birth, which is about 5.5% of the mean rate of twinning.[9] Diabetes and hypertension prior to pregnancy have standardized effects of 0.2 to 0.3 pp, while being obese or underweight prior to pregnancy has smaller effects of 0.04 and 0.16 pp, respectively. Height and education have larger standardized effects, of 0.61 and 0.8 pp, respectively. In appendix table A6, we remove potential outliers from the sample of mothers when considering height and the results are nearly entirely unchanged. Estimates for women using ART are presented in table A7 and are, with the exception of being underweight, larger and statistically significant for every indicator, underlining the additional sensitivity of birth outcomes in this group.

Estimates for Sweden, Avon (U.K.) and Chile. Analysis of birth registers from Sweden for 1993 to 2012 indicates strikingly similar standardized effect sizes for smoking, diabetes, height, and being underweight to those for U.S. women. There are, however, some differences: the standardized coefficient on obesity in Sweden is about three times as large, while the coefficient on hypertension is only half as large. The Swedish data additionally record asthma prior to conception, which we estimate reduces the risk of twin births by 0.015 pp. Survey data from Avon County for 1991 to 1992 and Chile 2006 to 2009 again exhibit patterns similar to those identified for Sweden and the United States for anthropometric indicators of health, risky behaviors, and prepregnancy illnesses. For instance, for the United Kingdom, estimates for being underweight, obese, or smoking before pregnancy are all very similar to the corresponding estimates for the United States. However, the standardized impact of hypertension before pregnancy is twice as large, and the associations with diabetes, height, and education are smaller. The U.K. data contain unique information on eating healthily during pregnancy, and our estimates indicate that the standardized effect of this is a 0.54 pp increase in the likelihood of having twins, which is the single largest coefficient among variables available for the United Kingdom. The coefficients in the Chilean data for being underweight and for smoking, drugs, and alcohol consumption during pregnancy lie between 0.16 and 0.33 pp, broadly similar to the coefficients for other countries, and the coefficient on obesity is considerably larger (0.26). Chile is the only country in our sample for which we have information on drug use during pregnancy and the standardized effect for this is similar to that for (frequent) alcohol consumption in pregnancy.

[9]Effects of smoking during pregnancy are larger, in the range of 0.20 to 0.24 percentage points, with smoking in the third trimester imposing the largest reduction, consistent with evidence that adverse effects of smoking on birthweight are largest in the third trimester. See Bernstein et al. (2005) and a similar pattern estimated on our data in appendix table A5.

TABLE 3.—TWINNING AND STRESS IN UTERO

Dependent Variable: Twins × 100	(1)	(2)	(3)
ETA bomb casualities first trimester of pregnancy	0.002	−0.002	−0.002
	(0.006)	(0.006)	(0.004)
ETA bomb casualities second trimester of pregnancy	−0.010***	−0.010***	−0.010***
	(0.004)	(0.004)	(0.004)
ETA bomb casualities third trimester of pregnancy	−0.012*	−0.013*	−0.013**
	(0.007)	(0.008)	(0.006)
Observations	6,793,890	6,759,120	6,759,120
Year × month and province FE	Yes	Yes	Yes
Sociodemographic controls		Yes	Yes
Provice-specific linear year-month trends			Yes

Data consist of the Quintana-Domeque and Ródenas-Serrano (2017) sample of live births conceived between January 1980 and February 2003. Treatment is defined as number of ETA bomb casualties in the province of conception. Standard errors are clustered at the level of the province (fifty provinces). * $p < 0.1$; ** $p < 0.05$; and *** $p < 0.01$.

Estimates for developing countries. In the sample that pools data for 68 developing countries for 1972 to 2012, we observe height, weight, body mass index, and local availability of prenatal care and access to medical professionals. Reproductive health service coverage is far from universal in low-income countries, although this is a leading global health priority.[10] After adjusting for demographic covariates as for the other samples, we observe again that taller and heavier women are more likely to twin. This is true even in the pre-ART period (see table A8). The effects of height, underweight, and education are all smaller than in richer countries, while the effects of obesity are larger than in all countries other than Chile.[11] We estimate that a 1 SD increase in availability of doctors or nurses is associated with a 0.092 pp and 0.06 pp increase in the likelihood of twins, respectively.

Quasi-experimental variation in a negative intrauterine shock: Spain. Using the methodology and data described in Quintana-Domeque and Ródenas-Serrano (2017), we estimated the impact of bombing by ETA, a Basque terrorist group, as a plausibly exogenous negative intrauterine shock that may cause fetal stress, a proxy for maternal health in pregnancy. We find that an additional bomb casualty in the province of residence of a pregnant woman decreases the likelihood that she will have a twin birth by 0.01% and 0.012% (see table 3). This effect is larger and statistically significant only during the second and third trimesters, similar to the effects of smoking by trimester documented in table 2.[12]

[10]These variables are all measured as the rate of health care access in the mother's cluster of residence since we are interested in availability rather than use to avoid the concern that mothers conceiving twins may be more likely to actively seek birth attendance.

