Alice Goisis, Hanna Remes, Kieron Barclay, Pekka Martikainen and Mikko Myrskylä

Paternal age and the risk of low birth weight and preterm delivery: a Finnish register-based study

Article (Accepted version)
(Refereed)

Paternal age and the risk of low birth weight and preterm delivery: a Finnish register-based study

Alice Goisis[1,2], Hanna Remes[3], Kieron Barclay[1,2,4], Pekka Martikainen[2,3], Mikko Myrskylä[1,2,3]

[1]Department of Social Policy, London School of Economics and Political Science, Houghton Street, London,WC2A 2AE, United Kingdom

[2]Max Planck Institute for Demographic Research, Konrad-Zuse-Str 1, 18057 Rostock, Germany

[3]Population Research Unit, Department of Social Research, University of Helsinki, P.O. Box 18, 00014, Finland

[4]Department of Sociology, Stockholm University, 106 91 Stockholm Sweden

Correspondence to:

Alice Goisis, Department of Social Policy, London School of Economics and Political Science, Houghton Street, London,WC2A 2AE, United Kingdom

e-mail a.goisis@lse.ac.uk tel: +44 (0) 2079556582

Abstract

Background Based on existing studies, there is no conclusive evidence as to whether and why paternal age matters for birth outcomes.

Methods We used Finnish Population Registers on 106,652 children born 1987-2000. We first document the unadjusted association between paternal age and the risk of low birth weight (<2500g; LBW) and preterm birth (<37 weeks gestation). Second, we investigate whether the unadjusted association is attenuated on adjustment for child's, maternal and parental socio-economic characteristics. Third, by adopting a within-family design which involves comparing children born to the same father at different ages, we additionally adjust for unobserved parental characteristics shared between siblings.

Results The unadjusted results show that being born to a father aged 40+, as opposed to a father 30-34, is associated with an increased risk of LBW of 0.96 percent (95% CI: 0.5%-1.3%) and to a younger father (<25) with a 1 percent (95% CI: 0.6%-1.3%) increased risk. The increased risk at younger paternal ages is halved on adjustment for the child's characteristics and fully attenuated on adjustment for child/parental characteristics. The increased risk at paternal ages 40+ is partially attenuated on adjustment for maternal characteristics (β=0.62%; 95% CI: 0.13%-1.1%). Adjustment for unobserved parental characteristics shared by siblings further attenuates the 40+ coefficient (β=0.4%; 95% CI: -0.5%-1.2%). Results for preterm delivery are similar.

Conclusions the results underscore the importance of considering paternal age as a potential risk factor for adverse birth outcomes and of expanding research on its role and the mechanisms linking it to birth outcomes.

Summary box

What is already known on this subject?
There is no conclusive evidence as to whether and why paternal age matters for birth outcomes. Some studies show that a young or old paternal age at birth is associated with worse birth outcomes, whilst others fail to find one.

What does this study add?
Using Finnish register data, we study the association between paternal age and the risk of poor birth outcomes (LBW and preterm delivery) by carefully comparing unadjusted and adjusted results, including unobserved parental characteristics shared by siblings. In the unadjusted analyses, children born to younger and older fathers were at higher risk of poor birth outcomes. The association was attenuated on adjustment for birth order and maternal age at birth. Adjustment for unobserved parental characteristics shared by siblings further attenuated but didn't entirely explain the association. The results underscore the importance of considering paternal age – alongside maternal age – as a potential risk factor for adverse birth outcomes and of expanding research on its role and the mechanisms linking it to birth outcomes.

Introduction

Prior literature has identified several social and physiological characteristics of mothers as well as fathers as predictors of birth weight and gestational age. These include factors such as maternal smoking, maternal race/ethnicity, the sex and birth order of the child, parental height.[1,2] In addition, a large number of studies have identified maternal age as an important determinant of birth outcomes and evidence shows that children born to younger and older mothers face higher risks of poorer birth outcomes, which are largely explained on adjustment for maternal characteristics.[3,4] In contrast, fewer studies have examined the association between paternal age and birth outcomes, possibly because of lack of data on fathers, and they suffer from various limitations.[5,6] The first limitation is that studies examining the association between paternal age and birth outcomes only present adjusted models that include possible confounding variables such as parental socio-demographic characteristics, maternal age and parity. Adjusted results may reveal whether paternal age is an independent risk factor for birth outcomes but not the actual prevalence of disadvantageous birth outcomes among children of younger/older fathers or how (observed and unobserved) parental characteristics might explain the association. Second, possibly because of the different sets of controls used by existing studies, the literature is inconclusive on the direction of this association: although most studies show that fathering a child at an advanced paternal age is not associated with adverse birth outcomes, [5,7-11] a subset of studies reveals a positive association. [12-15] The evidence on the association between a young paternal age at birth and birth outcomes is also inconclusive, with some studies showing a positive association and others a lack of one.[7,8,16] As a result of the fact that only a few studies have analysed the association between paternal age and birth outcomes and with some limitations, the

