



# **The Economic Approach to Fertility: A Causal Mediation Analysis**

Gauthier T. Kashalala and Steven F. Koch

**ERSA working paper 442**

**July 2014**

Economic Research Southern Africa (ERSA) is a research programme funded by the National Treasury of South Africa.

The views expressed are those of the author(s) and do not necessarily represent those of the funder, ERSA or the author's affiliated institution(s). ERSA shall not be liable to any person for inaccurate information or opinions contained herein.

# The Economic Approach to Fertility: A Causal Mediation Analysis \*

Gauthier T. Kashalala<sup>†</sup> and Steven F. Koch<sup>‡</sup>

July 10, 2014

## Abstract

This study develops an economic fertility model which explicitly incorporates both the costs of childrearing and contraception behaviour. In this setting, a couple capacity to procreate depends on their fecundity, as well as their contraception and sexual behaviours; and the ideal number of children is chosen by maximizing the utility of children, subject to a budget constraint reflecting the couple's income, and their specific explicit and implicit costs of rearing children. Using a nonparametric causal mediation framework (Pearl, 2009; Heckman and Pinto, 2013), our analysis explicitly explores the role of family planning services and the cost of children in mediating the causal effect of income on fertility, subject to unmeasurable fecundity and unobserved sexual risk taking behaviour. In particular, we discuss the definition, identification and estimation of a variety of causal effects, namely, the direct income effect, the contraception effect, and price effect.

*J.E.L.:* C14, D13, I38, J13, J38

---

\*This research is taken from Kashalala's PhD dissertation, supported by the William and Flora Hewlett Foundation under grant G15085485, and partly completed during a visit to the Economics Research Center (ERC) at the University of Chicago, whose hospitality is gratefully acknowledge. We are also grateful to James J. Heckman and his collaborators for useful discussions and comments, as well as seminar participants at the University of Pretoria.

<sup>†</sup>Department of Economics, University of Pretoria, Private Bag X20, Hatfield 0028, Republic of South Africa. [gauthier.kashalala@up.ac.za](mailto:gauthier.kashalala@up.ac.za)

<sup>‡</sup>Department of Economics, University of Pretoria, Pretoria, Republic of South Africa; (0) 27-12-420-5285, (F) 27-86-691-2749. [steve.koch@up.ac.za](mailto:steve.koch@up.ac.za)

# 1 Introduction

Since standard models of household fertility postulate that children are normal goods, one would expect a positive relationship between income and children. However, empirical evidence invariably suggests that within a given society fertility is often higher in poorer families (Becker, 1960; Jones and Tertilt, 2008). That negative relationship is also consistent across countries; those with higher average fertility have lower average levels of industrialization (Galor and Zang, 1997; Bloom et al., 2009). In other words, standard economic assumptions are not generally enough to explain the negative fertility-income relationship regularly documented in empirical studies; instead, special assumptions regarding the functional form of the household's utility or production functions have been necessary (Jones et al., 2011).

Therefore, economic models of fertility have been extended to incorporate other relevant aspects of the childrearing decision, such as the explicit costs of raising the child and implicit costs associated with parental time and effort. Each of these costs can be linked to wage rates in the labour market (Becker, 1965; Mincer and Polachek, 1974), and create tradeoffs between the quantity of children and the quality of children (Becker and Lewis, 1973; Leibenstein, 1975; Caldwell, 1976). Contemporary economic fertility theories, such as these, focus on the effects of parental income and the opportunity costs of child-rearing on completed family size. With a few exceptions (Becker, 1960; Heckman and Willis, 1976; Michael and Willis, 1976), these theories do not explicitly incorporate child production function inputs, such as fecundity and family planning services, despite the role of these two factors in shaping a woman's fertility history.

This study develops an integrated analysis of fertility choices (Easterlin, 1975), where a couples' capacity to procreate depends on their fecundity, contraception decisions and sexual behaviours (Becker, 1960; Heckman and Willis, 1976; Michael and Willis, 1976). At the same time, they are assumed to choose the ideal number of children by maximizing the utility of their children, subject to a budget constraint reflecting the couple's income, and their explicit and implicit costs of rearing those children. This approach is consistent with the demographic transition literature, which postulates that the following three prerequisites should prevail for there to be a sustained fertility transition: (i) fertility must be within the calculus of conscious

choice; (ii) effective techniques of fertility reduction must be accessible; and (iii) reduced fertility must be viewed as advantageous, Coale (1984).

Using a nonparametric causal mediation framework (Pearl, 2009; Heckman and Pinto, 2013), our analysis explores the roles of family planning services and the cost of children in mediating the causal effect of income on fertility, subject to unmeasurable fecundity and unobserved sexual risk-taking behaviour. In particular, we decompose the total causal effect of income on fertility into a direct income effect, a price effect and a contraception effect. This qualitative causal model produces the following testable quantitative causal claim about the effect of income on fertility: the sign and the magnitude of the causal relation between income and fertility depends on the relative sizes of the positive direct income effect and the negative indirect effects that are mediated through the cost of childrearing and contraception. In this context, the total effect of income on fertility will be negative if the direct income effect is lower than the mediated effects, and positive, otherwise.

Theoretically, our model is a modest contribution towards solving the ambiguity of the negative fertility-income relationship previously documented in the empirical population economics literature. The principal contribution of this study is to recast the *fertility -income* debate into the *income - cost of childrearing - contraception - fertility* space, which effortlessly relates reproductive health outcomes to labour market outcomes. We assume that fertility choices are linked to labour market outcomes through family planning and time allocation. In particular, a pure reproductive health investment model suggests that investment in family planning is driven by the couple's desire to maximize their lifetime net money income, subject to dynamics in their stock of reproductive health capital (Ben-Porath, 1967). Thus, within the human capital theory of demand for health (Grossman, 1972, 2000; Becker, 2007) and under a household money income production function (Schultz, 1961; Ben-Porath, 1967; Mincer and Polachek, 1974) which varies with contraceptive efficiency, one would expect access to reproductive health services to influence participation or performance in the labour market.

Our causal mediation analysis of fertility is not only of theoretical interest. Salient policy implications arise for a variety of countries at different stages of the demographic transition. In the context of an aging population, for example, it provides tools to evaluate family-friendly labour

market policies aimed at boosting fertility, through the reduction of the cost of childrearing. Our analysis also provides a sound theoretical background for the evaluation of population policies aimed at reducing unwanted childbearing, through the subsidization of family planning services. In what follows, we formalize the causal link between labour and reproductive health outcomes. In the end, we illustrate all the causal mechanisms using a causal path diagram (Pearl, 2009).

## 2 The Causal Path to Conception or Birth

The causal path that we outline is built upon a number of stylized ‘facts’ that are observed in the literature. We highlight the most important observations, first, before constructing the model.

### 2.1 Background

In the model we propose, education is allowed to influence the wage rate (Heckman et al., 2006), as well as family planning. Almost every discussion of fertility differentials refers to education, although often as a proxy for lifetime income (Jones and Tertilt, 2008). Theoretically, education’s role in the fertility decline has been analyzed with reference to Coale (1984)’s first two preconditions of fertility transition relative to availability and choice of contraceptive methods (Cleland and Wilson, 1987). Intuitively, growth in formal education empowers individuals with a sense of control over their destiny and pursuits, which changes attitudes towards and increases birth control propensity. Furthermore, maternal education achievement is often empirically associated with reduced infant mortality (Caldwell, 1979), which in turn reduces the need for a large family to compensate for possible not surviving children (Schultz, 1986; Preston and Hartnett, 2010).

Moreover, our analysis also explicitly explores the contraceptive efficiency pathway to fertility that had been made redundant in previous studies. There is a growing body of evidence suggesting that falling fertility levels since the 1990s are largely due to a rising proportion of women demanding fertility control (Caldwell et al., 1992; Johnson-Hanks, 2007; Moultrie et al., 2012), while contraceptive use has been related to much of the Sub-Saharan African fertility decline (Caldwell et al., 1992; Timaeus and Moultrie, 2008).

The demographic rationale for family planning is to avoid or at least reduce the risk of pregnancy associated, through a reduction in fecundity (Moultrie et al., 2012). However, despite the fact that some earlier work on fertility choice discusses the impact of differential contraceptive knowledge on fertility (Becker, 1960; Schultz, 1973), the role of contraception effectiveness (or efficiency) in reducing fecundity has received much less attention in both the theoretical and empirical literature. Rather than explicitly analyse the causal effects of contraception on fertility, a number of theoretical studies were often concerned with the determinants of contraceptive use (Coale, 1984; Cleland and Wilson, 1987; Easterlin, 1975; Bongaarts, 1993). Consequently, much of early models of fertility choices implicitly assumed perfect control over pregnancy (Easterlin, 1975; Easterlin et al., 1988; Bongaarts, 1993), thus failing to take into account the stochastic nature of the human reproduction process.

Finally, apart from behavioural factors, such as contraceptive use and sexual behaviour, fertility differentials might also be an indication of some latent differences in biological factors, such as human fecundity (Gini, 1926; Skakkebaek et al., 2006; Louis, 2011). Fecundity is generally defined as the biologic capacity for reproduction, during a normal menstrual cycle with sexual relations and no contraception (Gini, 1926; Leridon, 2007). Although the fecundity spotlight focused on females, male fecundity has also begun to receive attention (Carlsen et al., 1992; Skakkebaek et al., 2006). As such, fecundity outcomes are couple dependent, are a function of the timing and frequency of sexual intercourse and the biological reproductive capacity of the two partners. One of the main challenges to incorporating fecundity is that it is not a dichotomy; rather, it is a continuum, and its value is highest in the first few months couples are attempting to conceive, relative to later months (Tietze, 1959). This pattern suggests that couples would differ considerably in their ability to achieve pregnancy (Louis, 2011). In what follows our fertility model treats human fecundity as an unobservable covariate influencing both a woman's contraceptive behaviour and her fertility outcomes.

