The effect of water and sanitation on child mortality in Egypt

HALA ABOU-ALI
Environmental Economics Unit, Department of Economics, Göteborg University, P.O. Box 640. SE 40530 Göteborg, Sweden
Tel: (46) 31-773 4667; Fax: (46) 31-773 1043; e-mail: Hala.Abou-ali@economics.gu.se

Abstract

This paper assesses water and sanitation's impacts on child mortality in Egypt. The analysis is conducted using a three-part model specification, comprising discrete choice to model the child prospects of dying during the neonatal period. The remaining parts uses transition models to model infant and childhood risk of death where unobserved heterogeneity is accounted for. The results show that access to municipal water decreases the risk and sanitation is found to have a more pronounced impact on mortality than water. The results suggest that increasing awareness of the Egyptian population relative to health care and hygiene is an important feature to decrease child's mortality risk. Moreover, gender discrimination is found to be of an important effect beyond the neonatal period.

Keywords: Child mortality; Household environment; Transition models; Unobserved heterogeneity; Middle East; Egypt
JEL classification: C14; C41; I12

1. Introduction

The objective of this paper is to quantify the mortality risk decrease for children under the age of five years that would result from certain improvements in water and sanitation services, using the Demographic and Health Survey in Egypt from 1995 (DHS, 1995) and transition rate models. Child mortality is studied by grouping children after age. With regard to water it is generally expected that the mortality risk will decrease as societies upgrade from less accessible poor quality water to community facilities and finally to household or residential connections. Concerning sanitation, it is expected that mortality will be decreased with flush toilets, next best with pit latrines and worst without any facilities. In order to draw some conclusions on these issues it is important to model mortality taking into account the interrelationship between the affecting factors. Since death is a biological process, the initial health endowment of the child emerges as one of the key factors in determining the outcome. This will be taken into consideration in the form of household frailty. Thus, environmental factors and socio-economic conditions of the household have an indirect impact on mortality to the degree they speed up or slow down the biological process (Ridder and Tunali (1999)).

A duration model is estimated for the entire sample together with a three-part model. Neonatal mortality is first modeled by using a discrete dependent variable model and the mortality risk in the infant -up to his first birthday- and childhood –from the first birthday to fewer than 5 years- stages is modeled using non-parametric, semi-parametric and parametric duration models. In this particular application, this three-part model predicts mortality better than a duration model for the under five child mortality in general since it uncovers some interesting differences between the impacts of household environmental and socio-demographic determinants on the neonatal, infant and subsequent mortality risk.

It has been argued that water supply and sanitation in Egypt have a considerable effect on child mortality. For instance Ali et al. (1990) find that access to clean water and adequate sanitation decreases child mortality. According to the World Bank (2002) there is an annual average loss (cost) of 0.8 percent of Egypt's Gross Domestic Product (GDP) due to diarrheal diseases and mortality primarily affecting children, caused by lack of access to safe water and sanitation, and inadequate domestic, personal and food hygiene. Table 1 summarizes some of the features of several studies that are most relevant to this work. In Woldemicael (1998) the central question is whether access to piped water and a flush toilet affects the survival chance of children under five years of age in urban areas of Eritrea. His finding shows that the effect of household's environment (water supply and toilet facility) is large and statistically significant during the post-neonatal and childhood periods while the effect totally disappears during the neonatal period. Lee et al. (1997) set out a framework for estimating the effects of interventions that improve health infrastructure on child health taking into account three responses. First, changes in the allocation of nutrition to children. Second, changes in whom among children survive, and lastly, changes in the health of the children who survives net of family resource allocation. They find that estimates based on structural-equations and semi-parametric models applied to data describing households from rural areas of two low-income countries indicate that conventional reduced-form estimates understates the effectiveness of improved sanitation facilities. In Ridder and Tunali (1999) the aim is to assess whether empirical evidence supports the presence of family specific frailty components. Although child mortality differentials with respect to water supply and sanitation in many developing countries suggests that access to piped water and toilet facility may improve survival chances of children, Ridder and Tunali (1999) could not find any evidence supporting this relation. Lavy et al. (1996) analyze the effects of quality and accessibility of health services and other public infrastructure on the health of children in Ghana. Incorporating some community characteristics they constructed an indicator of poor water quality and sanitation. Focusing on child survival, height and weight, their results suggest an important role for public health policy in eliminating the rural-urban disparities, particularly in improving health status of rural children as well as reducing their mortality rates. Guilkey and Riphahn (1998) estimate a structural discrete time hazard model of the determinants of infant and child mortality in the Philippines in order to evaluate the effect of biological variables on mortality. They find that controlling for biological mechanisms, the birth order and parity no longer have a direct effect on mortality. However, breastfeeding is found to be one of the most important determinants of child survival. Trussell and Hammerslough (1983) provide a complete self-contained exposition of estimating a life table with covariates through the use of hazard models applied to child mortality in Sri Lanka. Their results show that the type of toilet facility, mother's and father's education, urban/rural estate residence, ethnicity, birth order, age of mother at birth and gender are strongly related factors with child mortality.

Casterline et al. (1989) scrutinizes the effects of income on infant and early childhood mortality at the household level in Egypt. They also incorporate socioeconomic and demographic variables in their logistic regression equation, where this type of model does not account for censored data. The main conclusions of Casterline et al. concerning income are: i) household income does not affect survival through infancy but the effects are pronounced during early childhood. ii) The data used suggests that the impact of income is somewhat greater for educated mothers, when the father is of higher socioeconomic status and where the household receives piped water. However, as far as the author knows this is the first study that uses a three-part model setting to comprehensively explain the determinants of child mortality in Egypt. Furthermore, although, the issue of gender discrimination is highly ranked on the discussion agenda in the developing world this is in our knowledge the first application that has been able to quantify and measure the effect of gender discrimination in child mortality beyond the neonatal period. Finally, one of the aspects that make this paper of interest is that the developing world lacks this type of application to understand and explain the high child mortality levels and undertake the correct measures to deal with this problem.

The aim of this paper is to understand and explain the determinants of under-five mortality. The three-part model used allows the control of censored data together with the accounting of household unobserved characteristic. This paper is arranged as follows, Section 2 describes the data. Section 3 details the econometric modeling. In Section 4 the results from the empirical investigation are compiled. Section 5 concludes.

2. The data

2.1 The survey

The data used in this study originates from the DHS conducted in Egypt between November 1995 and January 1996. DHS is a large cross sectional data set that is comparable across countries. Containing information obtained from ever-married women aged 15-49 years regardless of whether they had any preschool children. It is a very rich data set that offers more degrees of freedom and additional information on individual attributes. This makes DHS ideal for describing mortality differentials in different regions of the world and for estimating reduced form models of the determinants of infant and child mortality within a country (Guilkey et al. (1998)).

Administratively, Egypt is divided into 26 governorates. These in turn consist of 467 districts, forming the primary sampling units of the DHS-95. Each district contains Shiakhas or villages giving rise to 934 segments from which the household's selection was implemented. The DHS-95 consists of a sample of 16 000 households in order to meet the target of a random sample of 14 000 interviews with ever-married women aged 15-49 that realize the objectives of the survey. The 934 segments will form the clustering

Table 1: A summary of child mortality studies.

