

# Modelling the socio-economic determinants of fertility: a mediation analysis using the parametric $g$ -formula

Maarten J. Bijlsma

*Max Planck Institute for Demographic Research, Rostock, Germany*

and Ben Wilson

*Stockholm University, Sweden, and London School of Economics and Political Science, UK*

[Received December 2017. Final revision August 2019]

**Summary.** Theories predict that the timing of childbearing and number of children born are determined by multiple socio-economic factors. Despite this, many methods cannot investigate the interrelationships between these determinants, including the direct and indirect influence that they have on fertility over the life course. Here we use the parametric  $g$ -formula to examine the interdependent influences of time-varying socio-economic processes—education, employment status and partnership status—on fertility. To demonstrate this approach, we study a cohort of women who were born in the UK in 1970. Our results show that socio-economic processes play an important role in determining fertility, not only directly but also indirectly. We show that increasing attendance in higher education has a largely direct effect on early childbearing up to age 25 years, resulting in a substantial increase in childlessness. However, childbearing at later ages is dominated by an indirect effect of education on fertility, via partnership status and employment status, that is twice as large as the direct effect. We also use the  $g$ -formula to examine bias due to unobserved heterogeneity, and we demonstrate that our results appear to be robust. We conclude that the method provides a valuable tool for mediation analysis in studies of interdependent life course processes.

**Keywords:** Fertility;  $g$ -formula; Mediation; Socio-economic determinants

## 1. Introduction

The study of fertility can be seen as an on-going effort to understand the determinants of human reproduction (Balbo *et al.*, 2012; Becker *et al.*, 1960; Cleland and Wilson, 1987; Coale and Watkins, 1986; Davis and Blake, 1956; Hirschman, 1994). There are a variety of interrelated socio-economic determinants—including partnership status, education and labour market participation—that influence childbearing over the life course (Balbo *et al.*, 2012; Hirschman, 1994). To understand and predict fertility, it is important to examine the links between these processes, and the collective effect that they have on childbearing (Macura and Beets, 2002). Socio-economic determinants of childbearing are a crucial component of most theories that have been used to explain fertility trends over the last 50 years (Becker, 1981; Esping-Andersen and Billari, 2015; Goldscheider *et al.*, 2015; Hirschman, 1994; Lesthaeghe, 2010) and are

*Address for correspondence:* Maarten J. Bijlsma, Max Planck Institute for Demographic Research, Konrad-Zuse-Straße 1, Rostock 18057, Germany.  
E-mail: bijlsma@demogr.mpg.de

© 2019 The Authors Journal of the Royal Statistical Society: Series A (Statistics in Society) 0964–1998/20/183000  
Published by John Wiley & Sons Ltd on behalf of the Royal Statistical Society.  
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

often seen as essential mechanisms in the expression and mediation of theoretical concepts (e.g. Andersson (2004), Forste and Tienda (1996), Johnson-Hanks *et al.* (2011), Lesthaeghe (1983), Lorimer (1956), Milewski (2010), Neyer *et al.* (2013) and van de Kaa (1987)).

Empirical research shows that socio-economic factors are significantly associated with fertility trends in most countries of the world (e.g. Adserà (2005), Balbo *et al.* (2012), Bongaarts (2016), Bongaarts *et al.* (2017), Dribe *et al.* (2017), Heaton *et al.* (2002), Martin (1995) and Reher and Requena (2015)). However, the links between socio-economic factors and fertility vary across contexts and over time. For example, education is most often negatively associated with female-completed fertility (children ever born at the end of childbearing), but in some settings like the Nordic countries this association has weakened over time (Kravdal and Rindfuss, 2008) and appears to have disappeared (see Jalovaara *et al.* (2018), who also showed that the association between education and fertility is different for women and men). Similarly, marriage has long been associated with higher fertility, but at the population level fertility is now higher in northern Europe than in southern Europe, even though marriage rates are also lower (Hoem, 2008). Although these examples of heterogeneity are now well known, much less is understood about how they arise and how macrolevel trends are determined by microlevel processes, which is one reason why recent research has called for new applications of methods that can extend our existing understanding (Balbo *et al.*, 2012; Billari, 2015).

Empirical research also shows that the socio-economic determinants of fertility are path dependent, simultaneous and interrelated (e.g. Cohen *et al.* (2011), Gerster *et al.* (2014), Kulu and Steele (2013), Lillard and Waite (1993), Steele *et al.* (2005) and Upchurch *et al.* (2002)). For example, economic theories predict that individuals—in particular women—will face a series of interconnected socio-economic trade-offs over their life course, including with respect to the timing of education, employment, partnership and fertility (Billari and Kohler, 2004; Dribe *et al.*, 2014; Lee and Mason, 2010). Along similar lines, recent sociological research has also emphasized the links between socio-economic choices and the timing of childbearing, including with respect to later-life opportunities (Goisis and Sigle-Rushton, 2014; Mclanahan, 2004). Each socio-economic process is part of a complex web of time-varying relationships and may therefore have both direct and indirect effects on fertility, as well as mediating the effects of other socio-economic processes on fertility (Aassve *et al.*, 2006; Buhr and Huinink, 2014; Cohen *et al.*, 2011; Upchurch *et al.*, 2002). We therefore define socio-economic determinants as the socio-economic processes that influence child-bearing behaviour, either directly, indirectly or as mediators.

The complex interrelationships between socio-economic determinants raise some methodological issues for empirical research, in particular when studying childbearing over the life course (Balbo *et al.*, 2012). Socio-economic determinants are difficult to disentangle, and their influence on life course childbearing is difficult to study, even when drawing on the range of methods that are available to social scientists. At the heart of this complexity is the fact that fertility is a time-varying and path-dependent process (Buhr and Huinink, 2014; Huinink and Kohli, 2014). Researchers are often interested in both the timing of births (the ‘tempo’ of childbearing), as well as the number of children born (the ‘quantum’ of childbearing) (Balbo *et al.*, 2012). Moreover, not only is early life course childbearing (or a lack of early life course childbearing) likely to have a direct effect on future fertility, but it is also likely to have indirect effects on future fertility via various (mediating) socio-economic processes. For example, becoming a parent may disrupt education or employment (Buhr and Huinink, 2014; Upchurch *et al.*, 2002). Previous childbearing is therefore likely to have an effect on the same socio-economic processes that determine future childbearing. Sometimes referred to as reverse causality, this highlights the fact that time-varying socio-economic processes can act simultaneously as causes, effects

and mediators of fertility behaviour. For example, marriage is a determinant of fertility, but it may also be determined by fertility, or marriage may act as a mediator for the relationship between another determinant—such as education—and fertility later in life.

As a result of this complexity, previous empirical research has often focused on single determinants of fertility, rather than trying to estimate their interrelationships (Balbo *et al.*, 2012). A common strategy has been to design research that tries to estimate the influence of one determining process, such as partnership status, on fertility (e.g. Brien *et al.* (1999) and Steele *et al.* (2005)). However, when focusing on a single determinant (and treating other determinants merely as confounders), it is difficult to examine the interrelationships between different determinants. This is equally true of econometric approaches that seek to minimize bias due to endogeneity (selection and reverse causality) (Engelhardt *et al.*, 2009; Moffitt, 2003, 2005). In addition, many of these approaches sacrifice the capacity to generalize (external validity) for a reduction in the likelihood of bias due to confounding (Cartwright, 2011). Such approaches are well suited to answering research questions about single determinants, especially with respect to narrowly defined subsections of the population (such as compliers with an instrumental variable) (Moffitt, 2005). However, they are far less useful when attempting to study the interactions between multiple socio-economic determinants of fertility over the life course, or when researchers are keen to make generalizations about the fertility behaviour of broader populations.

In this study, we therefore propose a different approach for the study of multiple fertility determinants. Our aim is not to criticize existing approaches, but rather to expand the methodological toolbox of researchers. To do this, we make use of the *g*-formula, which is a method that has been developed by epidemiologists and biostatisticians over the last few decades (Robins, 1986; Robins and Hernán, 2009). Similarly to other causal inference approaches, the *g*-formula uses counterfactual theories of causal inference to enable researchers to estimate the effects of a cause (Greenland *et al.*, 1999), such as the hypothetical effect of treatment on the chance of experiencing a cardiac event in the presence of time-varying confounding (Daniel *et al.*, 2013). Another advantage of the *g*-formula is that it was developed to estimate time-varying effects in the presence of time-varying confounding and mediation (Robins, 1986) and is therefore capable of testing hypothetical changes in life course processes while accounting for their interdependence (Robins and Hernán, 2009; Daniel *et al.*, 2013; Bijlsma *et al.*, 2017).

This paper contributes to both the applied and the methodological literature. From an applied perspective, it is the first time that the *g*-formula has been used to study fertility. This method enables us to model the time-varying determinants of fertility, allowing for selection and reverse causality, while simultaneously investigating the effect of changes to these determinants—in the form of plausible counterfactual scenarios. Unlike previous studies, we estimate causal effects at the population level, rather than for a subset (e.g. compliers with a quasi-experiment), thereby providing findings that are more readily generalizable. Although this paper focuses on fertility, we believe that more generally it represents the first use of the *g*-formula to study the mediating interrelationships between time-varying life course processes (rather than merely using the method to control for time-varying confounding). One of our methodological contributions is therefore to show how the *g*-formula can be used to carry out dynamic mediation analysis (i.e. to examine how mediation changes over time), in particular for an outcome that is both highly path dependent and monotonic (because an individual's number of children ever born cannot decrease). In addition, we also make a further contribution to the methodological literature by demonstrating, apparently for the first time, how to simulate unobserved confounding to carry out sensitivity analysis when using the *g*-formula. To make these contributions, we carry out a case-study of fertility in the UK, and our guiding research question is 'What are the direct

and indirect (mediated) effects of employment status, education and partnership status on birth risks over the life course for a nationally representative birth cohort?'.

