Available online at www.sciencedirect.com
ScienceDirect
Journal of Econometrics 143 (2008) 274–290
JOURNAL OF Econometrics
www.elsevier.com/locate/jeconom

Birth-spacing, fertility and neonatal mortality in India: Dynamics, frailty, and fecundity ☆

Sonia Bhalotra[a,*], Arthur van Soest[b]

[a]*Department of Economics, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, UK*
[b]*Tilburg University, The Netherlands*

Available online 4 November 2007

Abstract

Using microdata on 30,000 childbirths in India and dynamic panel data models, we analyse causal effects of birth-spacing on subsequent neonatal mortality and of mortality on subsequent birth intervals, controlling for unobserved heterogeneity. Right censoring is accounted for by jointly estimating a fertility equation, identified by using data on sterilization. We find evidence of frailty, fecundity, and causal effects in both directions. Birth intervals explain only a limited share of the correlation between neonatal mortality of successive children in a family. We predict that for every neonatal death, 0.37 additional children are born, of whom 0.30 survive.

JEL classification: I12; J13; C33

Keywords: Fertility; Birth-spacing; Neonatal mortality; Health; Dynamic panel data models; Siblings; India

1. Introduction

Interest in the determinants of child mortality has recently surged, with the inclusion of targets for child mortality amongst the Millennium Development Goals (Lawn et al., 2005; UNDP, 2003), and short birth-spacing and high fertility are widely regarded as among the most important causes of early childhood death. However, reproductive behaviour is endogenous to mortality and both are influenced by characteristics and choices of families, some of which are difficult to observe. For these reasons, there is limited evidence of the true causal associations of these variables.

In developing countries, 30% of deaths are of children under 5, compared to less than 1% in rich countries (Cutler et al., 2006). Almost half of child deaths are in the neonatal period, the first month of life, when the tie between mortality and fertility is closest (Cleland and Sathar, 1984). About 4 million neonates died in 2000, 99% of them in developing countries, and 27% in India. The proportion of neonatal in under-5 deaths has

☆ We are grateful to Wiji Arulampalam for helpful discussions, and to John Geweke, an Associate Editor and an anonymous referee for their suggestions. We also received helpful comments from Julie DaVanzo, Helene Turon, and Mike Veall. The paper has benefited from presentations made at the World Congress of the Econometric Society, the ESRC Econometric Study Group, the PAA meetings, RAND, Bristol, Tilburg and the Cambridge Panel Data conference.

*Corresponding author. Tel.: +44 117 9288418.
E-mail addresses: s.bhalotra@bristol.ac.uk, avas@uvt.nl (S. Bhalotra).

0304-4076/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2007.10.005

increased, since interventions like immunization, control of acute respiratory infection, or oral rehydration have had more of an effect on post-neonatal death (Lawn et al., 2005). It is thus important to focus attention on the causes of neonatal death.

Despite a long-standing interest of economists and demographers in the relation between childhood mortality and reproductive behaviour, the literature is scarce in a complete microdata analysis of all inter-relations of these variables (Wolpin, 1997). The main contribution of this paper is to use panel data based on retrospective fertility histories to estimate causal effects of birth interval length on subsequent neonatal mortality risk *and* of neonatal mortality on subsequent birth interval length, controlling for unobserved heterogeneity in both processes (referred to as *frailty* and *fecundity*, respectively). It also provides estimates of the effects of expected mortality (hoarding) and realized mortality (replacement) on fertility. Third, we model mortality dynamics within families, estimating both the extent to which observed persistence in death risk is explained by state-dependence, and the contribution of endogenously determined birth-spacing to state-dependence effects. Other contributions are methodological, relating to the way in which we deal with right-censoring of birth intervals and with the initial conditions problem that arises in dynamic models with unobserved heterogeneity.

Understanding the way in which biological and behavioural factors shape the family-level relation between reproductive behaviour and childhood mortality is crucial to understand the demographic transition that has historically preceded economic growth (Kalemli-Ozcan, 2002), and the endogenous processes by which societies evolve past the Malthusian spectre (Galor and Weil, 2000). Time-series analyses of historical data for today's industrialized countries suggest that declining mortality stimulated fertility decline (e.g. Ben-Porath, 1976; Eckstein et al., 1999), and a similar tendency can be seen in recent data for developing countries (e.g. Nyarko et al., 2003). Cross-sectional studies using household survey data have emphasized the reverse direction of causation, namely high fertility, associated with close birth-spacing or an early start to childbearing, causes an increase in childhood mortality (e.g. Cleland and Sathar, 1984).

In families with multiple children, there is a recursive bi-causal relation of these variables. The death of a child is often followed by a shorter interval to the next birth, which may be explained either by volitional replacement (e.g. Olsen, 1988) or by the fact that the mother stops breastfeeding, enabling her to conceive the next child sooner than otherwise (e.g. Chen et al., 1974). A short birth interval, in turn, increases the mortality risk of the *next* child in the family, possibly because the mother has not recuperated from the previous birth (e.g. DaVanzo and Pebley, 1993). Thus vulnerable families are caught in a *death trap*, creating persistence in death risk within families. This mechanism operates by the endogenous shortening of intervening birth intervals. Of course a birth interval is only observed if the mother has another birth, and this fertility decision is also influenced by whether her previous birth survived or not. While these relationships have each been studied, their interactions have rarely been studied jointly, and unobserved heterogeneity, another potential source of correlation of death risks within a family, is often ignored.

The analysis in this paper provides estimates, using survey data from India, of a dynamic panel data model that describes the complete process of child survival and birth-spacing (and thus fertility), allowing for endowment heterogeneity, input endogeneity, right-censoring and accounting for the initial conditions problem. We find evidence that childhood mortality risk is influenced by the pattern of childbearing, that is, by the timing and spacing of births, and that birth-spacing and fertility are, in turn, a function of realized mortality. We find a replacement effect of 0.37, in line with the few available estimates in the literature. The results suggest that the full impact of family planning interventions extends to reducing mortality and that mortality-reducing interventions like provision of piped water also affect birth-spacing and fertility. Our finding of causal effects of sibling mortality on both mortality and reproductive behaviour implies that interventions that reduce mortality or lengthen birth intervals will have multiplier effects.

The paper is organized as follows. Section 2 summarizes related research. Section 3 describes the data. The econometric model is presented in Section 4 and estimation and simulation results are reported in Section 5. Section 6 summarizes and concludes.

2. Related literature and contributions

Previous demographic research provides estimates of some of the main effects analysed in this paper, although not in a unified framework: for example, see Curtis et al. (1993), Madise and Diamond (1995),

Hobcraft et al. (1985), and Whitworth and Stephenson (2002) for analysis of the effects of birth-spacing on mortality, and Olsen (1988), Zenger (1993) or Frankenberg (1998) for analysis of the effects of mortality on birth-spacing. The limitation of these studies is that their estimates cannot be given a causal interpretation (see Moffitt, 2003).

There is limited previous research in economics in this area. Bhargava (2003) estimates a single-equation probit model of infant mortality in India, and argues that endogeneity of birth-spacing is taken care of by controlling for the survival status of older siblings, which is instrumented using household possessions and number of previous births. Maitra and Pal (2004) estimate a simultaneous hazards model of birth-spacing and child mortality, relying upon similarly strong identifying assumptions. Rosenzweig and Schultz (1983b) estimate a model of infant mortality in which birth-spacing is instrumented using household incomes and local prices. However, as discussed in Rosenzweig and Wolpin (1988, 1995), the implied exclusion restrictions typically do not hold. Rosenzweig and Wolpin (1988, 1995) instead use sibling differences to eliminate the mother-specific endowment. To allow for differences across siblings in frailty, they instrument inputs into the health of the index child using inputs relating to older siblings and parental characteristics. The econometric strategy in our paper is similar in that it relies upon natural information restrictions associated with the sequencing of births.

