Estimating Changes in Orphanhood

after the January 2010 Haiti Earthquake

by

Thomas W. Pullum

Bureau for Global Health

U.S. Agency for International Development

tpullum@usaid.gov

and

Department of Sociology

The University of Texas at Austin

tom.pullum@mail.utexas.edu

September 10, 2010

Acknowledgments

The author is grateful to Jeffrey Hammer, Elizabeth Frankenberg, and Duncan Thomas for comments on an early draft, and to Sam Preston and Jessica Ho for sharing related bibliographies.

Estimating Changes in Orphanhood

after the January 2010 Haiti Earthquake

Abstract

Following an environmental disaster or emergency such as an earthquake, flood, or tsunami, estimates of the overall number of deaths and population affected are often developed rapidly, but with no detail on the age, sex, or household status of those who died. This paper develops a simple model to estimate the number of children who survived the emergency but lost one or both parents, or their household head. Such children are especially vulnerable to separation or abandonment, psychological trauma, and exploitation and have a special need for assistance. The model uses data that are often available from surveys and from international estimates and projections, and assumptions about the likely association between children and their caregivers in the probabilities of dying or surviving the emergency. The model is applied to the aftermath of the devastating earthquake in Haiti on January 12, 2010. It is estimated that approximately 101,000 children under 18 died, 125,000 children became single orphans, and 12,000 children became double orphans. About one-third of the new double orphans lost both parents in the earthquake. About two-thirds had previously lost one parent, and experienced the death of their only surviving parent.

Estimating Changes in Orphanhood

after the January 2010 Haiti Earthquake

Introduction

For children, many types of welfare—or, conversely, vulnerability—hinge critically on household and kinship structure. In the aftermath of a major disaster, such as an earthquake, tsunami, hurricane or typhoon, especially if it occurs against a backdrop of widespread poverty and low levels of education, many children become vulnerable, or *more* vulnerable, because of the deaths of parents or household heads. If a parent or household head dies, a child is at increased risk of being abandoned or separated from surviving kin and household members. A range of negative outcomes, both in the short term and the long term, become more likely.

Following such a disaster, two numbers are often estimated quickly, and subsequently updated: the number of deaths and the size of the affected population. The number of deaths typically cannot be disaggregated by age and sex. The affected population corresponds to the population living inside a geographic zone and can be interpreted as the population who experienced more than inconvenience and had some risk of displacement, serious injury, or death. Some background on general issues in estimating mortality and other impacts of a disaster are described by Checci and Roberts (2008) and Cutter (1996).

This paper will show how, with a limited amount of auxiliary data that are often available, it is possible to develop estimates of the number of children who survived the disaster but experienced a major change in their household or kinship structure. The model will be applied to the aftermath of the devastating earthquake that occurred in Haiti on January 12, 2010.

Only a limited amount of previous research appears to have been done on the effect of a disaster on household structure, particularly from the perspective of children. The most recent example of such research was by Frankenberg, Gillespie, Preston, Sikoki, and Thomas (2009) following the Indonesian tsunami of December 26, 2004. In the provinces of Aceh and North Sumatra, their analysis focused on the age-sex pattern of mortality and on the association between outcomes for members of the same household. Other analyses of mortality levels and differentials, as well as casualties and displacement, following the Indonesian tsunami were carried out by Doocy et al. (2007), Nishikiori et al. (2006), Rofi et al. (2006), and Yeh (2010).

Model

Consider, to begin with, children under age 18[1] who live in households, and not in institutions such as orphanages or residential care centers.[2] Each such child has one, and only one, household head, who may or may not be a parent of the child.[3] For each child we can conceptualize a pairwise relationship with a household head. Household heads may have no children in their household, or they may have several children. If there are several children in the household, then the death of the household head will produce several children whose household head has died. The head/child pairs, as units of analysis, are not statistically independent of one another.

We conceptualize a population at risk, in which the probability of death from the emergency is uniformly *p*, regardless of whether a person is an adult or a child, and regardless of place of residence or other characteristics. Below, this assumption will be relaxed.

First consider the sub-population of households that contain exactly two children. Say that H_2 is the number of such households; H_2 will also be the number of household heads for such households. The expected number of these household heads who will die in the emergency is pH_2, and the expected number of children whose head will die is $2pH_2$, because each head is associated with exactly two children.

Assume that the probability a child in the household will die is independent of whether the head has died, an assumption that will also be relaxed below. Then for each child whose household head dies, the probability that the child survives is 1-*p*. The expected number of surviving children, in all of the two-child households in which the head dies, will be $S_2 = (1-p)(2pH_2) = p(1-p)(2H_2) = p(1-p)C_2$, where $C_2 = 2H_2$ is the number of children in two-child households before the disaster. In general, for *k*=1,2,…, the expected number of surviving children in *k*-child households in which the head dies will be $S_k = p(1-p)C_k$.

Adding children in all sizes of households, we have

$$S = p(1-p)C \qquad (1)$$

where *C* is the total number of children in the household population and *S* is the expected number of surviving children whose head dies. Alternatively, (1) can be written as

$$S = p(1-p)Pc \qquad (2)$$

where P is the population at risk and c is the proportion of that population which consists of children of the specified type, because $C = Pc$.

This result can be generalized to roles other than the household head, so long as the role has at most one occupant for each child. For example, if C is the number of children who have a living mother before the emergency, then S, calculated from (1), will be the expected number *after* the emergency who survived but became maternal orphans.

The model applies a well-established practice of describing a population of individuals as a population of pairs of individuals of different types. With the Kermack-McKendrick (1927) equations, the progression of an epidemic, through contacts between infected and susceptible cases, is proportional to the number of possible pairings of these two types of cases. A similar approach was used by Coleman, Katz, and Menzel (1957) to model the social diffusion of new information or behavior. Kendall (1962) used all possible pairs of cases to define a generalized correlation coefficient. Pullum and Peri (1999) approached the measurement of homogamy by comparing the observed number of pairings of a man of a particular type and a woman of a particular type with all possible pairings of those types.