literature is inconclusive as to whether paternal age matters for birth outcomes. Since the current trend towards delayed childbearing means that both women and men have children at older ages,[17] it is important to revisit this question with high quality data and a careful research design.

Using data from the Finnish population registers, we aim to address the limitations of the existing literature as follows. First, we document the unadjusted association between paternal age and birth outcomes to reveal whether children born at younger and older paternal ages actually experience higher risks of poorer birth outcomes. Second, we investigate what potentially drives this association by showing whether the unadjusted association changes on adjustment for observed child and parental characteristics. Third, by comparing children born to the same father at different ages, a method so far never adopted in this literature, we additionally adjust for unobserved parental characteristics that are shared between siblings.

Materials and Methods

Study population

We use data from the Finnish population and other administrative registers. The base data are from a 20% random sample of households with at least one child aged 0-14 at the end of 2000 with individual-level information on all household members. Therefore, the data includes children who were born between 1987 and 2000. The individual-level linkages between different registers, maintained by Statistics Finland, Finland's National Institute for Health and Welfare and the Finnish Social Insurance Institution, were carried out by Statistics Finland using the unique personal identification numbers given to residents of Finland.

The data covered a sample of 170,621 children born between 1987 and 2000. Multiple births are excluded from the analyses (2.6%), since they are more likely to be low birth weight and/or preterm. Observations that had a missing value on any of the variables used in the analyses were also excluded (4.3% of the total sample). The prevalence of missing data was low and ranged from 0.02% (birth year) to 2.6% (smoking during pregnancy). Paternal age was missing for less than 1% of the sample. Formal ethical review was not required for this study because of it being an analysis of secondary data. Data were completely anonymised prior to use for research purposes.

Exposure and outcome

The key explanatory variable is paternal age at the birth of the child. We divide it into the following categories: <25, 25-29, 30-34, 35-39, 40 and above. We use ages 30-34 as the reference category because this is the most common age range for having children amongst fathers in our analytical sample and the one showing the lowest prevalence of LBW and preterm delivery.

We use two dependent variables: whether the child was born low birth weight (LBW, less than 2500 g at birth) and whether the child was delivered preterm (fewer than 37 weeks of gestation). Although the two indicators are strongly correlated as the majority of LBW (66%) children are born preterm, one third of the LBW children are not preterm. Of the preterm babies, 40% are LBW, and 60% are not, thus there is considerable variation in both dimensions and it is important to analyse both indicators.

Covariates

We adjust for the sex of the child in all models. We consider a range of child and parental characteristics that might be associated with both paternal age at birth and with the risk of LBW and/or preterm delivery. We adjust for the following child characteristics: birth order (1, 2, 3, 4, or more) and 5-year birth years (1987-1990; 1991-2004; 1995-2000). We also adjust for maternal age at birth (<20, 20-24, 25-29, 30-34, 35-39, 40+) and maternal smoking during pregnancy. To capture the socio-economic characteristics of the family, we adjust for deciles of family income (continuous) and the highest level of education in the household (basic, secondary, lower tertiary, higher tertiary). Income was measured for the year of each child's birth. Education was measured for the year of the first child's birth.

Statistical analyses

We estimate six linear probability models (LPM), such that the coefficients of the models are directly interpretable as marginal effects; i.e. percentage differences in LBW/preterm delivery with respect to the average level in the analytical sample, and to enable comparability across models.