## 2. Data

### 2.1. Setting

In common with most high income countries, the UK has experienced a considerable shift in demographic behaviours over the last 50 years. These include a rise in cohabitation, divorce and childbearing outside marriage (Beaujouan and Ní Bhrolcháin, 2011; Sigle-Rushton, 2008). Relative to the rest of Europe, fertility in the UK has remained high, although there are considerable socio-economic differences in both the quantum and the tempo of childbearing (Champion and Falkingham, 2016; Perelli-Harris *et al.*, 2010). This includes a strong association between birth timing and socio-economic status (postponement of birth for high socio-economic status groups), as well as higher completed fertility (children ever born at the end of childbearing) among women who are in the lowest and highest socio-economic groups (compared with those in between) (Sigle-Rushton, 2008). As noted in Section 1, these associations are not observed in every European country outside the UK.

Here we study women in the UK who were born in 1970. Since this birth cohort began their childbearing, there have been changes in the observed associations between childbearing and its most commonly studied socio-economic determinants (Berrington *et al.*, 2015; Berrington and Pattaro, 2014; Hobcraft, 1996; Kneale and Joshi, 2008; O'Leary *et al.*, 2010; Rendall and Smallwood, 2003; Steele *et al.*, 2005, 2006). The most prominent determinants of UK fertility trends over the last 30 years include an increased participation in higher education among women, an increase in female employment and a reduction in the prevalence of marriage (Aassve *et al.*, 2006; Berrington and Pattaro, 2014; Chamberlain and Gill, 2005; Neels *et al.*, 2017; Ní Bhrolcháin and Beaujouan, 2012; Office for National Statistics, 2016; Sigle-Rushton, 2008; Steele *et al.*, 2005, 2006).

### 2.2. Source of data

Our data come from the 1970 British Cohort Study, which follows the lives of around 17000 people who were born in a single week of 1970 (UK Data Archive, 2016). The sample covers Great Britain, which includes all of the constituent countries of the UK (England, Scotland and Wales), except for Northern Ireland. After dropping all women who have missing information on all variables at the start of our observation period (age 16 years), we are left with approximately 4100 women. After this, we drop all cases who have missing data on any one of our baseline covariates, which reduces the sample by around 21% to just over 3200 women (on-line appendix Fig. A1). There are no further changes to the analytical sample, such that we carry out complete-case analysis, with the addition that women are censored from the year in which they cease to have complete information on all time-varying variables in the analysis. We follow the childbearing life course of these women from age 16 up to age 38 years (because the next wave data at age 42 years were not yet readily available when we carried out our analysis). According to recently released statistics based on all registered births in England and Wales, the cohort of women who were born in 1970 had an average completed fertility of 1.9 children per woman at age 45 years, compared with 1.8 children ever born at age 38 years (Office for National Statistics, 2016).

### 2.3. Fertility

Fertility is the primary outcome variable of this study. It is modelled as a time-varying binary

variable, with a value of 1 in any year when a live birth occurs, and 0 otherwise. Birth events can result in multiple live-born children (i.e. twins or triplets). Although such events are rare for this sample, they are taken into account in the modelling procedure (see below). In addition to births, we include a variable that counts the number of live-born children up to time  $t$  (where  $t$  is measured in years, which is the same as age because we are studying a single birth cohort).

#### 2.4. Time-varying (confounding) mediators and determinants

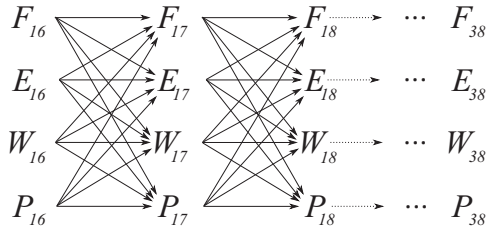
The main processes that we study are education, partnership status, employment status and socio-economic status (with socio-economic status measured only for those who are employed). Each process is measured by using annual time-varying indicators derived from the British Cohort Study data. We make use of two education variables: one measuring whether a woman is in full-time education or not, and the other counts the number of years spent in full-time education. Partnership status is a multinomial variable with three categories, indicating whether a woman is cohabiting, married or single (not in a partnership, regardless of marital history). Where we refer to 'partnership status' in the rest of this paper, we refer to this variable. Where we refer to being in a partnership, we mean that women are either cohabiting or married. Employment status is a multinomial variable indicating whether a woman is part-time employed, full-time employed, disabled or unemployed (i.e. not disabled or in any form of employment). The socio-economic status variable indicates whether a woman's employment (full time or part time) is of high or low socio-economic status, based on the standard classification of occupations in the UK (UK Data Archive, 2016).

#### 2.5. Time constant baseline characteristics

We included some (baseline) time constant variables as control variables, in addition to a time-dependent measure of age. To account for the intergenerational transmission of fertility we include mother's age at birth (when she gave birth to the woman who is being studied) and number of siblings (calculated by using information recorded at age 16 years). We note that the number of siblings may be slightly underestimated, because of unrecorded or unacknowledged siblings, or it may be overestimated because of the inclusion of step-siblings (although for 90% of respondents this variable is based on questions that ask about 'blood brothers and sisters'). As a measure of parental resources, we use household income at age 10 years in three categories (low, medium and high), and as a measure of spatial selection we use region of residence at age 16 years in six categories (Scotland, Wales, and England North, Midlands, South and London). Age was measured in three categories, 15–22, 23–29 and 30–38 years. In the modelling process, the use of a finer age categorization did not strongly impact our population-averaged results.

### 3. Method

Our method employs the  $g$ -formula, which was developed by Robins to estimate time-varying effects in the presence of time-varying confounding and mediation (Daniel *et al.*, 2013; Bijlsma *et al.*, 2017; Robins, 1986; Robins and Hernán, 2009). The  $g$ -formula is implemented in a series of steps. First, we formulate the directed acyclic graph (DAG) in Fig. 1, which shows the interrelationships that we investigate. Second, on the basis of the interrelationships assumed that we make explicit in Fig. 1, we estimate a series of multivariable generalized linear models for each outcome that is defined by our DAG, and by the variable categories in the 1970 British Cohort Study. Third, we define the counterfactual scenarios that are of interest, given our research question. Then, using estimates from these multivariable generalized linear models,



**Fig. 1.** DAG of fertility and time-varying mediating confounders (each letter represents a time-varying variable as follows: F, fertility; E, education; W, work (employment status); P, partnership status; each arrow represents a causal relationship in the direction indicated by the arrow; only the relationships between age 16 and 18 years are shown since the relationships are the same from each year to the next; time constant baseline characteristics—parental income, mother’s age at birth, number of siblings and region—are not shown because these are modelled to have a direct effect on every time-varying process in every year; the same is true of age; all models include interactions between age and partnership, fertility and education, fertility and partnership, and fertility and employment; fertility is modelled using ‘birth in a given year’, and the lagged equivalent as shown; this outcome is then adjusted to account for multiple births, which enables a cumulative count of ‘children ever born’ to be estimated and used as a control variable in each of our simultaneous equations)

we perform a series of microsimulations to produce each counterfactual scenario, including to estimate total, direct and indirect effects. In addition to these  $g$ -formula steps, we perform a sensitivity analysis to examine the influence of unobserved heterogeneity. The code for our analysis is available from

<https://rss.onlinelibrary.wiley.com/hub/journal/1467985x/series-a-datasets>,

and the data are available from UK Data Archive (2016).

### 3.1. Directed acyclic graph

The presence of time-varying variables, confounders, mediators and outcomes is illustrated by the DAG in Fig. 1. Control variables are allowed to affect all time-varying determinants but are not shown. The existence of time-varying determinants is evident from the arrows between each of the other processes and fertility. Reverse causality is evident from the paths that lead away from fertility in one year, and return to fertility in a future year. The relationship between fertility and itself is therefore mediated by each of the other time-varying processes.