Author(s)	*Dependent variable*	*Method*	*Source of data*	*Water supply or source of drinking water*	*Sanitation (toilet) facility*	*Other covariates*
Trussell and Hammer-slough (1983)	Infant mortality	Hazard model (all main effects vs. no interaction among covariates)	Sri Lanka World Fertility Survey (WFS) 1975	Not significant and has been excluded	Negatively related i.e., better sanitation less mortality (-)	Education of mother and father, age of mother at birth, sex of child, birth order of child, place of mother's current residence and ethnic group.
Casterline et al. (1989)	Infant and child mortality	Logit model	Egypt WFS 1980	Net impact on early childhood mortality	No significant effect	Net household income, region and type of place of residence, sex of child, paternal status, maternal education and demographic status and paternal kin relationship
Aly and Grabowski (1990)	Child death probability	Logit model	Egypt WFS 1980	Significant and negatively related with mortality (-)	Significant (-) and a non significant (+)	Breastfeeding, blood relation between spouses, number of pregnancy and husband and wife literacy.
Lavy et al. (1996)	Child survival and child anthropometric measures	Hazard regression model (Weibull specification)	Second Ghana Living Standard Survey (GLSS) 1988	Poor sanitation and water is positively related to survival in urban areas and male cases. It shows a negative relation in rural areas and female cases		Health service quality and availability, food prices, child sex, education of mother and father, height for age and weight for age.
Lee et al. (1997)	Child health	Structural equation semi-parametric model	1981-82 nutrition survey of rural Bangladesh and 84-85 IFPRI Bukidnon – Philippines	No significant effect	Most categories (parts or types) has no significant effect	Wealth, head or female (h/f) education, sex of child, (h/f) age and place of residence.
Guilkey and Riphahn (1998)	Child mortality up to 2 years	Structural discrete time hazard	Longitudinal data from-Metropolitan Cebu-Philippines from 1983 to 1986	Not included	Not included	Biological mechanism (e.g., birth weight and nutritional status), income, sex of child, and place of residence.
Woldemicael (1998)	Neonatal, post-neonatal and childhood mortality	Hazard regression model	Eritrea DHS 1995	Improvement in the provision of water supply and toilet facilities are likely to reduce mortality especially beyond the first month of life		Mother's education, place of residences, H. economic status and year of childbirth.
Ridder and Tunali (1999)	Child mortality risk	Stratified partial likelihood estimation	Malaysian family life survey 1976-77	No much impact	No much impact	Birth interval, child sex, mother's age at birth, birth weight, birth order, childcare and breastfeeding.

basis in the analysis to come. Furthermore, DHS data used here is subject to some problems as a source of information on the distribution of duration. In this survey, extended information is only collected for children that are born five years before the survey was conducted. This means that information on children born prior to this period is not collected. In this paper, the data set was rearranged into spell data in order to ease the estimations. Henceforth, the analysis is limited to children under the age of five. This then leaves us with a sub-sample of 8 017 women that gave birth to 12 104 children that are the interest of this study.

The biomedical perspective distinguishes between the causes of neonatal deaths, those occurring in the first 28 days, and post-neonatal deaths, those occurring between 29 days and one year. This age classification is attributed to biomedical research where the causes of child mortality have been divided into endogenous and exogenous categories.[1] Due to the DHS Egypt 1995, which is in a monthly format, neonatal is defined here to be infants up to one month of age. The infant stage is defined to be post-neonatal where the age of the infant is between one and 12 months. As for the childhood stage it is defined to be the age greater than 12 months and less than 60 months.

The analysis is restricted to the period 1991-1995 where it is assumed that no investments have been undertaken by the households in order to improve their living conditions. Hence, some covariates are assumed not to vary with time; For example, the water supply and toilet facility variables are constructed based on the state of affairs at the time of the interview. Thus, it may be wise to stick to the recent past that may correspond to the category of facility listed. The place of residence of the child while exposed to the risk of death assumed unchanged during the period of concern will be quite realistic since changing residence is not a very common phenomenon in Egypt.

[1] For an extensive discussion on this matter see Wolpin (1997) pp. 523-24.

2.2 Variable descriptions and hypotheses

The variables in this paper are classified into four groups: environmental, socio-economic, demographic and behavioral differences. Table 2 presents definitions and descriptive statistics of the environmental variables.[2] The choice of these variables was guided by the determinants of infant and child mortality literature such as in Trussell and Hammerslough (1983) and Ridder and Tunali (1999). The literature on endogenous fertility (e.g., Olsen et al. (1983) and Wolpin (1984)), poses some difficulties for the proximate determinant approach justifying the inclusion of regressors such as mother age at birth, birth interval and the total number of children ever born. These, as well as breast-feeding behavior, are to some extent choice variables and are likely to be influenced by parental norms and perceptions. As mentioned earlier improved water supply and higher quality sanitation facilities are epidemiologically directly related to lower mortality. Therefore, the improvement of water supply is expected to be inversely related to mortality risk. Since water's bacterial content is the one that affect health and since there is no available information about this variable, water supply as a source of drinking water will be used as a proxy variable, (see Wolpin (1997) for more details).

Communities without municipal water range between 23 and 36 percent. As concerns the lack of sanitation, the coverage is between 6 and 17 percent. The former communities rely on unimproved water supplies (e.g., wells, rivers, ponds, canals and unprotected springs) and the latter on unimproved sanitation facilities as holes in the ground, bushes and other places where human waste is not contained to prevent it from contaminating the environment. Those categories are treated as baselines in the analysis. Communities with improved water and sanitation do not all have the same services. It should be noted that the functioning or the improvements in sanitation facilities also depends on its connection to a sewer system. However, only some of the urban households have access to sewer systems, the data used here does not contain this type of information and the reader should be aware that this is not controlled for. There is a wide variation in the types of

[2] For an extended list of variables together with their definitions and descriptive statistics see Table A1 in Appendix A.

services, but for the purpose of this study, water services were classified into two categories: residential water facilities and public tap. For sanitation, a flush toilet was considered to be a modern type of facility. A traditional facility is defined such as a water-based system, pit latrine or a similar fecal disposal system.

Table 2: Sample statistics and the distribution (%) of environmental variables among children in Egypt

Variable	*Neonatal*	*Infant*	*Childhood*
Sample size (number of children)	12104	11379	8906
Mortality prevalence (number of children)	363	384	106
Mortality prevalence (%)	3	3.4	1.2
Tap water into residence (yes)†	62.9	63.1	63.5
Public tap water (yes)	9.4	9.4	9.4
Modern facility (yes)	18.8	19	19.2
Traditional facility (yes)	70.8	70.6	70.6

† (yes) refers to a dichotomous variable indicating that the value 1 is taken by the variable name (e.g.. Tap water into residence (yes) = Dichotomous variable indicating that the household has municipal water piped into residence).

Turning to the socio-economic variables, income is not available from the data but some proxy variables are used in order to avoid the ambiguity of some related variables (e.g., sanitation facilities and education) that pick up the effect of the omitted income variable and become difficult to interpret. For this purpose some of the household wealth ingredients in the DHS are used to construct a standard of living index (SLI). The SLI is a simple summation of the number of items present in the household. The items considered are the existence of electricity, refrigerator, radio, TV, electric fan, car and/or motorcycle and type of flooring. Hence, the SLI is a categorical variable indicating the number of possessions of the household ranging from zero to seven. It is expected that this variable will be inversely related to the mortality risk. Although this appears to be a reasonable proxy for income given the data, the reader should be aware of the measurement problems. Furthermore, even though the author has no reason to believe that the likely measurement errors will directly bias the coefficients in any particular direction, a caveat should be made that random measurement error biases may tend to bias coefficients downwards (see e.g. Greene, 2000).

Differentials by urban/rural residence have commonly been observed, with urban areas having more advantages. Using this partition, the characteristics and the administrative division of Egypt, the place of residence takes the form of six dummies; urban governorates, Lower Egypt urban and rural, Upper Egypt urban and rural and the frontier governorates, with the last category treated as omitted in the estimation.[3] The mother's education is included, it is believed to be inversely related to child mortality. Four dummy variables are constructed to capture different levels of education; Mothers with no education which is treated as the omitted variable, low education where the mother did some primary schooling, medium education is given for the ones who achieved primary schooling and /or continued through secondary and high education category encloses the ones who completed secondary school and higher.

As concerns the demographic variables, the patterns of mortality by maternal age and birth order are typically U-shaped; Children born to both relatively old and young women have higher mortality rates than others; the interpretation of the effect of maternal age at birth on infant mortality must be biological, i.e., it depends on reproductive maturity.[4] Moreover, first and higher order births also have higher mortality rates. Since, the birth order reflects the components of the child's biological endowments. As for the child's gender, it is widely believed that male mortality is higher due to biological disadvantages. However, the subsequent results clearly show higher female mortality indicating gender discrimination, i.e. girls have higher risk of dying in Egypt. This could be a cultural reflection of preference of male offspring. Where son's preferences have a powerful significant influence on fertility in some selected Arab countries (Al-Qudsi (1998)).