Several applications of the procedure to the situation in Haiti after the earthquake will be given below, but it may be helpful to give a specific example here. Since the latter part of February, 2010, the official estimate of the number of deaths due to the earthquake has been D=230,000 and the estimated size of the "affected" population has been P=3,000,000. Interpreting the "affected" population to be the population at some risk of dying, the overall probability of death (within this population) is estimated as $p = D/P = 0.0767$. The U.S. Census Bureau's International Data Base (IDB) estimates that immediately before the earthquake, a proportion .4407 of the total population of Haiti was age 0-17. Using the same proportion for the earthquake zone, the estimated number of children in the affected area who survived but lost their household head is $S = p(1-p)Pc = 93{,}590$. For reporting purposes, to avoid the appearance of unjustifiable precision, such a number would be rounded to 94,000.[4]

Extensions and sensitivity to assumptions

If additional data are available, the model can be refined and assumptions can be relaxed. By reviewing possible extensions, we can assess the model's sensitivity to its assumptions or to inaccurate data. The modifications described here could be used in various combinations.

Sensitivity to estimates of the population at risk

If the probability of dying is estimated as $p = D/P$, in which a fairly firm estimate of the number of deaths, *D*, is combined with a rather vaguely specified population at risk *P*, one might expect the estimate of *S* to be highly sensitive to the assumed value of *P*. However, this is not the case.

Figure 1 shows the estimated value of S, from formula (2), using D=230,000 deaths and *P* in a range from 2 million to 4 million. If, *P* were revised *downwards* by a third to 2 million, then *S* would only be reduced by 3.9%. If *P* were instead revised *upwards* by a third to 4 million, then *S* would be increased by 1.9%. Thus, the estimate of *S* is surprisingly robust within a wide range of the estimated population at risk.

The sensitivity of *S* to the estimate of *P* can also be described in terms of the elasticity, defined as $E = S'(P)(P/S)$, where $S'(D)$ is the first derivative of *S* with respect to *D*. Evaluated at *D*=230,000 and *P*=3,000,000, we find *E*=0.083. That is, in the vicinity of the specified values of *D* and *P*, a 1% change in the value of *P* will produce a change in *S* of only 0.083%.

Figures 1 and 2 about here

Sensitivity to estimates of the number of deaths

Estimates of *S* are much more sensitive to the estimated number of deaths, *D*. Following the same strategy as with *P*, figure 2 shows the estimated values of *S* for *D* in the range from 160,000 to 320,000 deaths. *S* is a quadratic function of *D*, but in this range is almost perfectly linear. Using the values *D*=230,000 and *P*=3,000,000, a one-third increase in *D* will cause *S* to increase by 30.0%. A one-third decrease in *D* will cause *S* to decrease by 31.6%. The elasticity at the specified values is *E*=0.917, close to 1.000. A good estimate of the total number of deaths is critically important.

The third quantity that must be specified in formula (2) is *c*, the proportion of the total population that is children of the specified type. S is directly proportional to c, and the elasticity is 1.000, so it is most important to have a good estimate of this number.

Sensitivity to differences in risk between children and adults

Two possible types of variation in risk can be incorporated into the estimates—if data are available. The first type of variation is age, with different probabilities of death for adults and children.

It is likely that adults and children have different probabilities of death. If, say, the data on deaths could possibly be disaggregated by the age of the victims, so that the numbers of deaths D could be expressed as the sum of deaths to adults and to children, $D = D_a + D_c$, then equation (1) would become

$$S = p_a(1 - p_c)C\,. \qquad (3)$$

If children are *more* likely to die than adults, that is, if $p_c > p > p_a$, then p_a will be less than p and $1 - p_c$ will be less than 1-p, so the product $p_a(1 - p_c)$ will always be less than p(1-p). In this case the failure to distinguish between the death rates of adults and children will lead to an *over*-estimate of S. Conversely, if children are *less* likely to die than adults, then the uncorrected formula will lead to an *under*-estimate of S.

There is evidence that in most natural disasters, children are more likely than adults to die, simply because they are not as strong as adults. If physical strength is an important factor, then the elderly and women will also tend to be more vulnerable. However, physical strength is not a crucial factor for all disasters, and sometimes children appear to have an advantage, because of special efforts to rescue them. In the case of the Haiti earthquake, the numbers of deaths cannot be disaggregated by age, and there is no basis for assuming a difference between the probabilities for adults and children, so (3) will not be used.

Sensitivity to heterogeneity in risk

A second source of variation in risk, affecting both adults and children, is related to geographical location, socio-economic characteristics, or different levels of frailty before the event. The impact of Haiti earthquake's followed concentric zones according to distance from the epicenter. Most other kinds of disasters would also have their impact according to some kind of geographic gradation. But even individuals at approximately the same location can have different levels of risk according to whether they are indoors or outdoors, whether or not their housing is well-constructed, whether they are already weak or disabled, and so on.

The available data on deaths or death rates are not spatially structured, and many sources of variation in risk cannot be specified. Therefore, as a strategy to assess the sensitivity of equation (1) to heterogeneity, the population will be artificially disaggregated into ten strata, ranging from lowest risk (stratum 1) highest risk (stratum 10). The probability of dying in stratum s is defined to be p_s (s=1,…,10). The probabilities of dying are assumed to follow a logistic pattern, such that

$$\log\left(\frac{p_s}{1-p_s}\right) = b_0 + b_1 s \,. \tag{4}$$

It is arbitrarily assumed that the odds of dying are ten times as great in stratum 10 as in stratum 1, that is, $\frac{p_{10}}{1-p_{10}} = 10\frac{p_1}{1-p_1}$, that 10 percent of the total population at risk, *P*, is in each of the 10 strata, and that the overall probability of dying is *p*, calculated earlier as $p = D/P = 0.0767$.[5]

The fitted values of p range from 0.0207 in stratum 1 to 0.1742 in stratum 10, averaging to p=0.0767. The value of S ranges from only 2,675 children in stratum 1 to 19,020 children in stratum 10, adding up to 90,409. This is smaller than the un-stratified estimate of S, given earlier as 93,590 children, but only 3.4% smaller. Thus stratification of risk, within the larger population at risk, produce a smaller estimate of the number of surviving children in the outcome category, but not a great deal smaller unless there is a very steep gradient in the risk. The estimate that ignores heterogeneity is an upper bound with respect to possible heterogeneity.

Extension to more than one adult role

The basic model is built around adult-child pairs, such as the household head and the child, the mother and the child, or the father and the child. For all of these dyads, one adult is paired with one child, and the goal is to estimate the number of pairs in which the child has survived but the adult has died.

The adult node can be expanded to include more occupants. For example, there may be an interest in estimating how many children with two living parents, before the emergency, survived but lost both of their parents and became double orphans. Let C be the number of children, within the population P, that consists of children with two living parents before the emergency, p be the probability of death, and again assume independence among the survivorship of the family members.