We begin with analysing the association between paternal age and birth outcomes using a standard *between-family* regression approach, which consists of comparing children born to different fathers at different ages, and by only adjusting for the child's sex (hereafter referred to as 'unadjusted' model or association). We then reveal the role that each category of covariates has in attenuating paternal age gradients in LBW or preterm delivery by adjusting for each category separately. Model 2 introduces controls for the child's characteristics (birth order and birth year).

Model 3 introduces controls for maternal characteristics (maternal age and smoking during pregnancy). Model 4 adjusts for family socio-economic characteristics (household income and education). Then, to assess to what extent the paternal age gradient is attenuated when all the observed child and parental characteristics are considered, Model 5 adjusts for all covariates simultaneously. But since not all possible confounders are observed in the data, to further adjust for unobserved paternal characteristics that are shared by siblings in Model 6 we use a *within-family comparison*, also known as sibling fixed effects. This model includes an indicator for each sibling group and identifies the association between paternal age and the risk of LBW/preterm birth from variation between siblings.[18] Compared to a standard between-family model, the advantage of the within-family model is that it enables us to account for parental characteristics which are unobserved in the data and do not vary between siblings; these could include the social backgrounds of the parents, shared genetic factors and health characteristics (see the on line Appendix for a more detailed discussion). Model 6 includes all covariates included in Model 5, with the exception of household education since there was little variation between siblings.

Sample selection

Since the within-family results are estimated from groups of full siblings (i.e. who share the same mother and father), we excluded only children and children whose siblings were born before 1987 and could therefore not contribute to the estimation of the results (n=51,887). The resulting sample size for the analytical sample was 106,652 observations from 45,537 sibling groups. We used this analytical sample to conduct both the between and within-family models to retain comparability between the two estimation strategies.

Results

Table 1 shows the descriptive characteristics. The most common paternal age group is 30-34 (34.1%). LBW and preterm delivery show a U-shaped association with paternal age and fathers aged 40 and above show the highest prevalence of LBW and preterm delivery and fathers 30-34 the lowest. The results also show that, on one side, men who father children at ages 24 and below appear to be more disadvantaged, in terms of household income and highest level of education, than fathers who have a child at ages 25 and above; on the other side they show that the advantages level off from age 30 onwards. Maternal age at birth increases monotonically with paternal age.

Low birth weight

Table 2 shows the paternal age coefficients for LBW. Appendix Table 1 shows the rest of the model coefficients.

Results from Model 1, where we compare children born to different fathers at different ages, show that both younger and older paternal ages are associated with an increased risk of LBW. In Model 1 being born to a father aged 40 and above, as opposed to a father aged 30-34 whose probability to have a LBW child is 1.8%, is associated with an increased risk of 0.9% (95% CI: 0.5%-1.4%). When we adjust for birth order and birth year in Model 2, the increased risk of LBW associated with advanced paternal ages is marginally increased compared to Model 1. The excess risk at paternal ages 35-39 is substantially smaller when we adjust for maternal characteristics in Model 3, whilst it is attenuated by 30% for paternal ages 40 and above (β=0.6%; 95% CI: 0.1%-1.1%). When we adjust for parental socio-economic

characteristics in Model 4 the excess risk is attenuated, but less than in Model 3. The fully adjusted Model 5 shows coefficients at older paternal ages which are very similar to the ones we see in Model 3 and additional analyses (not shown) reveal that maternal age at birth plays the largest role in attenuating the unadjusted association. Being born to a young father (<25), relative to a father aged 30-34, is associated with a 1% (95% CI: 0.6%-1.3%) increased risk of LBW. The increased risk is almost halved in Model 2 and 3, where we include adjustment for the child and maternal characteristics respectively, and reduced by one third in Model 4 where we include parental socio-economic characteristics. Additional analyses (not shown) show that adjustment for the child's birth order attenuates the unadjusted association the most, followed by maternal age at birth. The increased risk at young paternal ages is fully attenuated in Model 5 which includes adjustment for all covariates (β=0.2%; 95% CI: -0.2%-0.7%). Results from the within-family Model 6 i.e. where we compare siblings born to the same father at different ages, are highly similar to those of Model 5. Although the coefficient for paternal ages 40 and above is smaller in the within family model (β=0.4%; 95% CI: -0.5%-1.2%) than the between-family one (β=0.6%; 95% CI: 0.1%-1.1%; Model 5) and not significantly associated with LBW, the confidence intervals of the former are wide and overlap with the confidence intervals of the latter. Figure 1 shows the results for Models 1, 5 and 6 for LBW.