Finally of the behavioral variables, breast-feeding as one of the most important variables is associated with the exposure level of pathogens (organism or substance that causes disease). It is the most appropriate food for young infants because it is nutritious, sterile, may reduce the ingestion of other, often contaminated foods and confers immunity. Around 65 percent of the literature on child mortality make use of the birth weight variable and find that infant mortality steeply declines with its increase (see Buehler et al. (1987)). Failing to control for birth weight, as a prior of health condition is due to lack of

[3] The definition of the administrative division of Egypt together with the urban rural partitioning used here evolves from CAPMAS (1996).

information in the data. It could be argued that this omission does not pause serious problems.[5] Although rural clinics are widespread throughout Egypt, their impact is limited by the poor quality of care provided and the extent to which people can effectively make use of the services. This could lead to an ambiguous effect of the maternity care variable.

3. Econometric modeling

The effect of water and sanitation on child mortality is analyzed by classifying children under five into three age groups. According to Pebley and Stupp (1987) such classification in the analyses of the determinants of child mortality poses two technical problems. The first relates to censored observations, that is, not all live births have the chance to survive to the oldest age under investigation. The second is concerned with the rapid changing behavior of mortality from one age to the next during the child's life span. The use of hazard regression model permits inclusion of all births and reflects the changing rates of mortality. Many researchers such as Trussell and Hammerslough (1983), Woldemicael (1998) and Ridder and Tunali (1999) have used this type of model. Comparing standard regression procedures and transition models two main differences may be noted.[6] The first, in regression models it is assume that mortality occurs at the rate of exactly one per period. The model used in this paper assumes that mortality occurs randomly. Thus the transition model is intrinsically stochastic. The second difference is that a distinguishing feature of duration data is the possibility that some of the durations observed will be censored. Censoring is an event that occurs at some time, so the data consists of a measured spell length together with the information that the spell was censored or not. Standard regression procedures are not efficient because they do not use all the information in the sample. In general, estimation procedures that do not account for the censored nature of the data will produce biased and inconsistent estimates.

---

[4] Wolpin (1997) p. 542
[5] See an extensive discussion on this matter in Wolpin (1997).
[6] For an extensive discussion on the relative merits of regression vs. hazard model approaches see Flinn and Heckman (1982).

To study child mortality in Egypt a three-part model is estimated, comprising discrete choice to model the child prospects of dying during the neonatal period. The remaining parts uses transition model to model his risk of death beyond then. The chances of dying the first month are modeled using a probit model (see Maddala (1983) pp. 22-27). Proceeding with the duration analysis, as a first step, the non-parametric Kaplan-Meier survival functions are investigated. As discussed in Kiefer (1988), graphical methods are useful for displaying data on durations and for preliminary analyses perhaps to suggest functional forms of homogeneity being achieved roughly by grouping on explanatory variables. Here the product limit or the Kaplan-Meier estimator of the survival function is used. The survival probability *S(t),* is the contribution of a censored child in duration *t* to the likelihood. Censored child is defined to be that he/she is alive during the observation period. On the other hand, the hazard $\lambda(t)$ (which is the risk of completing a spell at duration *t* conditional upon the spell's reaching duration *t*) is, $\hat{\lambda}_i = d_i / N_i$. Where d_i denotes the number of deaths at duration t_i and N_i denote the number of children at risk just before t_i. In particular, the plot of the logarithm of the integrated hazard, which can be estimated as minus the logarithm of the Kaplan-Meier survival estimator, against duration provides a starting point for the parametric analysis.

Since the primary interest of this paper is to investigate the effect of the household environmental characteristics –i.e., water and sanitation- on the probability of child mortality, a multivariate model of the child life duration is implemented. Two popular methods of analyzing the effect of explanatory variables on the hazard rate are the proportional hazard model and the accelerated failure time. The former is based on semi-parametric model that is a very flexible method of estimation since the baseline hazard is estimated nonparametricly and eliminates the risk of corrupting the estimated hazard parameters while the effect of the covariates takes a particular functional form. Cox proportional hazard framework is used; it is based on the following hazard function for the distribution of living duration that is the transition rate. The probability of leaving life at any moment given that the child is still alive up to that moment is, $\lambda(t) = \lambda_0(t)\, e^{\beta' x}$. Where, $\lambda_0(t)$ is the baseline hazard and $e^{\beta' x}$ is the relative risk associated with the regressors x. In the proportional hazard specification the effect of the regressors is to multiply the hazard function itself by a scale factor. The vector of parameters β in this setting will be estimated without specifying the baseline hazard function i.e., $\lambda_0(t)$ is treated non-parametrically. Therefore a partial likelihood approach will be used to estimate the parameters (see Cox (1972)). One of the problems here is the possible existence of unobserved heterogeneity between children from different families since they potentially have a different duration distribution and the control for the effect of the related explanatory variables is incomplete. The result that holds generally about heterogeneity is that it leads to a downward biased estimate of duration dependence. Therefore a further step is taken by incorporating unobserved heterogeneity in to the model. And thus the hazard function will be of the following form $\lambda_i(t) = \lambda_0(t)\,\alpha_i e^{\beta' x_i}$ where α_i is the group i level frailty. The frailties are unobserved positive quantities with mean one and variance θ.

On the other hand, the parametric hazard models require more restrictive hypothesis than those for the models represented above. Here the baseline hazard needs to be modeled. Parametric models provide hazard rates that are monotonic, constant, non-monotonic or a mixture of these. A very flexible form of the exponential model $\lambda(t) = \lambda$ (constant over time) is used. The time axis is split into time periods. The transition rates are constant in each of these intervals and may change between them. This model is known to be piecewise constant exponential model. When variables are included, the baseline hazard is characterized by a period specific constant that can vary across time periods but the variables have the same proportional effects in each period. The literature contains an abundance of choices for parametric models; a popular one is the Weibull model. The hazard function of which is defined as $\lambda(t) = \lambda p(\lambda t)^{p-1}$ where $\lambda = e^{-\beta' x}$ and p is a scale parameter with $p<1$ indicating that the hazard falls continuously over time, while $p>1$ indicates the opposite (see Greene (2000) pp. 939-944). It is determined whether this is a proper model by means of the plot of the logarithm of minus the integrated hazard which should be linear.

As previously mentioned, households and their environment differ in so many respects that no set of measured covariates can possibly capture all the variation among them. It is well known that the estimated hazard from a model that neglects heterogeneity falls more steeply or rises more slowly than the true hazards for homogenous groups. Gail et al. (1984) showed that the unobserved heterogeneity tends to attenuate the estimated coefficients toward zero. On the other hand, standard errors and test statistics are not biased. For these reason a correction for the unobserved heterogeneity based on the gamma distribution of heterogeneity with mean one and variance θ is used. Incorporating the heterogeneity into the Weibull distribution results in the following hazard function $\lambda(t) = \lambda p(\lambda t)^{p-1}(S(t))^{\theta}$ where $S(t) = \left[1 + \theta(\lambda t)^{p}\right]^{-1/\theta}$ is the unconditional survival function, where the further θ deviates from zero, the greater the effect of the heterogeneity.

Different models are used here for the purpose of checking robustness. A plot of cumulative Cox- Snell residuals is used as a further step in specification checking and to assess the general fit. This method is based on examining the estimated values of the integrated hazard, which should look like a sample from the unit exponential distribution, if the model is correctly fitted. That is, the plot of the integrated hazard for generalized residuals should be a straight line with the slope equal to one (for a discussion of Cox-Snell residuals see Cox and Snell (1968) and Klein and Moeschberger (1997) pp. 329-332).