There may be an interest in how many of the children who had two living parents lost exactly one of them. Under an assumption of independent outcomes, the probability that the mother dies, the father survives, and the child survives will be $(p)(1-p)(1-p) = p(1-p)^2$. This is the same as the probability that the father dies, the mother survives, and the child survives. The expected number of surviving children who lost just the mother, or just the father, will be

$$S = p(1-p)^2 C. \qquad (5)$$

This must be doubled to get the number of new single orphans. Finally, for each child with two parents, the probability that both parents die and the child survives will be $(p)(p)(1-p) = p^2(1-p)$ and the expected number of surviving children who lost both parents will be

$$S = p^2(1-p)C. \qquad (6)$$

Standard errors and confidence intervals

The critical data for the estimates in this paper are the number of deaths, *D*, the population at risk, *P*, and the number of children (within the population at risk), *C*, who are of a specific type. *D* and *P* are probably subject to serious measurement error, but it will be assumed that they are not subject to sampling error.[6] The third number, *C*, will be estimated from a sample and therefore is subject to sampling error.

Define a proportion, *c*, calculated from a variable *Y* that is defined to be 1 for children of a specified type in a household survey and 0 for all other cases in the survey. If *n* is the number of cases in the survey, then the estimated standard error of *c* is simply $\sqrt{c(1-c)/n}$ and, with the usual normal approximation to the binomial distribution, a 95% confidence interval for the population value of the proportion will be $c \pm 1.96\sqrt{c(1-c)/n}$. The data to be described in the next section come from a multi-stage cluster sample with sampling weights, so a somewhat wider and more accurate confidence interval will be obtained from a logit regression of *Y* with no covariates and with adjustments for clustering and sample weights. The regression output will produce a coefficient and the lower and upper ends of a 95% confidence interval for the population value of that coefficient, on a logit scale. These three numbers can be referred to as *b*, *L*, and *U*, respectively. Converting the three numbers from logits to proportions gives $c = e^b/(1+e^b)$, $c_L = e^L/(1+e^L)$, and $c_U = e^U/(1+e^U)$, respectively, where *c* is the sample proportion and c_L and c_U are the lower and upper ends of a confidence interval for the population proportion. The lower and upper ends of a confidence interval for the number of surviving children of a specified type is thus obtained by substituting $C = Pc_L$ and $C = Pc_U$ into the respective formulas that use *C*. For example, when using equation (2), the lower and upper ends of a 95% confidence interval for S will be given by $S_L = p(1-p)Pc_L$ and $S_U = p(1-p)Pc_U$, respectively.

Possible association between risks of death

Child and one adult

In most emergencies, it is probably not reasonable to assume that the probability a child will die is independent of whether or not the household head--or another key member of the child's household or kin structure—has died. If the household members are in the same building at the time of the event, or even in the same immediate neighborhood, there could be a strong positive association between the two. The Haiti earthquake happened in the late afternoon, on a weekday, and different household members could have been relatively scattered, but at least some positive association would be plausible.

To describe such an association, we first require that the marginal probability of dying is the same for both a child and an adult, namely p. Represent the outcomes of survival and death with 0 and 1, respectively. Define p_0 to be the conditional probability that the adult dies, given that the child survives, and p_1 to be the conditional probability that the adult dies, given that the child dies.[7] The probabilities of possible pairs of outcomes are shown in table 1.

Expressing table 1 in terms of counts, rather than probabilities (that is, multiplying each entry by the number of adult-child pairs), and referring to the counts in the (0,0), (0,1), (1,0), and (1,1) cells as a, b, c, and d, respectively, a well-known formula for the product-moment correlation[8] between the two variables is $r = (ad - bc)/\sqrt{(a+b)(c+d)(a+c)(b+d)}$, which reduces to $r = p_1 - p_0$. That is, the correlation can be interpreted as the difference between the two conditional probabilities. If the correlation is positive, then the probability that the adult dies will be greater if the child dies than if the child survives.

Table 1 about here

With algebra it is then possible to express both p_1 and p_2 in terms of p and r: $p_1 = p + r(1-p)$, or $p_1 = p(1-r) + r$, and $p_0 = p(1-r)$. Therefore the estimated number of children who survived but lost the household head in the emergency, taking association in outcomes into account, is

$$S = (1-p)p_0C = p(1-p)(1-r)C\,, \tag{7}$$

where C is the number of children in the population at risk who had a household head before the emergency. Note that if the correlation is zero, then (7) simplifies to (2).

This formula could also be used to estimate the number of single orphans who lost their only parent, first with C as the number of children who were single orphans with surviving mother, and second with C as the number of children who were single orphans with surviving father, within the population at risk.

The net effect of a positive association will be to reduce the magnitude of S, without changing the numbers of adults who die or children who die. The child deaths will tend to be paired with adult deaths, so surviving children will less often be linked with an adult who died. In percentage terms, the reduction in S, if this correlation is incorporated, will be $100r$%. For example, if r=0.2, then an estimate of S that ignores the association should be reduced by 20 %. It seems likely that $r \geq 0$ in most contexts, implying that the value of S calculated *without* this adjustment will be, in effect, an upper bound.

Child and two adults

If the child has two living parents before the emergency, and the outcomes are potentially associated, then adjustments to (5) and (6) are needed to estimate the number of children who survive but become new single orphans or double orphans. Rather than a pair of individuals, such as a child and a household head, there is a triad, consisting of a child, a mother, and a father. There are eight possible combinations of survival or death for the three members of the triad. The three of interest are listed in figure 3.

Figure 3 about here

A particularly undesirable outcome is listed third in figure 3, whereby a surviving child loses *both* parents in the emergency. Ignoring possible associations, the expected number of children with this outcome was given by (6). The following steps will allow for possible associations among the outcomes for the mother, the father, and the child.

It will be assumed that the marginal probability of dying is the same for the mother, the father, and the child, namely p. It will also be assumed that the three possible pairwise or zero-order correlations—between the outcomes for the mother and the father, between the mother and the child, and between the father and the child—are the same, namely r. The conditional probability p_0, defined above, will be generalized to be defined as the probability that either parent in the child+father+mother triad dies, given that the child has survived.