### 3.2. Multivariable modelling

We model the interrelationships between four social processes, each measured by using a series of binary variables as follows: fertility (one binary variable), education (one binary variable), partnership status (three binary variables) and employment status (four binary variables). Let  $Y_{gt+1}$  denote a variable measured at time  $t + 1$ , where  $g$  is an index for members of the set  $G$ , which contains the following binary outcomes: birth in the last year; in full-time education; single; cohabiting; married; out of work; working part time; working full time; disabled; low socio-economic status; high socio-economic status.  $M_{ht}$  denotes the values of a mediating covariate measured at time  $t$ , where  $h$  is an index for members of set  $H$ , which contains the same variables as  $G$ , as well as total education (a cumulative count of years of education) and total births (a cumulative count of the number of live-born children). To define the other explanatory variables, let  $B_c$  be a time constant variable measured before age 16 years, where  $c$  is an index for members of set  $C$ , which contains parental income at age 10 years, age of mother at respondent’s birth,

number of siblings and region. Let  $A_{kt}$  be a categorical indicator variable for age (where  $k$  is the range 15–22, 23–29 or 30–38 years) measured at time  $t$ . Finally, let  $L_{jt}$  be a series of interaction terms, where  $j$  is an index for members of set  $J$ , containing employment status interacted with socio-economic status, age interacted with partnership status, and childlessness (total births = 0) separately interacted with education, all three partnership status variables and all four employment status variables. Conditionally on the explanatory variables,  $Y_{gt+1}$  is taken to follow a logistic distribution and is modelled as

$$\text{logit}(E[Y_{gt+1}]) = \eta + \alpha_k A_{kt} + \sum_{c \in C} \beta_c B_c + \sum_{h \in H} \mu_h M_{ht} + \sum_{j \in J} \lambda_j L_{jt} \quad (1)$$

where the coefficients  $\alpha$ ,  $\beta$ ,  $\mu$  and  $\lambda$  are estimated for each of their subscripts ( $k$ ,  $c$ ,  $h$  and  $j$  respectively), and then this estimation procedure is repeated for each of the outcomes  $Y$  (as indicated by the subscript  $g$ ). We also note that the distinct categories of each multinomial variable (e.g. the three partnership status categories) were modelled by using logistic regression in the simulation step of the  $g$ -formula. Compared with using multinomial regression, this made no difference to our results, but it did improve the efficiency of our code considerably.

### 3.3. Counterfactual scenarios

As summarized in Table 1, we examine the effect of three scenarios on fertility. In the first two, we set out to examine the effect of large-scale demographic changes on fertility. These scenarios are based on observed trends in the UK from 1970 until the present day: increases in higher education attendance among women (scenario 1) and a reduction in the prevalence of marriage (scenario 2). As well as these trend-based scenarios, we also examine the effect of a hypothetical policy intervention to increase the post-birth employment of women (scenario 3).

### 3.4. Total effect estimation

To determine the effect of each of the counterfactual scenarios, we first produce a scenario without a hypothetical intervention, which is commonly referred to as the ‘natural course’

**Table 1.** Estimated scenarios

Scenario	Details
1 Increase in higher education attendance	We randomly select a group of women from those who are still in education at age 16 years, and then keep them in education until age 22 years, so that at least 50% of all women remain in (higher) education until age 22 years
2 Reduction in the prevalence of marriage	We simulate a decrease in the proportion of women who are married, and a corresponding increase in the proportions single and cohabiting, equivalent to a 50% reduction in the stock of married women over the study period
3 Increase in post-birth employment	We ensure that all women, except those who are disabled, are employed full time in the year after their birth (after which time they are free to follow any employment trajectory)

scenario in the literature (Keil *et al.*, 2014). Each counterfactual scenario is then compared with the natural course to determine its total effect. The outcome of a scenario (natural course or intervention) is estimated by using the parametric  $g$ -formula:

$$E[Y_{\bar{x}}] = \sum_{(m_1, \dots, m_T)} \left\{ E[Y | \bar{X} = \bar{x}, M = m] \prod_{t=1}^T P(M_t = m_t | \bar{X}_{t-1} = \bar{x}_{t-1}, M_{t-1} = m_{t-1}) \right\} \quad (2)$$

where  $Y_{\bar{x}}$  is the outcome of interest (e.g. total births at the end of follow-up) when some intervention on  $x$  has taken place, and  $M$  represents the time-varying mediators that we do not actively change with the intervention. In the  $g$ -formula, we also condition on age  $A_{kt}$  and the baseline covariates  $B_c$  from equation (1), as well as including the interaction terms  $L_{jt}$  from the same equation. However, we have omitted these from equation (2) for brevity, and because they are either not causally affected by the intervention (in the case of  $A_{kt}$  and  $B_c$ ) or are deterministic functions of the mediators (in the case of  $L_{jt}$ ).

The intervention is here represented by  $\bar{x}$ , with the variable that takes the place of  $x$  dependent on the specific scenario. Since in complex settings this summation may be analytically intractable, a simulation step, i.e. Monte Carlo integration, is used instead (within a bootstrap to determine standard errors). This simulation is composed of the following substeps.

- Step 1:* draw one random sample of individuals with replacement from the data.
- Step 2:* (re-)estimate earlier specified multivariable models.
- Step 3:* draw a data set for starting values at baseline ('first' year of follow-up).
- Step 4:* use multivariable models to predict outcomes for the 'second' year of follow-up.
- Step 5:* repeat step 4 for all years of follow-up, given (predicted) values at the previous year.
- Step 6:* save the results and go back to step 1.
- Step 7:* repeat steps 1–6 for each iteration (999 times).
- Step 8:* report the estimates and corresponding uncertainty.

These eight substeps produce the natural course  $E[Y_x]$ . The natural course is not exactly the same as the empirical data, but should approximate it, and is essentially a simulation of the empirical data with uncertainty (substep 1 generates the uncertainty). In the simulation, whenever a birth is predicted, it has a chance of being a multiple birth (e.g. twins) based on empirical probabilities of a multiple birth (by the number of live-born children).

To produce the intervention scenarios, we perform these same eight substeps but between substeps 4 and 5 we perform the 'intervention' as described in Table 1. For example, in scenario 3—an increase in employment after birth—a woman who had a birth in the last year and is predicted to be unemployed (by the multivariable models) is mechanically set to be employed instead. Performing the substeps this way produces  $E[Y_{\bar{x}}]$ . To account for covariance, we perform both the substeps for the natural course scenario, and for an intervention scenario, and compare their outcomes within each bootstrap iteration. This produces the total effect TE:  $E[Y_x - Y_{\bar{x}}]$ . Note that  $E[Y_x - Y_{\bar{x}}]$  can also be written as  $E[Y_{xM_x} - Y_{\bar{x}M_{\bar{x}}}]$  to denote that the values of the (other) mediators also change as a function of intervening on  $x$ .

### 3.5. Mediation: direct and indirect effects

We estimate the direct and indirect effect of each intervention on fertility for both the education and the marriage scenarios (we omit the post-birth employment scenario from this part of the analysis because of its small effect size). Given our DAG, the indirect effect of marriage can be via education or employment status, whereas the indirect effect of education can be via partnership status or employment status. To estimate these indirect effects, we use effect decomposition



(Pearl, 2001; Robins and Greenland, 1992; VanderWeele, 2014; VanderWeele and Vansteelandt, 2009). More specifically, we decompose the total effect TE into the total direct effect TDE and the pure indirect effect PIE (Wang and Arah, 2015). TDE is determined like TE, except that the covariate values of mediating variables are drawn from the natural course, so that the effect of the scenario of interest on the outcome can no longer act via these mediators and is thus only a direct effect, i.e.  $E[Y_{xM_x} - Y_{\bar{x}M_x}]$ . For example, changes in the prevalence of marriage can now only affect fertility directly—as shown in our DAG by arrows from P to F—and not via employment status or education because values for these mediators are drawn from the natural course. Furthermore,  $PIE = TE - TDE = E[Y_{xM_x} - Y_{\bar{x}M_{\bar{x}}}] - E[Y_{xM_x} - Y_{\bar{x}M_x}] = E[Y_{\bar{x}M_x} - Y_{\bar{x}M_{\bar{x}}}]$ . We note that other mediation effect decompositions are also possible (Wang and Arah, 2015).

### 3.6. Sensitivity to unobserved confounding

The  $g$ -formula relies on the assumption that the most important sources of confounding are observed. As such, sensitivity analysis is particularly useful to test the susceptibility of results to unobserved confounding (unobserved heterogeneity). We focus here on the same scenarios as in the mediation analysis.

We draw on recently proposed approaches to simulate the influence that a hypothetical unobserved confounder would have on our results (Carnegie *et al.*, 2016; Lin *et al.*, 2017; VanderWeele, 2015) and extend them to the  $g$ -formula. This is done by constructing an artificial confounder  $C$ , as a function of the variables that it confounds ( $X$  and  $Y$ ), and then determining how much its inclusion changes the outcome of interest, as follows.

- (a) Assume an equation for confounder  $C$  as a function of the variables that it confounds, e.g.

$$C_i = \beta_0 + \beta_x X_i + \beta_y Y_i + \varepsilon_i. \quad (3)$$

- (b) Set values for  $\beta_x$ ,  $\beta_y$  and  $\varepsilon_i$  that seem reasonable.  
(c) Predict  $C$  on the basis of equation (3), i.e. including random  $\varepsilon_i$  for each prediction.  
(d) Fit the regression  $E[X|C] = \theta_0 + \theta_1 C$ , and fit the regression  $E[Y|X, C] = \gamma_0 + \gamma_1 C + \gamma_2 X$ .  
(e) Check  $\theta_1$  and  $\gamma_1$ ; these are the effects of  $C$  on  $X$  and  $Y$ . When the model for  $Y$  is collapsible, check  $\gamma_2$  to see immediately how the effect of the determinant of interest ( $X$ ) on  $Y$  changes because of control for the hypothetical confounder. If the model for  $Y$  is non-collapsible, such as in our case, a population averaging step (such as the scenario comparison step of the  $g$ -formula) is needed.  
(f) Depending on the findings of step (e), return to step (a) and adjust  $\beta_x$ ,  $\beta_y$  and/or  $\varepsilon_i$  to adjust  $\theta_1$  and  $\gamma_1$ , where  $\varepsilon \sim N(0, \sigma^2)$ .