[11]Recall these are standardized effects; unstandardized effects are in the appendix.

[12]Quintana-Domeque and Ródenas-Serrano (2017) find that the same exposure reduces average birthweight by approximately 0.3 grams (trimester

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/rest_a_00789 by guest on 28 March 2024

TABLE 4.—TEST OF HYPOTHESIS THAT WOMEN WHO BEAR TWINS HAVE BETTER PRIOR HEALTH

	(1) Birth 2	(2) Birth 3	(3) Birth 4
A. Developing country (Dependent variable = infant mortality rate × 100)			
Treated	0.211	−1.283***	−1.722***
	(0.183)	(0.154)	(0.148)
Mean value	9.983	10.443	11.159
Observations	542,186	422,498	312,350
B. U.S. birth certificates (dependent variable = miscarriage rate × 100)			
Treated	−0.727***	−0.238***	−0.063
	(0.050)	(0.053)	(0.067)
Mean value	10.880	10.519	9.911
Observations	4,945,728	2,657,239	1,131,971

Dependent variables are constructed as the proportion of any prior births that died in the first year of life (panel A) or resulted in miscarriage or fetal death (panel B). Regressions are run at the level of the mother, taking averages over all prior births/pregnancies. In panel A, only children who have been entirely exposed to the risk of infant mortality are included (those over 1 year of age). "Treated" refers to giving birth to twins (rather than singletons) at the birth order indicated in the column heading. A full description of these samples and the treatment variable is provided in section II. Regressions include mother's age and race fixed effects. Standard errors are robust to heteroskedasticity. * $p < 0.1$; ** $p < 0.05$; and *** $p < 0.01$.

Survival of pre-twins as a marker of mother's health. Here we discuss the alternative test of the quasi-randomness of twin births. Estimates of equation (2) are in table 4. In the developing country sample, mothers who went on to have third- and fourth-born twins had an infant mortality rate 1.3 to 1.7 pp lower among their prior births than women who had singletons at the same birth order. This is a natural measure of maternal health, capturing a woman's ability to produce surviving children, which is exactly what we hypothesize is challenged by carrying twins. We used birth size as a measure of child health that is less extreme than mortality. We used the DHS again, as it allows us to observe all children ordered within mother, and we find that earlier births of women who later have a twin birth are less likely to be small at birth than the corresponding births of women who have only singleton children (see appendix table A9).

Similarly, in the U.S. population, we observe that women who have twins are less likely to have suffered a miscarriage prior to the twin birth. Mothers who give birth to twins at second birth are 0.7 pp less likely to have suffered a miscarriage of their first conception, which is 6.7% of the baseline rate for this group. The rate of miscarriages in the population of all women who gave birth was approximately 10%. Parity-specific estimates and means are in table 4.

Specification check, including woman fixed effects. So as to control for any genetic characteristics of the mother, we sought data that follow women over time, recording multiple births per woman as well as time-varying measures of maternal health. Such data are scarce, but the National Longitudinal Survey of Young Women (NLSY) meets these requirements. A sample of 5,159 mothers aged between 14 and 24 in 1968 is followed until 1999, when the youngest are aged 45. The health variables measured consistently through this period are whether the mother has any physical limitation that restricts her ability to work, whether she smoked prior to the pregnancy, and whether she has had a prior cancer diagnosis. (More information on the data structure and summary statistics is in appendix E.) We estimate the probability that a birth is a twin as a function of these indicators of maternal health conditional on mother fixed effects and controlling also for a quadratic in family income, mother's age, birth order, and year of birth fixed effects. (Results are in appendix table A17.) We find large statistically significant negative effects of smoking and cancer on the probability of having a twin birth, and no significant impact of health limiting work.

Specification check using monozygotic twins. The risk of giving birth to dizygotic twins (DZ) is elevated among women with high levels of the follicle-stimulating hormone (FSH), which is often more prevalent among taller and heavier women (Li et al., 2003; Hall, 2003; Hoekstra et al., 2008). Since dizygotic twins constitute about two-thirds of all twins, this could in principle contribute to explaining the associations we document with height and BMI (note that the biomedical literature has not documented these associations in any population-level data, let alone across countries, time, and indicators). Although, as discussed, genetic predispositions cannot explain our finding that health behaviors or aspects of the health environment (stress or prenatal care availability) predict twinning, we investigate this further by exploiting the fact that MZ twins are necessarily same sex (and about half of DZ twins are same sex) and repeat the analysis removing mixed-sex twins from the data.[13] Results are in appendix table A10. We continue to find significant associations between proxies for maternal health and the chances of a twin versus a singleton birth, and the coefficients are not significantly different from those that obtain in the full sample.