4. Estimation results

The statistical model used in this study depends on the nature of the problem in hand since the child mortality outcome is characterized by rapid change. This was shown in conducting a non-parametric analysis (see Appendix B for details). Graphical presentation of the hazard rate $\hat{\lambda}$ gives an insight on the distribution of the hazard, enabling the selection of a functional form for the parametric approach. The average hazard dropped by around 92 percent when moving from neonatal to the infant period. The declining pattern continues through the childhood period, where it levels out as duration or the age of the child increases in Egypt. Notice, however, that during one and half, two and three years of age peaks in the hazard are estimated. This result may be due to heaping effect where respondents may have had a tendency to round off the duration of their child survival. Another reason for the choice of this model is that some births have the chance to survive up to the end of five years. These types of observations are called censored. The use of hazard regression allows the inclusion of censored and uncensored observations. The plot of the log [-log (Kaplan-Meier survival estimates)] against the logarithm of the duration in Figure 1 that is close to a straight line suggests the Weibull model although further inspection is necessary.

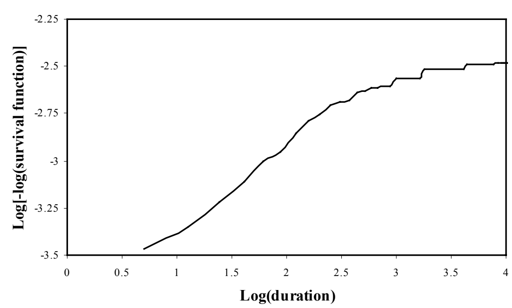
As mentioned earlier one of the problems facing the analysis of the determinants of child mortality is the rapid changing behavior of mortality pattern from one age to the next. The non-parametric analysis supports the fact of partitioning the age interval to neonatal and beyond, where the latter in turn can be split into infant and childhood periods. Starting by estimating a probit model for the neonatal case, this is because during the first month of age there is no variation in duration and thus the dependent variable is binary with the value zero if the infant survives and one if he or she dies during the first month. The results of this model are depicted in column I of Table 3. It should also be noted that the model was tested for heteroscedasticity following Greene (2000) pp. 829-831 where it was found that it does not pose a problem in our case. Thereafter, the analysis starts by a Cox type proportional hazard model with and without heterogeneity over sample clusters, for the full age interval, from more than one month to under 5 years and infant and childhood under the age of five.

In the parametric approach, piecewise constant exponential model (PCE) is used where time access i.e., age is split into three periods; zero to one, two to 12 and 13 to less than 60 months in order to study neonatal, infant and childhood mortality. Also a Weibull specification and a Weibull model with the correction for heterogeneity based on the gamma distribution are estimated.[7] In Tables 3 and 4 the parameter estimates, the

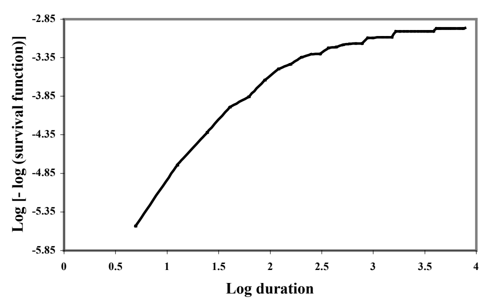
[7] Note that a log-logistic specification was also tried but the Weibull specification was found to be a better fit for the data.

standard errors, the marginal effects of the probit model and the relative risk of dying using various duration estimation methods are presented. When heterogeneity is not taken into consideration, the standard errors are estimated including the clustered nature of the sample (see Deaton (1997)). The error term is thus viewed as the sum of a cluster specific component and a child specific component. Nevertheless, although the error may be uncorrelated across children in different places of residence, it is unlikely to be uncorrelated across children in the same place of residence. For the Cox and the Weibull models, the standard errors are typically larger with the cluster option. The relative risk of dying is simply the exponential of the estimated parameter of the hazard model. It should be viewed as a scalar quantity that raises or lowers the underlying baseline hazard proportionally and this of course, depends on the sign of β.

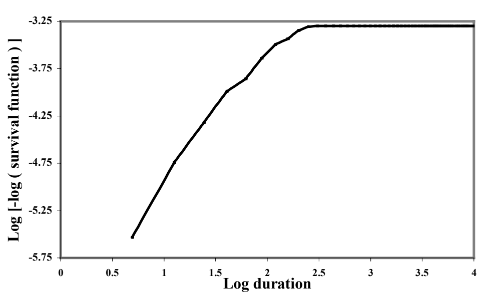
Figure 1: Log [-log (Kaplan-Meier survival estimates)]



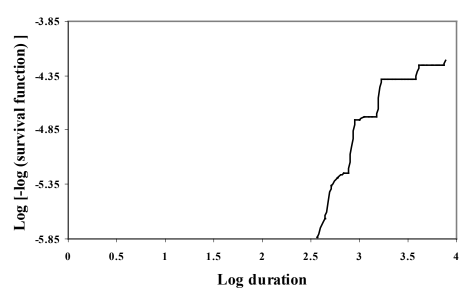
(a) Sample from birth to under five



(b) Sample from infant to under five



(c) Infant stage

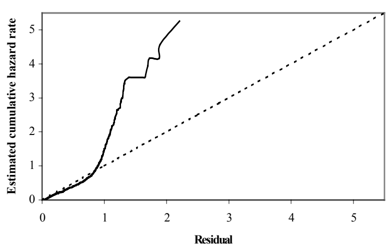


(d) Childhood stage

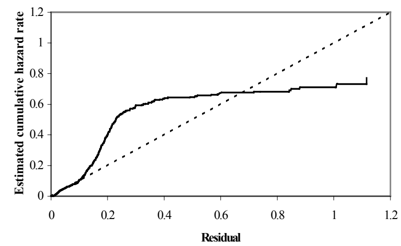
Evidence shows that the estimated value of the shape parameter θ in the Cox and Weibull regression with gamma heterogeneity deviates significantly from zero, suggesting that the data may contain unobserved heterogeneity. Thus incorporating heterogeneity into the Weibull regression is a more appropriate setting. In the Cox model the heterogeneity is considered on the cluster level. As for the Weibull, starting by taking into consideration heterogeneity on the household level, it is found that the correlation within families may be ignored in the childhood case but not in the infant case. So a step further is taken in the childhood period by considering heterogeneity on the cluster level and a significant frailty effect is detected. This result could be due to the fact that at a certain point in time the child outgrows the family specific frailty but is still affected by the unobserved characteristic of the place of residence. Checking the estimated integrated hazard for generalized residuals can further assess the goodness of the fit of the models. Noting that the Cox-Snell residuals are useful for assessing the fit of the parametric models, they are not very informative for Cox models estimated by partial likelihood (on this matter see Allison (1997) pp. 173-175). Figure 2 depicts the cumulative Cox-Snell residuals for the PCE, Weibull and Weibull with gamma heterogeneity models for various age grouping where a deviation from the 45º line indicates misspecification. The plots indicate that the Weibull with gamma heterogeneity for the sample from infant to under five years panel (e) of Figure 2 depicts a fairly better fit than its counter part using the full sample panel (c). This result further supports the consideration of the neonatal period separately. Panel (f) and (h) show that estimating the infant and the childhood period using a Weibull with gamma heterogeneity is quite a good fit.

A further diagnostic checking is called for to make the choice between semi parametric and parametric models. For this purpose a likelihood ratio test is used. Although models are not nested this test could still give an indication for the preferred models. The conclusion is that a Weibull with gamma heterogeneity of infant and childhood periods are preferred. Since the semi parametric and the parametric results give qualitatively the same results it could be also concluded that the restriction imposed by the chosen parametric model do explain fairly well the data and does not interrupt the results.

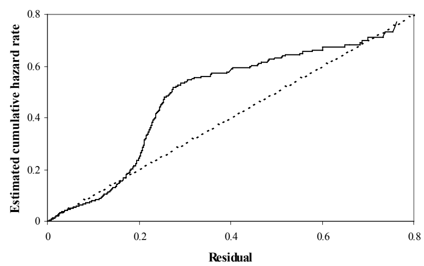
Figure 2: Cumulative Cox-Snell residuals.