Including this artificial  $C$  in a model for the effect of  $X$  on  $Y$  will change the estimated effect of  $X$ , and this change can be assumed to represent the effect of an (hypothetical) unobserved confounder (if it were instead observed and included in the analysis). Strictly speaking we have created a collider but, from the perspective of statistical associations, confounders and colliders are two sides of the same coin. Adjusting  $\beta_x$ ,  $\beta_y$  and  $\varepsilon_i$  will change the values of  $\theta_1$  and  $\gamma_1$ . Once  $\beta_x$  and  $\beta_y$  have been set, an easy way to adjust the overall strength of the confounder is then by adjusting  $\varepsilon$ . Increasing random variation weakens the relationship of the confounder with  $X$  and  $Y$ , and vice versa.

We note that  $C$  does not need to represent one confounding mechanism: it can also represent the joint effect of a number of confounders. We use an iterative procedure here because in non-linear settings it is difficult to determine the values of  $\beta_x$ ,  $\beta_y$  and  $\varepsilon_i$  that are required to produce statistically or theoretically plausible  $\theta_1$ - and  $\gamma_1$ -values. For various analytical solutions, see VanderWeele (2015).

This sensitivity analysis is applied to the education scenario and the marriage scenario, both of which may be affected by unobserved confounding. In the marriage scenario, we motivate our analysis by assuming—on the basis of the literature—that certain individuals have a more ‘traditional’ value system, obtained during childhood, which makes them more likely to marry and to have children (e.g. Guetto *et al.* (2015), Hirschman (1994), Inglehart and Baker (2000) and Lesthaeghe (1983)). Our analysis is motivated by the idea that traditional values of this kind most certainly existed for the 1970 birth cohort, at least for a subset of the population. If such a variable could be measured, it would function as a baseline confounder (and in some research contexts would be known as a selection effect). Indeed, it is also reasonable to ignore what this variable represents and instead to view it as a more general test of sensitivity to unmeasured baseline confounding.

We generate this confounder on the basis of each individual’s observed number of children and total years spent in marriage at age 38 years (i.e. both the outcome and exposure), setting the coefficient for the former to 3 and the latter to 1 (in step (a) above). Initially,  $\varepsilon_i$  was set to be normally distributed with a mean of 0 and standard deviation of 7. The confounder score  $C$  of an individual was then assigned to all observations of that individual (i.e. all ages before age 38 years, because we define the confounder to be a baseline or time constant confounder). This produces what we refer to as the ‘strong’ confounder. Then, to determine the strength of unobserved confounding that is needed to make our intervention effect completely non-significant, we generated even stronger confounders by changing  $\varepsilon_i$  to have a standard deviation of 3.5 (a ‘stronger’ confounder) and finally of 1 (the ‘strongest’ confounder).

For the education scenario, we assumed that traditional values may make a woman more likely to have children and less likely to be in education (e.g. Guetto *et al.* (2015), Hirschman (1994), Inglehart and Baker (2000) and Lesthaeghe (1983)). The confounder was constructed in a similar way to that in the marriage scenario, with a coefficient of 3 for total number of children,  $-1.95$  for total person-years in education,  $-1.15$  for total number of years in full-time employment and  $-0.65$  for women in part-time employment. Standard deviations of  $\varepsilon_i$  were 10, 5 and 1 for the strong, stronger and strongest confounding variants respectively.

## 4. Results

### 4.1. Multivariable models

Before presenting results that indicate the effects of each scenario, we briefly discuss some model diagnostics. Although the results from the multivariable models in equation (1) cannot be interpreted in full, because they represent only one step of the  $g$ -formula procedure, they are similar to those that would be obtained by using a traditional panel model. For this reason, some analysts use these results to check that estimated coefficients are coherently aligned with expectations. Unexpected values of parameters are not necessarily incorrect but may warrant further investigation. On-line appendix Table A1 shows that our models align with the results of previous research on fertility. For example, a decreased risk of birth is associated with being single, employed or in full-time education. Appendix Table A2 shows another form of model diagnostic: a comparison between the natural course and the empirical data. In our analysis, it appears that the  $g$ -formula does a good job of predicting cumulative fertility at age 38 years, although with some small differences from the empirical data. There is a slight underestimation of childlessness, although this is likely to be due, at least in part, to the censoring in our empirical subsample.

#### 4.2. Results from the *g*-formula

Fig. 2 shows the results of the three scenarios (described in Table 1), compared with the natural course. The first of these examines the effect of an increase in higher education attendance among women. Not only is there a significant negative effect on the quantum of fertility across most of the reproductive life course, but there also appears to be a strong effect on the tempo of fertility. Educational attendance has a suppressive effect on fertility, but this effect is most dominant while women are attending higher education. In our scenario, cumulative fertility is lowest (relative to the natural course) when women leave higher education—at around 75% of the level that it would be otherwise. From this point onwards, the women in this scenario catch up with women in the natural course—exhibiting an increase in fertility tempo. By age 38 years, their number of children ever born is around 90% of the level that it would be in the absence of the scenario.

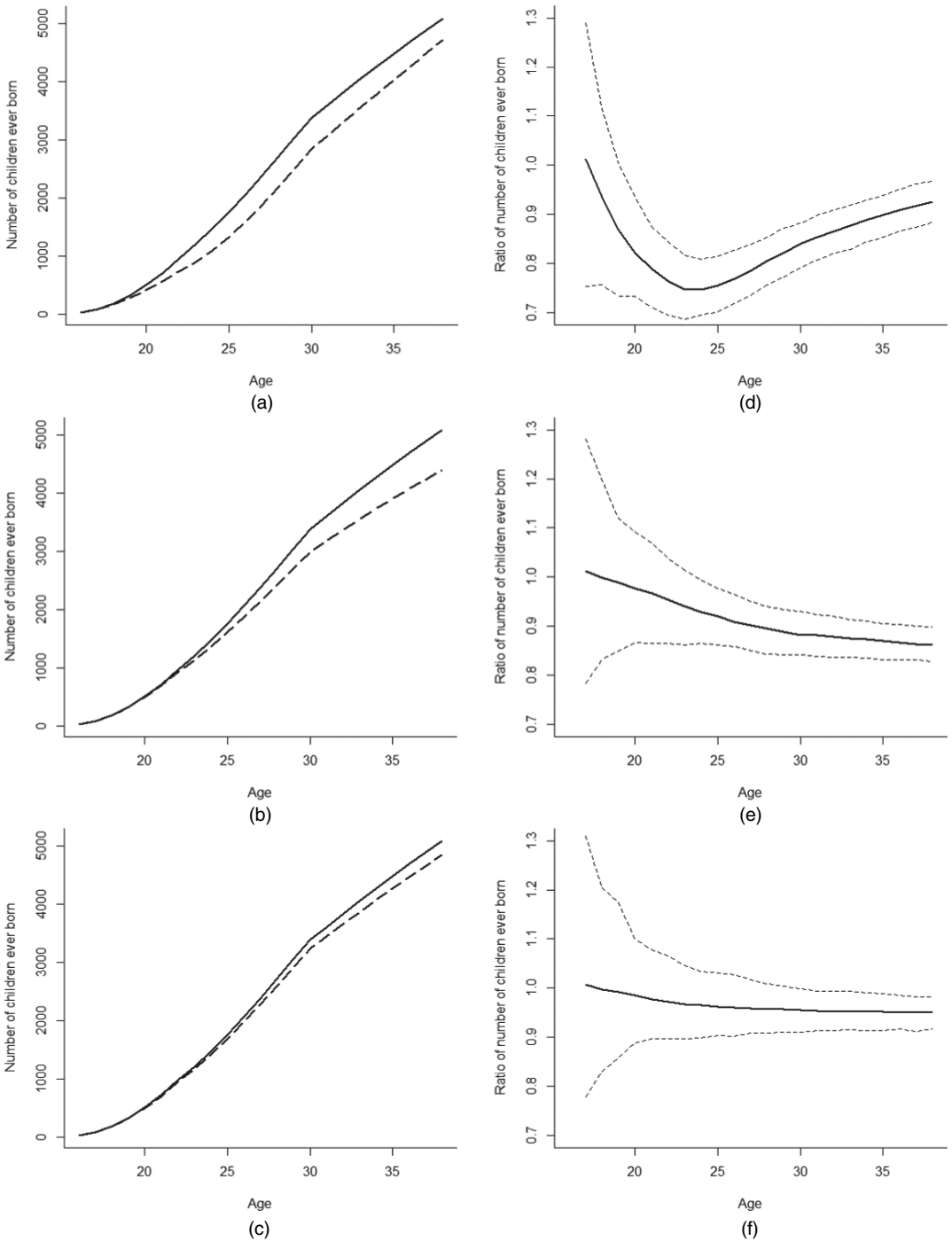
The second scenario examines the effect of a reduced prevalence of marriage. As with the education scenario, this is based on observed demographic trends, which in this scenario refer to a reduction in marriage that is accompanied by an increase in the prevalence of unmarried cohabitation or not living in a residential partnership (i.e. being single). The effect of this scenario is also significant, but with a steadily increasing effect over time. By age 38 years, this scenario leads to a reduction in children ever born of around 13%.