Preterm

Table 3 shows the paternal age coefficients for preterm. Appendix Table 2 shows the rest of the model coefficients. Results from the between-family Model 1 show that younger and older paternal ages are associated with an increased risk of preterm delivery. Being born to a father aged 40 and above is associated with 1.2% (95% CI:

0.6%-1.7%) higher risk of preterm delivery compared to being born to a father aged 30-34 (whose baseline risk is 3.6%). The results are very similar to the results for LBW and in particular, maternal age at birth attenuates the unadjusted association between an older paternal age and the risk of preterm delivery the most. Being born to a father younger than 25 is associated with an increased risk of preterm delivery of 0.7% (95% CI: 0.3%-1.1%). The increased risk at younger paternal ages is fully attenuated when we adjust for birth order and year (Model 2) and for maternal characteristics (Model 3). In the fully adjusted Model 5, the increased risk at younger paternal ages is fully explained. In Model 6 i.e. where we compare siblings born to the same father at different ages, there is no statistically significant association between paternal age 40 and above and preterm delivery, but their magnitude is similar to that of Model 5. Figure 2 shows the results from Models 1, 5 and 6 for preterm delivery.

Sensitivity analyses

We replicated the between family Models 1 and 5 using the full sample, thus also including only/single children and children whose siblings could not be included in the analyses because they were born before 1987 (Appendix Table 3). Model 1 on the full sample shows that the age gradient at advanced paternal ages is steeper possibly because in the full sample there is a higher proportion of firstborns, who are more likely to be born LBW/preterm. However, Model 5, which is adjusted for all covariates, shows model results that are very similar to the Model 5 between-family estimates presented on the sibling-only sample in Table 2 and Table 3. Moreover, we replicated the analyses using paternal age at conception (rather than at birth) and the results were qualitatively similar.

Discussion

Using Finnish Population Registers, we analysed the association between paternal age and the risk of LBW and preterm delivery by adjusting for both observed and unobserved child and parental characteristics, the latter by adopting a within-family design which is a method never used in this particular literature. The unadjusted results reveal that both children of younger and older fathers face higher risks of LBW and preterm delivery. The increased risk at young paternal ages was largely explained by the child's birth order since first births are more common at young paternal ages and to be LBW and/or preterm.[19] The increased risk of LBW or preterm delivery at paternal ages 35-39 was largely attenuated and at paternal ages 40 and above it was partially attenuated on adjustment for maternal age at birth. In the within-family model – which enabled us to additionally adjust for unobserved parental characteristics shared by siblings – the 40 and above coefficient was still positively associated with LBW and preterm, although confidence intervals were wide and the parameter was not precisely estimated. [18] Therefore, the results fail to provide evidence consistent with the hypothesis that a young paternal age at birth is a risk factor for LBW and/or preterm; conversely, they don't exclude the possibility that an advanced paternal age at birth could be an independent risk factor for LBW and/or preterm delivery. The potentially independent effect associated with having a child at advanced paternal ages could be attributed to age-related sperm abnormalities or chromosomal mutations which may affect fetal growth as well as the timing of childbearing for men. [7, 12, 16]

This paper has three main contributions. First, the results reveal the importance of showing both unadjusted and adjusted results when examining the association

between paternal age and birth outcomes. On one hand, the unadjusted association shows that, on average, children of younger and older fathers face higher risks of poorer birth outcomes. On the other hand, the adjusted results show that this association is attenuated by at least 50% when birth order and maternal age at birth are included in the model. Second, this study helps to reconcile contradictory findings from the previous literature. The literature could be inconclusive about the association between a young paternal age at birth and birth outcomes because the studies which documented an association did not adjust for birth order, which our findings suggest is the covariate with the highest explanatory power. [7,8,16] The lack of consistency about the association between advanced paternal age and birth outcomes can be explained by the fact that existing studies used different sets of control variables and lack the comprehensiveness of this study since they did not adjust for parental characteristics which were not observed in the data.[5] Third, whilst the literature has so far predominantly studied the role of maternal age, the results underscore the importance of acknowledging paternal age as a potential risk factor for birth outcomes. In a separate study, using the same Finnish register data, we investigate the association between maternal age and the risk of poorer birth outcomes.[4] The unadjusted association between paternal age and birth outcomes is smaller compared to the one documented for maternal age, but it's still not trivial. Conversely, the size of the association in the within-family model for paternal age is larger than the one documented for the association between maternal age and birth outcomes, which further supports the need to conduct more research on paternal age and its effects on the risk of poorer birth outcomes.