Under this assumption, the probability that the child survives and the father dies but the mother survives is $(1-p)p_0(1-p_0)$. If C is the number of children with two surviving parents before the emergency, then the expected number of children who survive but become single orphans with surviving mother will be

$$S = (1-p)p_0(1-p_0)C = p(1-p)(1-r)[1-p(1-r)]C\,. \tag{8}$$

Under the assumptions, (8) will also be the expected number of children who survive but become single orphans with surviving father.

The probability that the child survives and both parents die is $(1-p)p_0^2$, so the expected number who survive but become double orphans will be

$$S = (1-p)p_0^2C = p^2(1-p)(1-r)^2C\,. \tag{9}$$

If the correlation r is 0, then (8) and (9) simplify to (5) and (6), respectively.

Changes in the orphanhood distribution

There may be interest in estimating the net change, as a result of the emergency, in the numbers of orphans and the prevalence of orphanhood. We will sketch and apply a procedure to estimate these changes. Again use C to represent the number of children of a particular type in the affected population. Incorporate superscripts 0 and 1 to represent children before and after the emergency, respectively, and subscripts 1, 2, 3, and 4 to represent four complete and mutually exclusive categories of "Orphan Type": (1) non-orphans, (2) single orphans with surviving mother, (3) single orphans with surviving father, and (4) double orphans. Maternal orphans are the sum of categories 3 and 4; paternal orphans are the sum of categories 2 and 4; and single orphans are the sum of categories 2 and 3.

Changes in the distribution are determined by both the survivorship of children and the survivorship of their parents. That is, the number of children in a category of "Orphan Type" after the emergency will equal the number who were in that category before, minus the number who died, minus the number who survived but exited that category because of parental deaths, plus the number who survived and entered from another category because of parental deaths.

The number of non-orphans after the emergency, for example, will equal the number of previous non-orphans, minus those who died, minus the number who survived but lost one parent, minus the number who survived but lost both parents. The number of double orphans after the emergency will be the number of previous double orphans, minus those who died, plus previous non-orphans who survived but lost both parents, plus previous single orphans who survived but lost one parent (the previously surviving mother or father). The numbers of single orphans will change in similar ways.

Figure 4 gives the conditional probability of a transition from category *i* to category *j*, given that a child was in category *i* before the emergency, for each possible transition, that is, $p_{ij} = \Pr(Orphan\_type\_at\_time\_1 = j \mid Orphan\_type\_at\_time\_0 = i)$. Each probability includes a factor of (1-*p*), and the sum of the probabilities out of each initial category *i* is (1-*p*), the probability that the child will survive.[9]

Figure 4 about here

The frequencies after the emergency can be obtained from those before the emergency as

$$\begin{aligned}
C_1^1 &= (1-p)(1-p_0)^2 C_1^0 \\
C_2^1 &= (1-p)p_0(1-p_0)C_1^0 + (1-p)(1-p_0)C_2^0 \\
C_3^1 &= (1-p)p_0(1-p_0)C_1^0 + (1-p)(1-p_0)C_3^0 \\
C_4^1 &= (1-p)p_0^2 C_1^0 + (1-p)p_0 C_2^0 + (1-p)p_0 C_3^0 + (1-p)C_4^0 .
\end{aligned} \tag{10}$$

Now define a matrix *T*:

$$T = \begin{bmatrix} (1-p_0)^2 & 0 & 0 & 0 \\ p_0(1-p_0) & (1-p_0) & 0 & 0 \\ p_0(1-p_0) & 0 & (1-p_0) & 0 \\ p_0^2 & p_0 & p_0 & 1 \end{bmatrix}$$

where, as before, $p_0 = p(1-r)$.

Also define column vectors C^t, for the frequency distribution of Orphan Type before (t=0) and after (t=1) the emergency as

$$C^t = \begin{pmatrix} C_1^t \\ C_2^t \\ C_3^t \\ C_4^t \end{pmatrix}, \text{ for } t\text{=0 or 1.}$$

Then the frequency distribution after the emergency can be estimated from the frequency distribution before the emergency in matrix format as

$$C^1 = (1-p)TC^0. \tag{11}$$

The counts can be divided by their respective totals, at times 0 and 1, to obtain proportions in the four categories.

Note that the frequency distributions of children C^t, as given above, are restricted to the area or population affected by the emergency. The national distributions, if desired, must be augmented to include the children in the rest of the population.

Data

In an application of this simple model to the aftermath of the Haiti earthquake, four sources of data will be combined. The first is the officially estimated number of deaths, D, and population at risk, P.[10] The values D=230,000 and P=3,000,000 were given earlier.

The second source of data is the 2005/6 Demographic and Health Survey of Haiti, conducted by Macro International with funding primarily from the U.S. Agency for International Development. The median date of this survey was January 2006. For many purposes, the most important part of a DHS survey is the survey of women age 15-49, but this analysis uses only the household survey, whose main purpose is to identify eligible respondents for the survey of women. The household survey included 46,680 persons.[11]

The third type of data is the estimated age distribution of Haiti, obtained from the International Data Base (IDB) of the International Programs Center of the U.S. Census Bureau.[12] The estimated/projected age distribution for July 1, 2009, in single years of age, separately for males and females, was used to re-weight the DHS data as described below. It was also used to estimate the proportion of the population age 0-17, given earlier as 0.4407.

The household survey includes four questions pertaining to each child in the household below age 18, about the survivorship of the mother and father and whether (if living) they live in the same household as the child. The data are believed to be of generally good quality, although there is substantial heaping at some ages such as age 12. There is also clear evidence of displacement of 17-year-old males into age 18. In order to update the age distribution to January 2010 and to smooth it, the DHS sampling weights were inflated/deflated for each combination of sex and single year of age to match the U.S. Census Bureau's age distribution.

Table 2 cross-tabulates Orphan Type (*OT*) with Residence Type (*RT*), a variable that describes co-residence with parents. The four categories of Orphan Type were listed earlier. The four categories of Residence Type are (1) children living with both parents, (2) children living with the mother only, (3) children living with the father only, and (4) children living with neither parent. There are nine logically possible combinations (rather than 4 x 4 = 16) and the table gives the re-weighted percentage of children age 0-17 in each combination. It is obvious that in Haiti, prior to the earthquake, far more children lived separately from parents than were actually single or double orphans. 19.77% lived separately from parents and only 1.26% were double orphans.[13] The combinations of parental survival and coresidence in this table identify subpopulations of potential interest, but below we will only use the margins of the table.

