Child Mortality in Rural India

Bas van der Klaauw*
Limin Wang**

September 2004

Abstract
This paper focuses on infant and child mortality in rural areas of India. We use a flexible duration model framework, which allows for frailty at multiple levels and interactions between the child's age and individual, socio-economic and environmental characteristics. The model is estimated using the 1998/1999 wave of the Indian National Family and Health Survey. The estimation results show that socio-economic and environmental characteristics have significantly different impacts on mortality rates at different ages. These are particularly important immediately after birth. We use the estimated model for policy experiments. These indicate that child mortality can be reduced substantially, particularly by improving the education of women and reducing indoor air pollution caused by dirty cooking fuels. In addition, providing access to electricity and sanitation facilities can reduce under-five-years mortality rates significantly.

Keywords: duration analysis, mortality rates, heterogeneity, policy experiments.

We gratefully acknowledge the financial support from Swedish International Development Cooperation Agency (SIDA), TF024884. We thank Kirk Hamilton and Stephen Howes for their support and valuable suggestions. The findings, interpretations and conclusions expressed in this paper are entirely those of the authors. They do not represent the views of the World Bank, its Executive Directors, or the countries they represent.

* Free University Amsterdam, Tinbergen Institute and CEPR.
Address: Department of Economics, Free University Amsterdam, De Boelelaan 1105, NL--1081 HV Amsterdam, The Netherlands.
** World Bank.
Address: The World Bank, 1818 H Street, NW Washington DC 20433, USA.

1. Introduction

Reducing child mortality should be placed on the top policy priority in India. In 1999, while the total under-five child deaths among less developed countries amounted to about 10 million, India alone accounted for about one-fifth of the under-five deaths (2.1 million), the highest number within a single country (Cleason, Bos and Pathmanathan, 1999). Not only in absolute numbers is child death high in India, but also child mortality rates are substantially higher than in other low-income countries. In 1999 the under-five-years mortality rate (U5MR) in India was 90 (per 1,000 live births), while the U5MR for China, Brazil and Indonesia in the same year was 37, 40, and 51, respectively.

Within India, there are large variations in child mortality rates across states as well as between urban and rural areas within individual states. The U5MR in 1995-99 (estimated using the 1998/99 National Family Health Survey, NFHS) ranged from 18 in Kerala to 137 in Madhya Pradesh. During the same period the U5MR in rural India was 111, which was nearly twice as high as in urban areas (65 per 1,000 live births).

Not only is the initial child mortality rate high, but also the decline in child mortality in India has been slow. During the 1990s, the infant mortality rate decreased from 80 to 70. Several states with substantially higher child mortality in the early 1990s, such as Rajasthan, Madhya Pradesh, Uttar Pradesh and Bihar, saw little decline in mortality at the end of the 1990s.

There exists a large body of empirical studies that focus on analyzing the determinants of child mortality in India.[1] The key findings often indicate that household income, female education, access to health services, and immunization programs are important determinants of child mortality. Several studies also identify a strong interrelationship between mortality, fertility and gender bias. This evidence indicates that public policies emphasizing on improving access to health services, and in particular on improving female literacy and socio-economic participation, can play an important role in reducing child mortality. Several recent studies have also shown that environmental conditions, such as access to safe water and sanitation facilities, access to electricity, and use of clean cooking fuels, have an important health impact among young children (Pandey, Cheo, Luther, Sahu and Chand, 1998; Hughes, Lvovsky and Dunleavy, 2001; Jalan and Ravallion, 2003; and Mishra and Retherford, 1997).

In this paper we use a duration model framework to analyze the determinants of child mortality, focusing particularly on household environmental conditions. There exist several theoretical frameworks in the field of child mortality analysis (see for a survey Wolpin, 1997), but duration models have several advantages over the alternatives. First, duration models can directly deal with such data problems as right-censoring, which is a major concern in analyzing child mortality risks. Second, child mortality risk often depends on many characteristics of the child, the household and the village, and the impact of these covariates on child mortality is likely to vary with the age of children. For

[1] Pandey, Cheo, Luther, Sahu and Chand (1998), Rosenzweig and Schultz (1982), Murthi, Guio and Dreze (1995) and Gleason (2003).

example, Fikree, Il Azma and Berendes (2002) show that factors that affect neonatal deaths (in the first month before birth) are quite different from these for post-neonatal deaths. However, most empirical studies assume that these covariates have constant impacts on child mortality over the age of the child (e.g. Hughes, Lvovsky and Dunleavy, 2001; Guo and Rodriguez, 1992; Sastry, 1997; Ridder and Tunali, 1999). In contrast to the two studies on environmental health in India (Huges, Lvovsky and Dunleavy, 2001; and Jalan and Ravallion, 2003), we use a duration model that incorporates age-varying covariate effects, i.e. it allows the impact of socio-economic and environmental characteristics to change with children's age (see Van der Klaauw and Wang, 2004; for an extensive description). This modification is important as it allows us to identify the relevant health inputs at different ages of a child. The empirical results can thus provide critical information for quantifying the effectiveness of policy options for reducing child mortality through policy simulation.

We estimate the model using data from the 1998/99 wave of the Indian National Family and Health Survey (NFHS). We use the estimated survival functions to perform simulations of policy experiments. In particular, we simulate how interventions that improve socio-economic and environmental characteristics of a family affect the survival probabilities of children at different ages. We examine the impacts of both single and combined policy interventions. Also we use the flexibility of our model framework to show at which age the interventions are most effective.

The paper is organized as follows. Section 2 summarizes the 1998/99 NFHS. Section 3 describes the statistical model. Estimation results are discussed in Section 4. In Section 5 we use the estimated model for performing policy simulations. Section 6 concludes.

2. Data

In this section we briefly summarize the data used in the empirical analyses. The data are from the second wave of the National Family and Health Survey (NFHS-2), undertaken in 1998 and 1999. The NFHS is conducted using the same survey design as the Demographic and Health Surveys (DHS) for more than 60 low-income countries since 1985. It provides information on fertility, mortality, health issues, socio-economic and environmental conditions. The NFHS covers a nationally representative sample in 26 states.[2] The response rate of the interviews was over 95%. The survey consists of 3 separate questionnaires, a village questionnaire, a household questionnaire, and a woman questionnaire covering all ever-married women in a household between age 15 and age 49.

The data set contains information on 92,486 households. Since we are interested in child mortality in the rural areas, we abstract all 61,800 households living in rural areas. As is often the case with data on child mortality, information on child mortality comes from

[2] In Tripura due to local problems the data collection was delayed.

surveys among women (e.g. Guo and Rodriguez, 1992).[3] The households living in the rural areas include 62,248 ever-married women between 15 and 49 years old who are interviewed. In the rural areas, more than 95% of the women between ages 25 and 30 have been married. As we are interested in recent birth histories, we restrict the sample to women below age 32 at the moment of the interview, who gave birth to at least 1 live born child since January 1993.[4] This reduces the sample to 25,196 women, who gave in total birth to 43,185 children born alive since January 1993. The data show that in rural areas around 92.6% of the pregnancies result in a live birth. In urban areas this percentage is somewhat lower, mainly due to more abortions. The percentage of stillbirths is higher in rural areas.