## 2.2 Model Structure

Assume that any fecundable woman  $i$  is naturally endowed with a vector of unobserved abilities  $U_i$ , which consists of both cognitive and noncognitive skills affecting her schooling decisions, as

well as her labour and behavioural outcomes (Heckman et al., 2006). In particular, cognitive abilities, such as intelligence, indirectly affect the woman's probability to secure a wage rate  $W_i$  in the labour market through observed schooling decisions  $S_i$ . However, noncognitive traits such as motivation and persistence would also directly determine  $W_i$ . In terms of additional notation,  $\varepsilon_{ia}$  in the sequel is the individual-level idiosyncratic error term related to the variable  $a$ .

$$W_i = W_i(S_i, U_i, \varepsilon_{iw}). \quad (1)$$

On the reproductive health side, the demand for family planning is assumed to be an investment in reproductive health capital. Using the human capital theory of the demand for health (Grossman, 1972, 2000; Becker, 2007), we suppose that contraception increases the time a woman remains active in the labour market. US data, for example, suggests that the introduction of family planning was associated with significant and persistent reductions in fertility driven both by falling completed childbearing and childbearing delay (Kearney and Levine, 2009; Bailey, 2012), which in turn increased the time available to women for labour and non-labour activities. Thus, the time spent in the labour market  $H_i$  varies with the levels of family planning efficiency  $Z_i$ .

$$H_i = H_i(Z_i, \varepsilon_{ih}). \quad (2)$$

The number of hours spent in labour market activities is the main input in the money income production technology, or earnings function. Thus, Equations (1) and (2) imply the following equation for the woman's earnings, which might include other sources of income.

$$Y_i = Y_i(H_i, W_i, \varepsilon_{iy}) \quad (3)$$

Earnings and income, though, are assumed to have feedback effects. The wage rate influences the time cost of children (Mincer and Polachek, 1974). Similarly, it is necessary to allow for the growth in income to shift the household into a higher socio-economic status, which translates into higher expenses per child. Such a shift implies a direct relationship between income and the price of children (Leibenstein, 1975; Caldwell, 1976). In this context, the cost of childrearing

$X_i$  combines both expenditures on child related products, which varies with income, and the opportunity cost of parental time, which is measured by the wage rate.

$$X_i = X_i\left(Y_i(H_i, W_i, \varepsilon_{iy}), W_i(S_i, U_i, \varepsilon_{iw})\right) \quad (4)$$

The last component to incorporate is contraceptive behaviour. We consider a biological production process within which latent ability, including unobserved biological factors, is assumed to have a direct impact on contraception behaviour, as well as an indirect impact through education. For this analysis, we focus on the effectiveness (often referred to as efficiency, below) of contraception, because the effectiveness of contraceptive methods varies, partly due to consistency in application, which is more likely to capture the woman's behavioural decisions.

Thus, a woman's demand for contraception efficiency ( $Z$ ) is assumed to be affected by the above considerations, where  $V_i$  is latent fecundity, capturing both uncertainty in the human reproduction process and a variety of unobserved biological factors, including the quality of semen of the woman's partner. In this context, an infertile woman or a fecundable woman who is experiencing a temporary sterile period for a variety of reasons, such as lactational amenorrhea, is not likely to use contraception during intercourse.

$$Z_i = Z_i(S_i, Y_i, U_i, V_i, \varepsilon_{iz}) \quad (5)$$

Although a very limited number of studies have explicitly incorporated contraceptive efficiency in the theoretical analysis of fertility choices (Heckman and Willis, 1976; Michael and Willis, 1976). These models have tended to assume perfect fertility control. Our approach has the advantage of taking into account the impact of opportunity costs on the demand for family planning, while explicitly establishing the causal path between contraception use and fertility. Importantly, fertility control is not perfect.

It is clear that each variable in the model is generated by a nonparametric structural equation that relates that variable to its immediate causes, and to its idiosyncratic error,  $\varepsilon_{ia}$ , via some arbitrary deterministic function. All the nonparametric functions are invariant under external



manipulations of their arguments (Frisch, 1938). Further, these errors are mutually independent:  $\varepsilon_{ia} \perp \varepsilon_{ib}$ ,  $a \neq b$ . The random nature of these error terms implies that our fertility model is also stochastic.

To simplify the analysis, we drop the subscript  $i$  to ease notation and keep the error terms implicit. Thus, the causal assumptions conveyed above translate into the following fertility production function.

$$Q = f \left[ Y \left( H(Z), W(S, U) \right), X \left( Y \left( W(S, U) \right), W(S, U) \right), Z \left( S, Y \left( W(S, U) \right), U, V \right), v \right] \quad (6)$$

The diagrammatic description of this structural equation is given in Figure 1, which is a visual representation of our nonparametric structural equation fertility model. In particular, this cyclical causal model depicts the mediation structure, where some time lags are assumed to approach zero (Fisher, 1970).

For the purpose of this study, our version of Figure 1 assumes no simultaneity. We also abstract from the feedback effect of birth control on income, as mediated by the time spent in the labour market. This simplification produces the directed acyclical graph (DAG) in Figure 2, where no variable is a descendant of itself. The resulting model consists of eight random variables:  $Y$  denotes the level of income received,  $Z$  is the efficiency of the contraceptive strategy followed,  $Q$  represents fertility outcomes,  $S$  denotes the highest level of education achieved,  $W$  is the wage rate, and  $X$  is the per capita cost of childrearing; the latent variables  $U$  and  $V$  represent skills and fecundity, respectively.

This simplified DAG and associated structural equations for the simplified structural model translate into the following fertility production function.

$$q = f \left[ y \left( w(s, u) \right), x \left( y \left( w(s, u) \right), w(s, u) \right), z \left( s, y \left( w(s, u) \right), u, v \right), v \right]. \quad (7)$$

Under appropriate properties of continuity and derivability, the total effect of income on fertility

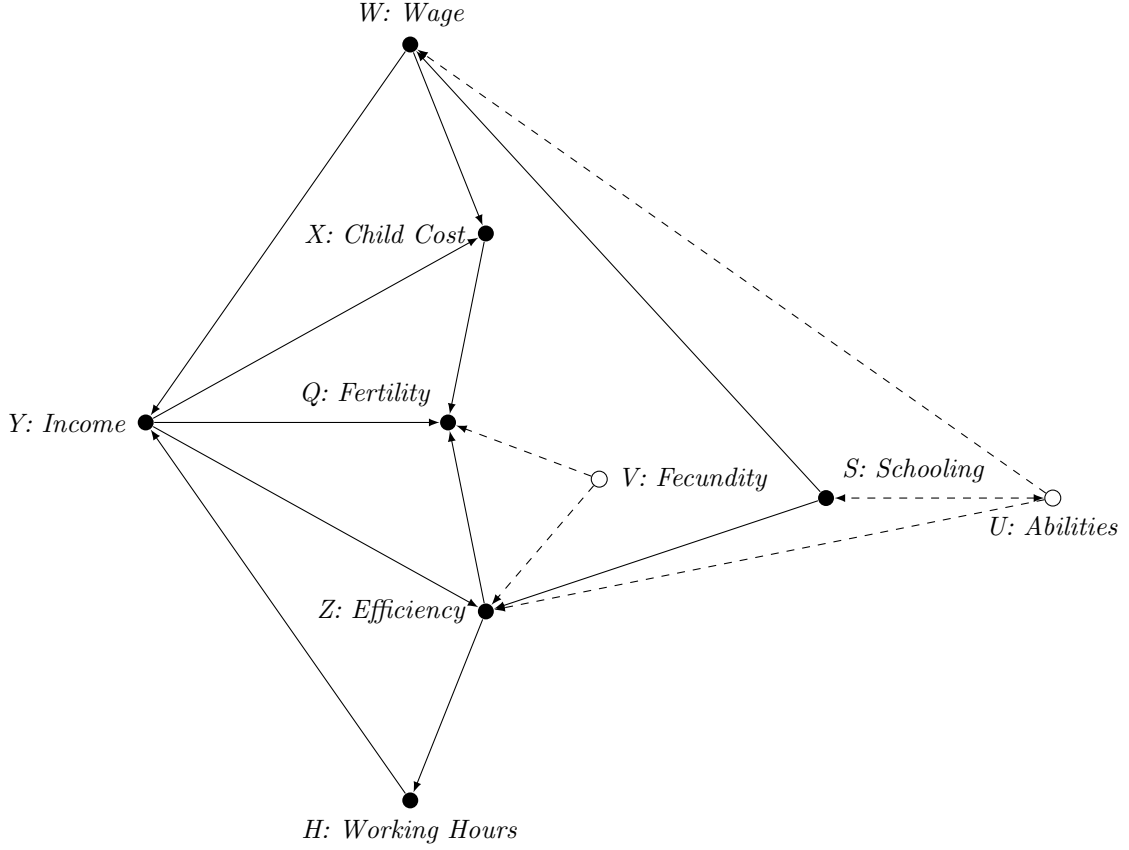


Figure 1: The Cyclical Causal Fertility Model. Solid bullets represent observed variables, hollow circles represent unobserved (or latent) variables.

related to Equation (7) can be determined.

$$\underbrace{\frac{dq}{dy}}_{\delta: \text{ Total Causal Effect}} = \underbrace{\frac{\partial f(\dots)}{\partial y}}_{\delta_y: \text{ Pure Income Effect}} + \underbrace{\frac{\partial f(\dots)}{\partial x} \frac{\partial x(\dots)}{\partial y}}_{\delta_x: \text{ Price Effect}} + \underbrace{\frac{\partial f(\dots)}{\partial z} \frac{\partial z(\dots)}{\partial y}}_{\delta_z: \text{ Contraceptive Effect}} \quad (8)$$

In Equation (8),  $\delta$  is the *total effect* of income on fertility, measured as the effect of a change in income. It is combined with the change in fertility mediated by the change in contraceptive efficiency or the cost of childrearing. The income effect  $\delta_y$  captures the *direct effect* of the change in income for a given level of contraceptive efficiency and cost of children. The price effect  $\delta_x$  is the *indirect effect* operating through the opportunity cost of children, and the contraceptive effect  $\delta_z$  is also an indirect effect of income via the demand for family planning services.