The post-birth employment scenario shows much weaker effects (although  $p < 0.05$  at age 38 years). In this scenario we examine the effect of a hypothetical policy intervention that increases post-birth employment. The results of this scenario are similar to that of the marriage scenario, but less strongly significant. This might be expected because we adjust employment rates in the year immediately after birth only, and we allow them to vary after that. However, it is also due to the relative strength of the (conditional) relationship between partnership status and fertility compared with employment status and fertility. Based on our assumptions, a scenario in which all women took on full-time employment the first year after birth would result in around 5% fewer births.

In addition to the effect on average fertility, the *g*-formula can be used to show the effect of each scenario on the distribution of births by age and parity (Table 2). Here we show only results for the middle and end of the child-bearing life course (age 28 and 38 years). The results concur broadly with those in Fig. 2. For example, they indicate the postponement of childbearing in the education scenario at age 28 years (a 9-percentage-point difference compared with the natural course). In addition, these results show that the relationship between parity and effect size is not necessarily straightforward. For example, the employment scenario makes no discernible difference to the level of childlessness at age 38 years, but instead leads to a reduction in fertility levels due to fewer births at higher parities (three or more children). This can be contrasted with the marriage scenario, which appears to have a strong effect on both childlessness and higher parity births.

#### 4.3. Mediation analysis

We perform a mediation analysis for the two scenarios with the most significant results: the education scenario (scenario 1) and marriage scenario (scenario 2). Given that we have annual data, and the ability of the *g*-formula to perform a mediation analysis at each time point, we can not only estimate mediation at the end of follow-up, but also at every year  $t$ . We decompose the total effect TE into the total direct effect TDE and pure indirect effect PIE, where PIE shows the profile of mediation at each age (Fig. 3). As a share of the total effect, the direct effect is initially large in both scenarios. However, the life course profiles of mediation are very different.



**Fig. 2.** Scenario effect on fertility over the life course: (a)–(c) ratio of the number of children ever born in our sample (scenario results (---)) compared with the natural course (—), by age; (d)–(f) scenario results (—) compared with the natural course, calculated as a risk ratio, and accompanied by 95% confidence intervals (----); (a), (d) education scenario; (b), (e) marriage scenario; (c), (f) post-birth employment scenario

**Table 2.** Parity distributions at ages 28 and 38 years†

Scenario	Parity distribution (%) among women age 28 years				Parity distribution (%) among women age 38 years			
	0	1	2	≥3	0	1	2	≥3
Natural course	47	26	22	5	22	20	39	19
1, Education scenario	56	22	17	5	28	19	35	18
2, Marriage scenario	51	26	20	3	29	21	35	15
3, Employment scenario	47	28	21	4	22	22	41	15

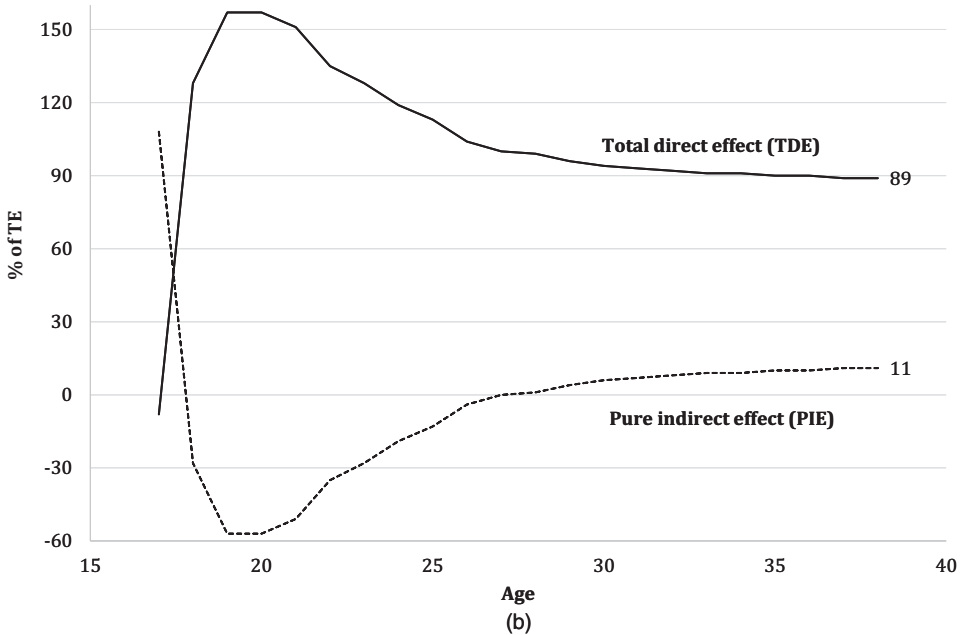
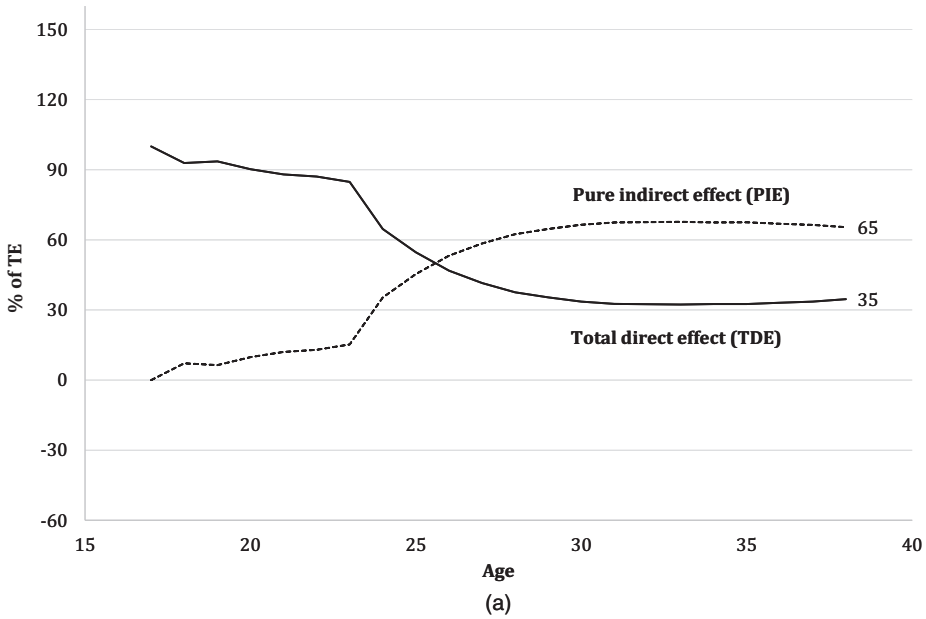
†These two ages are shown to illustrate the results, but the same comparison could be made at any age.

In the education scenario, the indirect effect is fairly small (5–15% of TE) during the years in which the scenario increases educational attendance (up to age 23 years), but it declines rapidly from ages 24 to 26 years, after which mediation becomes more dominant than the direct effect of education on fertility. This analysis shows that education has a strong indirect effect on fertility, mediated by partnership status and employment status (e.g. educational attendance reduces fertility indirectly by increasing employment and reducing marriage or cohabitation). It follows that these indirect effects are important for determining reductions in the number of children who are born at later child-bearing ages, and for the profile of recuperation that is observed in Fig. 2.

In the marriage scenario, we observe a very different pattern of mediation. Apart from age 17 years (when TE is itself fairly small), the direct effect is consistently larger than the indirect effect, and after age 25 years it is not only dominant but almost entirely responsible for the total effect. The only ages where mediation has a sizable effect are from 18 to 24 years. At ages 19 and 20 years, the size of the indirect effect is more than 50% of the total effect. Perhaps most interestingly, the direction of the indirect effect at younger ages is different from the direct effect, which in some cases—from ages 18 to 27 years—is more than 100% of the total effect. This is not an error but instead implies that mediation operates in the opposite direction from the total effect, thereby suppressing some of the direct effect. Since the total effect is a reduction in fertility, mediation from ages 18 to 27 years, via education and employment status, thereby serves to increase fertility (e.g. marriage increases fertility indirectly by reducing education and employment). As with the education scenario, this analysis shows that socio-economic processes are important determinants of fertility in their role as mediators, in addition to the direct relationships that they have. We are not aware of previous research that has estimated both the direct and the indirect (mediating) effect of multiple socio-economic processes on fertility, either for the UK, or in other contexts.