This study has two main strengths. First, large register data enabled us to analyse the association between paternal age and birth outcomes across the entire age range, including the youngest and oldest fathers. Second, the data enabled us to consider a set of observed child and parental characteristics and to account, for the first time in this literature, for unobserved parental or childhood characteristics which do not vary between siblings. This study has also some limitations. First, the main analyses excluded only children or children whose siblings could not be included in the data. This was done since the within-family model was estimated using sibling groups and we wanted to maintain comparability across models. Second, the results apply to the Finnish context only and are not necessarily generalizable to other contexts where the provision of health care services differs and where the selection process into paternal ages is different. Third, we cannot conclusively say whether in the fully adjusted model for the paternal ages 40 and above coefficient is associated with the risk of poor birth outcomes due to the large confidence interval. This could be because unobserved parental characteristics which don't vary between siblings attenuate the association between advanced paternal age and birth outcomes or because the parameter is not precisely estimated and results in wide confidence intervals.

To conclude, this is the first study to analyse the association between paternal age and birth outcomes by carefully comparing unadjusted and adjusted results, including unobserved parental characteristics. The results underscore the importance of considering paternal age as a potential risk factor for adverse birth outcomes and the need of expanding research on its role and the mechanisms which link it to birth outcomes.

Authors' contribution section
AG and MM designed the study. AG conducted the analyses. All authors contributed to the interpretation of the findings. AG wrote the first draft. All authors revised the paper and contributed to the writing.

References

1.   Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. Bullettin of the World Health Organization. 1987;65(5):663-737.
2.   Cnattingius S, Haglund B, Kramer MS. Differences in late fetal death rates in association with determinants of small for gestational age fetuses: population based cohort study. Bmj. 1998;316(7143):1483.
3.   Nwandison M, Bewley S. What is the right time to reproduce? Fetal and Maternal Medicine Review. 2006;17(3):185-204.
4.   Goisis A, Remes H, Barclay K, Martikainen P, Myrskylä M. Advanced maternal age is not a risk factor for low birth weight and preterm delivery. American Journal of Epidemiology. 2017;Forthcoming.
5.   Shah PS, Births KSGoDoPLB. Paternal factors and low birthweight, preterm, and small for gestational age births: a systematic review. American journal of obstetrics and gynecology. 2010;202(2):103-23.
6.   Sartorius GA, Nieschlag E. Paternal age and reproduction. Human reproduction update. 2009:dmp027.
7.   Chen X-K, Wen SW, Krewski D, Fleming N, Yang Q, Walker MC. Paternal age and adverse birth outcomes: teenager or 40+, who is at risk? Human Reproduction. 2008;23(6):1290-6.
8.   Abel EL, Kruger M, Burd L. Effects of maternal and paternal age on Caucasian and Native American preterm births and birth weights. American journal of perinatology. 2002;19(01):049-54.
9.   Basso O, Wilcox AJ. Paternal age and delivery before 32 weeks. Epidemiology. 2006;17(4):475-8.
10.   Nahum GG, Stanislaw H. Relationship of paternal factors to birth weight. The Journal of reproductive medicine. 2003;48(12):963-8.
11.   Tough SC, Faber AJ, Svenson LW, Johnston DW. Is paternal age associated with an increased risk of low birthweight, preterm delivery, and multiple birth? Canadian Journal of Public Health/Revue Canadienne de Sante'e Publique. 2003:88-92.
12.   Reichman N, Teitler JO. Paternal Age as a Risk Factor for Low Birthweight. Am J Public Health. 2006;96(5):862-6.
13.   Zhu JL, Madsen KM, Vestergaard M, Basso O, Olsen J. Paternal age and preterm birth. Epidemiology. 2005;16(2):259-62.
14.   Astolfi P, De Pasquale A, Zonta LA. Paternal age and preterm birth in Italy, 1990 to 1998. Epidemiology. 2006;17(2):218-21.
15.   Selvin S, Garfinkel J. The relationship between parental age and birth order with the percentage of low birth-weight infants. Human biology. 1972:501-9.
16.   Olshan AF, Ananth CV, Savitz DA. Intrauterine growth retardation as an endpoint in mutation epidemiology: an evaluation based on paternal age. Mutation Research/Genetic Toxicology. 1995;344(1):89-94.
17.   Bray I, Gunnell D, Davey Smith G. Advanced paternal age: How old is too old? Journal of Epidemiology and Community Health. 2006;60(10):851-3.
18.   Wooldridge J. Introductory econometrics: A modern approach: Cengage Learning; 2012.
19.   Shah PS. Parity and low birth weight and preterm birth: a systematic review and meta-analyses. Acta Obstetricia et Gynecologica Scandinavica. 2010;89(7):862-75.