Estimates of the effects of childhood mortality on subsequent birth-spacing and fertility have mostly relied on the implausible assumption that parents have no influence on the survival chances of their offspring (Ben-Porath, 1976; Wolpin, 1997; Cigno, 1998), although there are exceptions (Olsen, 1980, 1988; Olsen and Wolpin, 1983). Our approach is different, in that we use a dynamic panel data framework and provide estimates of causal effects in both directions, between mortality and reproductive behaviour, and for both birth-spacing and total fertility. The recent demographic literature has highlighted the widespread phenomenon of sibling death clustering, emphasizing the role of unobserved heterogeneity, estimated using multi-level models that incorporate a random effect at the mother-level (Guo, 1993; Zenger, 1993; Curtis et al., 1993; Sastry, 1997; Whitworth and Stephenson, 2002). Arulampalam and Bhalotra (2006a, b) contribute to this by introducing state-dependence in mortality. They identify state-dependence effects in 13 of 15 Indian states using a single-equation model for mortality. Otherwise, there is little previous research on state-dependence effects in analysis of sibling data, although sibling correlations in outcomes have been widely studied (e.g. Solon et al., 1991).

3. Data and descriptive statistics

The data are from the second round of the National Family Health Survey of India (NFHS-II) which recorded complete fertility histories for ever-married women aged 15–49 in 1998–1999, including the time and incidence of child deaths.[1] Mothers constitute the cross-sectional dimension of the data. As mothers are observed repeatedly, in relation to every birth, birth-order creates the time dimension of the (unbalanced) panel. We use data for Uttar Pradesh (UP), the largest Indian state, which, in the year 2000, contained 17.1% of the country's population (approximately 165 million people). It has social and demographic indicators that put it well below the Indian average (Drèze and Sen, 1997). After dropping mothers with at least one multiple birth, the sample contains 29,747 live births of 7286 mothers, that occurred between 1963 and 1999.[2]

Strictly, neonatal death refers to death in the first 4 weeks of life. We include deaths up to 1 month to allow for age-heaping. The birth interval is the interval between reported dates of birth, rather than the inter-conception interval. As a result, measured birth intervals will be shorter on account of premature births (Gribble, 1993) and longer on account of miscarriage or stillbirth (Madise and Diamond, 1995). The first problem is dealt with by removing intervals shorter than 9 months; the second is harder to address. Ignoring miscarriage and stillbirth may lead to under-estimation of the mortality-raising effect of short birth intervals if women who have these problems also tend to produce weaker live births, since then falsely long intervals will be associated with higher mortality. However, this bias may be expected to be small once we control for mother-specific frailty and fecundity.

[1] For details on sampling strategy and context, see IIPS and ORC Macro (2000).

[2] Elimination of multiple births is in line with the demographic literature on mortality. Children of a multiple birth face hugely higher odds of dying, other things equal.

Means and standard deviations of all variables used in the analysis are in the appendix Table A1. The incidence of neonatal death over the sample period in UP was 7.39%, compared with an all-India average of 5.21%. Previous research on developing country data suggests that preceding birth intervals less than 24, and especially 18, months raise mortality risk. In our sample, 17.5% of birth intervals are shorter than 18 months, 18.3% are 18–23 months long. The mean number of births per mother is 4.04, the median is 4, and the maximum is 14. The mean age of mothers at first birth is 18.4, and the mean age of mothers at (any) birth is 22.2. As many as 28.3% of all live births are to mothers under 19 and 14.3% to mothers under 18.

Although contraceptive prevalence is increasing, contributing to the fertility decline witnessed in India since the mid-1980s, it seems to have had little impact on neonatal mortality (James et al., 2000). This is because contraception is used primarily to limit fertility rather than to control early childbearing and lengthen birth intervals. At the time of the survey, women were asked what their current contraceptive method was. In the state of UP, 65.8% were using no method, and 19.6% reported female sterilization, which is the predominant form of contraception, as in other parts of India. In Section 4, we will argue that information on sterilization helps identify the fertility equation.

Fig. 1 is a non-parametric regression of neonatal death on the log of the preceding birth interval. The curve declines monotonically. At short birth intervals, the probability of neonatal death is highest, and the gains

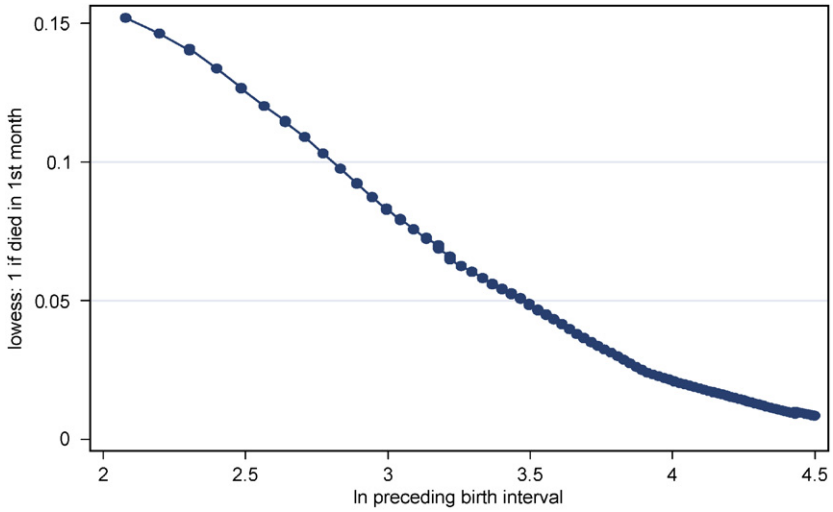

Fig. 1. Non-parametric (lowess) relation of (predicted) neonatal mortality and the preceding birth interval. *Notes*: The top 1% of observations were deleted.

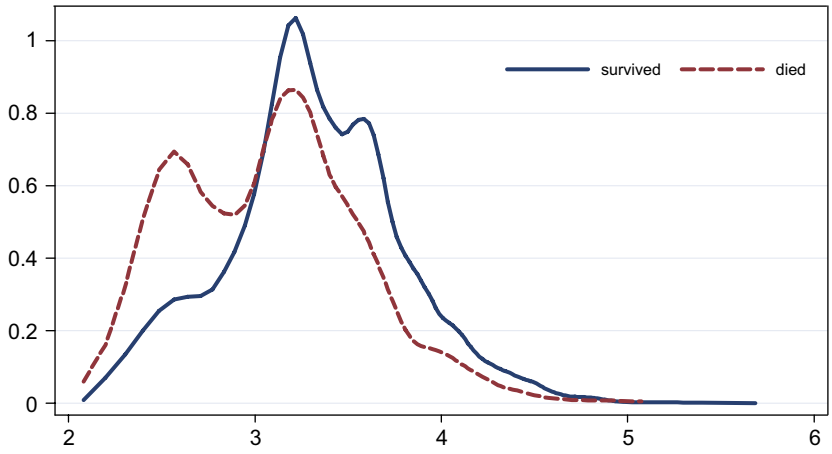

Fig. 2. Density of log birth interval by survival status of preceding sibling.

from an additional month's spacing are largest. Fig. 2 plots the kernel density function of the birth interval separately for whether or not the previous child in the family survived its first month. The birth interval distribution for the case where the preceding child has died clearly lies to the left of the other, with median birth intervals of 23 months after a neonatal death and 27 months otherwise (the means are 24.3 and 31.2 months). The raw data thus suggest the patterns that, as argued in Section 1, contribute to "*death traps*" at the family-level: short birth intervals raise subsequent mortality risk (Fig. 1), and mortality results in a shorter subsequent birth interval (Fig. 2).

In the sample of second- and higher-order children, the average probability of neonatal death is 6.28%. In the sub-sample in which the previous sibling survived, this probability is 5.20%, but amongst those whose previous sibling died, it is a remarkable 18.80%. Thus the death of a preceding sibling is associated with an increase in mortality risk of 13.6% points. This can be explained by both unobserved heterogeneity and genuine state-dependence, and state-dependence can, in turn, be explained by short birth-spacing or other mechanisms. Our analysis will disentangle these three explanations.

4. The model

The model has a recursive dynamic structure: the risk of neonatal mortality depends upon mortality of the previous sibling and on the preceding birth interval, while the birth interval depends upon survival of the preceding sibling. Similarly, the probability of continuing fertility depends on previous mortality and parity. Identification of the main causal effects rests on exploiting the natural sequencing of the birth and mortality processes, avoiding the need for exclusion restrictions. Amongst other covariates in the model are maternal age at birth of the child, and the year of birth of the child. Together with birth-order, these are endogenous because they depend upon the entire history of birth intervals and maternal age at first birth. The model accounts for this endogeneity. A limitation of our approach is that it does not readily extend to analysis of infant or under-5 mortality since, for e.g., infant mortality may occur after the next birth. As discussed in Section 1, this is not a strong restriction since neonatal mortality is particularly closely tied to reproductive behaviour.