B. *Mechanisms of Twin Selection*

We consider three alternative hypotheses for why maternal health may influence the probability of twinning, which relate to conception, gestation, and maternal survival. First, healthier mothers may be more likely to conceive twins on account of an underlying genetic or biological process. Second, conditional upon conceiving twins, healthier mothers may be more likely to take them to term. Third, conditional on conceiving twins and taking them to term, healthier mothers may be more likely to survive the birth, and hence appear in the available data.

1) and increases the likelihood of low birthweight by 0.14%. Placebo tests in support of their methodology, including examining the impact of bombs post-birth on birth outcomes, are presented in their paper.

[13] Note that this also addresses the elevation of twin birth rates among ART users for the DHS where ART use is not recorded, because most, if not all, of ART-generated twins are DZ. We implement this test using DHS data only, as in other administrative data sets, twins are not matched with their siblings to infer whether they are of the same sex.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/rest_a_00789 by guest on 28 March 2024

TABLE 5.—FETAL DEATHS, TWINNING, AND HEALTH BEHAVIORS

Dependent Variable: Fetal Death × 1,000	(1) Smokes	(2) Drinks	(3) No College	(4) Anemic	(5) No Smoking	(6) No Drinking	(7) Years of Education
A. Uninteracted Twin-Nontwin Difference							
Twin	9.979***	10.375***	10.397***	11.387***	9.971***	10.367***	10.397***
	(0.118)	(0.119)	(0.108)	(0.115)	(0.117)	(0.119)	(0.108)
Constant	5.344***	5.508***	5.172***	5.964***	5.337***	5.500***	5.172***
	(0.021)	(0.021)	(0.019)	(0.020)	(0.021)	(0.021)	(0.019)
B. Health, Twin and Twin × Health Interaction							
Twin	9.907***	10.368***	8.991***	11.337***	9.939***	10.354***	19.630***
	(0.123)	(0.119)	(0.145)	(0.117)	(0.121)	(0.119)	(0.552)
Health (Dis)amenity	1.394***	4.924***	1.683***	0.608***	0.108***	0.602***	−0.242***
	(0.066)	(0.260)	(0.038)	(0.131)	(0.005)	(0.038)	(0.007)
Twin × Health	1.154***	3.559**	3.573***	1.303**	0.061*	0.756***	−0.674***
	(0.416)	(1.754)	(0.218)	(0.641)	(0.032)	(0.206)	(0.040)
Constant	5.195***	5.476***	4.268***	5.949***	5.214***	5.482***	8.277***
	(0.022)	(0.021)	(0.028)	(0.020)	(0.021)	(0.021)	(0.088)
Observations	13,660,400	13,809,830	15,909,836	16,158,564	13,679,142	13,828,573	15,909,836

Each column in panel A represents a regression of whether a pregnancy ends in a fetal death (multiplied by 1,000) on whether the pregnancy is a twin pregnancy. Panel B augments the same regressions to include a health behavior or health stock, and the interaction between being a twin pregnancy and the health variable. The health variable in each column is indicated in the column title. Regressions including controls for mother's age, child birth year and total fertility fixed effects are presented in appendix table A11. * $p < 0.10$; ** $p < 0.05$; and *** $p < 0.01$.

Either of the first two processes is sufficient to violate the "as good as random" assumption insofar as they imply that observing twins will depend on possibly unmeasured maternal behaviors and characteristics. Since taller and heavier women and active smokers have higher levels of the FSH hormone associated with multiple births (Li et al., 2003; Hall, 2003; Hoekstra et al., 2008; Cramer et al., 1994), conception of twins may not be random. We cannot directly test the conception hypothesis since the required data are unavailable, but we now provide tests of the other two hypotheses and indicate the manner in which nonrandom conception will influence interpretation of our results.