Table 1: Descriptive statistics by Paternal age, for siblings born between 1987-2000

Paternal age in years	LBW (%)	Preterm (%)	Birth year 1987-1990	Birth year 1991 - 1994	Birth Year 1995-2000	Birth order (mean)	Household income decile (mean)	Mother smoked during pregnancy (%)	Household high education (%)	Maternal age (mean)	N	%
≤24	2.9	4.0	35.6	33.8	30.6	1.4	3.7	23.6	17.8	22.2	10774	10.1
25-29	2.0	3.5	31.2	36.0	32.8	1.7	5.2	12.8	37.0	26.2	34526	32.4
30-34	1.9	3.3	21.9	35.1	43.1	2.0	6.0	10.3	40.5	29.6	36320	34.1
35-39	2.2	3.7	16.5	31.6	52.0	2.4	6.1	9.7	39.3	32.7	17571	16.5
≥40	2.8	4.5	15.5	30.0	54.5	2.5	6.0	11.2	35.7	35.1	7461	7.0
Average	2.2	3.6	24.9	34.3	40.8	1.9	5.5	12.4	36.6	28.6	106,652	

Table 2: Change in the probability of low birth weight, with 95% CI (Linear models)

	Model 1: unadjusted (child's sex)		Model 2: Model 1 + child characteristics		Model 3: Model 1 + maternal characteristics		Model 4: Model 1 + hh sociodemographic variables		Model 5: fully adjusted		Model 6: Model 5 + sibling fixed effects	
	β	95% CI	B	95% CI	β	95% CI	β	95% CI	β	95% CI	B	95% CI
Paternal age ≤24	0.96	(0.60 - 1.31)	0.51	(0.14 - 0.88)	0.57	(0.14 - 1.00)	0.74	(0.37 - 1.11)	0.23	(-0.20 - 0.67)	-0.11	(-0.83 - 0.61)
Paternal age 25-29	0.14	(-0.07 - 0.34)	-0.08	(-0.29 - 0.13)	0.10	(-0.13 - 0.33)	0.09	(-0.12 - 0.29)	-0.06	(-0.29 - 0.17)	-0.17	(-0.56 - 0.22)
Paternal age 30-34	0.000		0.000		0.000		0.000		0.000		0.000	
Paternal age 35-39	0.25	(-0.01 - 0.50)	0.34	(0.08 - 0.60)	0.11	(-0.17 - 0.39)	0.24	(-0.01 - 0.50)	0.15	(-0.13 - 0.42)	0.10	(-0.35 - 0.54)
Paternal age ≥40	0.91	(0.48 - 1.35)	1.02	(0.58 - 1.46)	0.62	(0.13 - 1.11)	0.88	(0.44 - 1.31)	0.61	(0.12 - 1.09)	0.36	(-0.51 - 1.23)
Constant	1.78	(1.61 - 1.95)	2.81	(2.54 - 3.08)	1.52	(1.31 - 1.72)	1.82	(1.41 - 2.23)	2.66	(2.20 - 3.12)	2.47	(1.84 - 3.09)
Number of observations	106,652		106,652		106,652		106,652		106,652		106,652	
Number of sibling groups											45,537	

Note: Model 1 adjusted for the child sex. Model 2 adjusted for the child birth order and birth year. Model 3 adjusted for maternal age and maternal smoking during pregnancy. Model 4 adjusted for household income deciles and household level of education. Model 5 is fully adjusted. Model 6 adjusted for all covariates in Model 5 except household level of education since there was little variation between siblings. Standard errors are clustered at the family level.