The survey collects the birth history of all qualified women. . For each live born child the month of birth is recorded and whether or not the child is still alive at the moment of the interview. If a child died during the observation period, the age at which the child died is asked. The age of death is observed within intervals, in case a child died within a month after birth, the age of death is recorded in days, if the child died between 1 month and 2 years, it is recorded in months, and otherwise it is recorded in years. Because we are only interested in child mortality until age 5, we artificially right-censor at this age. Right-censoring can also occur if a child is alive at the moment of the interview and younger than 5 years old.

The data are recorded retrospectively and can therefore suffer from misreporting. One may argue that this problem of misreporting becomes larger when the time interval between the moment of the interview and the date at which a child was born or died increases. If this is actually the case, we can get some indication of retrospective misreporting by comparing child mortality rates computed from the first wave of the NFHS conducted in 1992/1993, with those computed from the 1998/1999 wave. This is shown in Figure 1 and Figure 2 for rural and urban areas, respectively. In case there would be substantial retrospective misreporting the estimated mortality rates (in the overlapping period) based on the 1998/1999 wave should be much lower than those based on the 1992/1993 wave. As can be seen in the figures the mortality rates based on both waves are almost similar. Furthermore, it can be seen from the graphs that rural child mortality rates dropped somewhat during the 1990s, while urban child mortality rates remained constant. But still rural child mortality rates are much higher than urban child mortality rates.

Table 1 provides for each state (except Tipura) the number of women and the number of children in the sample, and neonatal (first month), infant (first year) and child (first 5 years) mortality rates. In rural India out of every 1,000 live born children almost 100 die before reaching age 5 and of these almost half die within the first month after birth (see also Kaplan-Meier estimates provided in Figure 3. The table shows that there is a large

[3] Because we do not have any information about women who have never been married, we can also not analyze child mortality rates of this group of women. We also do not have any indication how substantial fertility among never-married women is.

[4] The restriction to focus only on women below age 32 does not exclude many women from the sample, as recent fertility rates for older age-groups are very low.

variation in child mortality rates across states. The ranking of the states by child mortality rates largely coincides with this ranking in 1960 (see Ravallion and Datt, 2002).

As mentioned above the survey consists of 3 questionnaires. This implies that we can distinguish 3 types of explanatory variables, child specific covariates, family specific covariates and village specific covariates.

3. Model

In this section we briefly describe our model for describing infant and child mortality, a more detailed description can be found in Van der Klaauw and Wang (2004). We focus on children that are born alive and model their mortality probabilities until reaching age 5. We use duration models to specify these mortality probabilities (see Van den Berg, 2001, for a recent survey on duration models). The model specifies mortality rates very flexibly, it extends the commonly used specifications by allowing socio-economic and environmental characteristics to have different impacts on mortality at different ages of the child. Child mortality risks not only depend on observed characteristics, but also on unobserved biological endowment or frailty. Not properly taking account of these unobserved characteristics or the relation between children within a family may lead to inconsistent and inefficient estimators. This issue is not only addressed in the economic literature (i.e. Ridder and Tunali, 1999), but also in other disciplines, including epidemiology and demography (Guo and Rodriguez, 1992; Sastry, 1997; Vaupel, Manton and Stallard, 1979). The model allows for frailty both at the level of the household and at the level of the child. These two types of frailty can be correlated with each other. The relevant outcome measure is age at death measured in months.

We observe J families, which are denoted by $j = 1,..., J$. Family j has some family specific characteristics that are described by a vector z_j, which includes for example religion of the family, asset ownership, parental education and availability of sanitation. A parameter s_j describes the fixed effect of the state in which the family resides. Obviously, two families living in the same state have the same state fixed effect. Furthermore, we allow for additional heterogeneity v_j that describes unobserved family specific characteristics. Both observed and unobserved family specific characteristics do not vary over time. During the observation period in family j there are I_j children born. The data describe a so-called inflow sample. We allow children born in the same family to differ in both observed and unobserved characteristics. The observed child specific characteristics such as gender and birth order, are captured in a vector x_{ij}. Additionally, there is some component w_{ij} that accounts for child specific frailty. We assume that all covariates are exogenous, i.e. the joint distribution of v_j and w_{ij} does not depend on z_j, s_j and x_{ij}.

For a child we can distinguish two possible observations, *(i)* the child is observed to die during the observation period before reaching age 5, and *(ii)* the child reached its fifth birthday alive or is still alive at the end of the observation period. In the first case, we observe that child i of family j died in some age interval $\left(\underline{t}_{ij}, \bar{t}_{ij}\right)$. In the second case, $\underline{t}_{ij}$

equals 60 months or the age of the child in months at the end of the observation period if the child did not reach its fifth birthday (and $\bar{t}_{ij} = \infty$). We introduce a dummy variable d_{ij} that takes value 1 if the individual died within the observation period (the first case) and the value 0 otherwise. All observations are thus artificially right-censored at age 5.

We define T_{ij} as the continuous random variable (with one month as the unit of time) that describes the age at which child i of family j dies. This stochastic variable can only take non-negative values. A common way to model these types of random variables is to specify the corresponding hazard rates θ_{ij}. In the context of child mortality, the hazard rate is often referred to as mortality rate (e.g. Ridder and Tunali, 1999). The mortality rate at age t can be interpreted as the intensity at which a child dies at this age, given that the child survived until age t. The mortality rate $\theta_{ij}(t)$ of child i living in family j at age t is specified as

$$\theta_{ij}(t \mid z_j, x_{ij}, v_j, w_{ij}) = \lambda(t)\exp(I(t \le 1)(z_j\gamma_1 + x_{ij}\beta_1) + I(1 < t \le 12)(z_j\gamma_2 + x_{ij}\beta_2) + I(12 < t \le 60)(z_j\gamma_3 + x_{ij}\beta_3) + s_j)v_j w_{ij}$$

The mortality rate can be decomposed into three parts. The first part $\lambda(t)$ is the baseline hazard that is similar for all children in all families. This baseline hazard captures age dependence. The second part is a regression function, where $I(.)$ is the indicator function taking the value 1 if its argument is true and 0 otherwise. Both the family specific effects z_j and the child specific x_{ij} are allowed to have different impacts on the mortality rate at three different age intervals, *(i)* during the first month after birth, *(ii)* from the second month until the first birthday, and (iii) after the first year until reaching age 5. The impact of the state fixed effect s_j does not vary over the child's age. The third part $v_j w_{ij}$ accounts for the effect of unobserved heterogeneity on the mortality rate.