Selective fetal death. The gestation hypothesis is that carrying twins to term is more demanding than carrying singletons to term, and so stressors of maternal health will lead to selective miscarriage of twins. It has been documented that the biological demands of twin pregnancies are higher than the demands of nontwin pregnancies (Shinagawa et al., 2005) and also that in general, healthier mothers are less likely to miscarry (García-Enguídanosa et al., 2002). What we contribute here is to test the natural intersection of these hypotheses and estimate the extent to which miscarriage is more frequent among less-healthy women carrying twins. The estimated equation is

$$FetalDeath_{ijt} = \gamma_0 + \gamma_1 Twin_{ijt} + \gamma_2 Health_{jt} + \gamma_3 Twin \times Health_{ijt} + \lambda_t + \phi_y + \mu_b + u_{ijt}. \quad (3)$$

$FetalDeath_{ijt}$ is a binary variable (multiplied by 1,000) indicating whether a birth was taken to term (coded as 0) or resulted in a miscarriage (coded as 1), i indicates a conception leading to birth or fetal death, j a mother, and t is year. *Health* is an indicator of the mother's health, *Twin* is an indicator for whether the conception is a twin or a singleton, and, as before, fixed effects for year (λ_t), birth order (μ_b), and mother's age (ϕ_y) are included. The coefficient of interest, γ_3, is the differential effect of the variable $Health_{jt}$ on twin conceptions.

Birth registers often do not include maternal health indicators, and if they do, it is unusual that they also also include information on fetal deaths, but the U.S. Vital Statistics data do.[14] We pooled all births and fetal deaths recorded from 1999 to 2002. We stopped in this year because from 2003, a considerable redefinition of birth certificate data meant that fetal death and birth data did not share similar controls. Prior to 2002, however, we are able to observe for all states whether a mother smokes or drinks during pregnancy, whether she suffered from anemia prior to pregnancy, and her educational level. The results, using the U.S. birth certificate and fetal death data, are in table 5. In panel A, we document the difference in the risk of fetal death for twin relative to singleton conceptions. The evidence confirms previous research showing that the spontaneous abortion rate among twins (at one in eight conceptions), is about three times that among singletons (Boklage, 1990). In panel B, we test how maternal health indicators modify this differential risk. We can consistently reject that the interaction term γ_3 is 0. In other words, twin fetal survival is more sensitive to mother's health than singleton survival. For example, a 1 SD increase in rates of smoking while carrying a singleton elevates the risk of miscarriage by 1.39 fetal deaths per 1,000 live births. The corresponding risk

[14] We would have liked to replicate this analysis in other data sets. However, we are not aware of other data that have all the details necessary to run such a test—in particular, maternal health outcomes, births, miscarriages, and an indicator of whether miscarried births were twin or singleton. For instance, in DHS data, miscarriages are recorded in certain surveys, such as for Nepal, but there is no record of whether they are for single or twin pregnancies.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/rest_a_00789 by guest on 28 March 2024

elevation among mothers pregnant with twins is an increase of 2.55 fetal deaths, almost twice the risk. Alcohol consumption is similarly almost twice as risky for women carrying twins, and the risks associated with anemia are about three times as high. We also see that a college education, which is a predictor of healthy behavior, modifies the difference in miscarriage probabilities more than three times as much when the mother is carrying twins than when she is carrying a singleton. Now it may be that one of two twins miscarries. In such cases, if the survivor is recorded as a singleton birth, then we tend to underestimate the importance of maternal condition. In other words, our contention holds a fortiori.

Overall, these results establish a plausible mechanism for the associations that we document in tables 2 to 4. Here we have modeled miscarriage conditional on the conception being twin or singleton. If in fact maternal health also raises the chances of a twin conception, then this will reinforce our contention. If, instead, maternal health is for some undocumented reason negatively associated with twin conception, then our findings hold despite this and are conservative.

Trivers and Willard (1973) made an argument similar to ours but pertaining to the distribution of sons across women (Almond & Edlund, 2007). They observed that since the male fetus is more vulnerable to adverse health conditions (Waldron, 1983), sons are more likely to be born of healthy mothers. As for twins, so for sons, selective miscarriage is the suggested mechanism. Intersecting our hypothesis with theirs, we investigated whether males are underrepresented among twins, other things equal. We used the large data sets in table 2 (data from the United States, Sweden, and developing countries). We find that twin births are approximately 0.1 to 0.3 pp more likely to be female ($p < 0.001$). This affords a further test of our hypothesis and a validation of the Trivers-Willard hypothesis (refer to appendix table A12).[15]

Our findings suggest that twin birth is a marker of fetal health. Our findings, which range across indicators and countries, highlight the relevance of maternal health to fetal health. Recent research demonstrating long-run socioeconomic returns to investing in fetal and infant health, improving the preschool environment, and raising parenting quality has stimulated policy interventions across the world that are motivated to enhance the potential for nurture to lift up the trajectories of children, especially when born into disadvantaged circumstances (Heckman et al., 2010; Almond & Currie, 2011; Carneiro, Løken, & Salvanes, 2015). Our results point to the significance of, for instance, nutrition, stress, and prenatal care for mothers in achieving these goals.