Another useful way to describe children's living arrangements is in terms of parental survival and the child's relationship to the head of the household, described as the father, the mother, or someone else. This tabulation is given for Haiti, using the 2005/6 DHS survey, in table 3. As the third column shows, about 36% of children under 18 have a household head who is someone other than a parent. In many cases, these children are living with the extended family and a parent is present but is not the household head. The application of the model includes some examples that describe children who lost a parent or lost a household head. Table 3 can be used to identify the overlap between these outcomes, because the parent may or may not be the household head.

Tables 2 and 3 about here

The fourth type of data pertains to the association between outcomes for children and the adults with whom they can be paired. There are no data from Haiti that allow a direct or even an indirect estimate of this association. The appendix to this paper describes how we have borrowed from the association observed in Indonesia by Frankenberg et al. (2009) following the 2005 tsunami, to justify the range for *r* used in this paper, 0.1 to 0.3.

Application of the model to the aftermath of the Haiti earthquake

Transitions

A large number of children *did not* survive the Haiti earthquake. Using the estimate of 230,000 total deaths, the assumption that the probability of dying was the same for children and adults, and the estimate that 44.07 % of the population consisted of children under 18, the estimated number of child deaths would be approximately 101,000, more than 2% of all the children in Haiti.

Turning to the children who did survive, and applying the model, it is first necessary to specify a pair-wise relationship between a child and an adult. Then we must calculate c, the proportion of the population P that consists of children with such a relationship.

The percentages in table 2 refer to children age 0-17, who accounted for 44.07% of the population. The proportions of the entire population who were children in the specific types or orphan and residence categories (the various values of c) are estimated by converting the relevant percentages in table 2 to proportions and multiplying by 0.4407.

Equations (2), (5), and (6) omit any possible association between the outcomes for children and associated adults. Those outcomes probably *are* associated, positively, and those equations will effectively give maximum possible numbers of surviving children of the different types. Equations (7), (8), and (9) incorporate a possible correlation. We somewhat arbitrarily assume a positive correlation of 0.2, with the effect of adjusting the estimates downwards. We will also give a plausible range of uncertainty if the correlation were actually in the range 0.2 ± 0.1.

Six specific estimates will be developed. They are listed in the order of number of children at risk, with the largest category of children at risk listed first. All of these categories refer to surviving children.

S_1: Children whose household head died

At risk: all children

S_2: Children with two living parents who became single orphans

At risk: children with two living parents, OT=1

S_3: Children with two living parents who became double orphans

At risk: children with two living parents, OT=1

S_4: Children with no parents in the household whose household head died

At risk: children with no parents in the household, RT=4

S_5: Children with one living parent who became double orphans

At risk: children with one living parent, OT=2 or 3

S_6: Children who were double orphans whose household head died

At risk: children with no living parents, OT=4 .

Using $p_0 = p(1-r)$ as the conditional probability of an adult death, given that the child survived, the proportion under 18 (.4407), the population at risk (P=3,000,000), and proportions drawn from the margins of table 2, the specific formulas for S_1 through S_6 are as follows:

$$S_1 = (1-p)*p_0(.4407)P$$

$$S_2 = (1-p)p_0(1-p_0)(2)(.8832)(.4407)P$$

$$S_3 = (1-p)p_0^2(.8832)(.4407)P$$

$$S_4 = (1-p)p_0(.1977)(.4407)P$$

$$S_5 = (1-p)p_0(.0709+.0297)(.4407)P$$

$$S_6 = (1-p)p_0(.0126)(.4407)P$$

Table 4 presents the results. The estimates in the third column omit possible associations. The fourth column gives the estimates if r=.2; the half-width of the range is given in the final column. Sampling error that arises from the use of sample data for the estimate of c will be ignored; it is much less important than the uncertainty about the value of r.

By far the most common transition was from two living parents to one, that is, single orphanhood. It is estimated that about 125,000 of the surviving children lost one parent. This category is so large because most children had two living parents. Another 4,100 non-orphans lost both parents. About 7,500 children were single orphans before the earthquake, and lost that parent, also becoming new double orphans.

The second largest category consisted of children who lost their household head, seen earlier in the paper. It is estimated that there were 75,000 such children. This number is large because all children in the household population have, by definition, a household head.

In Haiti, substantial numbers of children have surviving parents but do not live with them. DHS data indicate that most such children live with a grandparent, or aunt or uncle or, occasionally, an older sibling. Some children live separately from parents and have a household head with whom they have a more distant biological relationship or no kinship at all. Some of these children are restaveks, or unpaid household servants, who are often abused and deprived of schooling. Children who had no parents in the household, but whose household head died, are particularly vulnerable for negative outcomes after a disaster. About 15,000 surviving children are estimated to be in this category.

Probably the most vulnerable category in table 4 consists of children who were already double orphans and then lost their household head in the earthquake. Fortunately, this category is small, amounting to only about 900 children.

Table 4 about here

Net changes in the numbers of orphans

The mortality of children and parents leads to transitions across the four categories of "Orphan Type" that must be combined to obtain the net changes in the distributions. The distributions before and after the earthquake are given for the earthquake zone (that is, the affected population of P=3 million) in table 5, and for all of Haiti (that is, the total population of N=9.848 million) in table 6. In each table, the procedure is applied for a correlation r=0 and for three positive values:

0.1; 0.2; and 0.3. We suggest that a correlation *r*=0.2 is the most plausible specific value, but prefer estimates corresponding to the range between *r*=0.1 and *r*=0.3.

As seen in table 5, an increasing value of r produces a modest increase in the number of surviving non-orphans, because of the increasing chance that the child and adults survived together or died together. An increasing value of *r* leads to smaller numbers in all the other categories.

Within the population of 3 million people affected by the earthquake, there were substantial increases in the percentages who were single orphans. In terms of frequencies, this increase was partially offset by the reduced population of children. The percentage of children who were single orphans is estimated to have increased from 10.06% to19.66%, nearly doubling, and the frequency to have increased by 80%, from 133,127 to 240,013. The percentage who were double orphans increased from 1.26% to 2.21%, and the number increased by 62%, from 16,702 to 27,032.

Tables 5 and 6 about here

Combining the rest of the population with the affected population, as in table 6, the numerical changes are the same but the percentage changes are muted, because the denominator is larger. For example, the percentage of single orphans increased from 10.06% to 12.83% and the number of single orphans increased by 24%, from 449,815 to 559,073. The percentage of double orphans increased from 1.26% to 1.54%, and the number of double orphans increased by 19%, from 54,826 to 65,156.