The mortality equation can be regarded as a health production function in which the birth interval is an endogenous input, as in Rosenzweig and Schultz (1983a, b). The birth-spacing equation is an input equation, but it also describes an outcome that depends upon tastes and technology. These two equations are estimated jointly with an equation for continued fertility that accounts for right-censoring of birth intervals, and an equation for mortality risk of the first-born child that addresses the initial conditions problem.

The estimation allows for endowments (persistent mother-specific traits), unobservable by the econometrician but potentially known to the mother, and for the agency of the parent in influencing outcomes. The health endowment is referred to as *frailty*. We also incorporate inter-family unobserved heterogeneity in the birth-spacing and fertility equations (for convenience both of these terms are henceforth referred to as *fecundity*), and allow this to be correlated with frailty since, for e.g., women who are more careful about contraception may also better maintain the health of their children. Ignoring unobserved heterogeneity would bias estimates of the dynamics of each process (Heckman, 1981; Hyslop, 1999) and of the causal effect of each variable on the other (Alessie et al., 2004).

The econometric model is an extension of the univariate model of Heckman (1981) and is broadly similar to the bivariate model of Alessie et al. (2004). Our approach to dealing with right-censoring is new, exploiting data on sterilization. The way in which the initial conditions problem is addressed is also novel. To take account of the sampling design, we use random effects at the community (cluster) level. The model is estimated by simulated maximum likelihood.

Let n_i be the number of live births of mother i at the time of the survey. Let M_{ij} be an indicator variable with value 1 if child j in family i suffers neonatal death, and 0 otherwise. B_{ij} is the log of the length of the interval between the birth of child $j-1$ and child j in family i. Thus B_{ij} refers to the interval closed by the birth of child j. As it is the preceding birth interval for child j, it is, by definition, predetermined with respect to M_{ij}. The rest of this section describes the equations and the model in detail.

4.1. Neonatal mortality

For child j ($j = 2,\ldots,n_i$) in family i ($i = 1,2,\ldots, N$), the equation for neonatal mortality is

$$M^*_{ij} = g(\boldsymbol{x}_i, \boldsymbol{x}_{i1}, \boldsymbol{x}_{ij}, M_{i1}, \ldots, M_{i,j-1}, B_{i2}, \ldots, B_{ij}; \theta_{\mathrm{m}}) + \alpha_{\mathrm{m}i} + u_{\mathrm{m}ij};$$
$$M_{ij} = 1 \quad \text{if } M^*_{ij} > 0 \quad \text{and} \quad M_{ij} = 0 \quad \text{if } M^*_{ij} < 0. \tag{1}$$

To explain the assumptions needed for consistent estimation, it is initially written in a general form. $\alpha_{\mathrm{m}i}$ is mother-specific unobserved heterogeneity, reflecting the child's health endowment or "*frailty*", which may derive from genetic sources (e.g. maternal propensities to low birth weight and prematurity), environmental factors, or child-care behaviours (e.g. Sastry, 1997; Rosenzweig and Schultz, 1983a, b). As emphasized in Rosenzweig and Wolpin (1988), the fact that endogenous inputs like breastfeeding are not explicitly incorporated implies that the estimated mother-effect will also reflect inter-family heterogeneity in preferences. The error term $u_{\mathrm{m}ij}$ varies over mothers as well as children. It is revealed at the birth of child j and assumed not to influence parental inputs to child j in the one month of life during which parental choices can influence neonatal mortality risk. However, we allow $u_{\mathrm{m}ij-1}$ to influence parental inputs into child j through past mortality in the family, M_{ij-1}.

The vectors $\boldsymbol{x}_i$, $\boldsymbol{x}_{i1}$, and $\boldsymbol{x}_{ij}$ are exogenous explanatory variables, partitioned into variables that vary over children ($\boldsymbol{x}_{ij}$, $j = 2,\ldots,n$), are specific to the first child ($\boldsymbol{x}_{i1}$), or do not vary over children ($\boldsymbol{x}_i$). The vector of unknown parameters is denoted by θ_{m}. The variables $M_{i1},\ldots M_{i,j-1}$, $B_{i2},\ldots, B_{ij}$ are realized at or before the birth of child j.

We will specify g as a linear function of $\boldsymbol{x}_i$, $\boldsymbol{x}_{ij}$, $M_{i,j-1}$, B_{ij}, and include quadratic terms in the year of birth of the child, and in the age of the mother at birth of the index child, both of which are functions of $\boldsymbol{x}_{i1}$ and $B_{i2},\ldots, B_{ij}$.[3] Since the age of the mother at birth of child j depends upon her age at birth of child j-1 and the length of the intervening birth interval, B_{ij}, it is clear from recursivity of the model that maternal age at birth of j can be expressed as a function of maternal age at first birth (in $\boldsymbol{x}_{i1}$) and the history of birth intervals up until that date ($B_{i2},.. B_{ij}$). Thus, by allowing for endogeneity of birth intervals and conditioning on x_{i1}, we also allow for endogeneity of maternal age. Since our data include births that occurred across a span of about 30 years, a quadratic in the year of birth of the child is included to capture technological change in health production. This is also a function of the birth year of the first child and previous birth intervals.

We expect a negative effect of B_{ij} on M_{ij}, consistent with the hypotheses of maternal depletion (Section 1) and competition amongst closely spaced siblings (Cleland and Sathar, 1984). The effect of lagged mortality, $M_{i,j-1}$, on M_{ij} can be negative if learning or sibling-competition effects dominate, or positive if there is a strong role for factors like maternal depression. The first-order Markov assumption implicit in our specification of g is justified by the nature of the mechanisms driving state-dependence (that is, a causal effect of M_{ij-1} on M_{ij}): see Zenger (1993).

We assume that $\boldsymbol{x}_i$, $\boldsymbol{x}_{i1}$, and $\boldsymbol{x}_{ij}$ are independent of $\alpha_{\mathrm{m}i}$ and $u_{\mathrm{m}ij}$. Mean independence of ($\boldsymbol{x}_i$, $\boldsymbol{x}_{i1}$) and $\alpha_{\mathrm{m}i}$ is the usual assumption in a random effects model, needed for identification; the conditional mean of $\alpha_{\mathrm{m}i}$ given $\boldsymbol{x}_i$ and $\boldsymbol{x}_{i1}$ is subsumed in g. In $\boldsymbol{x}_i$, we include variables reflecting education levels of the mother and father, and caste and religion dummies. In $\boldsymbol{x}_{i1}$ we also include calendar year and age of mother at first birth.

A potential drawback of random effects models as compared with fixed effects models is the assumption that the "*time-varying*" (in this context, varying across siblings) regressors $\boldsymbol{x}_{ij}$ are assumed to be independent of the individual effects $\alpha_{\mathrm{m}i}$. In our case, however, the only variables included in $\boldsymbol{x}_{ij}$ are child gender and birth-order, which will be uncorrelated with mother-level frailty, so the independence assumption seems plausible.

4.2. Birth-spacing

The log length of the birth interval is modelled in a similar way as mortality:

$$B_{ij} = h(\boldsymbol{x}_i, \boldsymbol{x}_{i1}, \boldsymbol{x}_{i,j-1}, M_{i1}, \ldots M_{i,j-1}, B_{i2}, \ldots, B_{ij-1};\ \theta_{\mathrm{b}}) + \alpha_{\mathrm{b}i} + u_{\mathrm{b}ij}. \tag{2}$$

[3] Interactions and squares of other terms gave no significant improvement.