#### 4.4. Unobserved confounding

When added to our multivariable models, the strong unobserved confounder that we simulate for the marriage scenario increases the yearly odds of marriage by 4.0 and the odds of childbirth in each year by 1.6. Comparing these odds ratios with exponentiated coefficients from the multivariable models shown in the on-line Table A1, we note that the effect of this unobserved confounder (on marriage) is stronger than all observed conditional relationships, except for the lagged effect of marriage on itself. Similarly, its effect on birth is stronger than that of all baseline covariates and many of the time-varying covariates. Taken together, we therefore conclude that



**Fig. 3.** Total direct effect TDE (—) and pure indirect effect PIE (----) as a percentage of total effect TE (TDE is determined by drawing the values of mediating variables from the natural course; broadly speaking, mediators are set at their observed values, such that only the direct effect remains; PIE is calculated as  $(TE - TDE)/TE$ —and defined as the proportion of TE that would be eliminated by intervening to set the mediators at the values that occur under a given scenario—given this definition, sometimes (and not erroneously) TDE is more than 100% of TE; in the marriage scenario, this is so for ages 27 years and under, and implies that mediation operates in the opposite direction from the total effect, thereby suppressing some of the direct effect): (a) education scenario; (b) marriage scenario

this is a strong confounder. Conclusions for the education scenario are roughly similar. In the education scenario, the strong confounder increases the yearly odds of not being in education of 1.6 and the odds of childbirth in each year by 1.8 (when added to our multivariable models).

Despite our view that these are strong confounders, we also generate two even stronger confounders. For marriage, the first increases the odds of marriage by 9.0 and birth by 2.4, and the second increases the odds of marriage by 16.0 and birth by 3.0. To evaluate the sensitivity of our analysis (and conclusions) to unobserved confounding, we then control for this confounder in our analysis, once at each strength, and compare the natural course with the marriage scenario.

Our sensitivity analysis shows how our results would be changed in the marriage or education scenarios after accounting for (hypothetical) unobserved confounding (Fig. 4). The substantive results are only changed in the case of the strongest hypothetical confounder, suggesting that unobserved heterogeneity could only make a relevant difference in this case. The results remain statistically significant in the two other confounding analyses, at early child-bearing ages for the education scenario, and at later child-bearing ages for the marriage scenario.

We take these results to mean that our conclusions are robust to unmeasured confounding. This is based on our view that unobserved confounding is unlikely to be as strong in reality as in any of our confounding scenarios, in particular the strongest scenario. In reaching this conclusion, we acknowledge that there could be unobserved factors that are as strong as those that we generate. However, with the exception of the variables that are included in our analysis, we are not aware that such factors have been observed in the demographic and sociological literature on fertility.

## 5. Discussion

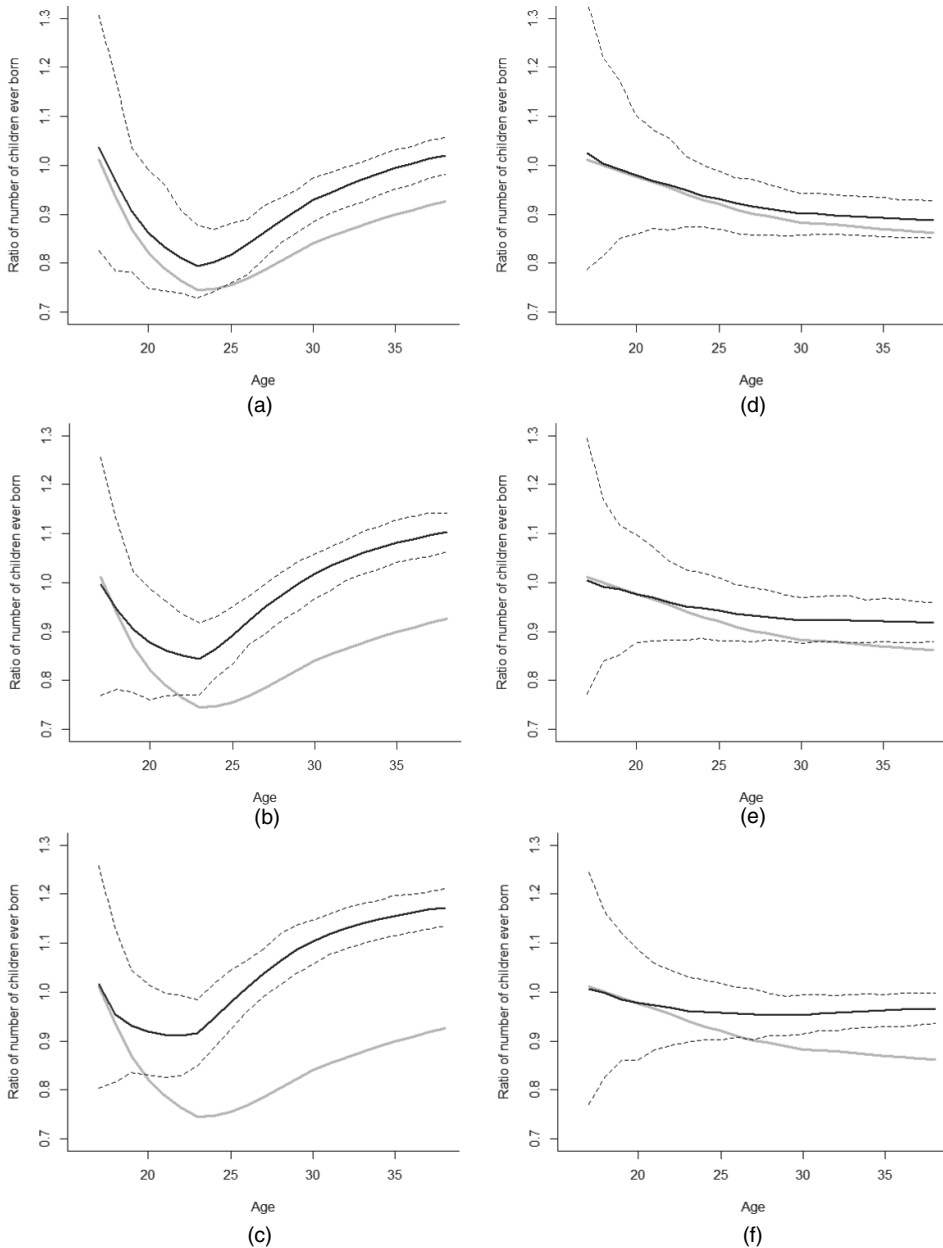
In this paper, we have shown that education, partnership status and employment status are all significant in determining fertility, but to different degrees, and with different profiles over the life course. Moreover, our results help to disentangle the direct and indirect (mediated) effects of these determinants, while showing how the *g*-formula can be used to examine bias due to unobserved heterogeneity. Previous research has focused almost exclusively on estimating the total effects. Here we calculate direct and indirect effects for two scenarios—an increase in higher education attendance, and a reduction in the prevalence of marriage—both of which represent realistic demographic changes. We show that the role of socio-economic mediation can be very different, dependent on the socio-economic processes and stage of the life course that are considered.

On the one hand, the effects of marriage and education are similar, in that they are both important determinants of fertility. Other studies of the UK have also found significant total effects for these factors (Berrington and Pattaro, 2014; Neels *et al.*, 2017; Ní Bhrolcháin and Beaujouan, 2012; Steele *et al.*, 2005, 2006). Similarly, the fact that we find only a small effect of employment (for our limited scenario) also aligns with prior research which shows that

‘different levels of labour force participation by females do not necessarily lead to large changes in fertility events’

(Aassve *et al.* (2006), page 781).

On the other hand, with respect to their direct and indirect effects, they are very different. The indirect effect of marriage, via education and employment status, is relatively small, compared with its direct effect on fertility at age 38 years. At early child-bearing ages, the indirect effect of marriage serves only to counteract its direct effect on fertility partly. In contrast, the indirect effect of education, via partnership status and employment status, is in the same direction as



**Fig. 4.** Sensitivity analysis for the effect of unobserved confounding on (a)–(c) the education and (d)–(f) marriage scenarios (the plot compares the effect of an unobserved confounder between education and fertility and between marriage and fertility which we generate at three levels of strength; for education, at each year of follow-up, the strong confounder has odds ratios on not being in education and on birth respectively of 1.6 and 1.8, the stronger confounder of 2.3 and 2.9, and the strongest of 2.9 and 5.6; for marriage, the odds ratios on marriage and births respectively of the strong confounder are 4 and 1.6, the stronger 9 and 2.4, and the strongest 16 and 3; the original estimates are the same as those found in Figs 2(d)–2(f) (—, original estimates; —, estimates with confounding; ····, 95% confidence bounds): (a) education scenario (strong confounder); (b) education scenario (stronger confounder); (c) education scenario (strongest confounder); (d) marriage scenario (strong confounder); (e) marriage scenario (stronger confounder); (f) marriage scenario (strongest confounder)



the direct effect at all ages. Moreover, the indirect effect of education is much larger than the direct effect after age 27 years. This implies that mediation plays a much more prominent role in determining the effect of education on fertility than the effect of marriage on fertility, at least at later child-bearing ages. Taken together, these results show that socio-economic processes play an important role in determining fertility, not only because of their direct and indirect effects, but also in their role as mediators. To the best of our knowledge, previous empirical research has yet to provide such evidence.

In the final stage of our analysis we examine the sensitivity of our results to unobserved confounding. This is an important extension because it addresses one of the main assumptions of our approach, namely that the most important sources of confounding are observed. There may be some confounders that are not included in our analysis, but the essential question is whether they would make a difference to our findings. On the basis of our sensitivity analysis, it appears that they would not. In reaching this conclusion, we recognize that there are other ways of specifying unobserved confounding, although we have attempted to choose a rigorous and transparent approach.