The sign of the different effects represents a stable economic relationship that only depends

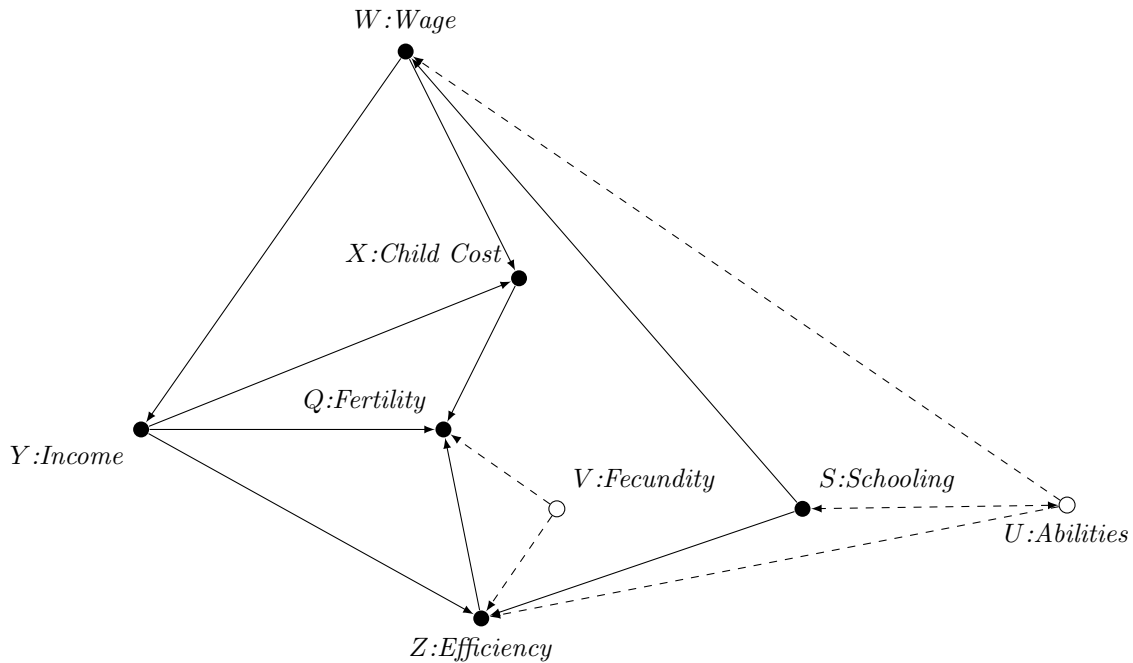


Figure 2: The Acyclical Causal Fertility Model. Solid bullets represent observed variables, hollow circles represent unobserved (or latent) variables.

on the economic nature of children, and is transportable across any cultural and sociological environment that may affect fertility. The direction of the total effect in Equation (8) would depend on the sign and the relative size of the direct income, price and contraceptive effects, respectively. The total effect of income on fertility will be positive, if the direct income effect is greater than the mediated price and contraceptive effects combined, and negative, if the opposite is true. Note that the magnitude of the total effect is also affected in the same fashion, despite the fact that the total effect may be weak or nonexistent, due to the cancellation of effects. Such effects might arise when a strong positive direct effect is cancelled by equally strong negative mediated effects. Thus, it is possible for the total effect to be weak or nonexistent, while direct and indirect effects are both strong.

The direct income effect,  $\delta_y$  should always be positive. If we assume that children are normal goods, children are an increasing function of income, *i.e.* richer parents would naturally have larger families. The price effect,  $\delta_x$ , is the indirect effect of income on fertility, mediated by factors indicative of the cost of having children. These factors include the opportunity cost of

time, as measured by the wage rate (Mincer and Polachek, 1974), and the social class-related expenditures on children, which are assumed to be positively related to the family’s income (Martineau, 1958; Hotz et al., 1997). It is clear that, by the Law of Demand, an increase in the cost of children will decrease the demand for children. Lastly, the contraceptive effect,  $\delta_z$ , describes the impact of income changes on the supply of children, via the demand for family planning. Assuming contraceptive efficiency is a normal good, higher earning couples will tend to use more efficient family planning services, which will result in a smaller completed family size, *ceteris paribus*.

### 3 Nonparametric Identification

This section reviews the conditions necessary for the identification and estimation of the causal effects of income on fertility outcomes, from the perspective of a nonparametric structural equation framework (Petersen et al., 2006; VanderWeele, 2009; Pearl, 2009, 2012*b*). We use contraceptive efficiency and cost of childbearing as mediator variables. In particular, we discuss the conditions under which the direct income, price and contraceptive effects, defined in the previous section, can be identified and estimated using observational studies.

#### 3.1 The Counterfactual Definition of Causal Effects

Causality is a thought experiment (Frisch, 1930; Haavelmo, 1943, 1944). In the counterfactual framework, the causal effect of treatment is defined as the difference in individual outcomes assuming both exposure to treatment and no exposure (Rubin, 1974; Holland, 1986). Since one can not observe the value of the outcome under treatment and non-treatment on the same individual, it is not possible to uncover the true causal effect (Holland, 1986). The hypothetical nature of causal analysis implies, in the context of structural causal inference (Pearl, 2009), that the various effects described in Equation (8) have causal meanings if and only if it was possible to independently manipulate the causal paths implied by the system in Equation (7). Following Frisch (1930, 1938) and Haavelmo (1943, 1944), the thought experiment subsumes hypothetical variation in treatment levels, in a way that does not affect other variables in the

model. In practice, this often entails variation in hypothetical exogenous inputs (Heckman and Pinto, 2013), and the hypothetical disablement of specific direct causal links (Pearl, 1995).

Using the hypothetical approach to causality we can define a range of causal effects as discussed in the mediation analysis literature (Pearl, 2012b). These effects include the *Total Effect* (TE); the *Natural Direct Effect* (NDE); as well as the *Natural Indirect Effect* (NIE). Although similar counterfactuals have been defined before in other fields of research, our analysis is a comprehensive application of a general counterfactual framework to fertility choices that integrates the utility maximization and production functions of the household. The framework is flexible enough to simultaneously accommodate several aspects of the economic approach to fertility, and can be used to evaluate a wide range of family policies.

The TE is the impact on fertility resulting from a move in income from  $y$  to  $y'$ , with the mediator variables  $X$  and  $Z$  being allowed to track the change in  $Y$  as dictated by the model in Equation (7).

$$\begin{aligned} TE &= E \{ f[y', x(y', w(s, u)), z(s, y', u, v), v] - f[y, x(y, w(s, u)), z(s, y, u, v), v] \} \\ &= E \{ Q(y') \} - E \{ Q(y) \} \end{aligned} \tag{9}$$

$Q(y)$  denotes the outcome  $Q$  when treatment  $Y$  is fixed at  $y$ .

The NDE is defined as the expected change in outcome, induced by moving treatment from  $y$  to  $y'$ , while setting the intermediate variables to whatever value they would have attained in the absence of treatment. This definition implies the need to disable the ability of mediating variables to respond to treatment, while allowing them to vary with the other individual characteristics in the model, as if the treatment did not change. In our model, this counterfactual definition translates into the following expression for the *Direct Income Effect* (DIE).

$$\begin{aligned} DIE &= E \{ f[y', x(y), z(s, y, u, v), v] - f[y, x(y), z(s, y, u, v), v] \} \\ &= E \{ Q(y', X_y, Z_y) \} - E \{ Q(y, X_y, Z_y) \}. \end{aligned} \tag{10}$$

Thus, the DIE is defined as the expected effect of income on fertility outcomes, while blocking the

impact of increases in income on the demand for contraceptive efficiency or the price of children. In other words, the DIE measures the expected impact of income on fertility, while setting the demand for contraceptive efficiency and the cost of childrearing for each individual to whatever value they would have attained prior to the change in income. This counterfactual definition of the DIE suggests that, other things being equal, growth in family income would have a positive impact on completed family size.