Summary and conclusions

In the immediate aftermath of a humanitarian disaster, there is an urgent need for estimates of the numbers of casualties—the numbers of people who have died, been injured, lost their homes, or been displaced. There is a particular concern for children, who may become separated from their caregivers and require special forms of protection.

In the race to develop estimates of the impact of the disaster, as part of the effort to mobilize aid, some numbers are typically developed very quickly, whereas others come much later. The earliest numbers are generally the estimates of deaths, not broken down by age, and the population affected. Here those numbers have been referred to as D and P, respectively. By contrast, the numbers of separated or unaccompanied children are much more difficult to estimate.[14]

This paper has shown how some indicators of the impact of the disaster on children, in particular, can be estimated rather quickly by combining D and P with pre-disaster data on parental survivorship and co-residence, often available from a census or survey. The estimates given here describe changes in the overall pattern of parental survivorship and co-residence for those children who survive the disaster.

The death of a parent or household head in a disaster will be emotionally traumatic and disruptive to the continuity of the child's social environment. Children who lose a parent or household head do not necessarily enter the status of being separated or unaccompanied, but they are much more likely to enter that status. Their post-disaster household structure is almost sure to change in important ways. If and when those children are individually identified, programs to assist them are desirable.

It is unlikely that the early estimates of *D* and *P* described here can be revised, because of the lack of registration data in Haiti and the continuing displacement of much of the population. Eventually it should be possible, through the next DHS survey or a census, to measure the net changes in the distribution of Orphan Type, with which the estimates given here could be compared to help validate the model and estimate the actual association between outcomes for paired children and adults.

References

Checci, F. and L. Roberts. 2008. Documenting mortality in crises: What keeps us from doing better? *PLoS Medicine 5*:1025-1032.

Coleman, James, Elihu Katz, and Herbert Menzel. 1957. The diffusion of an innovation among physicians. *Sociometry 20:* 253-270.

Cutter, S. 1996. Vulnerability to environmental hazards. *Progress in Human Geography 20*:529-539.

Doocy, S., A. Rofi, C. Moodie, E. Spring, S. Bradley, G. Burnham, and C. Robinson. 2007. "Tsunami Mortality in Aceh Province, Indonesia." *Bulletin of the World Health Organization 85*:273-278.

Frankenberg, Elizabeth; Thomas Gillespie, Samuel Preston, Bondan Sikoki, and Duncan Thomas. 2009. Mortality, the Family, and the Indian Ocean Tsunami. California Center for Population Research, On-Line Working Paper Series #CCPR-20009-028, UCLA. http://papers.ccpr.ucla.edu/papers/PWP-CCPR-2009-028/PWP-CCPR-2009-028.pdf. [Note: the final version will include the most up-to-date citation for this paper.]

Kendall, Maurice G. 1962. *Rank Correlation Methods (3rd edition)*. London: Griffin.

Kermack, W.O., and A.G. McKendrick. 1927. A contribution to the mathematical theory of epidemics, *Proceedings of the Royal Society of London. Series A, 115:*700-721.

Nishikiori, N., T. Abe, D. Costa, S. Dharmaratne, O. Kunii, and K. Moji. 2006. Who died as a result of the tsunami?—Risk factors of mortality among internally displaced persons in Sri Lanka: a retrospective cohort analysis. *BMC Public Health 6:*73.

Rofi, A., Doocy, S., Robinson, C. 2006. Tsunami mortality and displacement in Aceh province, Indonesia. *Disasters 30:*340–350.

Yeh, H. 2010. Gender and age factors in tsunami casualties.” *Natural Hazards Review. 11*:29-34.

Appendix. The association between outcomes for children and adults

An alternative approach to the two binary variables in table 2 uses log linear modeling rather than correlation to describe their association. This strategy was used by Frankenberg et al. (2009). Referring back to table 2, restated with frequencies *a*, *b*, *c*, and *d*, the log of the odds ratio is $\beta = \log\left(\frac{ad}{bc}\right)$. Here, β is the slope coefficient in a logit regression of "child died" on "adult died" (or vice versa).[15]

There is no simple way to re-state equations that include *r* in terms of β, rather than *r*. However, for any specified value of *p*, it is possible to iterate through the different possible combinations of relative frequencies and calculate both *r* and β, thereby establishing a correspondence between the two measures of association.

In the absence of data for Haiti that allow us to estimate the association between outcomes for children and adults, we have borrowed from the association observed in Indonesia (Frankenberg et al. 2009). In figure 3 of that paper, the "log odds of concordant survival outcomes for pairs of close kin" are shown for nine such pairs. The log odds (β), are presented in a bar graph and do not include specific numerical values, but seven of the nine values (dropping the smallest and largest values) appear to be in the interval 1.5 to 2.5. This range in the log odds appears to match roughly with the range for *r* used in this paper, 0.1 to 0.3

Log linear models are usually more appropriate than correlation to describe the association between two binary variables. However, as noted in the paper, the correlation equals the difference between two conditional probabilities, $r = p_1 - p_0$, which is easily interpreted.

Figure 1. Estimated values of *S* for *P* in the range 2 million to 4 million.