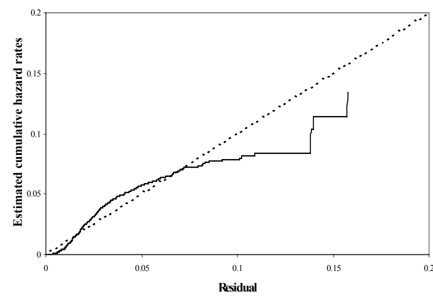
(a) Picewise constant exponential model, full sample



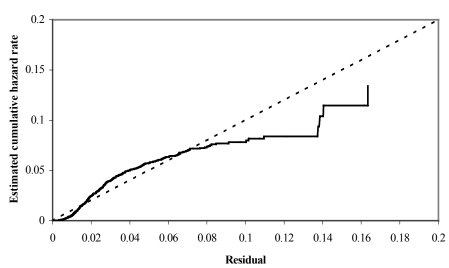
(b) Weibull, full sample



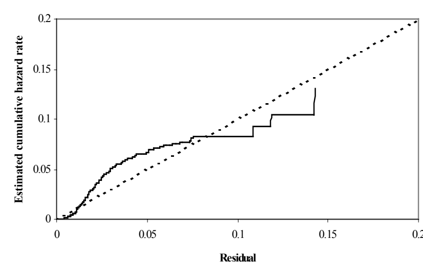
(c) Weibull with gamma heterogeneity, full sample



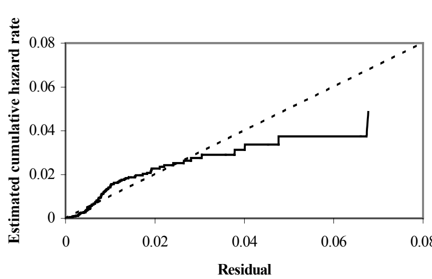
(d) Weibull, sample from infant to under five



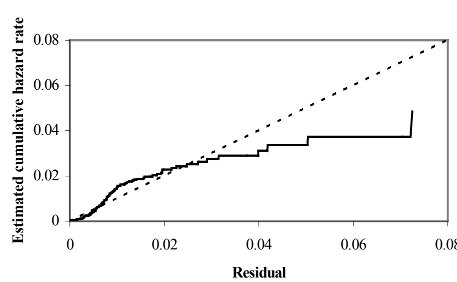
(e) Weibull with gamma heterogeneity, sample from infant to under five



(f) Weibull with gamma heterogeneity, infant stage



(g) Weibull, childhood stage



(h) Weibull with gamma heterogeneity, childhood stage

A number of potentially confounding variables (Table A1) were measured and controlled for in the analysis by first estimating a full model. Many of the effects are insignificant due to the small number of transitions compared to the number of observations to explain so many variables. The indicator for maternal employment activity is omitted since it failed to show impact in the analysis, possibly because the information available is not specific to the periods immediately preceding or following particular births. The birth order variables are correlated with the total number of children ever born. Another pair of variables that have shown some correlation is the preceding birth interval with the first birth variable. This was expected since the first order child has no preceding birth interval. Therefore a subset of covariates where selected following the literature. In this manner a reduced model was estimated. Only the results from the parsimonious model are presented below.[8]

The following discussion on the impacts of explanatory variables will be based on the three-part model composed of the probit estimation for the neonatal case and the Weibull regression with gamma heterogeneity for the infant and childhood cases which are depicted in column I Table3 and column II and III Table 4, respectively. This specification reveals some interesting differences between the impacts on the neonatal, infant and subsequent survival amongst certain environmental and socio-economic determinants. Furthermore, it has revealed to be a more robust setting. Notice however, that it seems as expected, that models not accounting for heterogeneity underestimates the effects of the regressors contrasting to the ones with heterogeneity. As shown in

[8] One of the variations that are considered for these models is through the construction of an alternative measure of the standard of living indicator (SLI). Dummy variables are specified for each distinct value of the SLI with the lowest value treated as the omitted category. This specification is used to allow the sum of the index item to exert a non-linear influence on mortality. The incorporation of these dummies does not seem to improve the overall results and the parameter estimates are insignificant. On the other hand another set of four dummies was constructed taking into consideration the type of flooring, electricity, refrigerator and a car/motorcycle. Where those possessions were varied to form the following dummies; (i) households with finished floor, no electricity, no refrigerator, and no car/motorcycle, (ii) finished floor, electricity, no refrigerator, and no car/motorcycle, (iii) finished floor, electricity, refrigerator, and no car/motorcycle, and (iv) finished floor, electricity, refrigerator, and car/motorcycle. This specification did not give any new insights to how wealth affects mortality. It is therefore chosen to only present models with the aggregated SLI variable. A caveat related to the categories of the mother's age is that those variables have some kind of correlation with each other. This division of the mother's age variable was meant to capture the U-shape pattern exhibited by these categories. The parameter estimates of these categories showed to be

Table 4, access to municipal water into the residence decreases infant mortality by around 27 percent. Access to public water decreases the risk of death by 27 percent in general and a decrease of 31 percent is encountered in the infant stage while its effect is significant in the neonatal case it is insignificant in the childhood case. Suggesting that the use or the availability of municipal water decreases the risk of mortality as opposed to a non-municipal water source. A counter intuitive result though with an insignificant effect reveals that access to a modern facility increases the changes in the neonatal mortality. This may be due to low hygienic awareness and the misuse of modern sanitation facilities.[9] This finding is supported by the results of Lavy et al. (1996) who found that access to poor water and sanitation in the urban Ghana decreases the risk of mortality. As for the childhood mortality the modern facility reduces it by 68 percent. Here, sanitation is found to have a more pronounced impact on mortality than water.

This is consistent with a study including eight countries Burundi, Bolivia, Ghana, Guatemala, Morocco, Sri Lanka, Togo and Uganda, Esrey (1996) found that the effects of improved water were less pronounced than those for sanitation. Benefits from improved water occurred only when sanitation was enhanced and only when optimal water was present. Similarly, Esrey et al. (1992) found that sanitation has a larger impact than water in improving health.

Living in urban areas decreases the mortality risk by around 30 percent as opposed to living in rural areas. Taking interactions between water supplies and urbanicity into account, it was found that residential water in urban areas has a significant role in mortality risk reduction in different age groups as compared to its role in rural areas. The mortality is reduced in the infant and childhood cases by 61 and 90 percent, respectively. This result suggests an important role of policy in eliminating the urban-rural disparities. A mother that has completed secondary school and higher reduces infant mortality by 63 percent as opposed to an illiterate one. In research concerning child mortality, education

insignificant. Therefore, the inclusion of a continuous variable of mother age was chosen for parsimonious reasons. However, this choice did not change much the estimation.

[9] A thorough examination of the characteristics of our sample shows that 10 percent of the households possessing a modern facility have no access to residential water.