Similarly, the hypothetical model can be used to define the NIE of treatment on outcome as the portion of total effect explained by the mediating variables, if the outcome had not responded to changes in treatment. Thus, the indirect effect involves nested counterfactuals, since it requires fixing the treatment itself, at its reference level, while allowing the mediators to change to new levels, the ones they could have attained under treatment (Pearl, 2001; Petersen et al., 2006). In our case, we have two types of NIEs: the *Price Effect* (PE) and the *Contraception Effect* (CE).

$$\begin{aligned}
CE &= E \{ f[y, x(y), z(s, y', u, v), v] - f[y, x(y), z(s, y, u, v), v] \} \\
&= E \{ Q(y, X_y, Z_{y'}) \} - E \{ Q(y, X_y, Z_y) \}.
\end{aligned} \tag{11}$$

The magnitude of CE is of particular interest for the evaluation of family planning programs aimed at reducing unintended childbearing by subsidizing contraception. However, the direction is not clear (Bailey, 2012); some models predict that family planning programs could reduce childbearing (Becker and Lewis, 1973; Michael and Willis, 1976; Kearney and Levine, 2009), but not all (Ananat et al., 2009). We address this inconclusiveness, assuming that contraception is linked to fertility through its impact on the probability ( $p$ ) of pregnancy (Heckman and Willis, 1976), which is affected by contraceptive efficiency ( $z$ ) and fecundity ( $v$ ).

$$p = (1 - z)v, \quad 0 < p, z, v < 1, \tag{12}$$

Assuming that contraception efficiency is a normal good, CE implicitly recognizes an inverse link between contraception use and birth intervals (Yeakey et al., 2009), which ultimately results in

a negative fertility response to a rise in efficiency levels, consistent, in our case, with an increase in income from  $y$  to  $y'$ . PE, on the other hand, flows through its effect on child costs.

$$\begin{aligned} PE &= E \{ f[y, x(y'), z(s, y, u, v), v] - f[y, x(y), z(s, y, u, v), v] \} \\ &= E \{ Q(y, X_{y'}, Z_y) \} - E \{ Q(y, X_y, Z_y) \}. \end{aligned} \tag{13}$$

The identification and estimation of the PE is not only of theoretical interest, it has significant policy implications in countries with an aging population. Decreasing fertility and increasing life expectancy has skewed the age distribution upwards in a number of developed countries, posing long-term risks to economic growth and the sustainability of the welfare state. The below-replacement fertility rate is partly blamed on the increased cost of raising children, due to economic development (Kalwij, 2010). A variety of family-friendly labor market policies, such as family allowances, childcare subsidies, and maternity-parental leave benefits, have been proposed to reconcile family life and employment. Using our model, one can evaluate the fertility response of these employment-oriented family policies, by estimating the PE. In this context, PE captures a key prediction of our model: fertility will respond to changes in the price of children.

### 3.2 Identifying the Causal Effects

In the analysis of causal effects, identification seeks to answer the fundamental question of whether, given a set of assumptions, the modified distribution required for the definition of counterfactuals can be estimated from observational data. Thus the goal of identification is to express the causal parameters from the *hypothetical* model using observed probabilities from the *empirical* model that governs the data-generating process (Pearl, 2001, 2009; Heckman and Pinto, 2013).

Consider a model describing the causal mechanisms amongst a set of variables  $\mathcal{L} = \{K_1, \dots, K_n\}$  associated with a set of mutually independent error terms  $\varepsilon = \{\varepsilon_1, \dots, \varepsilon_n\}$ . The mechanism is outlined through a system of autonomous structural equations  $\{g_1, \dots, g_n\}$ . We are interested in the causal relationships  $K_i = g_i(Pa(K_i), \varepsilon_i)$ , where the *endogenous* variable  $K_i$  is directly caused by its parents  $Pa(K_i) \subset \mathcal{L}$  and  $\varepsilon_i \in \varepsilon$ . Variables directly or indirectly causing  $K_i$  are

called ancestors  $An(K_i) = \{K_j \in \mathcal{L}; K_i \in De(K_j)\}$ . Similarly,  $De(K_i) = \{K'_i \in \mathcal{L}; K_i \in An(K'_i)\}$  are all descendants directly or indirectly caused by a variable  $K_i$ . If  $Pa(K_i) = \emptyset$  then  $K_i$  is an *exogenous* variable determined outside the model. By analogy, a set of variables that have  $K_i$  as a parent are called children of  $K_i$ ,  $Ch(K_i) = \{K'_i \in \mathcal{L}; K_i \in Pa(K'_i)\}$ .

### 3.2.1 Conditional Independence

The statistical information encoded in a causal model is completely captured by conditional independence relationships among the variables in the model. Consider a DAG which contains  $n$  variables,  $\{K_1, \dots, K_n\} = \mathcal{L}$ . The local Markov property states that, under mutually independent error terms, each variable in a recursive model is conditionally independent of its non-descendants, given its parents:

$$K \perp\!\!\!\perp \mathcal{L} \setminus [De(K) \cup Pa(K)] \mid Pa(K). \quad (14)$$

If the  $n$  variables in the model form a joint probability distribution  $P(k_1, \dots, k_n)$ , then this property is equivalent to the following Markov factorization property:

$$P(k_1, \dots, k_n) = \prod_{i=1}^n p(k_i \mid Pa(k_i)), \quad (15)$$

which, in turn, is equivalent to the global Markov property defined by the concept of directional separation, or *d*-separation (Pearl, 1988). *d*-separation is a graphical criteria allowing researchers to read, off their respective diagrams, all structural implications of a causal model.

**Definition 3.1.** (*d*-separation): *A path  $p$  is blocked by a conditioning set of variables  $B$  if either (1)  $p$  contains a chain  $i \rightarrow m \rightarrow j$  or a fork  $i \leftarrow m \rightarrow j$ , such that  $m \in B$ ; or (2)  $p$  contains a collider  $i \rightarrow m \leftarrow j$ , such that neither  $m$ , nor any descendant of  $m$ , is in  $B$ . If  $B$  blocks all paths from set  $X$  to set  $Y$ , it is said to “*d*-separate  $X$  and  $Y$ ”.*

The above definition gives rise to the *d*-separation theorem widely used in nonparametric causal mediation analysis (Verma and Pearl, 1991; Pearl, 2009).

**Theorem 3.1.** *d*-separation theorem: *If two sets of variables  $X$  and  $Y$  are *d*-separated by another*



set of variables  $B$  along all paths in a graphical causal model, then  $X$  is statistically independent of  $Y$ , conditional on  $B$ , in every distribution compatible with the graphical causal model:  $X \perp\!\!\!\perp Y|B$ .

In our approach, a different type of causal effect gives rise to a structurally different hypothetical model, characterized by a different set of pairwise marginal and conditional independencies. In what follows, we use the  $d$ -separation criteria to reveal the structure of associations and independencies, which could be observed, if the data were generated according to the causal assumptions encoded in each hypothetical causal fertility model. In this study, we assume that all testable implications are derived from the hypothetical manipulation of the causal graph in Figure 2.

Consider an intervention  $do(k)$  fixing the value of  $K$  to some constant  $k$ , irrespective of the usual behaviour of  $K$  given by the empirical joint distribution  $P(\mathcal{L})$ . Every such intervention in a statical model would contain testable causal implications, only if the resulting interventional distribution is expressible in terms of conditional distributions, where the local Markov assumption enables the move from the encoded causal assumptions to associations observable in the data.

$$P(\mathcal{L} \setminus K_j | do(k)) = \prod_{k_i \notin K_j} P(k_i | Pa(k_i)). \quad (16)$$

Equation (16) is the so-called truncation formula (Pearl, 2009), where the intervention has removed all  $P(k_j | Pa(k_j))$  terms from the Markov factorization, such that the parents of every variable are that variable's direct causes.

In order to operationalize hypothetical manipulations of the causal relationships, as required by the counterfactual causal effects definitions, we expand the empirical model with an hypothetical treatment variable  $\tilde{K}_j$  related to the treatment variable  $K_j$ . For consistency, we assume that assigning a value  $k$  to  $\tilde{K}_j$  is equivalent to observing  $K_j$  attaining a value  $k$ . We refer to this as a hypothetical variable.

**Definition 3.2.** (*The hypothetical variable*): We define an hypothetical variable  $\tilde{K}_j$ , associated with the treatment variable  $K_j$ , such that

$$(1) An(\tilde{K}_j) = \emptyset,$$

(2)  $\tilde{K}_j \in Pa(K_i)$ , and

(3)  $[De(\tilde{K}_j) \in \tilde{\mathcal{L}}] \subseteq [De(K_j) \in \mathcal{L}]$ ,

where  $K_i$  is any variable in the empirical model with parent  $K_j$ , and the lists of variables in the empirical and hypothetical models are  $\mathcal{L}$  and  $\tilde{\mathcal{L}} = \mathcal{L} \cup \{\tilde{K}_j\}$ , respectively.

