The Effect of the HIV/AIDS Epidemic on Orphanhood Probabilities and Kinship Structure in Zimbabwe.

Results Based on Formal Demography and Microsimulation*

Emilio Zagheni
Department of Demography
University of California, Berkeley
2232 Piedmont Avenue
Berkeley, CA 94720-2120
e-mail: emilioz@demog.berkeley.edu

March 2, 2010

*ACKNOWLEDGEMENTS: I have benefited from the valuable comments of Kenneth Wachter, Ronald Lee, James Holland Jones, Nicholas Jewell, Gretchen Donehower and the participants to the Graduate Students Seminar and the Brownbag Seminar at the Department of Demography, UC Berkeley. I am greatly indebted to Carl Mason, who wrote the source code for the latest version of SOCSIM, and provided crucial technical support with the microsimulation program.

Abstract

The paper analyzes the effect of the HIV/AIDS epidemic on orphanhood and kinship structure in Zimbabwe. An approach based on the formal demography of kinship is used to obtain insights on the effect of the epidemic on probabilities of maternal orphanhood. The approach relies on demographic rates that are widely available across countries and can thus be used for comparative purposes. A microsimulation model based on SOCSIM, with larger data requirements, is calibrated to the Zimbabwean setting to quantitatively evaluate the crisis of care. Estimates and predictions of number of orphans, and key kinship resources available to them, are obtained for the period 1980-2050. The paper gives a methodological contribution to the literature that analyzes the material basis of traditional kin relations in sub-Saharan Africa. Substantively, it documents the extent of the orphanhood crisis in Zimbabwe, its evolution over time, and the progressive erosion of kinship resources available to children.

Contents

1 Introduction

The HIV/AIDS epidemic has devastating effects on infected individuals, but the impact of the HIV/AIDS epidemic is not limited to the people who contracted the disease. The psycho-physical, emotional and economic consequences of the disease are felt by family members, as well as members of the extended family and the community at large (e.g., Palloni and Lee 1992; Bor and Elford 1998; Wachter et al. 2002).

Children are particularly vulnerable to the epidemic and are at great risk physically, emotionally and economically. They are affected by the epidemic both indirectly and directly (e.g., Foster and Williamson 2000). Children are affected indirectly when their communities are strained by the consequences of the HIV/AIDS epidemic. The services provided by the communities may be reduced both quantitatively and qualitatively with the onset of the epidemic. Doctors, nurses and teachers may suffer from the disease and this may have adverse consequences on the level of education and health care provided to children.

Children are directly affected by the HIV/AIDS epidemic in several different ways. Before the death of a parent, they may live with ill people and be requested to work and postpone their education in order to take care of the household. They may experience economic strain in the family and they may be subject to discrimination and stigma. Children can also become orphans and need care from the extended family.

It has been noted that with the rising levels of widowhood and orphanhood associated to the HIV/AIDS epidemic, the material basis of traditional kin relations may weaken to a point such that new forms of social relations emerge (e.g. Palloni and Lee 1992; Merli and Palloni 2006). Although the importance of kinship resources as a safety net in sub-Saharan Africa has been recognized, there has not been a comprehensive attempt to quantitatively evaluate the effect of the epidemic on kinship resources available to orphans. In this paper, I focus on the estimation of the evolution of kinship structure in Zimbabwe, one of the countries hit the hardest by the HIV/AIDS epidemic. I use both formal demography and microsimulation to quantify the extent of the orphanhood crisis and the kinship resources available to orphans.

First, I give some background on the consequences of the HIV/AIDS epidemic on children in sub-Saharan Africa and I discuss the methodologies that have been used in the literature to quantify the impact of the HIV/AIDS epidemic on the generation of orphans. Then I discuss the insights that can be obtained from the formal demography of kinship. I provide a method to estimate orphans that relies on a minimal set of aggregate demographic rates and that could thus be applied to a fairly large set of countries affected by the HIV/AIDS epidemic. Finally, I use microsimulation to reconstruct and predict the evolution of kinship structure in Zimbabwe for the period 1980-2050. I show some results which document the persistence of the orphanhood crisis for decades after the peak in the HIV prevalence. The increased mortality of young adults due to the epidemic, together with decreasing fertility, generate a steady erosion of kinship resources for orphans. Traditional caregivers for orphans, such as paternal uncles and aunts, become less and less available, shifting the burden on other members of the community.