The family-specific effect, α_{bi}, is referred to as "*fecundity*" though it will include not only biological fecundity but also, e.g., variation in preferences for family planning or desired fertility. A causal effect of mortality of child $j-1$ on the birth interval to child j is allowed through $M_{i,j-1}$. We include $\boldsymbol{x}_{i,j-1}$ since gender of child $(j-1)$ may affect the interval to the birth of child j. The function h is specified as a linear combination of $\boldsymbol{x}_i$, $\boldsymbol{x}_{i,j-1}$, $M_{i,j-1}$, and the calendar year and age of the mother at the time of the birth of child $j-1$ and their squares, as in Section 4.1. Biomedical and demographic research give no argument for a causal effect of B_{ij-1} on B_{ij}, conditional on α_{bi}, so we do not allow for this.[4] Assumptions on family-specific effects and errors u_{bij} are similar to those for Eq. (1). We assume that $\boldsymbol{x}_i$, $\boldsymbol{x}_{i1}$, and $\boldsymbol{x}_{ij}$ are independent of α_{bi} and u_{bij} and that u_{bij} is independent of the past.

We allow for correlation between the unobserved heterogeneity terms α_{bi} and α_{mi} in Eqs. (1) and (2). This allows an alternative, non-causal explanation for the correlation between birth interval lengths and mortality in the raw data and for the potential endogeneity of the preceding birth interval in Eq. (1), which may be correlated with frailty, α_{mi}. For example, parents with weak endowments may choose shorter birth intervals to meet their target number of children in a given time. Similarly, our model allows M_{ij-1} in Eq. (2) to be correlated with family-level fecundity, α_{bi}.

The distribution of the family effects $(\alpha_{mi}, \alpha_{bi})$ is assumed to be bivariate normal with mean zero, variances σ_m^2, σ_b^2, and covariance $\sigma_m\sigma_b\rho_\alpha$. The child-specific error terms u_{mij} and u_{bij} are assumed to be independent of α_{mi} and α_{fi} and normally distributed with mean zero. Without loss of generality, the variance of u_{mij} is set to 1.

4.3. Right-censoring

Inclusion of the birth-spacing equation (Eq. (2)), in the model demands a correction for right-censoring because some mothers will not have completed their fertility at the time of the survey. To account for this, we model the probability that mother i will have another child after the birth of child j, as follows:

$$F_{ij}^* = f(\boldsymbol{x}_i, \boldsymbol{x}_{i1}, \boldsymbol{x}_{ij}, M_{i1}, \ldots M_{i,j-1}, B_{i2}, \ldots, B_{ij};\ \theta_f) + \alpha_{fi} + u_{fij};$$
$$F_{ij} = 1 \quad \text{if } F_{ij}^* > 0 \quad \text{and} \quad F_{ij} = 0 \quad \text{if } F_{ij}^* < 0. \tag{3}$$

We specify f as a linear combination of $\boldsymbol{x}_i$, the calendar year and age of the mother at the time of the birth of child $j-1$ and their squares (functions of $\boldsymbol{x}_{i1}$ and $B_{i2}, \ldots, B_{ij-1}$), dummies for the presence of boys and girls in the family that did not suffer neonatal death, and the total numbers of boys and girls in the family who survived the neonatal period (functions of j, $M_{i1}, \ldots M_{i,j}$, and $B_{i2}, \ldots, B_{ij-1}$). The variables are gender-specific to allow for son-preference, of which there is considerable evidence for UP (e.g. Drèze and Gazdar, 1997). Endogeneity of the sibship variables is taken care of in the same way as in the other equations—they are functions of lagged dependent variables. Moreover, confounding unobserved factors are controlled for by allowing arbitrary correlations of α_{fi} with α_{mi} and α_{bi}, assuming joint trivariate normality with arbitrary covariance matrix and independence of exogenous variables. We make similar assumptions on u_{fij} as on the other error terms: normality, independence of individual effects and error terms for other birth-orders or other equations, and independence of exogenous variables.

The data contain information on whether a mother is sterilized at the time of the survey, which helps to estimate the parameters of the model more efficiently. For sterilized mothers (19.6% of the sample), the complete birth process is observed. Of the remaining mothers, some will have another child after the survey date, and others will not. Sterilization is an incomplete indicator of whether the mother will have another child; it is an dependent variable in our model but not a variable of interest as such, as it is the decision to have another child or not that is modelled here. To identify Eq. (3) with data on sterilization, we assume that women who have decided to have no more children get sterilized with a fixed probability κ (a nuisance parameter).

To be precise: If mother i has more than j children, we know she has given birth to another child after child j, and the likelihood will incorporate the probability that $F_{ij} = 1$. If the mother reports that she has had exactly j children and was sterilized after the birth of the jth child, then the likelihood will incorporate the probability

[4] Heckman et al. (1985) show, for a sample of (married) Swedish mothers, that there is no state-dependence in the birth spacing process once controls for unobserved heterogeneity are introduced.

that $F_{ij} = 0$ and the probability κ. If, at the time of the survey, the mother had j children but was not (yet) sterilized, then it is unclear whether child j is the last child or not—the birth interval after the birth of child j may extend beyond the time of the survey. The probability that this will happen, given that there will be another birth and given unobserved heterogeneity components, follows from (2) and is given by $\Phi([T-\{h(\boldsymbol{x}_i, \boldsymbol{x}_{i1}, \boldsymbol{x}_{i,j-1}, M_{i1}, \ldots M_{i,j-1}, B_{i2}, \ldots, B_{ij-1}; \theta_b)+\alpha_{bi}\}]/\sigma)$, where T is the length of the time interval elapsed between the birth of child j and the time of the survey, and σ is the standard deviation of the error term in (2). In this case, the likelihood (conditional on unobserved heterogeneity) contains a factor that accounts for the fact that we do not observe whether or not there will be another birth after birth j.[5]

The usual approach to right-censoring is to assume that the same process continues but is not observed after the time of the survey (e.g. Wooldridge, 2002, Chapter 20). This approach does not work well here since the fertility process is necessarily finite and ended well before the time of the survey for many women in the sample.[6] In the absence of information on sterilization, an alternative would be to assume that fertility stops at a given age for all mothers, or to estimate equation (3) without the sterilization information. In the latter case, the fertility equation would only be indirectly identified and the estimates are likely to be much less precise.

4.4. The initial conditions problem

"Lagged" mortality, M_{ij-1}, is endogenous in equation (1) by virtue of being correlated with frailty, α_{mi}. This creates the initial conditions problem common in this type of model (e.g. Heckman, 1981). This is addressed by formulating a separate equation for the mortality risk of the first-born child of every mother:

$$M_{i1}^* = g_1(\boldsymbol{x}_i, \boldsymbol{x}_{i1};\ \theta_{m,1}) + \lambda_m\alpha_{mi} + \lambda_b\alpha_{bi} + \lambda_f\alpha_{fi} + u_{mi1},$$
$$M_{i1} = 1 \quad \text{if } M_{i1}^* > 0 \quad \text{and} \quad M_{i1} = 0 \quad \text{if } M_{i1}^* < 0. \tag{4}$$

In most applications of this type of models (e.g. Hyslop, 1999) the true process is ongoing and the first observation is generated in the same way as later observations. Heckman et al. (1985) is an exception. They model birth-spacing and observe the process from its natural start, the start of menarche. Here, similarly, we observe the birth and mortality processes from their beginning for each mother, and the first child is a genuine starting point of that process. This makes Heckman's approach quite natural compared to, for example, the alternative approach of Wooldridge (2000).

We will work with a linear specification of g_1, in line with the specification of (1). It is likely that M_{i1} will be correlated with α_{mi} and we also allow it to be correlated with the other family-specific effects α_{bi} and α_{fi}. The error term u_{mi1} is assumed to be standard normal and independent of the other error terms in the model, of the individual effects, and of the exogenous regressors $\boldsymbol{x}_{ij}$ and $\boldsymbol{x}_i$. θ_{m1}, λ_m, λ_b and λ_f are auxiliary parameters. Eq. (4) is a flexible function of the exogenous variables. We do not impose restrictions on the relation of the parameters in (4) to those in (1).

4.5. Community effects and estimation

The data are collected in 333 geographical clusters ("communities") with, on average, 24.4 mothers per cluster. To allow for the possibility that mothers (and children) within a cluster share unobservable traits (for example, sanitation or social norms), we need to include a cluster-level term in the equation error.[7] As the large number of clusters makes it infeasible to use cluster dummies, we incorporate random cluster effects in Eqs. (1)–(3) in the same way as the mother-specific effects, with similar assumptions.[8] A linear combination of the cluster effects in (1)–(3) is added to Eq. (4), with three additional auxiliary parameters as coefficients. For identification, it is assumed that the cluster effects are independent of mother-specific effects.