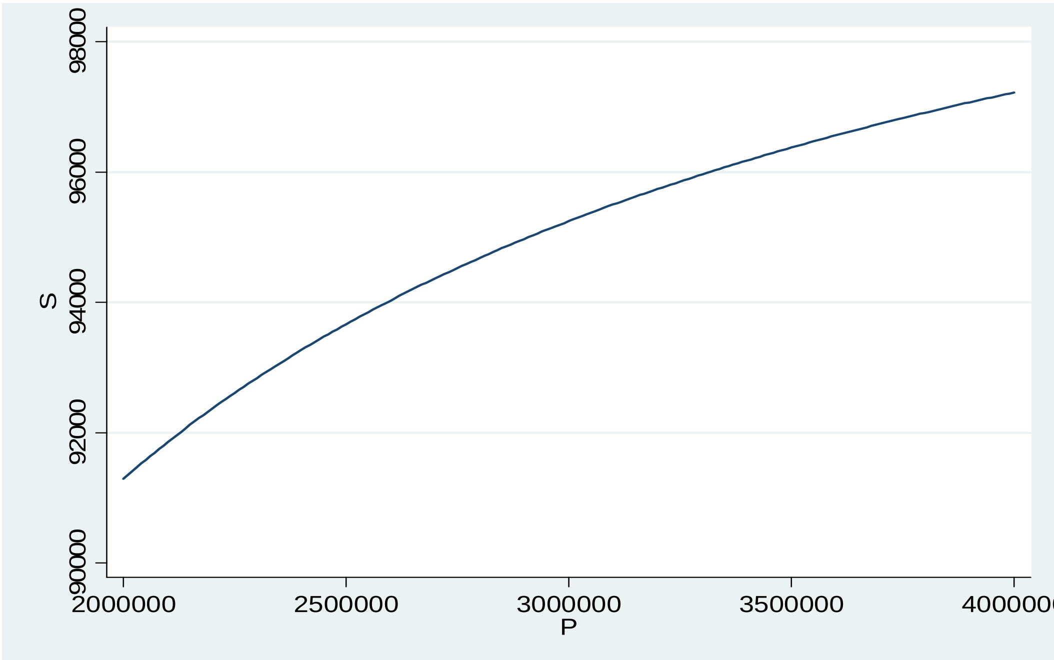

Figure 2. Estimated values of *S* for *D* in the range 160,000 to 320,000.

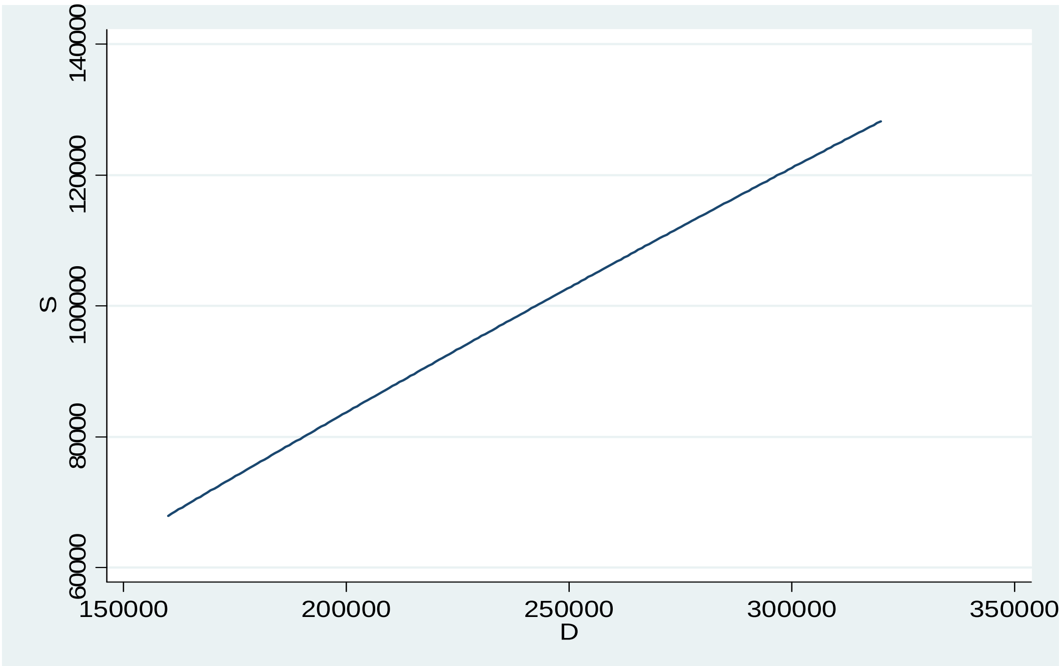

Figure 3. Possible transitions to orphanhood for children who had two surviving parents before the emergency

	Child Dies	Mother Dies	Father Dies
Non-orphans become single orphans with surviving mother:	No	Yes	No
Non-orphans become single orphans with surviving father:	No	No	Yes
Non-orphans become double orphans:	No	Yes	Yes

Figure 4. Possible transitions within the distribution of Orphan Type for children who had any surviving parents before the emergency.

Orphan Type Time 0 i	Orphan Type Time 1 j	Conditional Probability p_{ij}	Source Formula*
1	1	$(1-p)(1-p_0)^2$	
1	2	$(1-p)\,p_0(1-p_0)$	(8)
1	3	$(1-p)\,p_0(1-p_0)$	(8)
1	4	$(1-p)\,p_0^2$	(9)
2	2	$(1-p)(1-p_0)$	
2	4	$(1-p)\,p_0$	(7)
3	3	$(1-p)(1-p_0)$	
3	4	$(1-p)\,p_0$	(7)
4	4	$(1-p)$	

* The conditional probability of staying in the same category of Orphan Type is calculated as a residual, such that the sum of the conditional probabilities for each initial category is 1-*p*.

Table 1. Cross-tabulated probabilities of dying and surviving for paired children and adults, where p is the marginal probability of dying, for either a child or an adult; p_1 is the conditional probability that the adult dies if the child dies; and p_0 is the conditional probability that the adult dies if the child survives.

		Child died?		
		No (0)	Yes (1)	Total
Adult died?	No (0)	$(1-p)(1-p_0)$	$p(1-p_1)$	$1-p$
	Yes (1)	$(1-p)p_0$	pp_1	p
	Total	$1-p$	p	1

Table 2. Percentage of children age 0-17 in Haiti in the possible combinations of parental survival (Orphan Type) and coresidence with parents (Residence Type). Limited to children living in households. Source: 2005/6 Demographic and Health Survey of Haiti (median date: January 2006), re-weighted to match the IDB age distribution.

	Child is living with				
Parental Survival	Both parents	Mother only	Father only	Neither parent	Total
Both parents alive	45.06%	24.24%	4.84%	14.53%	88.67%
Only mother alive	-	4.94%	-	2.16%	7.09%
Only father alive	-	-	1.15%	1.82%	2.97%
Both parents dead	-	-	-	1.26%	1.26%
Total	45.06%	29.18%	6.00%	19.77%	100%

Note: “-“ indicates that a combination is logically impossible.

Table 3. Percentage of children age 0-17 in Haiti in the possible combinations of parental survival and household headship. Limited to children living in households. Source: 2005/6 Demographic and Health Survey of Haiti (median date: January 2006), re-weighted to match the IDB age distribution.