Empirical studies of child mortality usually do not allow the effects of socio-economic and environmental covariates to vary over the child's age. Guo and Rodriguez (1992) mention that covariate effects might be dependent on the age of the child, but they do not explicitly model it. In their application the covariate effects do not vary over the child's age. The identification of the model is straightforward. The mortality rates $\theta_{ij}(t)$ satisfy the mixed proportionality assumption and the time varying regressors $I(.)z_j$ and $I(.)x_{ij}$ change values exogenously. The frailty components are assumed to be mean one random effects, for which the joint distribution is independent of the regressors. These conditions ensure that the model is identified (e.g. Van den Berg, 2001).

There is a one-to-one relation between the hazard rate $\theta_{ij}(t)$ and the distribution function of T_{ij}. In duration analyses it is often more convenient to use `survivor' functions. The survivor function denotes the probability that a child survives up to a particular age,

$$S(t \mid z_j, x_{ij}, v_j, w_{ij}) = \Pr(T_{ij} > t \mid z_j, x_{ij}, v_j, w_{ij}) = \exp(-\int_o^t \theta_{ij}(s \mid z_j, x_{ij}, v_j, w_{ij})ds)$$

This survivor function will be used for the policy experiments performed in Section 5.

Before estimating the model we determine the parameterization of the baseline hazard and of the distribution of both frailty terms. The baseline hazard $\lambda(t)$ describes how the mortality rate changes with the age of the child. This is parameterized as a piecewise constant function.

We allow for frailty on multiple levels (see Sastry, 1997). In particular, our model framework allows for frailty both at the level of the child and the level of the family (i.e. frailty similar for all children living in the same household). To distinguish these types of frailty we exploit the fact that children within the same family are related, which implies that they might share some unobserved characteristics. In contrast to Sastry (1997), who also includes two independent frailty terms, which both follow a gamma distribution, our model specification does not impose the restriction that different types of frailty are independent of each other. Both family specific and child specific frailty are modeled using distributions with discrete mass-points. Distributions with discrete mass-points are flexible and attractive from a computational point of view. Such a distribution also allows easily for dependency between the two frailty components (see Van der Klaauw and Wang, 2004).

4. Empirical results

In this section we discuss the estimation results. Since our model contains many parameters, we will only discuss some policy relevant parameters. All parameter estimates are presented in Van der Klaauw and Wang (2004). In the next section we discuss some policy experiment, which summarize our main empirical results. The model allowed for both child specific frailty and household specific frailty. We could not find any significant and substantial child specific frailty. The family specific frailty has two points of support. 78% of the probability mass is assigned to a mass point that is almost 13 times higher than the other mass point. Allowing for a third mass point did not improve the estimation results.

The fixed effects for the states differ significantly from each other, the *p*-value for a Wald-test for joint significance is less than 0.0001. This implies that after controlling for child, family and village specific effects, there are still significant differences in child mortality rates between states.

The pattern of the baseline hazard decreases over the age of the child, which implies that mortality rates decrease as a child gets older. However, this duration dependence is not the only source of age dependence, as observed child characteristics and the household's socio-economic and environmental characteristics are allowed to have different impact on the mortality rates at different ages of the child. Also this source of age dependence in mortality rates is significant, a Wald-test rejects the null hypothesis that these covariates effects do not change over the age of the child (the *p*-value of this test is almost 0). This implies that the commonly used specifications in empirical research of child mortality, which impose that observed characteristics have the same effect on the mortality rate at all ages of the child are not sufficiently flexible to capture all relevant changes in child mortality rates.

Socio-economic and environmental characteristics are particularly important in the first month after birth. For each child we can compute $\exp(z_j\hat{\gamma}_k + x_{ij}\hat{\beta}_k)$ for *k=1,2,3*, and we can use these to compute the variance within the population of children. A large variance of $\exp(z_j\hat{\gamma}_k + x_{ij}\hat{\beta}_k)$ indicates that the covariates are relatively important in explaining mortality rates. The computed variances equal 0.039, 0.0073 and 0.0051, for *k=1,2,3* respectively. This variance is thus highest in the first month after birth and decreases afterwards, indicating that socio-economic and environmental characteristics become less important in explaining mortality rates as a child gets older.

Within the first month after birth boys have higher mortality rates than girls, and this is particularly true if he is the first born child. After the first month, there is no significant gender effect on mortality rates among first-born children. However, among children who are not the first born female children are significantly more likely to die than their male counterparts. Cleason, Bos and Pathmanathan (1999) suggest that due to social norms families have preferences for sons, less money is spent on girls, girls are taken to hospital in a later stage of illness than boys and girls are taken to worse doctors. Reducing such gender discrimination might substantially reduce child mortality rates. A child is more likely to survive if the mother was between age 20 and 25 when the child was born than children whose mother was younger or older at the moment of birth. A longer preceding birth interval significantly reduces mortality rates.[5] The National Population Policy aims at delaying childbearing (of young couples) and increased spacing of children, both of which can be useful in reducing child mortality. Children that are part of a twin have much higher mortality rates than single born children, particularly during the first month after birth. After that the difference in mortality rates decreases, but remains substantial and significant. We have tried to separate between the first born of a twin and later born, but there is no difference.

The education attainment of the parents is important to determine children's survival prospect in rural India. Mortality rates at any age are lower for children born into families with parents of finished primary education. This effect is stronger for the mother than for the father. The effect of land, livestock and asset ownership on mortality rates is negligible. Only some of the covariates are significant, particularly on neonatal mortality rates, but quantitatively the variables are not very relevant.

Most of the covariates describing the source of drinking water do not have a significant effect on mortality rates, but jointly these covariates have significant impacts. Water purification decreases mortality rates when the source of drinking water is not piped water, but increases mortality rates in case the household has access to piped drinking water (although the latter increase is insignificant). Such findings are counterintuitive as one expects water purification to decrease child mortality rates. To purify water most households in rural areas strain water by cloth. The empirical result might indicate that either straining water by cloth lowers the quality of the piped drinking water or that households that purify water have access to lower quality piped drinking water than

[5] Interacting the length of the preceding birth interval with gender shows that the covariate effect of the preceding birth interval is similar for boys and for girls.

households that do not purify. Mortality rates are lower within families with some type of toilet facility, and access to electricity. The use of clean cooking fuels and having a separate kitchen also reduces mortality rates. We return to this in the next section.

Access to medical facilities does not seem to be very important in reducing mortality rates, with the exception for doctor availability in the village. Having a doctor in the village reduces the mortality rates of children after their first birthday. Neonatal mortality rates are somewhat higher in larger villages and in villages that are closer to the nearest town. After the first month these effects disappear.