By assumption, both the hypothetical and empirical models share common features and encode the same sets of conditional independence relations (Verma and Pearl, 1991; Ali et al., 2009). In order to preserve equivalence, the hypothetical manipulation should not create or destroy any independence in the modified model (Pearl, 2012a). Thus,  $d$ -separation can also be used to test whether a given hypothetical manipulation equivalence preserving. A hypothetical modification of the data generating process maintains equivalence if it neither creates nor destroys any  $d$ -separation condition in the resulting hypothetical model (Verma and Pearl, 1991; Pearl, 2009). In our analysis, equivalence is preserved through a requirement of common support for  $(K_j, \tilde{K}_j)$ , invariant structural equations and the stationary distribution of error terms  $\varepsilon$ . The following theorem links the probability measures of our hypothetical and empirical fertility models, and shows that these two causal models are equivalent (Verma and Pearl, 1991; Ali et al., 2009):

**Theorem 3.2.** *Let  $\tilde{K}_j \in \tilde{\mathcal{L}}$  be the hypothetical variable in the hypothetical model associated with treatment variable  $K_j$ . Assume  $\tilde{K}_j$  is uniformly distributed over the support of treatment variable  $K_j$ . Then, the empirical and hypothetical models define the same probability distribution  $P(\tilde{\mathcal{L}}|K_j = \tilde{K}_j) = P(\mathcal{L})$ .*

*Proof.*

$$\begin{aligned}
P(\tilde{\mathcal{L}}|K_j = \tilde{K}_j) &= P(\mathcal{L} \cup \tilde{K}_j | K_j = \tilde{K}_j) \\
&= P(\mathcal{L} | K_j = \tilde{K}_j) + P(\tilde{K}_j | K_j = \tilde{K}_j) - P(\mathcal{L} \cap \tilde{K}_j | K_j = \tilde{K}_j) \\
&= \sum_{k \in \text{supp}(K_j)} P(\mathcal{L} | K_j = k, \tilde{K}_j = k) \frac{P(K_j = k, \tilde{K}_j = k)}{\sum_{k \in \text{supp}(K_j)} P(K_j = k, \tilde{K}_j = k)} \\
&= \sum_{k \in \text{supp}(K_j)} P(\mathcal{L} | K_j = k, \tilde{K}_j = k) \frac{P(K_j = k) P(\tilde{K}_j = k)}{P(\tilde{K}_j = k) \sum_{k \in \text{supp}(K_j)} P(K_j = k)} \\
&= \sum_{k \in \text{supp}(K_j)} P(\mathcal{L} | K_j = k, \tilde{K}_j = k) P(K_j = k) \\
&= \sum_{k \in \text{supp}(K_j)} P(\mathcal{L} | K_j = k) P(K_j = k) \\
&= P(\mathcal{L}).
\end{aligned}$$

The second equality follows an application of the inclusion-exclusion principle. The third equality stems from the fact that  $P(\mathcal{L} \cap \tilde{K}_j | K_j = \tilde{K}_j) = P(\tilde{K}_j | K_j = \tilde{K}_j)$ . The fourth equality comes from the local Markov property in Equation (14), since  $\tilde{K}_j \notin \text{De}(K_j)$ . The fifth equality results from the fact that  $\sum_{k \in \text{supp}(K_j)} P(K_j = k) = 1$  and the assumption that the hypothetical variable  $\tilde{K}_j$  is uniformly distributed, such that  $P(\tilde{K}_j = k) = c$ , a constant. The sixth equality originates from the so-called consistency assumption guiding the definition of the hypothetical variable.  $\square$

Equivalence between the hypothetical and the empirical distributions implies that the causal effects defined in the hypothetical model can be identified using data generated by the empirical model. In particular, equivalence between the two models provides an opportunity to identify a number of causal effects in the model by replacing at least one disabled causal link originating from  $K_j$ , with a corresponding hypothetical  $\tilde{K}_j$ -input, and evaluating the impact of this hypothetical modification on the resulting conditional probabilities. Symbolically,  $\tilde{K}_j = k$  translates the causal operation of setting each  $\tilde{K}_j \in \tilde{\mathcal{L}}$  to a value of  $k$ . With this translation, the Markov factorization property, thus, captures the mental action of intervening in the system, in a way

that affects only the distribution of  $Ch(\tilde{K}_j)$  (Pearl, 2001, 2009; Heckman and Pinto, 2013).

$$P(\tilde{\mathcal{L}} \setminus \{\tilde{K}_j\} | \tilde{K}_j = k) = \prod_{K \in \tilde{\mathcal{L}} \setminus \{\tilde{K}_j\} \cup Ch(\tilde{K}_j)} P(K | Pa(K)) \prod_{K \in Ch(\tilde{K}_j)} P(K | Pa(K) \setminus \{K_j\}, \tilde{K}_j = k), \quad (17)$$

### 3.2.2 Identification

In establishing identification, we start with the total effect of income on fertility as defined in Equation (19), before the mediated effects through the cost of childbearing and contraceptive behaviour; from that point, we derive the direct effect of income as a residual.

Our previous discussion suggests that the move from interventional distributions to conditional distributions relies on testable implications, those described by marginal and conditional independencies implied by a given hypothetical model. Therefore, we present, for each causal effect, some testable implications, based on the application of the  $d$ -separation criteria to the relevant graphical hypothetical model. These marginal and conditional independencies are used to clarify the conditions, under which, the intervention distribution  $P(Q | \tilde{Y} = y)$  can be represented as a function of conditional distributions. We assume consistency, in the sense that if  $A$  is observed to attain a value  $a$ ,  $\tilde{A} = a$  and  $A$  are the same (Shpitser, 2013). Our analysis is illustrated in terms of discrete variables, where we measure the change in  $Q$  produced by a unit change in  $Y$ , say from  $Y = 0$  to  $Y = 1$ ; it extends to continuous variables by replacing sums with integrals (Imai et al., 2010).

Graphically, the capacity of a path to carry associations depends both on the orientation of its arrows and on the conditioning variables. Causal paths are sequences of adjacent arrows, all pointing away from the treatment toward the outcome; all other paths are considered noncausal. Consequently, not all associations between treatment and outcome in a causal model are causal. Some paths may transmit spurious associations. Assuming the necessary variables are observed, the challenge is to ensure that all spurious associations between treatment and outcome are stopped, without blocking any causal path between treatment and outcome. In other words, the identification of causal effects reduces to identifying all potentially spurious associations, and

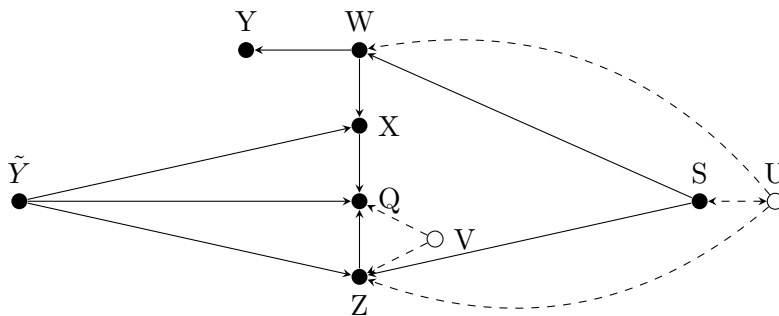


Figure 3: Hypothetical Model for the *Total Effect*. Solid bullets represent observed variables, hollow circles represent unobserved (or latent) variables.

choosing a conditioning set of variables able to block them.

Within a graphical causal model, there are three elementary sources of spurious association that need to be blocked; not doing so may lead to estimation bias (Elwert, 2013). The first results in *overcontrol bias*, which arises from too many intercepting paths, such that the intercepting variable manages to remove any causal effect that may exist between two variables of interest. The second is the presence of confounding factors, which graphically correspond to a divergent path. Failing to condition on a common cause creates *common cause confounding bias*. The third is the common effect, as described by a convergent path. Conditioning on a common effect, or a descendant of a common effect, produces an *endogenous selection bias*. With these concerns in mind, we turn to the causal effects of interest in the model.

### 3.2.3 Total Effect

The marginal and conditional independencies arise from an application of the  $d$ -separation criteria to the Total Effect hypothetical model in Figure 3.

**Lemma 3.1.** *In the hypothetical model for the Total Effect of  $Y$  on  $Q$ , (1)  $\tilde{Y} \perp\!\!\!\perp W$  and (2)  $Q \perp\!\!\!\perp Y|W$ .*

*Proof.* In Figure 3, each and every path connecting  $\tilde{Y}$  to  $W$  contains a collider; thus, the two variables are marginally independent according to part 1 of the  $d$ -separation criteria. The Variable  $Y$  is connected by a single arrow to the rest of the model through  $W$ ; therefore,  $Q \perp\!\!\!\perp Y|W$ .  $\square$

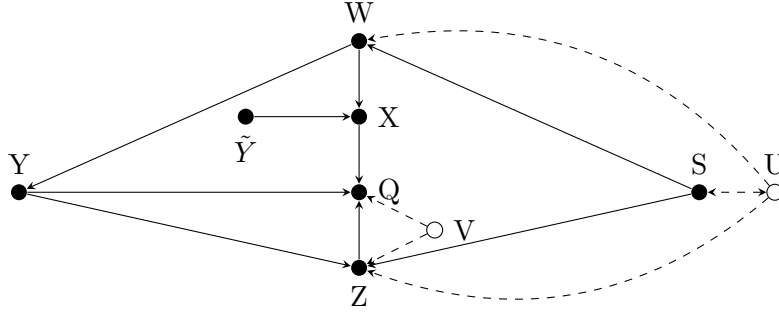


Figure 4: Hypothetical Model for the *Price Effect Fertility Model*. Solid bullets represent observed variables, hollow circles represent unobserved (or latent) variables.

Applying these results,

$$\begin{aligned}
 P(Q|\tilde{Y} = y) &= \sum_W P(Q|\tilde{Y} = y, W)P(W|\tilde{Y} = y) \\
 &= \sum_W P(Q|\tilde{Y} = y, Y = y, W)P(W) \\
 &= \sum_W P(Q|Y = y, W)P(W).
 \end{aligned} \tag{18}$$

The second equality stems from relationships (1)  $\tilde{Y} \perp\!\!\!\perp W$  and (2)  $Q \perp\!\!\!\perp Y|W$  of Lemma 3.1. The third equality comes from applying the consistency assumption. Thus, the *Total Effect* is given by

$$TE = \sum_W \left[ E(Q|Y = 1, W) - E(Q|Y = 0, W) \right] P(W). \tag{19}$$

### 3.2.4 Price Effect: PE

A similar process yields the price effect.