Table 3: Neonatal probit model Cox proportional hazard estimation

Variables	***Probit model Neonatal***	***Cox Proportional Hazard Model***							
		Without heterogeneity				With gamma heterogeneity			
		Infant	Child	All	1-<60	Infant	Child	All	1-<60
	I								
Constant/ θ	-0.51***					0.31***	0.68**	0.34	0.32***
	(0.154)					(0.112)	(0.415)	(0.067)	(0.094)
	-0.024								
I. Environmental conditions variables									
Tap water	-0.08	-0.26***	-0.315	-0.265	-0.29***	-0.3**	-0.336	-0.27***	-0.3***
into residence	(0.06)	(0.094)	(0.26)	(0.214)	(0.095)	(0.126)	(0.24)	(0.089)	(0.113)
(yes)‡		*0.77*			*0.752*	*0.744*		*0.766*	*0.737*
Public tap	-0.185**	-0.31***	-0.143	-0.25**	-0.29**	-0.35*	-0.135	-0.29**	-0.29*
water (yes)	(0.099)	(0.06)	(0.4)	(0.11)	(0.13)	(0.19)	(0.318)	(0.135)	(0.165)
	- 0.0072	*0.732*		*0.776*	*0.749*	*0.708*		*0.748*	*0.749*
Modern	0.114	-0.31***	-1.11*	-0.041	-0.28**	-0.095	-1.14*	-0.085	-0.274
facility (yes)	(0.13)	(0.059)	(0.65)	(0.045)	(0.121)	(0.287)	(0.64)	(0.19)	(0.259)
		0.735	0.33		*0.753*		*0.319*		*0.735*
Traditional	0.217**	0.153	0.008	0.27***	0.117	0.155	-0.05	0.174	0.076
facility (yes)	(0.095)	(0.107)	(0.07)	(0.04)	(0.08)	(0.17)	(0.285)	(0.122)	(0.149)
	0.0091			*1.303*					
II. Socioeconomic variables									
Urban	-0.0023	-0.36***	-0.23**	-0.27***	-0.31***	-0.36**	-0.23	-0.2**	-0.33**
residence	(0.066)	(0.047)	(0.11)	(0.066)	(0.088)	(0.154)	(0.294)	(0.105)	(0.139)
(yes)		*0.696*	*0.798*	*0.764*	*0.731*	*0.696*		*0.82*	*0.721*
Low	-0.163**	-0.011	-0.044	-0.188	0.092	0.123	-0.023	-0.066	0.108
education	(0.076)	(0.21)	(0.37)	(0.17)	(0.243)	(0.13)	(0.27)	(0.095)	(0.12)
(yes)	*-0.007*								
Medium	0.0018	-0.38***	0.15	-0.181	-0.151	-0.221	0.158	-0.071	-0.133
education	(0.09)	(0.112)	(0.62)	(0.203)	(0.227)	(0.197)	(0.365)	(0.128)	(0.174)
(yes)		*0.686*							
High	0.076	-0.99***	-0.522	-0.52***	-0.89***	-0.97***	-0.522	-0.48***	-0.87***
education	(0.086)	(0.117)	(0.39)	(0.017)	(0.139)	(0.225)	(0.42)	(0.128)	(0.2)
(yes)		*0.373*		*0.592*	*0.409*	*0.377*		*0.616*	*0.42*
Standard of	-0.05***	-0.05***	-0.2***	-0.06***	-0.04***	-0.012	-0.16**	-0.06***	-0.04
living	(0.018)	(0.013)	(0.006)	(0.011)	(0.005)	(0.035)	(0.066)	(0.024)	(0.031)
indicator	*-0.0023*	*0.952*	*0.856*	*0.943*	*0.959*		*0.853*	*0.937*	
III. Demographic variables									
Mother age at	0.0013	-0.02***	-0.005	-0.0009	-0.0043	-0.002	-0.005	-0.0001	-0.004
birth	(0.004)	(0.008)	(0.02)	(0.008)	(0.006)	(0.008)	(0.016)	(0.006)	(0.007)
		0.977							
Gender	0.115**	-0.35*	-0.4***	-0.129	-0.34***	-0.32***	-0.41**	-0.127*	-0.35***
(male=yes)	(0.052)	(0.2)	(0.14)	(0.12)	(0.12)	(0.103)	(0.2)	(0.07)	(0.09)
	0.005	*0.705*	*0.669*		*0.709*	*0.724*	*0.666*	*0.881*	*0.703*
IV. Behavioral variables									
Breast-	-1.61***	-1.37***	-1.1***	-2.83***	-0.82***	-0.95***	-1.14***	-2.15***	-0.86***
feeding (yes)	(0.062)	(0.008)	(0.1)	(0.054)	(0.183)	(0.192)	(0.2)	(0.07)	(0.178)
	-0.274	*0.254*	*0.336*	*0.059*	*0.443*	*0.388*	*0.32*	*0.116*	*0.424*
Log likelihood	-1306	-2791.1	-907.4	-7221	-4430.2	-3496.8	-905.13	-7614	-4419.6

† The first number in a cell is the parameter estimates; number in parenthesis is the standard error and Italic number stands for the marginal effects in case of the probit specification and mortality risk elsewhere.

‡ (yes) refers to a dichotomous variable indicating that the value 1 is taken by the variable name (e.g.. Tap water into residence (yes) = Dichotomous variable indicating that the household has municipal water piped into residence).

*** Means that the estimate is significant at 1 percent

** Means that the estimate is significant at 5 percent

* Means that the estimate is significant at 10 percent.

Table 4: Parametric estimation

Variables	*Weibull without heterogeneity*				*Weibull With gamma heterogeneity*			
	Infant	Child	All	1-<60	Infant	Child	All	1-<60
					II	**III**		
Constant	-3.8***	-7.12***	-1.4***	-3.74***	-3.69***	-7.06***	-0.387	-3.65***
	(0.199)	(0.669)	(0.025)	(0.162)	(0.383)	(0.669)	(0.285)	(0.342)
P	0.553***	1.302	0.487	0.642***	0.588***	1.316	0.522	0.677***
	(0)	(0.046)	(0.052)	(0.003)	(0.029)	(0.116)	(0.0465)	(0.029)
Θ					3.44***	0.81***	6.026***	2.46***
					(0.84)	(0.45)	(0.33)	(0.58)
I. Environmental conditions variables								
Tap water into residence (yes)‡	-0.28***	-0.311	-0.278	-0.28***	-0.32**	-0.336	-0.362	-0.32***
	(0.059)	(0.256)	(0.246)	(0.1)	(0.137)	(0.244)	(0.266)	(0.12)
	0.758			*0.754*	*0.727*			*0.729*
Public tap water (yes)	-0.34***	-0.139	-0.26***	-0.29**	-0.374*	-0.126	-0.355*	-0.32*
	(0.035)	(0.31)	(0.11)	(0.135)	(0.207)	(0.32)	(0.19)	(0.035)
	0.713		*0.773*	*0.749*	*0.688*		*0.701*	*0.726*
Modern facility (yes)	-0.112	-1.12**	-0.035	-0.29**	-0.192	-1.14*	-0.092	-0.333
	(0.248)	(0.57)	(0.044)	(0.127)	(0.308)	(0.64)	(0.11)	(0.274)
		0.319		*0.748*		*0.319*		
Traditional facility (yes)	0.19**	-0.014	0.291***	0.12	0.182	-0.037	0.328***	0.122
	(0.08)	(0.305)	(0.012)	(0.076)	(0.19)	(0.286)	(0.63)	(0.16)
	1.206		*1.338*				*1.389*	
II. Socioeconomic variables								
Urban residence (yes)	-0.35***	-0.228	-0.28***	-0.31***	-0.36**	-0.233	-0.33***	-0.34**
	(0.084)	(0.289)	(0.074)	(0.093)	(0.158084)	(0.296)	(0.073)	(0.139)
	0.706		*0.758*	*0.731*	*0.697*		*0.719*	*0.712*
Low education (yes)	0.12	-0.052	-0.232	0.091	0.14	-0.03	-0.407	0.115
	(0.21)	(0.27)	(0.178)	(0.24)	(0.147)	(0.27)	(0.251)	(0.13)
Medium education (yes)	-0.194*	0.175	-0.168	-0.121	-0.207	0.19	-0.252	-0.137
	(0.107)	(0.358)	(0.216)	(0.23)	(0.215)	(0.366)	(0.249)	(0.188)
	0.823							
High education (yes)	-0.97***	-0.5	-0.53***	-0.87***	-0.99***	-0.5	-0.676***	-0.9***
	(0.037)	(0.389)	(0.008)	(0.133)	(0.235)	(0.421)	(0.076)	(0.207)
	0.377		*0.587*	*0.418*	*0.372*		*0.509*	*0.408*
Standard of living indicator	-0.02***	-0.16**	-0.07***	-0.05***	-0.017	-0.16**	-0.0673***	-0.044
	(0.004)	(0.073)	(0.013)	(0.003)	(0.039)	(0.066)	(0.003)	(0.034)
	0.981	*0.85*	*0.933*	*0.956*		*0.85*	*0.935*	
III. Demographic variables								
Mother age at birth	-0.002	-0.005	0.001	-0.004	-0.003	-0.005	0.0009	-0.005
	(0.003)	(0.015)	(0.008)	(0.006)	(0.009)	(0.016)	(0.007)	(0.008)
Gender (male=yes)	-0.321*	-0.41**	-0.114	-0.35***	-0.362***	-0.41**	-0.279***	-0.39***
	(0.103)	(0.187)	(0.123)	(0.118)	(0.11)	(0.199)	(0.106)	(0.098)
	0.725	*0.664*		*0.707*	*0.696*	*0.664*	*0.757*	*0.675*
IV. Behavioral variables								
Breast-feeding (yes)	-0.87***	-1.08***	-2.94***	-0.8***	-1.01***	-1.15***	-5.127***	-0.91***
	(0.222)	(0.2)	(0.063)	(0.183)	(0.224)	(0.2)	(0.335)	(0.203)
	0.408	*0.316*	*0.053*	*0.448*	*0.363*	*0.316*	*0.006*	*0.403*
Log likelihood	-2211.8	-603.46	-3859.8	-2617.5	-2188.99	-600.62	-3703.65	-2596.3

† The first number in a cell is the parameter estimates; number in parenthesis is the standard error and Italic number stands for the mortality risk.