Next we investigate the fit of the model. The model predictions on child mortality are based on the estimated survivor function for child *i* in family *j*

$$S(t \mid z_j, x_{ij}) = \sum_{k=1}^{K} \Pr(V_j = v^k) \sum_{l=1}^{L_k} \Pr(W_{ij} = w^{lk} \mid V_j = v^k) \exp(-\int_o^t \theta_{ij}(s \mid z_j, x_{ij}, v^k, w^{lk}) ds)$$

which we evaluate at *t=1,2,...,60*. We weight the individual survivor function with the sample weights to make them nationally representative for rural India. Obviously, the estimated mortality rates directly follow from the estimated survivor functions.[6] Figure 4 shows how the mortality rates predicted by our estimated model coincide with the observed data. It should be noted that the 95%-confidence interval around the model predictions is relatively tight, the model predictions are very precise. The mortality rates computed from the actual data always falls within the model's confidence interval, indicating that the model is sufficiently flexible to describe the data accurately. However, a good fit of the model does not guarantee that the parameter estimates are true partial effects, i.e. it might be that some covariates account for child, family or environmental characteristics that are not included in the model. Failing to correct for all relevant heterogeneity may lead to serious biases in the estimated partial effects. For the policy experiments performed in the next subsection it is crucial that the parameter estimates are actually the true partial effects. Therefore, we have investigated the sensitivity of the parameter estimates with respect to including additional regressors. Obviously, such sensitivity analysis can never fully guarantee that all relevant heterogeneity is included in the model.

As mentioned earlier, India experienced a period of economic growth during the 1990s. Therefore, we have tried to include a polynomial describing calendar time variation in the model. Such a time trend does not have a significant impact on mortality rates (all parameter estimates of the polynomial were close to zero). The decreasing trend in mortality rates during the 1990s can thus be explained by improvements in observed socio-economic and environmental characteristics during the 1990s. For example, the fraction of the population that is either illiterate or did not finish primary education is steadily decreasing. In the sample of ever-married women, the fraction of women that is illiterate decreased from 72% in the 1992/1993 wave of the NFHS to 67% in the 1998/1999 wave. Also more households have access to electricity and safe water sources.

[6] We have used the Delta method to construct standard errors around the predicted mortality rates.

Next, we tried to include additional observed heterogeneity to the model. In particular, we tried to include additional variables describing asset ownership, house ownership, size of land ownership, type of work performed by the parents, additional variables describing the medical facilities in area around the village, and village-level variables such as the percentage of households with access to piped water, electricity, and etc. The variables that we added to the model did not have any significant effect on mortality rates and did not change the covariate effects of the other variables.

5. Policy experiments

In this section we discuss some policy experiments based on the estimated model. We are interested in two measures, infant and child mortality rates. Table 2 provides the effects on the policy experiments on the infant mortality rate and Table 3 provides these effects on the child mortality rate. Both tables present the number of deaths averted as a result of the policy interventions discussed below. All parameters are treated as causal effects and behavioral changes within a household as response to the policy experiments are ignored. Lee, Rosenzweig and Pitt (1997) mention that this ceteris paribus assumption may be too strong. We provide the percentage of household affected by the policy. If detailed cost data are available we also give the costs per household to actually implement the policy. This allows to compute the costs per death averted as a result of a policy interventions, i.e. the cost-effectiveness analysis.

The number of lives saved per 1,000 live births is based on the estimated survivor functions (see the previous section). We compute the estimated survivor function twice, first without any changes and second after changing the explanatory variables affected by the policy experiment. The difference between these survivor functions provides the increase in probability that a child survives until a given age. We use the sample weights to make these differences in survival probabilities nationally representative and use the Delta method for computing standard errors.

The costs of policy interventions are estimates supplied by a World Bank staff member in India and are presented in Hughes, Lvovsky and Dunleavy (2001). Unfortunately, only for a small number of policy interventions costs information is available. The costs describe the situation in 1998. Annual costs are computed as the sum of recurrent costs and a discounted annuity on capita over the lifetime of the installation. In this computation the average family within rural areas is set to 5.4 and the discount rate is set to 12%.

Source of drinking water: In the first set of policy experiments we consider the source of drinking water. Unsafe drinking water is considered to be a major cause for diarrhea among children and yearly many children die from diarrhea (Jalan and Ravallion, 2003). In rural areas piped water is a safer source of drinking water than the alternatives. Around 74% of the rural households does not have access to piped water (either private or public). In the first policy experiment, we consider the case where the type of water is changed to piped water. The distance to the water source is unaffected and households do not change their water purification behavior. The reduction in neonatal, infant and child mortality rates due to this policy experiment is small and insignificant. These results coincide with that of Jalan and Ravallion (2003) who argue that access to piped water alone is not a sufficient condition for improving the child's health status. Their empirical results based on propensity score matching methods show that access to piped water does not reduce the incidence of diarrhea of children in poor families. Also Ridder and Tunali (1999) do not find any significant effect of having access to piped water on child mortality rates in Malaysia.

Recall from the previous section that water purification is not particularly useful if a family has access to piped water. Therefore, we modify the policy experiment such that all households quit purifying water after having access to piped water. This policy experiment affects 84% of the households in our sample, 10% of the households only quit purifying, 17% quit purifying and obtain access to piped water and 57% only gets access to piped water. The reduction in mortality rates is somewhat larger than in the previous policy experiments, out of every 1,000 live births almost 7 more children are expected to survive up to their fifth birthday. However, this number of death averted is not significant.

Next, we modify the policy experiment even more by giving each household access to private piped water. We maintain that household do not purify their piped water. Of all households in rural India only 11% has access to private piped water. Connecting a household to private piped water costs annually around 100 rupees if connecting is made from a public well or public piped water in the community. This is the case in 38% of the households. In the remaining 51% the connection to private piped water will cost approximately 210 rupees per year. This implies that the costs of implementing this policy is around 145 rupees per household annually (1 US dollar equals around 42 rupees). The birth rate in the rural areas is around 26.2 live born children per 1,000 individuals yearly. The mean household size in the rural areas equals 5.5. Per 1,000 live born children the average costs of providing private piped water is slightly over 1 million rupees.[7] The policy experiment shows that per 1000 live born children 10 additional children reach age 5 (and that all benefits from having private piped water occurs in the first year after birth).[8] The costs of providing all households with private piped water per saved live thus equal 95,900 rupees.

[7] Since the birth rate equals 26.2 live born children per 1,000 inhabitant, 1,000 children are born alive per 38,168 individuals. These 38,168 individuals live in around 6,940 households. Supplying a single household of private piped water costs around 145 rupees, thus supplying these 6,940 households costs around 1,006,940 rupees.