Nevertheless, unobserved confounding is not the only limitation of our analysis. In common with many previous studies, our analysis is limited to only those women who had complete data for all variables at baseline and in the first year of observation. As such, our analysis is susceptible to assumptions regarding missing data, including that cases are missing conditionally at random and that censoring is non-informative. However, given the similarities between the final analytical sample and the (larger) sample of all cases for which (pre-)baseline variables are observed, we do not believe that a multiple-imputation procedure to account for missingness on time-varying variables would result in substantially different findings (on-line appendix Table A3). Another limitation, or *caveat*, concerns the interpretation of our results. Like most inference methods, the results of the  $g$ -formula are driven by observed relationships in the empirical data. For example, we show a significant effect of changes in marriage behaviour on fertility behaviour. However, we do not model a change in the underlying meaning of marriage. Similarly, our post-birth employment scenario does not include the anticipation of full-time employment after birth that would accompany any real world policy. We note that this limitation is unlikely to be surmounted by an alternative method but instead requires different data, such as (accurate) data on attitudes and intentions, or data on policies that are implemented in the real world. That said, a more nuanced application of the same method may be achieved by using more fine-grained (e.g. monthly) data, if such data were available, e.g. to test the specific effects of monthly variation in maternity leave policies.

One benefit of the  $g$ -formula is that it enables us to make generalizations at the population level. When the interest is in making statements about the population, as is common for example in sociology and demography, this provides advantages over alternative methods that sacrifice external validity. In addition, our example could be extended fairly easily to model additional (or more nuanced) scenarios, e.g. the effect of targeted policies, or the combined effect of changes to more than one socio-economic determinant. As is almost always the case in research on fertility, our analysis focuses on women, but it could also be repeated for men, or for specific subgroups. In addition, future research could examine the extent to which the findings that are shown here can be generalized over time (to other UK birth cohorts) or across space (to other countries).

In settings with larger analytical samples, it may be possible to model the fertility process, and its determinants, with greater complexity. As well as enabling the inclusion of a richer set of covariates, this could enable the links between fertility and its determinants to be modelled by using lags of more than 1 year (similarly to an auto-regressive analysis of time series data: AR(2),

AR(3), etc.). The outcome, births, could be modelled with greater nuance, e.g. by including stillbirths and conceptions not resulting in pregnancy. It would be informative to incorporate the effect of adoptions and child mortality, as well as the proximate determinants of fertility that are not considered here, in particular, contraception, abortion, miscarriages and fecundability. At the moment, all these factors are absorbed into the time-varying mediating processes, but if a sufficiently rich data source were used they could be investigated to gain a deeper understanding of mediation.

To conclude, our study not only shows the many potential benefits of using the  $g$ -formula, but also how it can be developed beyond previous applications. Most methods that have been used to study fertility do not enable researchers to estimate the direct and indirect effects of time-varying determinants, or their role as mediating factors. Indeed, we believe that the same can be said for most methods that are used more generally to study life course processes. In this paper, we have shown how the  $g$ -formula can be used to carry out dynamic mediation analysis, as well as demonstrating for the first time how to carry out sensitivity analysis for unobserved confounding when using the  $g$ -formula. Our analysis highlights the potential benefits of the method for a range of life course topics. For example, it could be usefully applied to the study of linked couple level decision making based on individual level data. Such an analysis is often prevented, or at least made difficult, because of the mediating interrelationships between the life course of both partners. Along similar lines, this approach could be used to study migration, which is a mediating process in the life course of most individuals. Time-varying mediation is methodologically complex and yet, as we have shown, the  $g$ -formula provides social scientists with a sophisticated and flexible tool for understanding this complexity, while also providing meaningful results that are generalizable to the population.

## Acknowledgements

Maarten and Ben thank the many colleagues who provided helpful comments in the development of this research. This includes the organizers and attendees of the International Union for the Scientific Study of Population international seminar on ‘Causal mediation analysis in health and work’ in September 2016. Ben is grateful for financial support received from the Swedish Research Council (Vetenskapsrådet), project 2017-01021 and grant 340-2013-5164 (the Swedish initiative for Research on Microdata in the Social and Medical Sciences), and from the Strategic Research Council of the Academy of Finland, decision 293103 (the research consortium Tackling Inequality in Time of Austerity). Finally, we thank Joshua Perleberg for the initial data management.

## References

- Aassve, A., Burgess, S., Propper, C. and Dickson, M. (2006) Employment, family union and childbearing decisions in Great Britain. *J. R. Statist. Soc. A*, **169**, 781–804.
- Adserà, A. (2005) Vanishing children: from high unemployment to low fertility in developed countries. *Am. Econ. Rev.*, **95**, 189–193.
- Andersson, G. (2004) Childbearing after migration: fertility patterns of foreign-born women in Sweden. *Int. Migr. Rev.*, **38**, 747–774.
- Balbo, N., Billari, F. C. and Mills, M. (2012) Fertility in advanced societies: a review of research. *Eur. J. Popul.*, **29**, 1–38.
- Beaujouan, É. and Ní Bhrolcháin, M. (2011) Cohabitation and marriage in Britain since the 1970s. *Popul. Trends*, **145**, 35–59.
- Becker, G. S. (1981) *A Treatise on the Family*. Cambridge: Harvard University Press.
- Becker, G. S., Duesenberry, J. S. and Okun, B. (1960) *An Economic Analysis of Fertility*. New York: Columbia University Press.

- Berrington, A. and Pattaro, S. (2014) Educational differences in fertility desires, intentions and behaviour: a life course perspective. *Adv. Lif. Course Res.*, **21**, 10–27.
- Berrington, A., Stone, J. and Beaujouan, E. (2015) Educational differences in timing and quantum of childbearing in Britain: a study of cohorts born 1940–1969. *Demog. Res.*, **33**, 733–764.
- Bijlsma, M. J., Tarkiainen, L., Myrskylä, M. and Martikainen, P. (2017) Unemployment and subsequent depression: a mediation analysis using the parametric G-formula. *Soc. Sci. Med.*, **194**, 142–150.
- Billari, F. C. (2015) Integrating macro- and micro-level approaches in the explanation of population change. *Popln Stud.*, **69**, suppl. 1, S11–S20.
- Billari, F. C. and Kohler, H.-P. (2004) Patterns of low and lowest-low fertility in Europe. *Popln Stud.*, **58**, 161–176.
- Bongaarts, J. (2016) Africa's unique fertility transition. *Popln Devlpmnt Rev.*, **43**, no. S1.
- Bongaarts, J., Mensch, B. S. and Blanc, A. K. (2017) Trends in the age at reproductive transitions in the developing world: the role of education. *Popln Stud.*, **71**, 139–154.
- Brien, M. J., Lillard, L. A. and Waite, L. J. (1999) Interrelated family-building behaviors: cohabitation, marriage, and nonmarital conception. *Demography*, **36**, 535–551.
- Buhr, P. and Huinink, J. (2014) Fertility analysis from a life course perspective. *Adv. Lif. Course Res.*, **21**, 1–9.
- Carnegie, N. B., Harada, M. and Hill, J. L. (2016) Assessing sensitivity to unmeasured confounding using a simulated potential confounder. *J. Res. Educ. Effect.*, **9**, 395–420.
- Cartwright, N. (2011) *Evidence, External Validity and Explanatory Relevance*. Oxford: Oxford University Press.
- Chamberlain, J. and Gill, B. (2005) Fertility and mortality. In *Focus on People and Migration* (ed. R. Chappell), pp. 71–89. Berlin: Springer.
- Champion, A. G. and Falkingham, J. (eds) (2016) *Population Change in the United Kingdom*. London: Rowman and Littlefield International.
- Cleland, J. and Wilson, C. (1987) Demand theories of the fertility transition: an iconoclastic view. *Popln Stud.*, **41**, 5–30.
- Coale, A. J. and Watkins, S. C. (1986) *The Decline of Fertility in Europe: the Revised Proceedings of a Conference on the Princeton European Fertility Project*. Princeton: Princeton University Press.
- Cohen, J. E., Kravdal, O. and Keilman, N. (2011) Childbearing impeded education more than education impeded childbearing among Norwegian women. *Proc. Natn. Acad. Sci. USA*, **108**, 11830–11835.
- Daniel, R. M., Cousens, S. N., De Stavola, B. L., Kenward, M. G. and Sterne, J. A. C. (2013) Methods for dealing with time-dependent confounding. *Statist. Med.*, **32**, 1584–1618.
- Davis, K. and Blake, J. (1956) Social structure and fertility: an analytic framework. *Econ. Devlpmnt Cult. Change*, **4**, 211–235.
- Drbe, M., Breschi, M., Gagnon, A., Gauvreau, D., Hanson, H. A., Maloney, T. N., Mazzoni, S., Molitoris, J., Pozzi, L., Smith, K. R. and Vézina, H. (2017) Socio-economic status and fertility decline: insights from historical transitions in Europe and North America. *Popln Stud.*, **71**, 3–21.
- Drbe, M., Oris, M. and Pozzi, L. (2014) Socioeconomic status and fertility before, during, and after the demographic transition: an introduction. *Demog. Res.*, **31**, 161–182.
- Engelhardt, H., Kohler, H.-P. and Prskawetz, A. (2009) *Causal Analysis in Population Studies: Concepts, Methods, Applications*. Berlin: Springer.
- Esping-Andersen, G. and Billari, F. C. (2015) Re-theorizing family demographics. *Popln Devlpmnt Rev.*, **41**, 1–31.
- Forste, R. and Tienda, M. (1996) What's behind racial and ethnic fertility differentials? *Popln Devlpmnt Rev.*, **22**, 109–133.
- Gerster, M., Ejrnæs, M. and Keiding, N. (2014) The causal effect of educational attainment on completed fertility for a cohort of Danish women—does feedback play a role? *Statist. Biosci.*, **6**, 204–222.
- Goisis, A. and Sigle-Rushton, W. (2014) Childbearing postponement and child well-being: a complex and varied relationship? *Demography*, **51**, 1821–1841.
- Goldscheider, F., Bernhardt, E. and Lappegård, T. (2015) The gender revolution: a framework for understanding changing family and demographic behavior. *Popln Devlpmnt Rev.*, **41**, 207–239.
- Greenland, S., Pearl, J. and Robins, J. M. (1999) Causal diagrams for epidemiologic research. *Epidemiology*, **10**, 37–48.
- Guetto, R., Luijckx, R. and Scherer, S. (2015) Religiosity, gender attitudes and women's labour market participation and fertility decisions in Europe. *Acta Sociol.*, **58**, 155–172.
- Heaton, T. B., Forste, R. and Otterstrom, S. M. (2002) Family transitions in Latin America: first intercourse, first union and first birth. *Int. J. Popln Geog.*, **8**, 1–15.
- Hirschman, C. (1994) Why fertility changes. *A. Rev. Sociol.*, **20**, 203–233.
- Hobcraft, J. (1996) Fertility in England and Wales: a fifty-year perspective. *Popln Stud.*, **50**, 485–524.
- Hoem, J. M. (2008) Childbearing trends and policies in Europe. *Demogr. Res.*, **19**, 1–4.
- Huinink, J. and Kohli, M. (2014) A life-course approach to fertility. *Demogr. Res.*, **30**, 1293–1326.
- Inglehart, R. and Baker, W. E. (2000) Modernization, cultural change, and the persistence of traditional values. *Am. Sociol. Rev.*, **65**, 19–51.
- Jalovaara, M., Neyer, G., Andersson, G., Dahlberg, J., Dommermuth, L., Fallesen, P. and Lappegård, T. (2018) Education, gender, and cohort fertility in the Nordic countries. *Eur. J. Popln*, **35**, 563–586.