‡ (yes) refers to a dichotomous variable indicating that the value 1 is taken by the variable name (e.g.. Tap water into residence (yes) = Dichotomous variable indicating that the household has municipal water piped into residence).

*** Means that the estimate is significant at 1 percent

** Means that the estimate is significant at 5 percent

* Means that the estimate is significant at 10 percent.

is thought to be one of the most important factors in reducing the level of child mortality. Cochrane (1979, pp. 93-98) found, from surveying 16 studies on the subject, that female education increases the chances of infant and child survival in two thirds of the studies. The SLI generally marks a significant effect on the neonatal and childhood mortality reduction.

Moreover, it is widely believed that male mortality is higher due to biological disadvantages. This is apparent in the neonatal case where the gender effect is significant and positive marking the male disadvantages. However, on the whole our sample shows higher female mortality indicating gender discrimination together within the infant and childhood cases. When the offspring is male he has an advantage of a decreased risk of death than being a female ranging from 30 to 34 percent. In the absence of social security males are seen to be a long-term investment to depend on in older days (Dasgupta (2002)).

Many of the factors listed in Tables 3 and 4 serve as proxies for health behavior and wealth together with the consideration of the unobserved heterogeneity. Although controlling for these and other potential confounders, the main results on water and sanitation apply. Thus, it is unlikely that the differences found are due to some inherent characteristic of health or wealth associated with water quality or having a latrine.

5. Discussion and concluding remarks

Using DHS Egypt 1995 and a three-part model while studying child mortality, this paper assesses whether improvements in water and sanitation services are leading to a decrease in mortality of children under the age of five. The analysis is conducted using a probit model specification for the neonatal case. Non-parametric, semi parametric and parametric duration model specifications are used for different age intervals including infant and children less than five years. One of the problems in studying child mortality is that biological and social entities usually differ in ways that are not fully captured by the model. This unobserved heterogeneity can produce misleading estimates of hazard functions and attenuated estimates of covariate effects. Thus, the transition regression models are estimated with and without heterogeneity. Assessment of model fit by means of Cox-Snell residuals and likelihood ratio tests reveals that a Weibull model with gamma heterogeneity better explains the second and the third part of the model.

The three-part model uncovers some interesting differences between the impacts on the neonatal, infant and subsequent survival amongst certain environmental and socio-economic determinants. The results show a negative relationship between access to municipal water and mortality. The advantages of having a modern facility, prevails in the childhood case with a 68 percent reduction in mortality risk. In a quite similar setting with an application to Eritrea Woldemicael (1998) shows that household environment (water supply and toilet facility) is large and statistically significant during the post-neonatal and childhood periods while the effect totally disappears during the neonatal period. Trussell and Hammerslough (1983) in a hazard model analysis for child mortality in Sri Lanka, found that improvement in the type of toilet facility reduces mortality while the source of water supply was found insignificant and hence was excluded from their analysis. Ridder and Tunali (1999) found that access to piped water and toilet facilities did not appear to have much of an impact on the child mortality risk in Malaysia.

Moreover, the analysis with the extended dummies of the place of residence show that living in Upper Egypt implies a relative risk of dying ranging between 88 and 160 percent higher than living in another area. Knowing the characteristics of the region this is not a surprising result. First of all, the services in that area are not well taken care of or may not exist. Taking water supply for instance, municipal water is only available to 67 percent of the residents of this area against 99 and 86 percent in urban governorates and Lower Egypt residents, respectively. Second, this region also marks the highest rate of illiteracy together with low standard of living. Last but not least, the recently completed report 'Poverty reduction in Egypt-Diagnosis and strategy' indicates that about 17 percent of the Egyptian population was poor in year 2000. By region, only five percent of the population of Cairo was poor, while poverty rates in several governorates in Upper Egypt exceeded 30 percent (GOE/World Bank (2002)). In general, there is a disparity in rural-urban mortality that largely results from a combination of poor water and sanitation services, higher illiteracy rate and lower standards of living. Mother education showed to be of importance to reduce child mortality. The results also exhibit gender discrimination with preference to male offspring. This finding confirms the ones of Ammar (1954), Ayrout (1963) and others that Egypt is a strongly patriarchal society in which the male children receive preferential treatment from an early age. Though Casterline et al. (1989) discuss this phenomenon their sample does not show evidence to confirm it.

In the light of these results increasing awareness of the Egyptian population relative to health care and hygiene is a prerequisite to decrease child mortality risk. More concern should be given to the region of Upper Egypt. Furthermore, if water infrastructure is extended as much as possible to the whole country then child mortality would be reduced.

In addition to the direct benefits in terms of increased health and reduced mortality, there are also possible indirect effects, such as a reduced fertility and population growth. Indeed, in line with the theory of the effect of infant and child mortality on fertility as argued by Wolpin (1997), a high child mortality may lead to a high fertility where mothers give birth to many children suspecting that the probability of losing some of them is high. This fact is corroborated with Al-Qudsi's (1998) findings for the Arab countries that infant mortality has a positive influence on fertility.

References

Al-Qudsi S. (1998), 'The demand for children in Arab countries: Evidence from panel and count data models', *Journal of Population Economics,* **11**: 435-452.

Allison P. D. (1997), *Survival analysis using the SAS system: A practical guide*, Cary, NC: SAS Institute Inc.

Aly H.Y., R. Grabowski (1990), 'Education and child mortality in Egypt', *World Development,* **18**: 733-742.

Ammar H. (1954), *Growing up in an Egyptian village,* London, Routledge and Kegan Paul.

Ayrout H.H. (1963), The Egyptian peasant, Boston, Beacon Press.

Buehler J.W., J.C. Kleinman, C.J.R. Hogue, L.T. Strauss and J.C. Smith (1987), 'Birth weight-specific infant mortality, United States, 1960 and 1980', *Public Health Reports,* **102:** 151-161.

Casterlin J.B., E.C. Cooksey and A.F. Ismail (1989). 'Household income and child survival in Egypt', *Demography,* **26**: 15-35.

Central Agency for Public Mobilization and Statistics (CAPMAS), (1996), *Estimate of population size, Arab republic of Egypt on 1-1-96* [in Arabic], Cairo: Central Agency for Public Mobilization and Statistics.

Cochrane S. (1979), *Fertility and education: What do we really know?* Baltimore MD: the Johns Hopkins University Press.

Cox D.R. (1972), 'Regression models and life-tables', *Journal of the Royal Statistical Society, Series B*, **74**: 187-220.

Cox D.R. (1975), 'Partial likelihood', *Biometrica,* **60**: 269-276.

Cox D.R. and D. Oakes (1984), *Analysis of survival data,* Chapman and Hall, London.

Cox D. R. and J. Snell (1968), 'A General Definition of Residuals', *Journal of the Royal Statistical Society. Series B*, 30: 248-275.

Dasgupta P. (2002), 'A model of fertility transition', in B. Kriström, P. Dasgupta K. -G. Löfgren eds., *Economic Theory for the Environment Essays in Honor of Karl-Göran Mäler*, Edward Elgar Publishing Ltd, 119-128.

Deaton A. (1997), *The analysis of household surveys: a microeconometric approach to development policy*, Baltimore M.D., Johns Hopkins University Press

Esrey S.A. and J.P. Habicht (1986), ' Epidemiologic evidence for health benefits from improved water and sanitation in developing countries', *Epidemiology Review,* **8**: 117-128.