**Lemma 3.2.** *In the hypothetical model for the Price Effect of  $Y$  on  $Q$ , (1)  $\tilde{Y} \perp\!\!\!\perp Q|X, W$ , and (2)  $X \perp\!\!\!\perp Y|W$ .*

*Proof.* In Figure 4,  $\tilde{Y}$  and  $Q$  are  $d$ -separated by  $X, W$  because  $X$  is a collider on the path from  $\tilde{Y}$  to  $Q$  traversing  $W$ . Since  $Q$  is a collider, only information on  $W$  is needed in order to  $d$ -separate  $X$  from  $Y$ .  $\square$

Applying the results from Lemma 3.2 yields

$$\begin{aligned}
P(Q|\tilde{Y} = y) &= \sum_X \sum_W P(Q|\tilde{Y} = y, X, W)P(X|W, \tilde{Y} = y)P(W|\tilde{Y} = y) & (20) \\
&= \sum_X \sum_W P(Q|Y, X, W)P(X|\tilde{Y} = y, W)P(W|\tilde{Y} = y) \\
&= \sum_X \sum_W P(Q|Y, X, W)P(X|\tilde{Y} = y, W)P(W) \\
&= \sum_X \sum_W P(Q|Y, X, W)P(X|\tilde{Y} = y, Y = y, W)P(W) \\
&= \sum_X \sum_W P(Q|Y, X, W)P(X|Y = y, W)P(W).
\end{aligned}$$

In the derivations above, the second equality exploits the fact that conditioning on  $\{X, W\}$   $d$ -separates all non-causal paths from  $Y$  to  $Q$ . In this case, the remaining source of dependence left between the two is causal. The third equality comes from the fact that  $W$  causes  $Y$ , which implies that intervening on  $Y$  will not influence  $W$ . The fourth equality stems from the conditional independence in (2) –  $X \perp\!\!\!\perp Y|W$  – of Lemma 3.2. The fifth equality arises from the consistency assumption. Thus, we identify the *Price Effect*.

$$PE = \sum_X \sum_W E(Q|Y, X, W)P(W) \left[ P(X|Y = 1, W) - P(X|Y = 0, W) \right] \quad (21)$$

### 3.2.5 Contraception Effect: CE

The final causal effect, the contraception effect, is developed in the same fashion as the previous causal effects.

**Lemma 3.3.** *In the hypothetical model for the Contraception Effect of  $Y$  on  $Q$ , (1)  $\tilde{Y} \perp\!\!\!\perp W$ , and (2)  $Y \perp\!\!\!\perp Z|W$ .*

*Proof.* In Figure 5, the path from  $\tilde{Y}$  to  $W$  contains at least one collider implying that  $\tilde{Y} \perp\!\!\!\perp W$ . Furthermore, since  $Q$  is a collider, only information on  $W$  is needed to  $d$ -separate  $X$  from  $Y$ .  $\square$

Applying results from Lemma 3.3 yields:

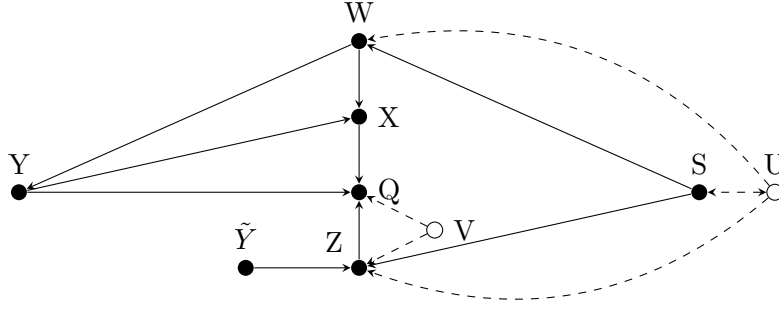


Figure 5: Hypothetical Model for the *Contraception Effect*. Solid bullets represent observed variables, hollow circles represent unobserved (or latent) variables.

$$\begin{aligned}
P(Q|\tilde{Y} = y) &= \sum_Z \sum_W P(Q|Z, W, \tilde{Y} = y)P(Z|W, \tilde{Y} = y)P(W|\tilde{Y} = y) & (22) \\
&= \sum_Z \sum_W P(Q|Z, W, Y)P(Z|W, \tilde{Y} = y)P(W|\tilde{Y} = y) \\
&= \sum_Z \sum_W P(Q|Z, W, Y)P(Z|W, \tilde{Y} = y)P(W) \\
&= \sum_Z \sum_W P(Q|Z, W, Y)P(Z|W, \tilde{Y} = y, Y = y)P(W) \\
&= \sum_Z \sum_W P(Q|Z, W, Y)P(Z|W, Y = y)P(W).
\end{aligned}$$

The second equality comes from applying the *back-door* criterion (Pearl, 1993), which implies that, when all spurious paths from  $Y$  to  $Q$  are disabled, leaving intact all other causal paths between the two, fixing the value of a covariate is the same as conditioning on it. The third equality comes from the relationship in (1) –  $\tilde{Y} \perp\!\!\!\perp W$  – of Lemma 3.3. The fourth equality stems from  $Y \perp\!\!\!\perp Z|W$ , described in Lemma 3.3, while the fifth equality results from the consistency assumption. Thus, the *contraception effect* in our fertility model is identified, and is a function of the observed data.

$$CE = \sum_Z \sum_W E(Q|Y, Z, W)P(W) \left[ P(Z|Y = 1, W) - P(Z|Y = 0, W) \right] \quad (23)$$



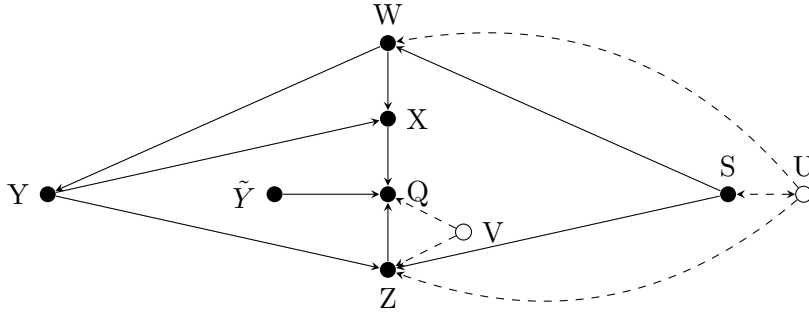


Figure 6: Hypothetical Model for the *Direct Effect*. Solid bullets represent observed variables, hollow circles represent unobserved (or latent) variables.

### 3.2.6 Direct Income Effect: DIE

The identification of the *Direct Income Effect* is relatively more involved, compared to the causal effects identified to this point. As illustrated in Figure 6, the direct impact of income on fertility is quantified by the level of fertility in a hypothetical situation, wherein we change an individual's level of income, while forcing the cost of childbearing and contraception to behave as if the latter did not change.

**Lemma 3.4.** *In the hypothetical model for the Direct Effect of  $Y$  on  $Q$ , (1)  $\tilde{Y} \perp\!\!\!\perp (X, V, Z)$ , and (2)  $Q \perp\!\!\!\perp Y | X, V, Z$ .*

*Proof.* In Figure 6, the flow between  $Y$  and  $Q$  is blocked by the vector of mediator variables  $\{X, Z\}$ , and unobserved fecundity  $V$ . Moreover, the variables  $X, W, Z$  and  $V$  are marginally independent from  $\tilde{Y}$  because the variable  $Q$  connecting  $\tilde{Y}$  to the rest of the model is a collider in all four directions. □

Using the results in Lemma 3.4 yields,

$$\begin{aligned}
P(Q|\tilde{Y} = y) &= \sum_{X,Z,V} P(Q|\tilde{Y} = y, X, Z, V)P(X, Z, V|\tilde{Y} = y) \\
&= \sum_{X,Z,V} P(Q|\tilde{Y} = y, X, Z, V)P(X, Z, V) \\
&= \sum_{X,Z,V} P(Q|\tilde{Y} = y, Y = y, X, Z, V)P(X, Z, V) \\
&= \sum_{X,Z,V} P(Q|Y = y, X, Z, V)P(X, Z, V) \\
&= \sum_{X,Z,V} P(Q|Y = y, X, Z, V)P(X|Y, W)P(Z|Y, S, U, V)P(V).
\end{aligned} \tag{24}$$

The second equality comes from the independence of fertility and income, conditional on wage rates, fecundity and contraception, as illustrated in relationship (1) –  $\tilde{Y} \perp\!\!\!\perp X, V, Z$  – of Lemma 3.4, while the third equality comes from relationship (2) –  $Q \perp\!\!\!\perp Y|X, V, Z$  – of the same lemma. The fourth equation comes from the assumption of consistency. The fifth equality results from the application of the Markov Factorization property, as defined in Equation (15).