[5] An appendix with likelihood details is available at http://www.efm.bris.ac.uk/www/ecsrb/bhalotra.htm

[6] Initial experimentation with our data showed that the usual procedure produces a poor fit, being unable to explain why so many women suddenly completely stop having children.

[7] The data have some information on community characteristics at the time of the survey. We did not use this since it may not reflect community characteristics at the time of birth.

[8] That is, trivariate normal with arbitrary covariance structure to be estimated, independent of exogenous variables and error terms.

The complete model can be estimated by maximum likelihood, including the nuisance parameters of the initial conditions equation, and the fertility equation.[9] Conditional on the random effects, the likelihood contribution of a given mother can be written as a product of univariate normal probabilities and densities over all births of a mother, and the likelihood for a given cluster can be written as the product over all mothers in that cluster. The actual likelihood contribution is the expected value of the conditional likelihood contribution, taking the expectation over all (unobserved) random effects (three in the model without cluster effects, six in the model with cluster effects). This is a three or six-dimensional integral, which could in principle be approximated numerically using, for example, the Gauss–Hermite quadrature.

In this paper, we instead use (smooth) simulated ML, drawing multivariate errors from $N(0,I_3)$, transformed into draws of the random effects using the parameters of the random effects distribution. The conditional likelihood contribution is averaged over R independent draws. If $R\to\infty$ with the number of clusters, this gives a consistent estimator; if draws are independent across households and $R/\sqrt{N}\to\infty$, the estimator is asymptotically equivalent to exact ML (see, e.g. Hajivassiliou and Ruud, 1994). We use Halton draws, which give more accurate results for given R than independent random draws (Train, 2003). The results we present use $R = 100$. Using $R = 50$ gives very similar results.

5. Results

This section presents the results of the complete "benchmark" model (Table 1).

5.1. Neonatal mortality

The left hand panel of Table 1 reports the estimates of the equation for neonatal mortality. In the discussion, we focus on the implied marginal effects for the second child, assuming that the first child survived the first month of life, and setting all family characteristics to benchmark values (boy, Hindu, not of a backward caste, maternal and paternal education zero average date of birth (1986.8), age of mother at birth (21.3 years), and previous log birth interval (3.32)). The estimated probability of neonatal mortality for this benchmark child is 4.84%.

A 10% increase in the length of the preceding birth interval reduces the probability of death by about 0.45 percentage-points in the benchmark case (and the marginal effect is similar for higher birth-orders). Some studies have found that the deleterious effects of short birth intervals are enhanced if the previous sibling has survived (e.g. Zenger, 1993; Cleland and Sathar, 1984). We therefore included an interaction of lagged mortality with the birth interval. This was insignificant—a result that contrasts with Whitworth and Stephenson (2002). Our results suggest that maternal depletion is more important than sibling-competition in explaining the mortality-increasing effects of short birth intervals. Maternal depletion will be especially pronounced amongst poor women who need longer to replenish stocks of nutrients like calcium and iron that are needed to support a healthy pregnancy.

Neonatal mortality of the previous sibling makes neonatal death significantly more likely for the index child, even with the birth interval held constant. For the benchmark second child, the estimated difference is 4.16 percentage-points. Similar effects are found for the third and later children.[10] This suggests that, for neonatal mortality, learning effects (a mother is better able to avoid a further child death once she has experienced one) or reduced competition for scarce resources are dominated by state-dependence mechanisms that create a positive association of sibling deaths and do not operate via birth-spacing. We hypothesize that the loss of a child may create psychological effects that the mother may not have recovered from by the time she conceives her next child, as a result of which there may be physiological effects that make this child more vulnerable both in the womb and after birth. Several studies report negative effects of depression on pregnancy outcomes (Steer et al., 1992) and early childhood mortality (Chung et al., 2004; Drewett et al.,

[9] An explicit specification of the likelihood function is in the online appendix referred to in footnote 5.

[10] A similar positive effect of lagged mortality on current mortality is reported in Whitworth and Stephenson (2002), who do not interpret this finding. They use earlier data for all India and control for unobserved heterogeneity and birth intervals, but take birth intervals to be exogenous.

Table 1
Model parameter estimates

	Neonatal mortality		Log birth interval		Prob (further birth)	
	Parameter	s.e.	Parameter	s.e.	Parameter	s.e.
Lagged mortality	0.320*	0.068	−0.237*	0.017		
Log birth interval	−0.447*	0.050				
Religion						
Muslim	−0.197*	0.065	−0.076*	0.014	0.379*	0.045
Other	−0.043	0.338	−0.098	0.070	−0.404	0.242
Caste						
Scheduled caste	0.098	0.136	0.025	0.034	0.101	0.088
Scheduled tribe	0.100	0.054	0.001	0.013	−0.044	0.039
Other backward caste	−0.061	0.052	−0.005	0.012	−0.158*	0.038
Caste missing	0.063	0.099	0.001	0.024	−0.038	0.074
Maternal education						
Incomplete primary	−0.029	0.094	0.011	0.024	−0.115	0.070
Complete primary	−0.197*	0.093	0.034	0.021	−0.136*	0.059
Incomplete secondary	−0.093	0.099	0.013	0.024	−0.266*	0.067
Secondary and higher	−0.297*	0.144	0.034	0.022	−0.544*	0.062
Paternal education						
Incomplete primary	−0.004	0.083	0.013	0.020	0.116*	0.057
Complete primary	−0.109	0.076	0.002	0.015	−0.035	0.049
Incomplete secondary	−0.103	0.059	0.004	0.015	−0.098*	0.042
Complete secondary	−0.134*	0.067	−0.001	0.016	−0.198*	0.048
Higher than secondary	−0.031	0.066	0.010	0.016	−0.228*	0.047
Gender						
Female	−0.043	0.038	−0.028*	0.008		
Birth year of child						
Year of birth of child/10	−0.028	0.62	0.415*	0.121	14.462*	0.850
(year/10) squared	−0.008	0.036	−0.026*	0.007	−0.895*	0.051
Maternal age						
Maternal age at birth/10	−0.817*	0.341	0.308*	0.072	−1.195*	0.190
(age/10) squared	0.143*	0.063	−0.054*	0.015	0.126*	0.033
Child birth-order						
Birth-order	0.044	0.049	−0.025*	0.008		
Square of birth-order	−0.001	0.004	0.003*	0.001		
Surviving children						
1 if no boys					0.206*	0.049
1 if no girls					0.147*	0.041
Number of boys					−0.249*	0.025
Number of girls					−0.072*	0.020
Constant	1.633	2.706	1.402*	0.509	−53.872*	3.420
Sigma error			0.454*	0.002		

Notes: s.e. denotes standard error. *The parameter is significant at the two-sided 5% level.

2004). Rahman et al. (2004) find that maternal depression in the prenatal and postnatal periods is a risk factor for malnutrition and illness in infants in Pakistan. Mental health is increasingly recognized as an important health problem that is often neglected in poor countries (WHO, 2001). Overall, depression is a plausible causal mechanism, but further research is merited to investigate this and other potential pathways.

Conditional on the other covariates, gender and birth-order are insignificant. Over the period 1963–1999, neonatal mortality exhibits a trend reduction of 0.16 percentage-points per year. It is U-shaped in mother's

age at birth, a familiar pattern in developing country data. The minimum is at about 29 years of age. On average, mothers are much younger than this when giving birth to their second child (21.3 years). Mortality risk is decreasing in both maternal and paternal education, with larger effects of maternal education. We find no significant differences between castes. A striking result, that deserves further investigation, is that children of Muslim families (who, on average, have higher fertility, shorter birth-spacing and lower socio-economic status) are significantly less likely to die in the first month than Hindu children, with an estimated difference of about 1.7%-points.

Estimates of the "reduced form" probit equation for mortality of *first-born* children (Eq. (4)) are available in the online appendix (see footnote 5). The female dummy is now negative and significant at the two-sided 10% level, consistent with the fact that girls are born with a survival advantage, and with research that shows that discrimination against girls is smallest for first-borns (DasGupta, 1990). Other effects are broadly similar to those obtained for second- and higher-order births.