Esrey S.A. and J.P. Habicht (1988), 'Maternal literacy modifies the effect of toilets and piped water on infant survival in Malaysia', *American Journal of Epidemiology,* **127**: 1079-1087.

Esrey S.A., J.P. Habicht and G. Casella (1992), 'The complementary effect of latrines and increased water usage on the growth of infant in rural Lesotho', *American Journal of Epidemiology,* **135**: 659-666.

Esrey S.A. (1996), 'Water, waste and well-being: A multicountry study', *American Journal of Epidemiology,* **143**: 608-623.

Flinn C. and J. Heckman (1982), 'New methods for analyzing structural models of labor force dynamics', *Journal of Econometrics,* **18**: 115-168.

Gail M.H., S. Wieand and S. Piantadosi (1984), 'Biased estimates of treatment effect in randomized experiments with nonlinear regression and omitted covariates', *Biometrika*, **71**: 431-444.

GOE/World Bank, (2002) *Poverty reduction in Egypt-Diagnosis and strategy*.

Greene W.H. (2000), *Econometric Analysis*, 4th ed. Prentice-Hall, London.

Guilkey D.K. and R.T. Riphahn (1998), 'The determinants of child mortality in the Philippines: Estimation of a structural model', *Journal of Development Economics*, **56**: 281-305.

Kiefer N.M. (1988), 'Economic duration data and hazard functions' *Journal of Economic Literature*, **XXVI**: 646-679.

Klein, J. P. and M. L. M Moeschberger (1997), *Survival Analysis*, 1st ed. New York, Springer.

Lavy V., J. Strauss, D. Thomas and P. de Vreyer (1996), 'Quality of health care, survival and health outcomes in Ghana', *Journal of Health Economics,* **15**: 333-357.

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

Maddala G.S. (1993), Limited-dependent and qualitative variables in econometrics, *Econometric Society Monographs*, **3**. New York, Cambridge University Press.

Olsen R.G. and K. Wolpin (1983), 'The impact of exogenous child mortality on fertility: a waiting time regression with dynamic regressors', *Econometrica,* **51**: 731-749.

Pebley A.R. and P.W. Stupp (1987), 'Reproductive patterns and child mortality in Guatemala', *Demography,* **24**: 43-60.

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

Trussell J. and S. Preston (1981), 'Estimating the covariates of childhood mortality from retrospective reports of mothers' Seminar on Methodology and Data Collection in Mortality Studies, *International Union for the Scientific Study of Population*. Djakarta, Senegal.

Trussell J. and C. Hammerslough (1983), 'A hazards-model analysis of the covariates of infant and child mortality in Sri Lanka', *Demography,* **20**: 1-26.

Woldemicael G. (1998), 'The effects of water supply and sanitation on childhood mortality in urban Eritrea', Stockholm Research Reports in Demography No. 127, Demography unit, Stockholm University.

Wolpin K. (1984), 'An estimable dynamic stochastic model of fertility and child mortality', *Journal of Political Economy,* **92**: 852-874.

Wolpin K. (1997), 'Determinants and consequences of the mortality and health of infants and children', in M. Rosenzweig and O. Stark, eds., *Handbook of Population and Family Economics*, Elsevier Science B.V., 483-557.

World Bank (2002), 'Arab republic of Egypt cost assessment of environmental degradation', Report No. 25175-EGT.

World Health organization (2002), *The world health report 2002: Reducing risks, promoting healthy life*, Geneva, the World Health Organization.

Appendix A

Table A1: Distribution (%) of potentially confounding variables among children in Egypt

Variable	*Neonatal*	*Infant*	*Childhood*
Socioeconomic variables			
Urban Residence (yes) †	40.2	41	41.4
Urban governorates (yes)	15.4	15.8	16.3
Lower Egypt urban (yes)	8.9	9.2	9.2
Lower Egypt rural (yes)	22.3	22.1	22.7
Upper Egypt urban (yes)	10.2	10.1	9.9
Upper Egypt rural (yes)	34.0	33.5	32.3
Frontier governorates (yes)	9.2	9.4	9.6
No education	35.1	34.5	34.1
Low education	12.8	12.9	12.7
Medium education	15.4	15.3	14.9
High education	36.7	37.4	38.3
Mother's employment (yes)	17	17.1	17.8
Standard of living indicator	(4.3)‡	(4.3)	(4.3)
Mean person per room	(1.7)	(1.7)	(1.7)
Demographic variables			
Mother age at birth	(22.14)	(22.14)	(21.97)
Mother's age ≤ 19*	33.7 (18)	33.5 (18)	34.7 (17.9)
Mother's age 20-34*	65.4 (24.1)	65.6 (24.1)	64.4 (23.9)
Mother's age ≥ 35*	0.9 (37.3)	0.9 (37.3)	0.9 (37.5)
Divorced or widow (yes)	1.4	1.4	1.5
First birth	65.2	66.1	72
2nd birth	29.5	29.2	25.8
≥ 3rd birth	5.3	4.7	2.1
Gender (male=yes)	51.9	51.9	51.8
Behavioral variables			
Birth interval	(8.6)	(8.3)	(6.3)
Breast-feeding (yes)	94.2	95.5	71
Number of children	(1.8)	(1.8)	(1.8)
Maternity care (yes)	64.6	64.7	64.6

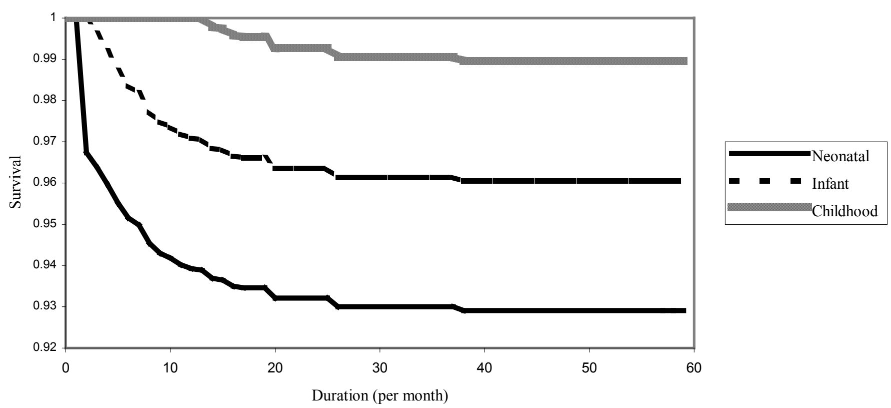
† (yes) refers to a dichotomous variable indicating that the value 1 is taken by the variable name (e.g.. Tap water into residence (yes) = Dichotomous variable indicating that the household has municipal water piped into residence).

‡ Numbers in parenthesis indicates the mean of the variable not the percentage of occurrence.

Appendix B: Kaplan-Meier estimates of the survival rates

Figure B1 illustrate the Kaplan-Meier estimates of the survival rate $\hat{S}(t)$ of the age grouping. The deaths represent the number of completed spells of duration t_i. During the first month 132 deaths have occurred which give rise to a risk equivalent to 3.32 percent of ending the spell. The survival rate of a spell lasting 1 month or more is 96.74 percent as could be seen from Figure B1. In Table B1 the average hazard dropped by 92 percent when moving from neonatal to the infant period.

Figure B1: The Survival Function of the Age Grouping



In Figure B1, the survival probability during the childhood period is higher than in the infant period, which is in turn higher than the survival probability during the neonatal period. This is an expected result since the hazard risk of dying is decreasing with age. Taking as an example the 20th month the childhood survival function indicates around 99 percent probability of surviving given that the child survives the infant age. This probability is 96 and 93 percent for the child given that he/she survives the neonatal period and the general chance of survival, respectively.

Table B1: Sample size and average hazard rate.

	Neonatal	***Infant***	***Childhood***
Sample size (No. of children)	4165	3854	2761
Mortality prevalence (No. of children)	132	102	23
Mortality prevalence (%)	3.2	2.7	0.8
Average hazard rate			
Neonatal	3.32	-	-
Infant	-	0.27	-
Childhood	-	-	0.02