[8] Recall that we rejected a model where covariate effects are similar over the child's age. However, this restriction on the model is often imposed in literature on child mortality. Therefore, it is interesting to see what would be the results of this policy experiment in such restricted model. The restricted model predicts that child mortality decreases with slightly over 11 children (out of every 1,000), which is close to the prediction of our model. However, it predicts that neonatal mortality (significantly) reduces with 5.4 children, which is more than twice a much as the 2.1 death averted predicted by our framework with age varying covariate effects.

Electricity: About 51% of the households in the 1998/99 NFHS have access to electricity. Providing access to electricity to the other 49% of the households reduces neonatal mortality rate by 2.6 children per 1000 births, infant mortality rate by 4 and child mortality rate by 5.5. The reductions in infant and child mortality are statistically significant. Providing 1% of the households with access to electricity reduces child mortality with approximately 0.11, which is similar to providing private piped water to 1% of the households. Ridder and Tunali (1999) find for Malaysia that having access to electricity reduces child mortality. And also based on cross-country analysis of over 60 DHS surveys, Wang (2003) finds that access to electricity is the most important factor in reducing child mortality rates in low-income countries.

Mother's education: The survey shows that about 67% of mothers in rural India did not finish primary school. A policy that would ensure that all mothers complete primary education could reduce child mortality rates significantly at any age of the child. The model predicts that infant mortality rates reduce from 73 death per 1,000 live born to less than 60 and the child mortality rates reduces from 99 to 76. These results confirm Strauss and Thomas (1995), who find that the education of the parents and in particular the mother is important for health outcomes of children. Educated mothers are usually healthier and give birth to healthier babies. They also provide a healthier environment to children and are more likely to have more knowledge about care taking. Also Pritchett and Summers (1996) argue that education is often mentioned as having a strong effect on reducing infant and child mortality. In India women have a low social status. Clearly this has a negative impact on the number of girls that attends school, which leads to high illiteracy levels among women. This is also expressed in the reason for never having attended school by women. Compared to men, women often have never attended school because education was not considered to be necessary or because they were required for household work or taking care of siblings.

Sanitation facilities: About 19% rural households have some type of sanitation facility.[9] The annual costs of providing a household with a private toilet is approximately 250 rupees. Having sanitation facilities significantly reduces the child mortality rates. This result is in agreement with Bhagrave (2003) who finds that access to sanitation facilities significantly reduces infant mortality using data for Utter Pradesh. About 11 children out of 1,000 live births would be saved if all households had sanitation facilities. The costs per death averted of providing each household with a private toilet equals 131,340 rupees. This is almost 37% higher compared with polices focusing on providing private piped water.

Cooking fuels: Most households use wood as the main cooking fuel, only about 9% of the households use clean cooking fuels. As mentioned in the previous subsection using clean cooking fuels reduces mortality rates at all ages of the child. Our policy experiment shows that if all households would switch to using clean cooking fuels, infant mortality rates reduce with almost 16 children per 1,000 and child mortality with about 26 children. These estimates are significant. Kerosene is considered to be a clean cooking fuel and

[9] Note that in our sample containing families with a recent child born, the percentage households with sanitation facilities equals 23%, which is slightly higher than the overall percentage in the rural areas.

shifting from a dirty cooking fuel to kerosene costs annually 419 rupees per household. This implies that the costs to avert one death before age 5 by providing clean cooking fuels are about 99,855 rupees.

Also the use of dirty cooking fuels might cause air pollution if the cooking takes place in the living area of the house. Therefore, we also investigate an alternative policy experiment, where households not only switch to clean cooking fuels, but they also use a separate room as kitchen. Having a separate room as kitchen seems to be particularly relevant during the first year after birth. The infant mortality rate is reduced by 27 per 1000 births with clean cooking fuel and separate kitchen, compared to 16 when only clean cooking fuels are used.

Doctor: Around half of the households live in villages where a doctor is available permanently. In this policy experiment we investigate how child mortality is affected if there would be a doctor available in each village. As mentioned in the previous subsection, living in a village with a doctor does not affect infant mortality rates. This is also expressed in the policy experiment. A doctor in the village reduces child mortality rates after the first birthday. If all villages would have a doctor on average 3.4 children per 1,000 die less before reaching age 5.

All interventions: So far we have investigated effects of single policy interventions. To get an idea to what extent infant and child mortality can be reduced by integrated public policies, we perform a policy experiment where we consider all policy experiments mentioned above jointly. Around 99% of the households will be affected by this joint policy experiment. Neonatal mortality rates can be reduced from around 47 death per 1,000 live born children to 22, infant mortality rates drop from 73 to 28 and child mortality rates reduce from about 100 to 37 children. Clearly improvements in household's socio-economic and environmental characteristics can reduces the mortality rates at any age of the child by more than half.

6. Conclusions

In this paper we have used a flexible parametric framework for analyzing infant and child mortality. The model allows individual characteristics and household's socio-economic and environmental characteristics to have different impacts on infant and child mortality at different ages. This modeling strategy is particularly relevant in describing infant and child mortality, because child-specific, and household socio-economic and environmental characteristics have significantly different impacts on mortality rates at different ages of the child.

We have used the estimated parameters to perform a number of policy experiments. The results from the policy experiments should be interpreted with care, as the proposed policy interventions may cause behavioral changes within the household that are not taken into account in the model. The policy experiments show that a significant number of under 5 years deaths can be averted by providing households with access to electricity, improving the women's education, providing sanitation facilities and access to clean cooking fuels. In particular, reducing indoor air pollution, through a combination of clean fuels and having a separate room as kitchen, seems to be particularly effective for reducing mortality risks during the first year after birth. However, providing clean cooking fuels to a household is a relatively expensive policy, compared with other choices such as providing households with sanitation facilities or safe drinking water. From the cost-benefit perspective, providing access to safe drinking water is a slightly cheaper option (per under 5 years death averted) than providing clean cooking fuels.

The National Population Policy was developed in 2000 to establish population stabilization and focused particularly on health and education of women and children in India. Our results provide strong evidence, alone with many other empirical studies on India, which supports the gender focused of population policy. Another dimension of the population policies focuses on encouraging women to start childbearing at later ages and increase the spacing between children. Our results indicate that these policy emphases can effectively achieve the objective of population stabilization through reducing infant and child mortality rates which are commonly believed to be inversely correlated with fertility rates. The relation between infant and child mortality, and fertility decisions is not investigated in this paper.

It is important to stress that reducing child mortality risks in rural India cannot be achieved with a narrow set of specific programs. It requires that empirical evidence on determinants of child mortality be brought to the forefront of public debate to inform the design of specific and well integrated programs and policies in rural areas. It also requires both political commitment from the government, public pressure, and well implemented programs on health, education, environment and gender. This paper only addresses a very narrow set of issues relevant to child mortality in rural India and the results from this study should be viewed mainly as adding empirical evidence to the body of our knowledge on this important topic.