5.2. Birth-spacing

Estimates of the birth-spacing equation are in the second panel of Table 1. Since the dependent variable is in logs, the interpretation of the parameters is in terms of percentage changes in the expected length of the birth interval. Note that all covariates in this model refer to the preceding child (i.e. the child born at the start of the birth interval). Neonatal death of the previous child reduces the subsequent birth interval by about 21%, consistent with replacement behaviour (e.g. Ben-Porath, 1976). Feeding this into equation (1), we find that the effect of M_{ij-1} operating via B_{ij} results in an increase in M_{ij} of about 1.06%-points. Since the direct effect of M_{ij-1} on B_{ij} in Eq. (1) is 4.16%-points, total state-dependence increases the risk of death by 5.2%-points. Thus genuine state-dependence accounts for 37% of the clustering of sibling deaths ("raw persistence"), which was 14%-points on average (Section 2). The residual 63% is (observed and unobserved) heterogeneity.

The gender of the last-born child is significant and its sign is consistent with son-preference. If the last birth was a girl, the expected birth interval is about 3% shorter than if it was a boy. There is a significant hump-shaped trend in birth-spacing, with a maximum in 1980. That birth intervals got shorter in the last 20 years might be because rising nutritional standards for mothers allow them to support shorter intervals. Alternatively, the rise in women's labour force participation may have encouraged clustering of fertility. Birth-spacing is hump-shaped in maternal age, with a maximum at about 29 years. For the average mother, birth intervals increase until the sixth child is born. Parental education and caste have no significant effect. Birth intervals of Muslim families are 7.6% shorter than or similar Hindu families. Other things equal, birth-order exhibits a non-monotonic pattern, with the shortest birth intervals preceding the birth of the fourth child.[11]

5.3. Fertility equation

Table 3 presents estimates of the probability of having another child after each birth. Of particular interest are the family composition variables, which indicate son-preference. The probability of continued fertility is decreasing in the number of surviving children, but more than three times as rapidly in the number of surviving boys than in the number of surviving girls. Similar results have been reported for other countries in Asia and North Africa (e.g. Rahman and DaVanzo, 1993; Nyarko et al., 2003).[12]

Fertility is hump-shaped in time, with a maximum at about 1981. The quadratic in mother's age is decreasing over the age range, until age 47. Fertility falls with the level of education of both mother and father, with mother's education having larger effects. Muslims show a higher tendency to continue fertility. Mothers in backward castes other than scheduled castes and tribes have lower fertility than others, not in the raw data, but after conditioning upon covariates.

[11] It is well-known that breastfeeding lengthens birth intervals (Habicht et al., 1985), and its effects are, in this study, subsumed in the birth interval effect. Direct investigation of the role of breastfeeding is difficult because of incomplete data and since breastfeeding is a choice variable.

[12] Angrist and Evans (1998) find no such asymmetry for the US.

Table 2
Unobserved heterogeneity: mother plus community level effects

	Mortality	Birth interval	Fertility
Covariance matrix			
Mortality	0.173		
Birth interval	−0.000	0.025	
Fertility	0.003	−0.049	0.181
Correlation matrix			
Mortality	1.000		
Birth interval	−0.004	1.000	
Fertility	0.015	−0.725	1.000

Notes: See Section 5.4 of the text.

5.4. Unobserved heterogeneity

We find that the covariance structure of the mother and community-specific effects is sensitive to specification choices (like the number of draws in simulated ML), but the covariance structure of the *sum* of these terms is not. We therefore focus on the latter (see Table 2). There is significant evidence of both effects in the mortality and birth interval equations, but only community-specific effects are significant in the fertility equation.

Overall, the heterogeneity terms are statistically significant but small compared to the idiosyncratic errors. Compared to the idiosyncratic noise term (with variance 1), the two heterogeneity terms in the mortality equation capture about 15% of the total unsystematic variation in M_{ij}^* $(0.173/(1+0.173))$. More than half of this is heterogeneity across communities. In the birth interval equation, the idiosyncratic noise term has estimated variance 0.206, and total heterogeneity is about 11% of the total unsystematic variation. The estimated covariance between the total unobserved heterogeneity terms in the birth-spacing and mortality equations is virtually zero, implying a correlation coefficient of −0.004.

The heterogeneity terms in the fertility equation explain about 15% of the unsystematic variation in F_{ij}^*, but this estimate is not very accurate. We find a large negative correlation between the total heterogeneity terms in the fertility and birth interval equations of −0.73. This suggests that mothers who desire many children tend to use shorter birth intervals to achieve this, other (observed) explanatory variables constant. This is consistent with replacement behaviour in, for example, the target fertility model (see Wolpin, 1997). On the other hand, the small correlations between unobserved heterogeneity in the mortality equation and both the birth interval and the fertility equations suggest that hoarding does not play much of a role: there is hardly any evidence that mothers who perceive their children to have relatively high mortality risk react *ex ante* by having persistently shorter birth intervals.

5.5. Robustness of the state-dependence effect and other specification checks

A challenging finding is the strong positive effect of lagged mortality on mortality in Eq. (1), even after unobserved heterogeneity and the length of the preceding birth interval are controlled for. There are three potential sources of positive correlation between lagged mortality and current mortality: unobserved heterogeneity, the causal mechanism operating through the birth interval, and other sources of state-dependence. We would probably overestimate the importance of the last one if we underestimated the importance of the first two because of model misspecification. We therefore investigated model extensions generalizing either the specification of unobserved heterogeneity, or the mechanism through which birth intervals affect mortality.

In the first category, we considered auto-correlated error terms, which can be seen as unobserved heterogeneity that changes gradually over time, and a more general distribution of unobserved heterogeneity. In the second category, we considered correlation between errors in birth interval and mortality equations (which may be due to measurement error in reported birth intervals), other functional forms of the

relationship between birth intervals and neonatal mortality, and interactions of the birth interval with socio-demographics. Finally, we also looked at interactions between lagged mortality and socio-demographics in order to see whether the positive effect of lagged mortality can be attributed to specific groups. The changes in the estimates are discussed in detail in the online appendix (see footnote 5). The upshot is that the coefficient on lagged mortality is robust to all these specification checks.

We also investigated robustness of the main coefficients of interest (those on the lagged endogenous variables) to several simplifications of the model. The results are generally as expected. For example, failing to control for unobserved heterogeneity or not controlling for the birth interval leads to overestimation of state-dependence in Eq. (1). See the online appendix for details.

5.6. Simulations

Table 3 shows the results of simulations performed with the benchmark model, to investigate the effects of neonatal mortality on birth intervals and fertility, and the importance of state-dependence and hoarding. Column 1 is the benchmark simulation where all mechanisms at work in the estimated model are active (see the online appendix for details). Other columns present percentage deviations from the benchmark for scenarios in which some behavioural or non-behavioural mechanisms are "switched off."

Switching off the effect of previous mortality on the birth interval (col. 2) increases average birth interval length by 1.7%. Since longer birth intervals depress subsequent mortality, neonatal mortality falls by almost 3.2%. The longer birth intervals also imply that women have their children later and, through the negative effect of mother's age on the decision to have another child, this reduces the number of births by 0.72%. Due to the fall in mortality, the reduction in the number of surviving children is smaller (0.47%).

The next simulation (col. 3) shows what happens if mortality affects neither birth intervals, nor the probability of having another child. Mainly because of longer birth intervals, mortality falls by 4.4%. Comparing these results with the benchmark, the total size of the *replacement* effect can be estimated: total simulated mortality is 7.08% of births. Births as a result of replacement are about 2.6% of all births, that is, 0.37 births for every neonatal death. Because replacement increases mortality, the replacement effect on the number of surviving children is smaller, about 0.30 surviving replacement children for every death. These estimates, although they refer to neonates only, are in line with existing estimates, which lie between 0.2 and 0.5 (Schultz, 1997, pp. 384–385). In particular, Olsen (1988) finds a replacement effect of 0.35 using Malaysian data, and shows that replacement is greatest for children who die soon after birth.