Note that the last equality in Equation (24) cannot be identified using only observed data, since it contains unobservables such as skills,  $U$ , and fecundity,  $V$ . However, one can still identify the IE as a residual using the following pseudo-additive relation, which captures the decomposition of the total effect into price effect, contraception effect and direct income effect, where  $PE_r$  and  $CE_r$  are the  $PE$  and  $CE$  for the reversed transition, from  $Y = 1$  to  $Y = 0$  (Pearl, 2012b).

$$DIE = TE + PE_r + CE_r \tag{25}$$

Thus, manipulating expressions in Equations (23) and (21), and combining them with (19) yields the direct income effect.

$$\begin{aligned}
DIE = & \sum_W \left[ E(Q|Y = 1, W) - E(Q|Y = 0, W) \right] P(W) \\
& + \sum_X \sum_W E(Q|Y, X, W) P(W) \left[ P(X|Y = 0, W) - P(X|Y = 1, W) \right] \\
& + \sum_Z \sum_W E(Q|Y, Z, W) P(W) \left[ P(Z|Y = 0, W) - P(Z|Y = 1, W) \right] \quad (26)
\end{aligned}$$

It is also possible, in the particular case of continuous treatment, to identify directly the direct income effect using control variable methods, such as those discussed in Florens et al. (2008) and Imbens and Newey (2009). Assume, for simplicity, that a woman's heterogeneous skills,  $U$ , influence the efficiency of her contraception behaviour,  $Z$ , only through the level of education  $S$ . Then, under a choice function monotonic in  $V$ , the direct income effect could be identified as

$$\begin{aligned}
DIE = & \int \int \int \left[ E(Q|Y = y', X, Z, V) - \right. \\
& \left. E(Q|Y = y, X, Z, V) \right] dF_V(v) dF_{X|Y,W}(x) dF_{Z|Y,S,V}(z), \quad (27)
\end{aligned}$$

where  $dF_A(\cdot)$  and  $dF_{A|B}(\cdot)$  represent the distribution function of a random variable  $A$  and the conditional distribution function of  $A$  given  $B$ .

## 4 Estimation

The causal effects identified in the previous section can be empirically estimated using a multi-stage nonparametric procedure. In this study, we are interested in the average increase in fertility  $Q$  expected when transitioning from  $Y = y$  to  $Y = y'$ , and focus our attention on that average.

### 4.1 Total Effect

To estimate the total effect, we assume that the cost of childbearing,  $X$ , and the efficiency of the contraception method,  $Z$ , are allowed to track the change in income,  $Y$ . Thus, in the first step,

estimate the conditional expectation for every  $(y', w)$  and  $(y, w)$  cell.

$$E(Q|Y = y', w) = g_{y'}(w) \quad \text{and} \quad E(Q|Y = y, w) = g_y(w) \quad (28)$$

In the second step, estimate the expected value of  $g_{y'}(w)$  and  $g_y(w)$ , respectively, and take the difference.

$$TE_{y,y'}(Q) = E_W \left[ g_{y'}(w) \right] - E_W \left[ g_y(w) \right] \quad (29)$$

## 4.2 Price Effect

Turning to the price effect, we employ the thought experiment of holding  $Y$  constant, at  $Y = y$ , but changing  $X$  to the value it would have attained if  $Y = y'$ . This counterfactual definition of the price effect calls for a three-stage regression. First, according to Equation (21), estimate the conditional expectation for every  $(y, x, w)$  cell.

$$E(Q|Y, X, W) = g(y, x, w) \quad (30)$$

Second, sort the estimated conditional expectation with respect to  $\{Y, X\}$  so that  $g(y, x, w)$  is a function  $g_{y,x}(w)$  of  $W$ , and estimate the expected value of  $g_{y,x}(w)$ .

$$f(y, x) = E_W \left[ g_{y,x}(w) \right] \quad (31)$$

Third, fix  $y$  and regard  $f(y, x)$  as the function  $f_y(x)$  of  $X$ , then estimate the conditional expectation of  $f_y(x)$ , conditional on  $Y = y'$  and  $Y = y$ , respectively. Take the difference of those resulting estimates.

$$PE_{y,y'} = E_{X|Y,W} \left[ f_y(x) | y', w \right] - E_{X|Y,W} \left[ f_y(x) | y, w \right] \quad (32)$$

### 4.3 Contraception Effect

Similar to estimating the PE, the CE is estimated over a number of steps. First estimate the conditional expectation in Equation (23) for every  $(y, z, w)$  cell.

$$E(Q|Y, Z, W) = k(y, z, w) \quad (33)$$

Second, fix  $(y, z)$  so that  $k(y, z, w)$  is seen as a function  $k_{y,z}(w)$  of  $W$ . Subsequently, estimate the expected value of  $k_{y,z}(w)$ .

$$m(y, z) = E_w \left[ k_{y,z}(w) \right] \quad (34)$$

Third, sort  $m(y, z)$  with respect to  $y$  yielding  $m_y(z)$ , then estimate the expectation of  $m_y(x)$ , conditional on  $Y = y'$  and  $Y = y$ , respectively. Finally, take the difference.

$$CE_{y',y} = E_{Z|Y,W} \left[ m_y(z)|y', w \right] - E_{Z|Y,W} \left[ m_y(z)|y, w \right]. \quad (35)$$

### 4.4 Direct Income Effect

Now consider the estimated expected change in  $Q$  induced by varying income from  $y$  to  $y'$  with a constant cost of childrearing,  $X$ , but setting contraception efficiency,  $Z$ , to its values before the intervention. Conceptually, based on Equations (29), (32) and (35), the direct income effect (DIE) can be derived as a residual. Begin by estimating the reverse causal effects from  $y$  to  $y'$  for  $PE$  and  $CE$ , respectively, following the preceding discussion.

$$CE_{y',y} = E_{Z|Y,W} \left[ m_{y'}(z)|y, w \right] - E_{Z|Y,W} \left[ m_{y'}(z)|y', w \right], \quad (36)$$

$$PE_{y',y} = E_{X|Y,W} \left[ f_{y'}(x)|y, w \right] - E_{X|Y,W} \left[ f_{y'}(x)|y', w \right]. \quad (37)$$

With these terms in place, the residual can be calculated.

$$DIE = TE_{y,y'} + PE_{y',y} + CE_{y',y}, \quad (38)$$

## 5 Conclusion

This study has introduced an economic fertility model integrating both the utility maximization and the production functions of the household, based on the cost of childrearing and contraception behaviour, subject to unobserved fecundity (Heckman and Willis, 1976) and unmeasured cognitive and noncognitive abilities (Heckman et al., 2006). In particular, the analysis captured the link between labour market and fertility outcomes, through family planning and productive time allocation. This is formalized in the concept of a household earnings production function (Schultz, 1961; Ben-Porath, 1967; Mincer and Polachek, 1974), which varies with contraceptive efficiency, in line with the human capital theory of demand for health (Grossman, 1972, 2000; Becker, 2007). Thus, the model links access to reproductive health services to participation and/or performance in the labour market.

Using a counterfactual structural estimation framework (Pearl, 2009; Heckman and Pinto, 2013), we are able to nonparametrically identify a number of causal effects: the total effect causal of income on fertility, the direct causal effect of income on fertility, and the indirect causal effects mediated through the cost of children and contraception efficiency. Further, we outline the steps underpinning the nonparametric estimation of those effects. Within the model, the sign and magnitude of the total effect of income on fertility depends on the magnitude of the direct income effect, which is always positive, and the size of the two indirect effects, which are assumed to be negative. The first of these effects represents the standard assumption of goods normality, while the latter two are intuitively similar to pure substitution effects arising from price changes.

## References

- Ali, R. A., Richardson, T. S. and Spirtes, P. (2009), ‘Markov Equivalence for Ancestral Graphs’, *The Annals of Statistics* **37**(5B), 2808–2837.
- Ananat, E. O., Gruber, J., Levine, P. B. and Staiger, D. (2009), ‘Abortion and selection’, *The Review of Economics and Statistics* **91**(1), 124–136.
- Bailey, M. J. (2012), ‘Reexamining the Impact of Family Planning Programs on US Fertility: Evidence from the War on Poverty and the Early Years of Title X’, *American Economic Journal: Applied Economics* **4**(2), 62–97.
- Becker, G. S. (1960), An Economic Analysis of Fertility, in ‘Demographics and Economic Change in Developed Countries’, Princeton: National Bureau of Economic Research, pp. 209–231.
- Becker, G. S. (1965), ‘A Theory of the Allocation of Time’, *The Economic Journal* **75**(299), 493–517.
- Becker, G. S. (2007), ‘Health as Human Capital: Synthesis and Extensions’, *Oxford Economic Papers* **59**, 379–410.
- Becker, G. S. and Lewis, H. G. (1973), ‘On the Interaction between the Quantity and Quality of Children’, *Journal of Political Economy* **81**(2), S279–S288.
- Ben-Porath, Y. (1967), ‘The Production of Human Capital and the Life Cycle of Earnings’, *Journal of Political Economy* **75**(4), 352–365.
- Bloom, D. E., Canning, D., Fink, G. and Finlay, J. E. (2009), ‘Fertility, Female Labor Force Participation, and the Demographic Dividend’, *Journal of Economic Growth* **14**, 79–101.
- Bongaarts, J. (1993), ‘The Supply-Demand Framework for the Determinants of Fertility: An Alternative Implementation’, *Population Studies: A Journal of Demography* **47**(3), 437–456.
- Caldwell, J. C. (1976), ‘Toward the Restatement of Demographic Transition Theory’, *Population and Development Review* **2**(3/4), 321–366.
- Caldwell, J. C. (1979), ‘Education as a Factor in Mortality Decline: An Examination of Nigerian Data’, *Population Studies* **33**(3), 395–413.
- Caldwell, J. C., Orubuloye, I. and Caldwell, P. (1992), ‘Fertility Decline in Africa: A New Type of Transition?’, *Population and Development Review* **18**(2), 211–242.
- Carlsen, E., Giwercman, A., Keiding, N. and Skakkebaek, N. (1992), ‘Evidence for Decreasing Quality of Semen during past 50 Years’, *British Medical Journal* **305**(6854), 609–613.
- Cleland, J. and Wilson, C. (1987), ‘Demand Theories of the Fertility Transition: An Iconoclastic View’, *Population Studies* **41**, 5–30.
- Coale, A. J. (1984), ‘The Demographic Transition’, *The Pakistan Development Review* **23**(4), 531–552.