References

Bhargava, A. (2003), Family planning, gender difference and infant mortality: evidence from Uttar Pradesh, India', *Journal of Econometrics* 112, 225-240.

Claeson, M., E. Bos and I. Pathmanathan (1999), Reducing child mortality in India, keeping up the pace, World Bank, Health, Nutrition and Population (HNP) Discussion Paper, Washington.

Fikree, F., S. Il Azma and H. Berendes (2002), Time to focus child survival programs on the newborn: assessment of levels and causes of infant mortality in rural Pakistan, *Bulletin of the World health Organization* 80, 271-276.

Gleason, S. (2003), Publicly provided goods and infrafamily resource allocation: female child survival in India, *Review of Development Economics* 7, 71-85.

Guo, G. and G. Rodriguez (1992), Estimating a multivariate proportional hazards model for clustered data using the EM algorithm, with an application to child survival in Guatemala, *Journal of the American Statistical Association* 87, 969-976.

Hughes, G., K. Lvovsky and M. Dunleavy (2001), Environmental health in India, priorities in Andhra Pradesh, World bank report, Washington.

Jalan, J. and M. Ravallion (2003), Does piped water reduce diarrhea for children in rural India, *Journal of Econometrics* 112, 153-173.

Lee, L-F., M.R. Rosenzweig and M.M. Pitt (1997), The effects of improved nutrition, sanitation, and water quality on child health in high-mortality populations, *Journal of Econometrics* 77, 209-235.

Mishra, V. and R.D. Retherford (1997), Cooking smoke increases the risk of accurate respiratory infection in children, *National Family and Health Survey Bulletin* 8.

Murthi, M., A. Guio and J. Dreze (1995), Mortality, fertility and gender bias in India, *Population and Development Review* 21, 745-782.

Pandey, A., M.K. Cheo, N.Y. Luther, D. Sahu and J. Chand (1998), Infant and child mortality in India, National Family Health Survey Subject Report 11, East-West Center Program on Population, Hawaii.

Pritchett, L. and L.H. Summers (1996), Wealthier is healthier, *Journal of Human Resources* 31, 841-868.

Ravallion, M. and G. Datt (2002), Why has economic growth been more pro-poor in some states of India than in others?, *Journal of Development Economics* 68, 381-400.

Ridder, G. and I. Tunali (1999), Stratified partial likelihood estimation, *Journal of Econometrics* 92, 193-232.

Rosenzweig M. and T.P. Schultz (1982), Market opportunities, genetic endowments, and intrafamily resource distribution: child survival in rural India, *American Economic Review* 72, 803-815.

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.

Strauss, J. and D. Thomas (1995), Human resources: empirical modelling of household and family decisions, in J. Behrman and T.N. Srinivasan, *Handbook of Development Economics, Volume 3A*, North-Holland, Amsterdam.

Van den Berg, G.J. (2001), Duration models: specification, identification, and multiple duration, in J.J. Heckman and E.E. Leamer (eds.), *Handbook of Econometrics, Volume 5*, North-Holland, Amsterdam.

Van der Klaauw, B. and L. Wang (2004), Child mortality in rural India, World bank, Health & Population working paper 3281, Washington.

Vaupel, J.W., K.G. Manton and E. Stallard (1979), The impact of heterogeneity in individual frailty on the dynamics of mortality, *Demography* 16, 439-454.

Wang, L. (2003), Determinants of child mortality in LDCs: empirical findings from demographic and health surveys, *Health Policy* 65, 277-299.

Wolpin, K.I. (1997), Determinants and consequences of the mortality and health of infants and children, in M.R. Rosenzweig and O. Stark (eds.), *Handbook of Population and Family Economics, Volume 1A*, Elsevier, Amsterdam

Table 1: Cumulative child mortality rates stratified by state.

			Mortality (no. death per 1000 births)		
Region	**Women**	**Children**	**≤ 1month**	**≤ 12month**	**≤ 60month**
India	25196	43185	47.2	73.3	99.9
North					
Delhi	79	144	34.5	42.5	42.5
Haryana	872	1467	36.9	61.9	76.6
Himachal Pradesh	783	1293	27.0	38.6	47.6
Jammu & Kasmir	822	1489	40.7	61.0	76.1
Punjab	678	1164	36.3	64.8	83.1
Rajasthan	2323	4163	52.6	83.6	115.2
Central					
Madhya Pradesh	2272	4122	65.0	102.7	158.8
Uttar Pradesh	3333	6032	57.5	93.9	128.3
East					
Bihar	2738	4652	50.4	74.8	100.8
Orissa	1437	2360	49.3	83.9	107.1
West Bengal	965	1479	35.7	52.5	71.8
Northeast					
Aranachal Pradesh	487	841	31.3	50.7	84.5
Assam	1109	1823	35.5	62.2	82.8
Manipur	375	678	23.6	46.8	77.6
Meghalaya	412	825	59.6	100.9	131.4
Mizoram	200	392	28.9	61.2	74.6
Nagaland	296	610	27.6	44.4	80.4
Sikkim	407	633	31.3	51.4	73.1
West					
Goa	200	288	27.5	34.6	41.3
Gujarat	818	1425	47.0	72.1	81.3
Maharashtra	871	1515	38.2	52.6	68.0
South					
Andhra Pradesh	1057	1589	42.4	67.9	88.3
Karnataka	1110	1888	41.5	58.9	77.1
Kerala	629	912	11.9	15.6	17.0
Tamil Nadu	923	1401	37.3	47.6	71.9

Explanatory note: Women provides the number of women in the sample with at least 1 child born since January 1993. Children is the total number of live born children. The mortality rates are per 1000 live born children, the number of children that die within 1 month, within 1 year, and before reaching age 5.

Table 2: Reduction in infant (under 1 year) mortality rates from policy changes: achieving universal access to basic services in rural India (standard errors in parentheses).