In addition to suppressing the effects of mortality on birth interval and fertility, the next simulation (col. 4) also suppresses the effect of previous mortality on index child mortality (M_{ij-1} on M_{ij}), so that *all* effects of previous mortality are eliminated. Mainly because of suppressing the replacement effects, the average length of

Table 3
Simulations

	(1)	(2)	(3)	(4)	(5)
Neonatal mortality (%)	7.404	−3.17	−4.40	−8.92	−10.86
Birth interval (months)	30.586	1.70	1.66	1.66	2.63
Number of births (fertility)	4.125	−0.72	−2.59	−2.59	−3.34
Number of survivors	3.819	−0.47	−2.25	−1.90	−2.50

Notes: Column 1 presents sample averages of the simulated outcomes for the benchmark model. Columns 2–5 show percentage deviations from the benchmark that arise when selected mechanisms are "switched off" as follows:
Column 2: no effect of mortality on birth interval.
Column 3: no effect of mortality on birth interval or probability of having another child.
Column 4: no effect of mortality on birth interval, probability of having another child, or next child's mortality.
Column 5: no effect of mortality on birth interval, probability of having another child, or next child's mortality; no hoarding (i.e. unobserved heterogeneity in the mortality equation is not correlated with unobserved heterogeneity in the equations for birth spacing and fertility).

birth intervals increases by 0.50 months and the total number of children born falls by 2.6%. Neonatal mortality falls by about 0.66%-points, because of suppressing state-dependence and the longer birth intervals. Since the negative fertility effect dominates the positive mortality effect on surviving children, the total number of surviving children falls.

In column 5, we further eliminate *hoarding*. High frailty mothers are now not expected, *ex ante*, to have shorter birth intervals or higher probabilities of having another child. Thus all behavioural and non-behavioural relations between mortality and birth-spacing/fertility are eliminated. The total reduction in neonatal mortality compared to the benchmark is 10.9%, most of which is due to eliminating state-dependence. On average, birth intervals are 2.6% longer, mainly because the replacement effect is eliminated. The total number of children born falls by 3.3%; the number surviving falls by less, because of the reduced mortality probability.

6. Summary and conclusions

The main findings are as follows. The difference in neonatal mortality according to whether or not the preceding sibling died in the neonatal period is 14%-points. This is enormous, given that the average risk of neonatal death in our sample is 7.4%. We estimate that genuine state-dependence accounts for 37% of this, the remaining 63% being explained by inter-family heterogeneity. The analysis confirms that endogenously determined birth-spacing is a mechanism generating state-dependence in mortality, but it explains only about a fourth of total state-dependence. Identification of the mechanisms driving the remaining state-dependence is an important avenue for further research, not addressed in the demographic literature. We suggest maternal depression as a possibility.

We find direct evidence of replacement behaviour: a child death results in a shortening of the interval to the next birth, and also increases the probability of a next birth. Our model simulations imply that, accounting for direct and indirect effects, 37 in 100 children who die during the neonatal period are replaced by new births. Of these, about 30 survive. There is no evidence that frailty is correlated with fecundity. This suggests that couples do not practice hoarding, i.e., we find no evidence that women who know that their children are at relatively large risk of neonatal death anticipate this by reducing the length of their birth intervals *ex ante*.

As a measure of the importance of allowing for the joint determination of death risk and reproductive behaviour, we estimated the effects on the main outcomes of eliminating all behavioural and non-behavioural relations between the mortality process and the birth interval and fertility processes. The predicted reduction in neonatal mortality is 10.9%, most of which is due to eliminating state-dependence. On average, birth intervals are 2.6% longer, mainly because the replacement effect is eliminated. The total number of children born falls by 3.3%, while the total number of children surviving the neonatal period falls by 2.5% (because of the reduced mortality probability).

Our estimates of fertility behaviour are consistent with son-preference. These are probably the first estimates of son-preference in fertility that allow for endogeneity of mortality. We find that mortality falls with maternal age for most of range of the sample, indicating benefits to interventions that delay first birth (and lengthen birth intervals). The literature is scarce in estimates of maternal age effects on mortality that account for its endogeneity. Maternal education decreases both mortality and fertility, but has no effect on birth-spacing. Paternal education depresses the probability of another birth but has no significant effect on the other endogenous variables. There are strong religion effects: Muslims exhibit higher fertility, shorter birth intervals and, yet, lower mortality. In contrast, fixed effects associated with caste are weak. Conditional upon all covariates, we estimated a trend reduction in mortality of 0.16%-points per annum during 1963–1999, which is about 3.3% of the benchmark probability. Fertility decline in India seems to have set in from 1981. Despite this, birth intervals have got shorter since about then.

Future work could extend the framework to analyse infant or under-5 mortality. This creates the additional complication that mortality events and births can take place in overlapping time periods, requiring a different modelling approach. These results are for one Indian state, albeit a state with the largest population (166 million in 2001) and the highest neonatal death rate in India. Extension of the analysis to consider other Indian states or other developing countries will lend important insight into the mechanisms driving the key relationships analysed here. Many countries in sub-Saharan Africa have higher rates of neonatal mortality and

persistently higher fertility than in India. It would be interesting to investigate this ''demographic trap''. And this is easily done. The Indian survey used in this paper is one of about 70 Demographic and Health Surveys conducted across the developing world (see www.measuredhs.com). The methods used in this paper are therefore immediately applicable to a vast array of countries with different profiles of the structural processes.

Acknowledgments

Sonia Bhalotra acknowledges an ESRC research grant for work on sibling death clustering. Part of this work was carried out when Arthur van Soest was working at RAND and was funded through a RAND IR&B grant.

Appendix

Pleas see Table A1.

Table A1
Variable definitions and summary statistics

Variable	Mean	Std. Dev.	Min	Max
Neonatal mortality	0.069		0.0	1.0
Lagged neonatal mortality	0.060		0.0	1.0
Log birth interval*	3.306	0.486	2.1	5.7
Hindu	0.824		0.0	1.0
Muslim	0.168		0.0	1.0
Other religions	0.007		0.0	1.0
Not backward caste	0.454		0.0	1.0
Scheduled caste	0.196		0.0	1.0
Scheduled tribe	0.022		0.0	1.0
Other backward caste	0.276		0.0	1.0
Mother has no education	0.753		0.0	1.0
Ma has incomplete primary	0.045		0.0	1.0
Ma has completed primary	0.075		0.0	1.0
Ma has incomplete secondary	0.061		0.0	1.0
Ma has secondary or higher	0.064		0.0	1.0
Father has no education	0.334		0.0	1.0
Pa has incomplete primary	0.068		0.0	1.0
Pa has completed primary	0.110		0.0	1.0
Pa has incomplete secondary	0.195		0.0	1.0
Pa has completed secondary	0.125		0.0	1.0
Pa has higher than secondary	0.164		0.0	1.0
Female	0.475		0.0	1.0
Year of birth of child*	86.992	7.394	630	99.0
Maternal age at birth*	23.224	5.539	12.0	47.0
Birth-order*	3.179	2.051	1.0	14.0
Dummy no surviving boys	0.122		0.0	1.0
Dummy no surviving girls	0.188		0.0	1.0
Number of surviving boys*	1.962	1.386	0.0	8.0
Number of surviving girls*	1.782	1.461	0.0	10.0

Notes: All variables other than those with a * are dummies. Lagged mortality refers to the mortality status of the preceding sibling. Italics indicate reference category omitted in the regressions.

References

Alessie, R., Hochguertel, S., van Soest, A., 2004. Ownership of stocks and mutual funds: a panel data analysis. Review of Economics and Statistics 86, 783–796.

Angrist, J.D., Evans, B., 1998. Children and their parents' labor supply: evidence from exogenous variation in family size. American Economic Review 88, 450–477.

Arulampalam, W., Bhalotra, S., 2006a. Sibling death clustering in India: genuine scarring vs unobserved heterogeneity. Journal of the Royal Statistical Society, Series A 169 (4), 829–848.

Arulampalan, W., Bhalotra, S., 2006b. Persistence in infant mortality: evidence for the Indian states. Discussion Paper 2488, IZA, Bonn.

Ben-Porath, Y., 1976. Fertility response to child mortality: micro data from Israel. Journal of Political Economy 84, S163–S178.

Bhargava, A., 2003. Family planning, gender differences and infant mortality: evidence from Uttar Pradesh, India. Journal of Econometrics 112, 225–240.

Chen, L., Ahmed, S., Gesche, M., Mosley, W., 1974. A prospective study of birth interval dynamics in rural Bangladesh. Population Studies 28, 277–297.