- Easterlin, R. A. (1975), 'An Economic Framework for Fertility Analysis', *Studies in Family Planning* **6**, 54–63.
- Easterlin, R. A., Wongboonsin, K. and Ahmed, M. A. (1988), 'The Demand for Family Planning: A New Approach', *Studies in Family Planning* **19**(5), 257–269.
- Elwert, F. (2013), Graphical Causal Models, in S. Morgan, ed., 'Handbook of Causal Analysis for Social Research', Springer.
- Fisher, F. M. (1970), 'A Correspondence Principle for Simultaneous Equation Models', *Econometrica* **38**, 73–92.
- Florens, J., Heckman, J., Meghir, C. and Vytlacil, E. (2008), 'Identification of Treatment Effects using Control Functions in Models with Continuous, Endogenous Treatment and Heterogeneous Effects', *Econometrica* **76**, 1191–1206.
- Frisch, R. (1930), *A Dynamic Approach to Economic Theory: Lectures by Ragnar Frish at Yale University*, Routledge. Published 2010.
- Frisch, R. (1938), Autonomy of Economic Relations (League of Nations Memorandum, 1938), in D. F. Hendry and M. S. Morgan, eds, 'The Foundations of Econometric Analysis', Cambridge University Press. Published 1995.
- Galor, O. and Zang, H. (1997), 'Fertility, Income Distribution, and Economic Growth: Theory and Cross-Country Evidence', *Japan and the World Economy* **9**(2), 197–229.
- Gini, C. (1926), 'Decline in the birth-rate and "fecundability" of woman', *The Eugenics Review* **17**(4), 258–274.
- Grossman, M. (1972), 'On the Concept of Health Capital and the Demand for Health', *Journal of Political Economy* **80**(2), 223–255.
- Grossman, M. (2000), The Human Capital Model, in A. Culyer and J. Newhouse, eds, 'Handbook of Health Economics', Vol. 1, Elsevier.
- Haavelmo, T. (1943), 'The Statistical Implications of a System of Simultaneous Equations', *Econometrica* **11**(1), 1–12.
- Haavelmo, T. (1944), 'The Probability Approach in Econometrics', *Econometrica* **12**, iii–vi+1–115. Supplement.
- Heckman, J. J. and Pinto, R. (2013), 'Causal Analysis After Haavelmo', *NBER Working Paper Series* **19453**, 1–51.
- Heckman, J. J., Stixrud, J. and Urzua, S. (2006), 'The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior', *Journal of Labor Economics* **24**(3), 411–482.
- Heckman, J. J. and Willis, R. J. (1976), Estimation of a stochastic model of reproduction: An econometric approach, in N. E. Terleckyj, ed., 'Household Production and Consumption', Columbia University Press (for NBER), pp. 99–146.



- Holland, P. W. (1986), ‘Statistics and Causal Inference’, *Journal of the American Statistical Association* **81**(396), 945–960.
- Hotz, V. J., Klerman, J. A. and Willis, R. J. (1997), The Economics of Fertility in Developed Countries, *in* M. Rosenzweig and O. Stark, eds, ‘Handbook of Population and Family Economics’, Vol. 1, Elsevier Science, chapter 7, pp. 275–347.
- Imai, K., Keele, L. and Yamamoto, T. (2010), ‘Identification, Inference and Sensitivity Analysis for Causal Mediation Effects’, *Statistical Science* **25**(1), 51–71.
- Imbens, G. W. and Newey, W. K. (2009), ‘Identification and Estimation of Triangular Simultaneous Equations Models without Additivity’, *Econometrica* **77**(5), 1481–1512.
- Johnson-Hanks, J. (2007), ‘Natural Intentions: Fertility Decline in the African Demographic and Health Surveys’, *American Journal of Sociology* **112**(4), 1008–1043.
- Jones, L. E., Schoonbroodt, A. and Tertilt, M. (2011), Fertility Theories: Can they Explain the Negative Fertility-Income Relationship?, *in* J. B. Shoven, ed., ‘Demography and Economy’, University of Chicago Press.
- Jones, L. E. and Tertilt, M. (2008), An Economic History of Fertility in the United States: 1826-1960, *in* P. Rupert, ed., ‘Frontiers of Family Economics’, Vol. 1, Emerald, pp. 165–230.
- Kalwij, A. (2010), ‘The Impact of Family Policy Expenditure on Fertility in Western Europe’, *Demography* **47**(2), 503–519.
- Kearney, M. S. and Levine, P. B. (2009), ‘Subsidised Contraception, Fertility, and Sexual Behavior’, *The Review of Economics and Statistics* **91**(1), 137–151.
- Leibenstein, H. (1975), ‘The Economic Theory of Fertility’, *The Quarterly Journal of Economics* **89**(1), 1–31.
- Leridon, H. (2007), ‘Studies of Fertility and Fecundity: Comparative Approaches from Demography and Epidemiology’, *Comptes Rendus Biologies* **330**, 339–346.
- Louis, G. M. B. (2011), Fecundity and Fertility, *in* G. M. B. Louis and R. W. Platt, eds, ‘Reproductive and Perinatal Epidemiology’, Oxford University Press.
- Martineau, P. (1958), ‘Social Classes and Spending Behavior’, *Journal of Marketing* **23**(2), 121–130.
- Michael, R. T. and Willis, R. J. (1976), Contraception and Fertility: Household Production under Uncertainty, *in* N. E. Terleckyj, ed., ‘Household Production and Uncertainty’, NBER, pp. 25–98.
- Mincer, J. and Polachek, S. (1974), ‘Family Investments in Human Capital: Earnings of Women’, *The Journal of Political Economy* **82**(2), 76–108.
- Moultrie, T. A., Sayi, T. S. and Timaeus, I. M. (2012), ‘Birth Intervals, Postponement, and Fertility Decline in Africa: A New Type of Transition?’, *Population Studies: A Journal of Demography* **66**(3), 241–258.

- Pearl, J. (1988), *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, San Mateo, CA.
- Pearl, J. (1993), ‘Comment: Graphical Models, Causality, and Intervention’, *Statistical Science* **8**, 266–269.
- Pearl, J. (1995), ‘Causal Diagrams for Empirical Research’, *Biometrika* **82**(4), 669–688.
- Pearl, J. (2001), Direct and Indirect Effects, in J. Breese and D. Koller, eds, ‘Uncertainty in Artificial Intelligence, Proceedings of the Seventeenth Conference’, Morgan Kaufmann, San Francisco, pp. 411–420.
- Pearl, J. (2009), *Causality: Models, Reasoning, and Inference*, Cambridge University Press.
- Pearl, J. (2012a), The Causal Foundations of Structural Equation Modeling, in R. H. Hoyle, ed., ‘Handbook of Structural Equation Modeling’, Guilford Press, New-York.
- Pearl, J. (2012b), ‘The Causal Mediation Formula: A Guide to the Assessment of Pathways and Mechanisms’, *Prevention Science* **13**(4), 426–436.
- Petersen, M. L., Sinisi, S. E. and van der Laan, M. J. (2006), ‘Estimation of Direct Causal Effects’, *Epidemiology* **17**(3), 276–284.
- Preston, S. H. and Hartnett, C. S. (2010), The Future of American Fertility, in J. B. Shoven, ed., ‘Demography and the Economy’, University of Chicago Press.
- Rubin, D. B. (1974), ‘Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies’, *Journal of Educational Psychology* **66**(5), 688–701.
- Schultz, T. P. (1986), ‘The Value and Allocation of Time in High-Income Countries: Implications for Fertility’, *Population and Development Review* **12**(S), 87–108.
- Schultz, T. W. (1961), ‘Investment in Human Capital’, *The American Economic Review* **51**(1), 1–17.
- Schultz, T. W. (1973), ‘The Value of Children: An Economic Perspective’, *Journal of Political Economy* **81**(2), S2–S13.
- Shpitser, I. (2013), ‘Counterfactual Graphical Models for Longitudinal Mediation Analysis with Unobserved Confounding’, *Cognitive Science* **37**, 1011–1035.
- Skakkebaek, N. E., Jorgensen, N., Main, K. M., Meyts, E. R.-D., Leffers, H., Andersson, A.-M., Juul, A., Carlsen, E., Mortensen, G. K., Jensen, T. K. and Toppari, J. (2006), ‘Is Human Fecundity Declining?’, *International Journal of Andrology* **29**, 2–11.
- Tietze, C. (1959), ‘Differential Fecundity and Effectiveness of Contraception’, *The Eugenics Review* **50**(4), 231–237.
- Timaeus, I. M. and Moultrie, T. A. (2008), ‘On Postponment and Birth Intervals’, *Population and Development Review* **34**(3), 483–510.

- VanderWeele, T. J. (2009), ‘Marginal Structural Models for the Estimation of Direct and Indirect Effects’, *Epidemiology* **20**, 18–26.
- Verma, T. and Pearl, J. (1991), Equivalence Model Synthesis of Causal Models, *in* P. Bonissone, M. Henrion, L. Kanal and J. Lemmer, eds, ‘Uncertainty in Artificial Intelligence, Proceedings of the Sixth Conference’, Elsevier.
- Yeakey, M. P., Muntifering, C. J., Ramachandran, D. V., Myint, Y., Creanga, A. A., and Tsui, A. O. (2009), ‘How Contraceptive Use Affects Birth Intervals: Results of a Literature Review’, *Studies in Family Planning* **40**(3), 205–214.