Selective maternal survival. A potential concern is that the less healthy women among those who delivered twins died in childbirth, and data sets like the DHS that obtain birth histories from mothers will not contain those women. In such cases, our findings could arise from selective maternal survival. This concern does not apply to the administrative U.S. and Swedish data where all births are recorded and where we see clear associations of twinning and maternal health, so it cannot be the only explanation of those findings. Similarly, in the U.K. and Chile data sets, the survey design ensures that representative coverage is not affected by maternal death.[16] The lifetime risk of a maternal death is 1 in 41 in low-income countries as compared with 1 in 3,300 in high-income countries. If twins only spuriously appear to be born of healthier mothers due to selective maternal death, then as mothers become more likely to survive childbirth (i.e., as maternal mortality declines), the associations should dissipate. The fact that they do not also undermines the relevance of selection.

We assess the magnitude of selection bias in our DHS estimates, following Alderman, Lokshin, and Radyakin (2011). We simulate the presence of the women who died and test whether correcting for maternal survival selection causes the association of twin births and maternal health to disappear.

A data challenge is that we do not observe the health of women who died in childbirth; indeed, the original problem is that we do not observe these women at all. We address this by using the maternal mortality status of all sisters of every female respondent.[17] We assume that the respondent's health (indicated by height and BMI) proxies the health of her sisters and validate this (figure A1).[18] We put our results to the harshest test by assuming that less healthy women who died in childbirth were all carrying twins, and healthier women who died in childbirth were not carrying twins, and the results stand up to this (see table 6). We test the sensitivity of the adjusted estimates to a range of different binary distinctions of healthy versus less healthy. Overall, these results establish that maternal mortality selection does not drive the DHS results.

V. Conclusion and Discussion

Twin births are not random. We show that mothers who have twin births are healthier prior to the occurrence of the

[15] We found an older biological literature that recognizes that boys are underrepresented among twins and even more underrepresented among triplets (James, 1975; Bulmer, 1970), but this literature does not explicitly link in with Trivers-Willard. When interacting twin births by maternal characteristics in table A12, most coefficients are not significantly different by the gender of the twins. However, two coefficients are significantly larger (more negative) for boys, consistent with the male fetus being more sensitive to fetal health.

[16] In data from the United Kingdom, women were prospectively enrolled when pregnant entirely before exposure to considerable maternal mortality risk and children were subsequently followed over their lives. In the data from Chile, a representative sample was chosen after birth; however, the sampling unit was at the level of the child rather than the mother, so children would be represented even in cases where their mother was no longer alive.

[17] Most DHS countries are in Africa, where fertility and maternal mortality are high.

[18] Maternal mortality is significantly higher among sisters of women with lower height or BMI, conditional on country and year fixed effects, a quadratic in mother's age, and age at first birth. Sisters of women shorter than the mean height of 155.5 cm are considerably more likely to have suffered maternal death, and there is a sharper gradient for women shorter than 145 cm.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/rest_a_00789 by guest on 28 March 2024

TABLE 6.—CAN SELECTIVE MATERNAL SURVIVAL EXPLAIN TWINNING RATES?

Dependent Variable: Twins × 100	MMR Sample	<140 cm or BMI <16	<145 cm or BMI <16.5	<150 cm or BMI <17	<155 cm or BMI <17.5
Height	0.065***	0.063***	0.058***	0.051***	0.044***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
BMI	0.048***	0.045***	0.044***	0.041***	0.042***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Observations	844,552	848,552	848,492	848,468	848,507
R^2	0.024	0.024	0.023	0.023	0.022

Each column presents a regression of maternal characteristics on twinning following specification 1. Column 1 includes the full sample of women surveyed in countries where the DHS maternal mortality module is applied. In columns 2 to 5, we inflate the sample by the number of women who, according to our sister method calculations, would exist in the sample if it were not for the fact that they died in childbirth (this match assumes that a woman's health is a good proxy for her sister's health, and estimates will be less precise if this proxy is weak). However, our measure of (sister) maternal mortality is very clearly decreasing in (respondent) height (see figure A1). We then examine the coefficients of interest in the estimates of equation (1) under the extreme assumption that all less healthy women who died were pregnant with twins, while all healthy women who died were not. We create a range of different binary distinctions of healthy versus less healthy, using the available individual data on height and BMI, with cutoffs described in column headings. Heteroskedasticity robust standard errors are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.1$.

twin birth. The findings in this paper have implications for identification strategies in economics and a number of other fields of research, and they extend the existing social science and biomedical literature on twinning. Here we delineate these contributions, using a list format for clarity.