- Johnson-Hanks, J., Bachrach, C. A., Morgan, S. P. and Kohler, H.-P. (2011) *Understanding Family Change and Variation: Toward a Theory of Conjunctural Action*. Dordrecht: Springer.
- van de Kaa, D. J. (1987) Europe's second demographic transition. *Popln Bull.*, **42**, 1–59.
- Keil, A. P., Edwards, J. K., Richardson, D. R., Naimi, A. I. and Cole, S. R. (2014) The parametric G-formula for time-to-event data: towards intuition with a worked example. *Epidemiology*, **25**, 889–897.
- Kneale, D. and Joshi, H. (2008) Postponement and childlessness—evidence from two British cohorts. *Demogr. Res.*, **19**, 1935–1968.
- Kraval, Ø. and Rindfuss, R. R. (2008) Changing relationships between education and fertility: a study of women and men born 1940 to 1964. *Am. Sociol. Rev.*, **73**, 854–873.
- Kulu, H. and Steele, F. (2013) Interrelationships between childbearing and housing transitions in the family life course. *Demography*, **50**, 1687–1714.
- Lee, R. and Mason, A. (2010) Fertility, human capital, and economic growth over the demographic transition. *Eur. J. Popln*, **26**, 159–182.
- Lesthaeghe, R. (1983) A century of demographic and cultural change in Western Europe: an exploration of underlying dimensions. *Popln Devlpmnt Rev.*, **9**, 411–435.
- Lesthaeghe, R. (2010) The unfolding story of the second demographic transition. *Popln Devlpmnt Rev.*, **36**, 211–251.
- Lillard, L. A. and Waite, L. J. (1993) A joint model of marital childbearing and marital disruption. *Demography*, **30**, 653–681.
- Lin, S.-H., Young, J., Logan, R., Tchetgen Tchetgen, E. J. and VanderWeele, T. J. (2017) Parametric mediational g-formula approach to mediation analysis with time-varying exposures, mediators, and confounders. *Epidemiology*, **28**, 266–274.
- Lorimer, F. (1956) *Culture and Human Fertility*. Paris: United Nations Educational, Scientific and Cultural Organization.
- Macura, M. and Beets, G. (eds) (2002) *Dynamics of Fertility and Partnership in Europe: Insights and Lessons from Comparative Research*, vol. I. New York: United Nations.
- Martin, T. C. (1995) Women's education and fertility: results from 26 Demographic and Health Surveys. *Stud. Famly Planng*, **26**, 187–202.
- McLanahan, S. (2004) Diverging destinies: how children are faring under the second demographic transition. *Demography*, **41**, 607–627.
- Milewski, N. (2010) *Fertility of Immigrants: a Two-generational Approach in Germany*. Berlin: Springer.
- Moffitt, R. (2003) Causal analysis in population research: an economist's perspective. *Popln Devlpmnt Rev.*, **29**, 448–458.
- Moffitt, R. (2005) Remarks on the analysis of causal relationships in population research. *Demography*, **42**, 91–108.
- Neels, K., Murphy, M., Ní Bhrolcháin, M. and Beaujouan, É. (2017) Rising educational participation and the trend to later childbearing. *Popln Devlpmnt Rev.*, **43**, 667–693.
- Neyer, G., Lappegård, T. and Vignoli, D. (2013) Gender equality and fertility: which equality matters? *Eur. J. Popln*, **29**, 245–272.
- Ní Bhrolcháin, M. and Beaujouan, É. (2012) Fertility postponement is largely due to rising educational enrolment. *Popln Stud.*, **66**, 1–17.
- Office for National Statistics (2016) Childbearing for women born in different years, England and Wales: 2015. Office for National Statistics. Newport. (Available from <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/conceptionandfertilityrates/bulletins/childbearingforwomenbornindifferentyearsenglandandwales/2015/pdf>.)
- O'Leary, L., Natamba, E., Jefferies, J. and Wilson, B. (2010) Fertility and partnership status in the last two decades. *Popln Trends*, **140**, 5–35.
- Pearl, J. (2001) Direct and indirect effects. In *Proc. 17th Conf. Uncertainty in Artificial Intelligence* (eds J. Breese and D. Koller), pp. 411–420. San Francisco: Morgan Kaufmann.
- Perelli-Harris, B., Sigle-Rushton, W., Kreyenfeld, M., Lappegård, T., Keizer, R. and Berghammer, C. (2010) The educational gradient of childbearing within cohabitation in Europe. *Popln Devlpmnt Rev.*, **36**, 775–801.
- Reher, D. and Requena, M. (2015) The mid-twentieth century fertility boom from a global perspective. *Hist. Famly*, **20**, 420–445.
- Rendall, M. S. and Smallwood, S. (2003) Higher qualifications, first-birth timing and further childbearing in England and Wales. *Popln Trends*, **111**, 18–26.
- Robins, J. M. (1986) A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math. Modllng*, **7**, 1393–1512.
- Robins, J. M. and Greenland S. (1992) Identifiability and exchangeability for direct and indirect effects. *Epidemiology*, **3**, 143–155.
- Robins, J. M. and Hernán, M. A. (2009) Estimation of the causal effects of time-varying exposures. In *Longitudinal Data Analysis* (eds G. Fitzmaurice, M. Davidian, G. Verbeke and G. Molenberghs), ch. 23, pp. 553–599. Boca Raton: Chapman and Hall—CRC.
- Sigle-Rushton, W. (2008) England and Wales: stable fertility and pronounced social status differences. *Demogr. Res.*, **19**, 455–502.

- Steele, F., Joshi, H., Kallis, C. and Goldstein, H. (2006) Changing compatibility of cohabitation and childbearing between young British women born in 1958 and 1970. *Popln Stud.*, **60**, 137–152.
- Steele, F., Kallis, C., Goldstein, H. and Joshi, H. (2005) The relationship between childbearing and transitions from marriage and cohabitation in Britain. *Demography*, **42**, 647–673.
- UK Data Archive (2016) *1970 British Cohort Study*. Colchester: Institute for Social and Economic Research, University of Essex.
- Upchurch, D. M., Lillard, L. A. and Panis, C. W. (2002) Nonmarital childbearing: influences of education, marriage, and fertility. *Demography*, **39**, 311–329.
- VanderWeele, T. J. (2014) A unification of mediation and interaction: a 4-way decomposition. *Epidemiology*, **25**, 749–761.
- VanderWeele, T. J. (2015) *Explanation in Causal Inference: Methods for Mediation and Interaction*, 1st edn. New York: Oxford University Press.
- Vanderweele, T. J. and Vansteelandt, S. (2009) Conceptual issues concerning mediation, interventions and composition. *Statist. Interf.*, **2**, 457–468.
- Wang, A. and Arah, O. A. (2015) G-computation demonstration in causal mediation analysis. *Eur. J. Epidem.*, **30**, 1119–1127.

#### *Supporting information*

Additional 'supporting information' may be found in the on-line version of this article:  
'Online supplementary material'.