Parental Survival	Household head: Father	Mother	Other	Total
Both parents alive	38.94%	20.58%	29.14%	88.67%
Only mother alive	-	3.53%	3.57%	7.09%
Only father alive	0.91%	-	2.07%	2.97%
Both parents dead	-	-	1.26%	1.26%
Total	39.85%	24.11%	36.04%	100%

Note: “-“ indicates that a combination is logically impossible.

Table 4. Estimated expected number of surviving children age 0-17 who lost a parent or household head in the 2010 Haiti earthquake, based on official estimates of 230,000 deaths and an affected population of three million. The correlation between outcomes for children and adults is assumed to be in the range *r*=.2±.1.

Prop. of children at risk	Prop. of population at risk (c)	Maximum number (r=0)	Adjusted number (r=.2)	Half width of range *r*=.2±.1
Children whose household head died				
1.0000	0.4407	93,590	74,872	9,359
Children with two living parents who became single orphans				
0.8832	0.3892	153,243	124,631	14,561
Children with two living parents who became double orphans				
0.8832	0.3892	6,362	4,072	1,018
Children with no parents in the household whose household head died				
0.1977	0.0871	18,502	14,802	1,860
Children with one living parent who became double orphans				
0.1006	0.0443	9,424	7,539	943
Children who were double orphans whose household head died				
0.0126	0.0055	1,181	946	118

Table 5. The estimated distributions of children according to orphan type, before and after the 2010 Haiti earthquake, *within the earthquake area*, based on official estimates of 230,000 deaths and an affected population of three million. Totals are affected by rounding.

Orphan Type	Before the Earthquake	After the Earthquake; Correlation *r*			
		0	.1	.2	.3
Non-orphans	1,172,270 88.67%	922,791 75.59%	938,179 76.85%	953,695 78.12%	969,337 79.41%
Mother alive Father dead	93,802 7.09%	156,592 12.83%	150,166 12.30%	143,614 11.76%	136,934 11.22%
Mother dead Father alive	39,325 2.97%	110,148 9.02%	103,337 8.47%	96,399 7.90%	89,333 7.32%
Double Orphans	16,701 1.26%	31,207 2.56%	29,056 2.38%	27,032 2.21%	25,135 2.06%
Total	1,322,100 100%	1,220,738 100%	1,220,738 100%	1,220,740 100%	1,220,739 100%

Table 6. The estimated distributions of children according to orphan type, before and after the 2010 Haiti earthquake, *in all of Haiti*, based on official estimates of 230,000 deaths, an affected population of three million, and a total population of 9.848 million before the earthquake. Totals are affected by rounding.

Orphan Type	Before the Earthquake	After the Earthquake; Correlation *r*			
		0	.1	.2	.3
Non-orphans	3,848,175 88.67%	3,598,695 84.90%	3,614,083 85.26%	3,629,598 85.63%	3,645,241 86.00%
Mother alive Father dead	307,921 7.09%	370,711 8.75%	364,285 8.59%	357,732 8.44%	351,052 8.28%
Mother dead Father alive	129,092 2.97%	199,915 4.72%	193,104 4.56%	186,166 4.39%	179,100 4.23%
Double Orphans	54,826 1.26%	69,331 1.64%	67,180 1.58%	65,156 1.54%	63,260 1.49%
Total	4,340,014 100%	4,238,652 100%	4,238,652 100%	4,238,652 100%	4,238,653 100%

[1] Unless stated otherwise, the term "children" always refers to ages 0-17.

[2] In many countries, most of the children living in these institutions have at least one surviving parent, and the terms "orphan" and "orphanage" can be misleading.

[3] Household surveys in developing countries generally use the term "household head" for the first person listed in the household roster. It is rare for a household head to be a child under age 18, but if such households occur, they should be excluded.

[4] All calculations were done with more decimal places than shown, but throughout the paper, numbers are given with more significant digits than is warranted. This is done simply to facilitate replication and avoid the propagation of rounding error. For any practical use of these estimates, they should be rounded to two or at most three significant digits. Computations were done with Stata 10.

[5] Under these assumptions, b_1=0.2558 and b_0=-4.1145.

[6] It is possible that the estimate of D arises from some kind of sampling process but we have no information about that.

[7] It may seem more natural to condition the child's outcome on the adult's outcome, but the interest here is in surviving children, and the results are symmetric with respect to outcomes for children and adults.

[8] This is the usual Pearson product-moment correlation, not the tetrachoric correlation sometimes used for binary outcomes.

[9] It would be possible to add a fifth category to the post-emergency distribution, an absorbing state for death, with a probability p of a transition to that state, so that the conditional probabilities from category i would add to 1 rather than $1-p$.

[10] The official numbers have been distributed in USAID fact sheets, with estimated number of deaths attributed to the Government of Haiti and the estimated "affected" population attributed to the United Nations. Some UNICEF fact sheets have estimated a slightly smaller number of deaths, 222,157. The true number will never be known.

[11] According to the DHS report on this survey, there were 45,936 household members who resided in the household "last night" regardless of whether they were a permanent resident. We were unable to reproduce this number. 46,680 is the number of cases in the household file with hv102=1 or 9 (usual residents), minus 5 cases with age>96.

[12] The online IDB provides data in five-year age intervals; the single-year intervals were provided through personal correspondence. In mid-2010 the U.S. Census Bureau modified its

pre-earthquake population estimates upwards, bringing them into better alignment with the U.N. Population Division's estimates and the official Haitian estimates. The post-earthquake population projections were also altered, to take account of both the pre-earthquake revisions and the earthquake deaths. When the model developed in this paper was first applied, the population and age distribution immediately prior to the earthquake were estimated by averaging the IDB distributions for July 1, 2009, and July 1, 2010, resulting in an estimated total population of 9,120,000 on January 1, 2010. Following the revisions to the IDB, this paper uses the new age distribution for July 1, 2009, totaling 9,778,000, inflated upward for half a year of population increase to a projected total of 9,848,000 on January 1, 2010.

[13] This table does not include children in institutions, many of which are referred to as 'orphanages' but include many children who have surviving parents (see note 2). UNICEF has estimated that 50,000 children were in these residential care centers before the earthquake.

[14] The numbers of separated or unaccompanied children are typically at a maximum soon after the disaster, and then steadily decline, even without interventions, making the estimates more elusive. By contrast, estimated numbers of deaths and injuries tend to stabilize soon after the event.

[15] Log linear models are usually more appropriate than correlation to describe the association between two binary variables. However, as noted in the paper, the correlation equals the difference between two conditional probabilities and is easily interpreted.