a. The biomedical literature has identified an association of twinning with the height, weight, and smoking status of the mother and attributed this to hormonal variation. This is the first study to demonstrate that these associations hold in representative population-level data in several richer and poorer countries and across several years. We also show that these associations hold conditional not only on age and parity (known predictors) but also on the mother's socioeconomic status and a range of other indicators of her health.
b. This is the first paper to demonstrate associations of twinning with other indicators of maternal health. These include a range of (prepregnancy and pregnancy) morbidities, the health of her lower-order births, the mother's health-related behaviors (before and during pregnancy), availability of reproductive health services, and indicators of the mother's exposure to environmental stress in pregnancy. The last three are clearly not genetic or hormonal associations. We nevertheless show associations of maternal health and twinning conditional on woman fixed effects that purge genetic differences between women.
c. Since it is known that twins are more likely among ART-assisted births and that ART users tend to be more educated, we show that associations of maternal health with twinning hold in ART-purged and pre-ART data samples. We also show the first systematic evidence that the education of the mother is positively associated with twinning in these samples, consistent with educated women being more likely to engage in health-seeking behaviors.
d. Our findings indicate no clear tendency for the association of maternal health with twinning to dissipate with economic development. Although intrinsic maternal health and access to public health services tend to improve with economic development, it is unclear that all relevant indicators (hypertension, obesity, diabetes) improve, and differences between rich and poor countries in age, parity, and race will also modify this relationship.
e. We are able to demonstrate that maternal health determines fetal selection conditional on conception of twins. The biomedical literature has discussed hormonal (FSH) predictors; our hypothesis that it is selectively healthy women who are able to mount the challenge of carrying twins to birth is new.
f. In the economics literature, the validity of several studies investigating the hypothesis that fertility has a causal effect on investments in children, or on women's labor supply, rests on the assumption that twin births are random (at least conditional on age, parity, and education). Twin births are used as an instrument because OLS estimates tend to be biased upward on account of negative selection of women into fertility. Our findings suggest that twin-IV estimates will tend to be biased downward on account of positive selection into twin birth. This is important because recent prominent studies cited in section I find that the trade-off is frequently not statistically different from 0, and in principle, this could be explained by a downward bias in the estimates.[19]

Educational attainment has risen considerably, while completed and desired fertility have fallen sharply over the past fifty years (see Hanushek, 1992). It is of considerable relevance to researchers and policymakers to determine whether these trends contain a causal component. Similarly, the fertility-work trade-off for women is topical again as educational attainments of women are overtaking those of men and transforming the work-family balance, with consequences for women's autonomy, marital stability, and child outcomes (Newman & Olivetti, 2016; Lundberg et al., 2016).

[19]For the estimates to be biased, we require not only that the probability of a twin birth is nonrandom and associated with potentially unobservable maternal health, but also that maternal health is correlated with the propensity to invest in children or to participate in the labor force, as the case may be. There is evidence of the latter.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/rest_a_00789 by guest on 28 March 2024

REFERENCES

Akresh, R., S. Bhalotra, M. Leone, and U. Osili, "War and Stature: Growing Up during the Nigerian Civil War," *American Economic Review (Papers and Proceedings)* 102 (2012), 273–277.

Alderman, H., M. Lokshin, and S. Radyakin, "Tall Claims: Mortality Selection and the Height of Children," World Bank policy research working paper 5846 (2011).

Almond, D., and J. Currie, "Killing Me Softly: The Fetal Origins Hypothesis," *Journal of Economic Perspectives* 25 (2011), 153–172.

Almond, D., and L. Edlund, "Trivers-Willard at Birth and One Year: Evidence from US Natality Data 1983–2001," *Proceedings of the Royal Society of London B: Biological Sciences* 274:1624 (2007), 2491–2496.

Angrist, J. D., and W. N. Evans, "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size," *American Economic Review* 88 (1998), 450–477.

Angrist, J., V. Lavy, and A. Schlosser, "Multiple Experiments for the Causal Link between the Quantity and Quality of Children," *Journal of Labor Economics* 28 (2010), 773–824.

Åslund, O., and H. Grönqvist, "Family Size and Child Outcomes: Is There Really No Trade-Off?" *Labour Economics* 17:1 (2010), 130–139.

Becker, G. S., and H. G. Lewis, "On the Interaction between the Quantity and Quality of Children," *Journal of Political Economy* 81 (1973), S279–S288.

Bernstein, I. M., J. A. Mongeon, G. J. Badger, L. Solomon, S. H. Heil, and S. T. Higgins, "Maternal Smoking and Its Association with Birth Weight," *Obstetrics and Gynecology* 106 (2005), 986–991.

Bhalotra, S., and S. Rawlings, "Gradients of Intergenerational Transmission of Health in Developing Countries," this REVIEW 95 (2013), 660–672.