	Current level of access	Under 1 year death averted (per 1000 births) from reaching 100% access	Costs per life saved (in rupees)
Access to piped water	26%	4.0 (3.5)	
Access to piped water & no purification	16%	6.1 (3.8)	
Access to private piped water & no purification	6%	10.9 (5.1)	92380
Access to electricity	51%	4.0 (2.0)	
Mother finished primary school	33%	13.7 (2.5)	
Has toilet facility	23%	6.1 (3.8)	230373
Uses clean cooking fuel	9%	15.7 (5.6)	156943
Has separate kitchen & uses clean cooking fuel	7%	27.3 (5.7)	
Doctor available in village	50%	0.0 (1.4)	
Access to all of the above	1%	45.2 (4.4)	

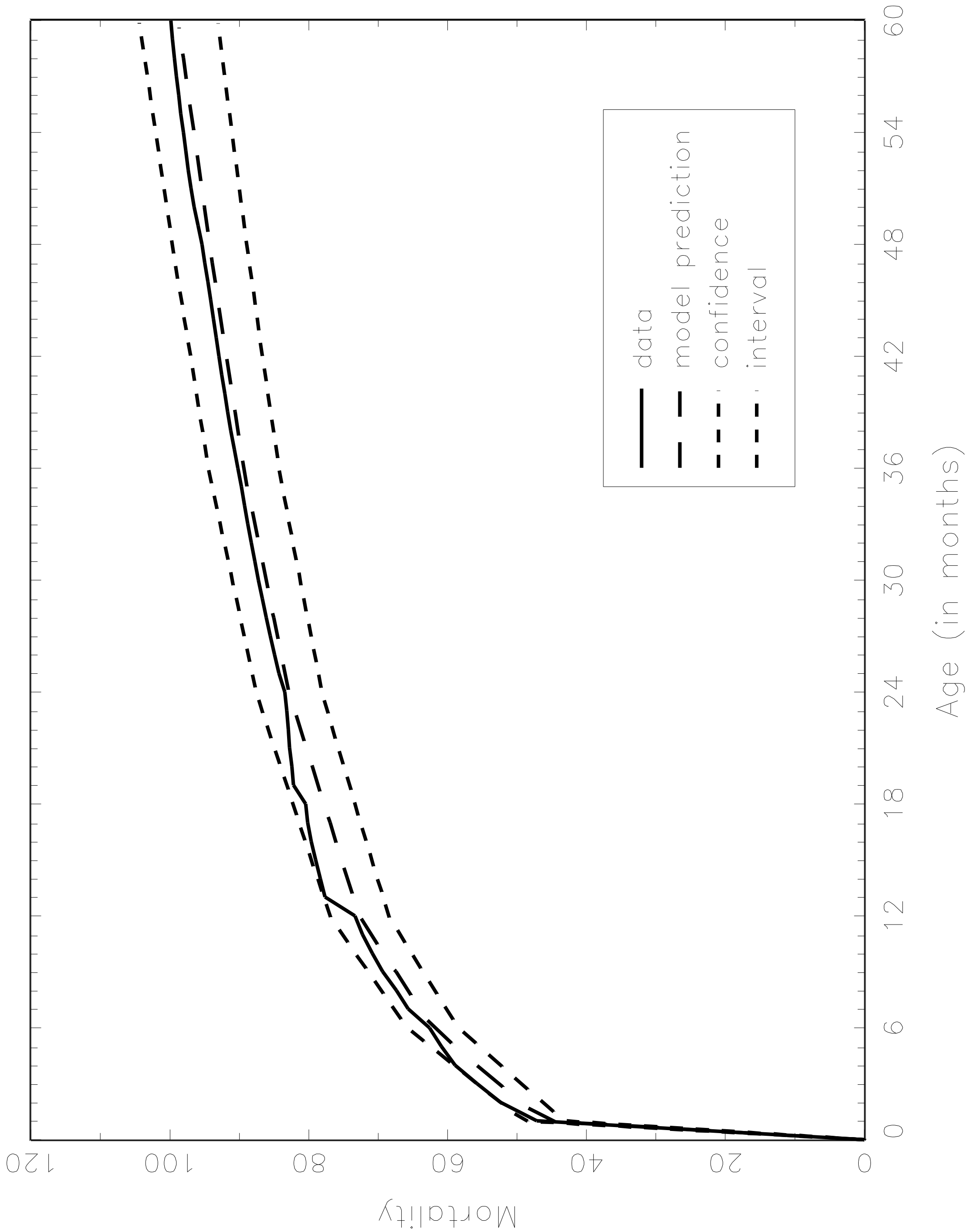
Explanatory note: Clean cooking fuels are kerosene. The annual costs of providing access to piped water is 145 rupees, the annual costs of providing toilet facilities is approximately 250 rupees and the annual costs of providing clean cooking fuels is 419 rupees per household.

Table 3: Reduction in child (under 5 years) mortality rates from policy changes: achieving universal access to basic services in rural India (standard errors in parentheses).

	Current level of access	Under 5 death averted (per 1000 births) from reaching 100% access	Costs per life saved (in rupees)
Access to piped water	26%	4.4 (4.6)	
Access to piped water & no purification	16%	6.8 (5.1)	
Access to private piped water & no purification	6%	10.4 (7.0)	96821
Access to electricity	51%	5.5 (2.6)	
Mother finished primary school	33%	22.9 (3.3)	
Has toilet facility	23%	10.7 (4.9)	131334
Uses clean cooking fuel	9%	26.5 (6.9)	99849
Has separate kitchen & uses clean cooking fuel	7%	33.6 (9.2)	
Doctor available in village	50%	3.4 (1.8)	
Access to all of the above	1%	62.8 (5.8)	

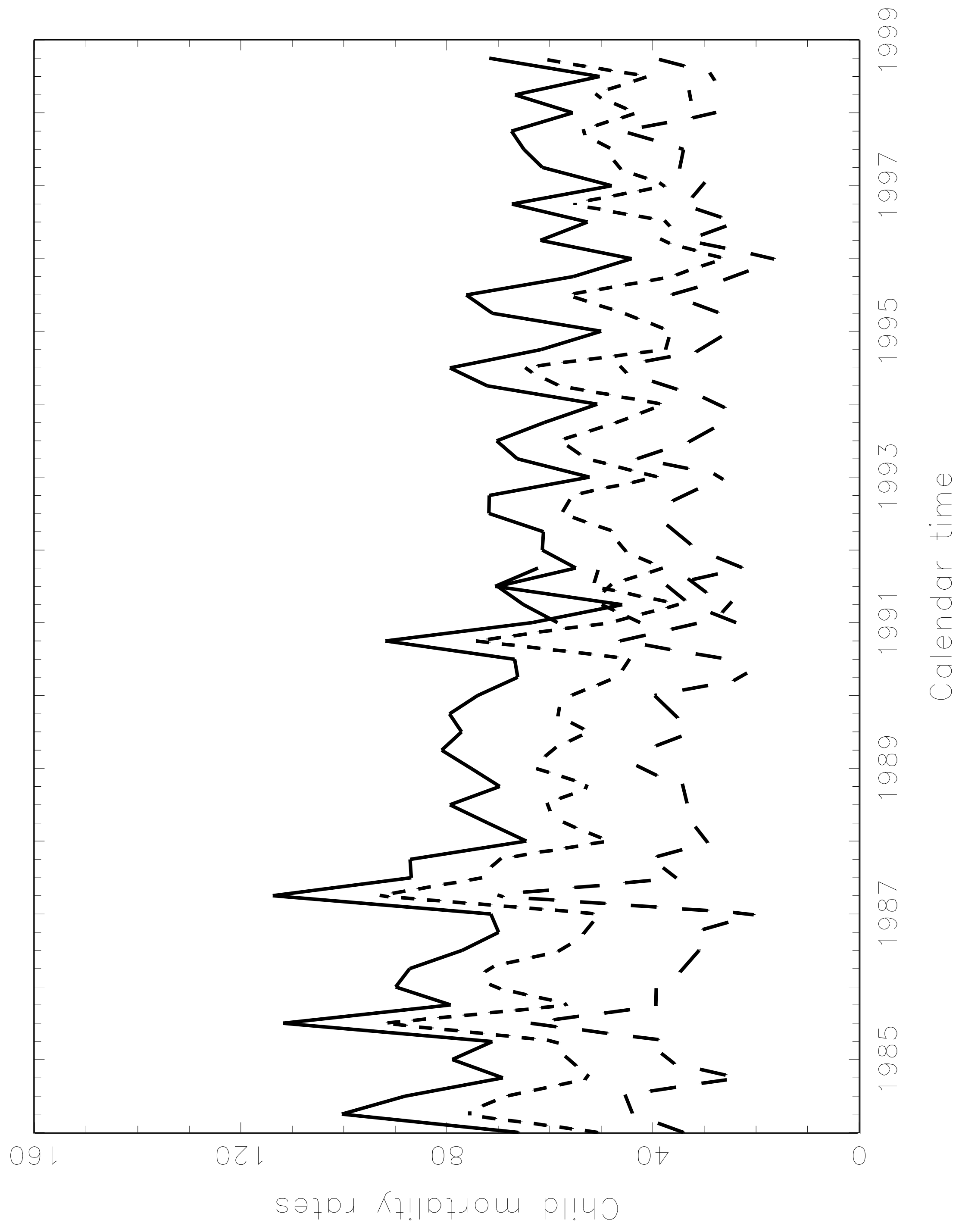
Explanatory note: Clean cooking fuels are kerosene. The annual costs of providing access to piped water is 145 rupees, the annual costs of providing toilet facilities is approximately 250 rupees and the annual costs of providing clean cooking fuels is 419 rupees per household.