Chung, E.K., McCollum, K.F., Elo, I.T., Lee, H.J., Culhane, J.F., 2004. Maternal depressive symptoms and infant health practices among low-income women. Pediatrics 113 (6), 523–529.

Cigno, A., 1998. Fertility decisions when infant survival is endogenous. Journal of Population Economics 11, 21–28.

Cleland, J., Sathar, Z.A., 1984. The effect of birth spacing on childhood mortality in Pakistan. Population Studies 38, 401–418.

Curtis, S.L., Diamond, I., McDonald, J.W., 1993. Birth interval and family effects on postneonatal mortality in Brazil. Demography 33, 33–43.

Cutler, D., Deaton, A., Lleras-Muney, A., 2006. The determinants of mortality. Journal of Economic Perspectives 20 (3), 97–120.

DasGupta, M., 1990. Death clustering, mothers education and the determinants of child-mortality in rural Punjab, India. Population Studies 44, 489–505.

DaVanzo, J., Pebley, A.R., 1993. Maternal depletion and child survival in Guatemala and Malaysia. Labor and Population Program Working Paper 93-18, RAND.

Drewett, R., Blair, P., Emmett, P., Emond, A., ALSPAC Study Team, 2004. Failure to thrive in the term and preterm infants of mothers depressed in the postnatal period: a population-based birth cohort study. Journal of Child Psychology and Psychiatry 45 (2), 359–366.

Drèze, J., Gazdar, H., 1997. Uttar Pradesh: the burden of inertia. In: Drèze, J., Sen, A. (Eds.), Indian Development: Selected Regional Perspectives. Oxford University Press, Oxford, UK, pp. 33–128.

Drèze, J., Sen, A., 1997. Indian Indian Development: Selected Regional Perspectives. Clarendon Press, Oxford, UK.

Eckstein, Z., Mira, P., Wolpin, K., 1999. A quantitative analysis of Swedish fertility dynamics, 1751–1990. Review of Economic Dynamics 2, 137–165.

Frankenberg, E., 1998. The relationship between infant and child mortality and subsequent fertility in Indonesia, 1971–1991. In: Montgomery, M., Cohen, B. (Eds.), From Death to Birth: Mortality Decline and Reproductive Change, National Research Council. National Academy Press, Washington, DC, pp. 254–315.

Galor, O., Weil, D., 2000. Population, technology, and growth: from Malthusian stagnation to the demographic transition and beyond. American Economic Review 90, 806–828.

Gribble, J.N., 1993. Birth intervals, gestational age and low birth weight: are the relationships confounded? Population Studies 47, 133–146.

Guo, G., 1993. Use of sibling data to estimate family mortality effects in Guatemala. Demography 30, 15–32.

Habicht, J.-P., DaVanzo, J., Butz, W.P., Meyers, L., 1985. The contraceptive role of breastfeeding. Population Studies 39, 213–232.

Hajivassiliou, V., Ruud, P., 1994. Classical estimation methods for LDV models using simulation. In: Engle, R., McFadden, D. (Eds.), Handbook of Econometrics, vol. IV. North-Holland, New York, pp. 2384–2443.

Heckman, J.J., 1981. The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. In: Manski, C.F., McFadden, D.L. (Eds.), Structural Analysis of Discrete Data with Econometric Applications. MIT Press, London, pp. 179–195.

Heckman, J., Hotz, V.J., Walker, J., 1985. New evidence on the timing and spacing of births. American Economic Review 75, 179–184.

Hobcraft, J.N., McDonald, J.W., Rutstein, S.O., 1985. Demographic determinants of infant and early child mortality: a comparative analysis. Population Studies 38, 193–223.

Hyslop, D.R., 1999. State dependence, serial correlation and heterogeneity in inter-temporal labor force participation of married women. Econometrica 67, 1255–1294.

IIPS and ORC Macro, 2000. National Family Health Survey (NFHS-2) 1998-9: India. International Institute for Population Sciences (IIPS), Mumbai.

James, K.S., Aitken, I., Subramanian, S.V., 2000. Neonatal mortality in India: Emerging paradoxes. Harvard Center for Population and Development Studies Working Paper Series, vol. 10, 13pp.

Kalemli-Ozcan, S., 2002. Does mortality decline promote economic growth? Journal of Economic Growth 7, 411–439.

Lawn, J.E., Cousens, S., Zupan, J., 2005. Neonatal survival 1–4 million neonatal deaths: when? where? why? Lancet 365, 891–900.

Madise, N.J., Diamond, I., 1995. Determinants of infant mortality in Malawi: an analysis to control for death clustering within families. Journal of Biosocial Science 27, 95–106.

Maitra, P., Pal, S., 2004. Birth spacing and child survival: comparative evidence from India and Pakistan. Working paper, Monash University, Melbourne.

Moffitt, R., 2003. Causal analysis in population research: an economist's perspective. Population and Development Review 29, 448–458.
Nyarko, P., Madise, N.J., Diamond, I., 2003. Child loss and fertility behaviour in Ghana. Social Statistics Research Centre Working Paper, A03/08, University of Southampton.
Olsen, R., 1980. Estimating the effect of child mortality on the number of births. Demography 17, 429–443.
Olsen, R., 1988. Cross-sectional methods for estimating the replacement of infant deaths. Research in Population Economics 6, 111–136.
Olsen, R., Wolpin, K., 1983. The impact of exogenous child mortality on fertility: a waiting time regression with exogenous regressors. Econometrica 51, 731–749.
Rahman, M., DaVanzo, J., 1993. Gender preference and birth spacing in Matlab, Bangladesh. Demography 30, 315–332.
Rahman, A., Iqbal, Z., Bunn, J., Lovel, H., Harrington, R., 2004. Impact of maternal depression on infant nutritional status and illness. Archives of General Psychiatry 61, 946–952.
Rosenzweig, M., Schultz, T.P., 1983a. Consumer demand and household production: the relationship between fertility and child mortality. American Economic Review 73, 38–42.
Rosenzweig, M., Schultz, T.P., 1983b. Estimating a household production function: heterogeneity, the demand for health inputs, and their effects on birth weight. Journal of Political Economy 91, 723–746.
Rosenzweig, M., Wolpin, K., 1988. Heterogeneity, intrafamily distribution and child health. Journal of Human Resources 23, 437–461.
Rosenzweig, M., Wolpin, K., 1995. Sisters, siblings and mothers: the effect of teenage childbearing on birth outcomes in a dynamic family context. Econometrica 63, 303–326.
Sastry, N., 1997. A nested frailty model for survival data, with an application to the study of child survival in Northeast Brazil. Journal of the American Statistical Association 92, 426–435.
Schultz, T.P., 1997. Demand for children in low income countries. In: Rosenzweig, M., Stark, O. (Eds.), Handbook of Population and Family Economics, vol. 1A. North-Holland, Amsterdam, pp. 349–432.
Solon, G., Corcoran, M., Gordon, R., Laren, D., 1991. A longitudinal analysis of sibling correlations in economic-status. Journal of Human Resources 26, 509–534.
Steer, R.A., Scholl, T.O., Hediger, M.L., Fischer, R.L., 1992. Self-reported depression and negative pregnancy outcomes. Journal of Clinical Epidemiology 45, 1093–1099.
Train, K., 2003. Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge, UK.
UNDP, 2003. Human Development Report: Millennium Development Goals: A Compact Among Nations to End Human Poverty. UNDP, Geneva.
WHO, 2001. The World Health Report. Mental health: New Understanding, New Hope. World Health Organization, Geneva.
Whitworth, A., Stephenson, R., 2002. Birth spacing, sibling rivalry and child mortality in India. Social Science & Medicine 55, 2107–2119.
Wolpin, K., 1997. Determinants and consequences of the mortality and health of infants and children. In: Rosenzweig, M., Stark, O. (Eds.), Handbook of Population and Family Economics, vol. 1A. North-Holland, Amsterdam, pp. 483–558.
Wooldridge, J., 2000. A framework for estimating dynamic, unobserved effects panel data models with possible feedback to future explanatory variables. Economics Letters 6, 245–250.
Wooldridge, J., 2002. Econometric Analysis of Cross-Section and Panel Data. MIT Press, Cambridge, MA.
Zenger, E., 1993. Siblings' neonatal mortality risks and birth spacing in Bangladesh. Demography 30, 477–488.