Biroli, P., "Health and Skill Formation in Early Childhood," University of Zurich, UBS International Center of Economics in Society working paper 17 (2016).

Black, S. E., P. J. Devereux, and K. G. Salvanes, "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education," *Quarterly Journal of Economics* 120 (2005), 669–700.

——— "Small Family, Smart Family? Family Size and the IQ Scores of Young Men," *Journal of Human Resources* 45 (2010), 33–58.

——— "Does Grief Transfer across Generations? Bereavements during Pregnancy and Child Outcomes," *American Economic Journal: Applied Economics* 8 (2016), 193–223.

Bloom, D. E., M. Kuhn, and K. Prettner, "The Contribution of Female Health to Economic Development," NBER working paper 21411 (2015).

Boklage, C. E., "Survival Probability of Human Conceptions from Fertilization to Term," *International Journal of Fertility* 35 (1990), 79–94.

Bozzoli, C., A. Deaton, and C. Quintana-Domeque, "Adult Height and Childhood Disease," *Demography* 46 (2009), 647–669.

Bronars, S. G., and J. Grogger, "The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment," *American Economic Review* 84 (1994), 1141–1156.

Bulmer, M. G., *The Biology of Twinning in Man* (Oxford: Clarendon Press, 1970).

Cáceres-Delpiano, J., "The Impacts of Family Size on Investment in Child Quality," *Journal of Human Resources* 41 (2006), 738–754.

——— "Can We Still Learn Something from the Relationship between Fertility and Mother's Employment? Evidence from Developing Countries," *Demography* 49 (2012), 151–174.

Campbell, D. M., "Epidemiology of Twinning," *Current Obstetrics and Gynaecology* 8 (1998), 126–134.

Campbell, D. M., A. J. Campbell, and I. MacGillivary, "Maternal Characteristics of Women Having Twin Pregnancies," *Journal of Biosocial Science* 6 (1974), 463–470.

Carneiro, P., K. Løken, and K. G. Salvanes, "A Flying Start? Maternity Leave Benefits and Long-Run Outcomes of Children," *Journal of Political Economy* 123 (2015), 365–412.

Chay, K., and M. Greenstone, "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession," *Quarterly Journal of Economics* 118 (2003), 1121–1167.

Clarke, D., "Children and Their Parents: A Review of Fertility and Causality," *Journal of Economic Surveys* 32 (2018), 518–540.

Cramer, D. W., R. L. Barbieri, H. Xu, and J. K. Reichardt, "Determinants of Basal Follicle-Stimulating Hormone Levels in Premenopausal Women," *Journal of Clinical Endocrinology and Metabolism* 79 (1994), 1105–1109.

Currie, J., and E. Moretti, "Biology as Destiny? Short- and Long-Run Determinants of Intergenerational Transmission of Birth Weight," *Journal of Labor Economics* 25 (2007), 231–264.

Deaton, A., *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy* (Baltimore: Johns Hopkins University Press, 1997).

Fitzsimons, E., and B. Malde, "Empirically Probing the Quantity-Quality Model," *Journal of Population Economics* 27 (2014), 33–68.

García-Enguídanosa, A., M. Calleb, J. Valeroc, S. Lunaa, and V. Domínguez-Roja, "Risk Factors in Miscarriage: A Review," *European Journal of Obstetrics and Gynecology and Reproductive Biology* 102 (2002), 111–119.

Hall, J. G., "Twinning," *Lancet* 362:9385 (2003), 735–743.

Hanushek, E. A., "The Trade-Off between Child Quantity and Quality," *Journal of Political Economy* 100 (1992), 84–117.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz, "The Rate of Return to the HighScope Perry Preschool Program," *Journal of Public Economics* 94 (2010), 114–128.

Hoekstra, C., Z. Z. Zhao, C. B. Lambalk, G. Willemsen, N. G. Martin, D. I. Boomsma, and G. W. Montgomery, "Dizygotic Twinning," *Human Reproduction Update* 14 (2008), 37–47.

Jacobsen, J. P., J. W. Pearce, III, and J. L. Rosenbloom, "The Effects of Childbearing on Married Women's Labor Supply and Earnings: Using Twin Births as a Natural Experiment," *Journal of Human Resources* 34 (1999), 449–474.

James, W. H., "Sex Ratio in Twin Births," *Annals of Human Biology* 2 (1975), 365–378.

Kenkel, D. S., "Health Behavior, Health Knowledge, and Schooling," *Journal of Political Economy* 99 (1991), 287–305.

Li, H., J. Zhang, and Y. Zhu, "The Quantity-Quality Trade-Off of Children in a Developing Country: Identification Using Chinese Twins," *Demography* 45 (2008), 223–243.