Figure 1: Rural child mortality rates over calendar time.



Explanatory note: The lower line is the mortality rate in the first month after birth, the middle line the mortality rate during the first year after birth and the upper line the mortality rate during the first 5 years after birth. The left-hand side of the figure is based on the 1992 wave of the NFHS, the right hand side on the 1998/1999 wave. For 1991 measures using both surveys are available.

Figure 2: Urban child mortality rates over calendar time.



Explanatory note: The lower line is the mortality rate in the first month after birth, the middle line the mortality rate during the first year after birth and the upper line the mortality rate during the first 5 years after birth. The left-hand side of the figure is based on the 1992 wave of the NFHS, the right hand side on the 1998/1999 wave. For 1991 measures using both surveys are available.

Figure 3: Kaplan-Meier estimates for the survival of 1000 live born children.

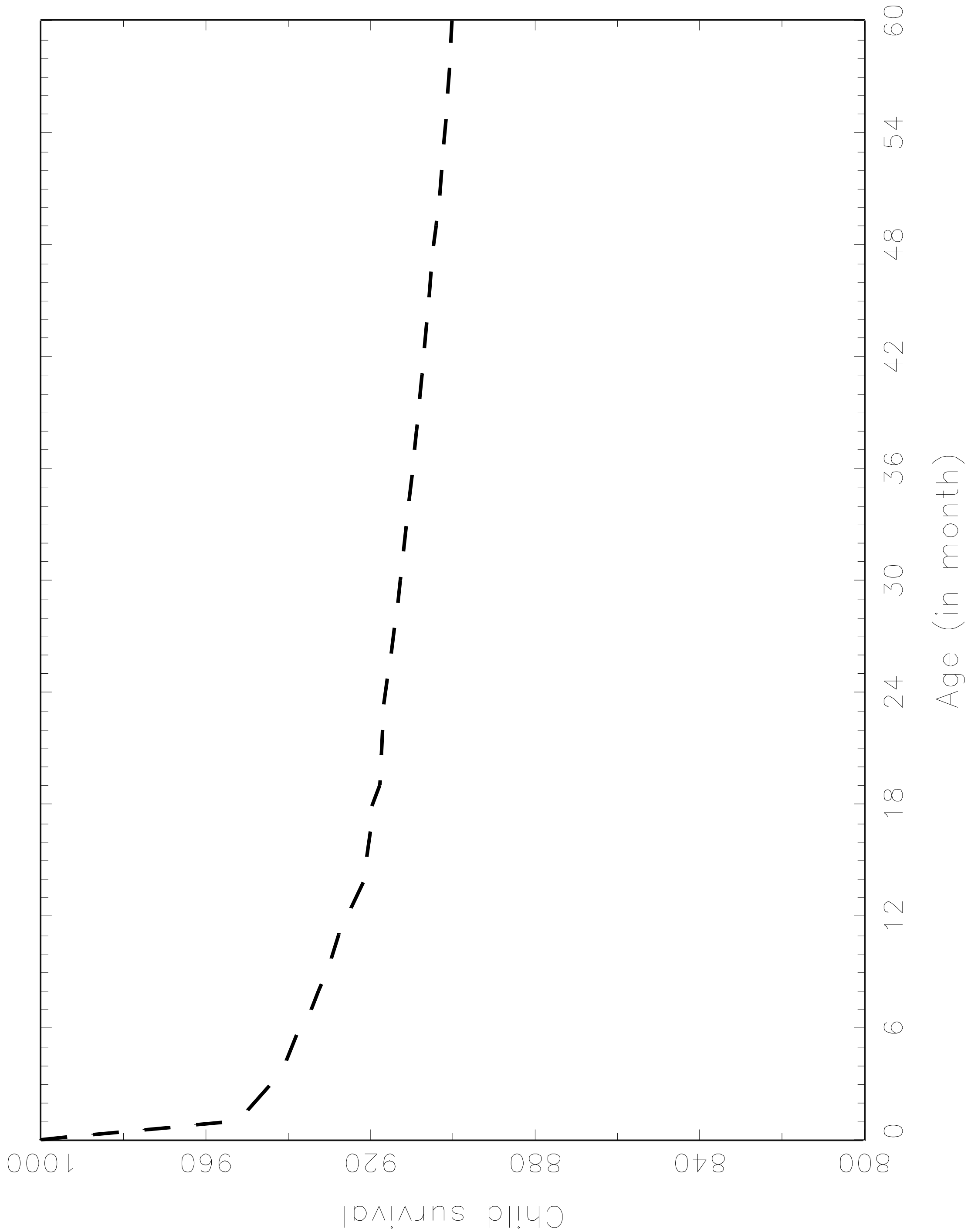


Figure 4: Size of child mortality before a particular age (out of 1000 live born). Both the data and the model predictions are weighted to make them representative.