Table 3: Change in the probability of preterm delivery, with 95% CI (Linear models)

	Model 1: unadjusted (child's sex)		Model 2: Model 1 + child characteristics		Model 3: Model 1 + maternal characteristics		Model 4: Model 1 + hh sociodemographic variables		Model 5: fully adjusted		Model 6: Model 5+ sibling fixed effects	
	B	95% CI	β	95% CI	B	95% CI	β	95% CI	β	95% CI	β	95% CI
Paternal age ≤24	0.70	(0.26 - 1.13)	0.16	(-0.29 - 0.61)	0.21	(-0.33 - 0.74)	0.51	(0.06 - 0.96)	-0.19	(-0.74 - 0.35)	-0.45	(-1.37 - 0.46)

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
Paternal age 25-29	0.23	(-0.04 - 0.50)	0.04	(-0.32 - 0.23)	0.1	(-0.20 - 0.40)	0.17	(-0.10 - 0.45)	-0.1	(-0.40 - 0.21)	0.01	(-0.48 - 0.50)
Paternal age 30-34		0.000		0.000						0.000		0.000
Paternal age 35-39	0.37	(0.03 - 0.71)	0.5	(0.16 - 0.84)	0.25	(-0.11 - 0.62)	0.38	(0.04 - 0.71)	0.31	(-0.06 - 0.67)	0.41	(-0.16 - 0.97)
Paternal age ≥40	1.19	(0.64 - 1.74)	1.32	(0.77 - 1.88)	0.79	(0.18 - 1.40)	1.17	(0.63 - 1.72)	0.79	(0.18 - 1.40)	0.67	(-0.42 - 1.77)
Constant	3.62	(3.39 - 3.84)	4.8	(4.45 - 5.15)	3.5	(3.22 - 3.78)	3.59	(3.05 - 4.13)	4.67	(4.05 - 5.30)	3.83	(3.04 - 4.62)
Number of observations	106,652		106,652		106,652		106,652		106,652		106,652	
Number of siblings											45,537	

Note: Model 1 adjusted for the child sex. Model 2 adjusted for the child birth order and birth year. Model 3 adjusted for maternal age and maternal smoking during pregnancy. Model 4 adjusted for household income deciles and household level of education. Model 5 is fully adjusted. Model 6 adjusted for all covariates in Model 5 except household level of education since there was little variation between siblings. Standard errors are clustered at the family level.
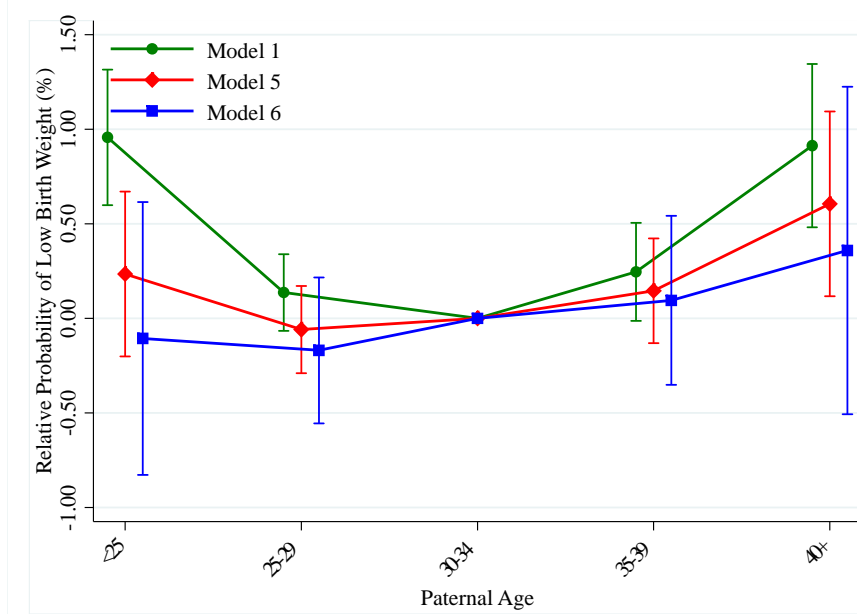
Figure 1 Change in the probability of LBW from between and within-family models (Table 2)

Figure 2 Change in the probability of preterm delivery from between and within-family models (Table 3)