Li, Z., J. Gindler, and H. Wang, "Folic Acid Supplements during Early Pregnancy and Likelihood of Multiple Births: A Population-Based Cohort Study," *Lancet* 361 (2003), 380–384.

Lleras-Muney, A., and D. Cutler, "Understanding Differences in Health Behaviors by Education," *Journal of Health Economics* 29 (2010), 1–28.

Lundberg, S., R. A. Pollak, and J. Stearns, "Family Inequality: Diverging Patterns in Marriage, Cohabitation, and Childbearing," *Journal of Economic Perspectives* 30 (2016), 79–102.

Lundborg, P., E. Plug, and A. W. Rasmussen, "Can Women Have Children and a Career? IV Evidence from IVF Treatments," *American Economic Review* 107 (2017), 1611–1637.

Marteleto, L. J., and L. R. de Souza, "The Changing Impact of Family Size on Adolescents' Schooling: Assessing the Exogenous Variation in Fertility Using Twins in Brazil," *Demography* 49 (2012), 1453–1477.

Mazumder, B., and Z. Seeskin, "Breakfast Skipping, Extreme Commutes and the Sex Composition at Birth," *Biodemography and Social Biology* 61 (2015), 187–208.

Meulemans, W. J., C. M. Lewis, D. I. Boomsma, C. A. Derom, H. Van den Berghe, J. F. Orlebeke, R. F. Vlietinck, and R. M. Derom, "Genetic Modelling of Dizygotic Twinning in Pedigrees of Spontaneous Dizygotic Twins," *American Journal of Medical Genetics* 61 (1996), 258–263.

Mogstad, M., and M. Wiswall, "Testing the Quantity-Quality Model of Fertility: Linearity, Marginal Effects, and Total Effects," *Quantitative Economics* 7 (2016), 157–192.

Newman, A. F., and C. Olivetti, "Career Women and the Durability of Marriage," mimeograph, Boston University (2016).

Polderman, T. J. C., B. Benyamin, C. A. de Leeux, P. F. Sullivan, A. van Bochoven, P. M. Visscher, and D. Posthuma, "Meta-Analysis of the Heritability of Human Traits Based on Fifty Years of Twin Studies," *Nature Genetics* 47 (2015), 702–709.

Ponczek, V., and A. P. Souza, "New Evidence of the Causal Effect of Family Size on Child Quality in a Developing Country," *Journal of Human Resources* 47 (2012), 64–106.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/rest_a_00789 by guest on 28 March 2024

Quintana-Domeque, C., and P. Ródenas-Serrano, "The Hidden Costs of Terrorism: The Effects on Health at Birth," *Journal of Health Economics* 56 (2017), 47–60.
Rosenzweig, M. R., and K. I. Wolpin, "Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment," *Econometrica* 48 (1980a), 227–240.
——— "Life-Cycle Labor Supply and Fertility: Causal Inferences from Household Models," *Journal of Political Economy* 88 (1980b), 328–348.
——— "Natural 'Natural Experiments' in Economics," *Journal of Economic Literature* 38 (2000), 827–874.
——— "Do Population Control Policies Induce More Human Capital Investment? Twins, Birth Weight and China's One-Child Policy," *Review of Economic Studies* 76 (2009), 1149–1174.
Shinagawa, S., S. Suzuki, H. Chihara, Y. Otsubo, T. Takeshita, and T. Araki, "Maternal Basal Metabolic Rate in Twin Pregnancy," *Gynecologic and Obstetric Investigation* 60 (2005), 145–148.
Silventoinen, K., "Determinants of Variation in Adult Body Height," *Journal of Biosocial Science* 35 (2003), 263–285.
Thorndike, E. L., "Measurement of Twins," *Journal of Philosophy, Psychology and Scientific Methods*, 2 (1905), 547–553.
Trivers, R. L., and D. E. Willard, "Natural Selection of Parental Ability to Vary the Sex Ratio of Offspring," *Science* 179:4068 (1973), 90–92.
Vitthala, S., T. A. Gelbaya, D. R. Brison, C. T. Fitzgerald, and L. G. Nardo, "The Risk of Monozygotic Twins after Assisted Reproductive Technology: A Systematic Review and Meta-Analysis," *Human Reproduction Update* 15 (2009), 45–55.
Waldron, I., "The Role of Genetic and Biological Factors in Sex Differences in Mortality," (pp. 141–164), in A. D. Lopez and L. T. Ruzicka, eds., *Sex Differentials in Mortality: Trends, Determinants and Consequences* (Canberra: Australian National University, Department of Demography, 1983).

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/rest_a_00789 by guest on 28 March 2024