2 Background on the impact of HIV/AIDS on children in sub-Saharan Africa

*"It wasn't supposed to be like this. These children's parents were supposed to be taking care of me. Now they are dead and I am nursing their children."**
-Akeyo, 74 years old,
looking after 10 grandchildren in Kenya

The HIV/AIDS epidemic has a pervasive negative effect on the multiple spheres of children's lives. Children may be affected only indirectly, at a community level, or they may experience the burden of the disease directly in their households. Children may have to drop out of school to help with household work or to care for ill parents, and may suffer both psycho-social distress and material hardship following the death of a parent.

The degree to which children are affected by the HIV/AIDS epidemic depends on several interrelated factors. For instance, some important elements that may mitigate or worsen the impact of the disease are the overall level of incidence of the epidemic, the economic situation of the community and the households affected by the disease, the gender of the children and their age when a household member gets infected, the efficiency of safety nets, etc.

In this section, I review studies on the impact of the HIV/AIDS epidemic on children. I discuss the economic impact of the epidemic, as well as the effects of the epidemic on children's education and health.

2.1 Economic impact

Children start feeling material hardship before they become orphans. When a parent develops HIV-related symptoms, children may have to take care of the household and the ill parents or young siblings. They may have to dedicate their time to activities such as cooking, cleaning, carrying water, care giving, etc.

The division of workload among children is gender-specific. Girls are more frequently care-giving providers for female relatives and take more responsibilities for domestic work (Robson 2000). Boys are more involved in agricultural and income generating activities that help with medical expenses and compensate the reduced amount of work of the parents.

When children become orphaned, their workload may increase, either because their household has been impoverished by the death of a parent (e.g. due to the loss of a parent's income and the high costs of a funeral), or because the orphans move to the household of a relative, where their workload may be greater than the one of non-orphans living in the same household (Foster et al. 1997).

Children may risk to lose the properties of their family. For instance, only a very small proportion of families in Zimbabwe write a will prior to death (Drew et al. 1996). In some cases properties are inherited by paternal relatives and there may be instances of "property-grabbing". The extent of these practices varies from region to region. The results of a survey study in Zimbabwe show that property is usually inherited by children, with 15% of respondents reporting "property-grabbing" (Drew et al. 1996).

The poorer people in a community, especially women, are the ones that usually take care of orphaned children in sub-Saharan Africa (Foster and Williamson 2000). This is related to

*Source: HelpAge International and International HIV/AIDS Alliance, *Forgotten Families: Older People as Carers of Orphans and Vulnerable Children*, HelpAge International / International AIDS Alliance, Brighton, 2003. As it appears in UNICEF (2006).

the fact that orphans tend to live more frequently in larger households with a less favorable dependency ratio and where the caregiver is much older than the child (Monasch and Boerna 2004).

Better-off families, on the other hand, tend to find their economic reserves depleted since they are continuously asked to provide economic resources to relatives affected by AIDS (Foster and Williamson 2000).

2.2 Impact on education

When a parent becomes sick, his/her children's education is often disrupted. With the financial strain of the disease and the reduced resources available for the household, there may not be enough funds for children to go to school, or the children's caregivers may have less interest in the children's welfare. Thus children might have to do either domestic work or income generating activities and miss opportunities in education in terms of lack of enrollment, interrupted schooling and poor performance while in school (UNICEF 2006).

Although there are significant variations across countries, several studies provide evidence that the enrollment rate for orphans is significantly lower than the one for non-orphans. For instance, a study based on data from eastern Africa shows that double orphans in the age group 6-10 are half as likely to be at the correct educational level, compared to non-orphan children of the same age (Bicego et al. 2003).

In some circumstances, young girls may be more at risk of being denied education than boys. However, it is not clear whether the gender gap is more prominent in orphans that in non-orphans (UNICEF 2006). In certain studies, probabilities of enrollment appear to be negatively correlated with certain characteristics of the children, such as being a girl orphan, an AIDS-related orphan, living in a rural or poor household or in a household headed by a man (World Bank 1997; Foster and Williamson 2000).

One important factor in the determination of educational outcomes for orphans is the relationship between the child and the head of household. The closer the biological tie, the more likely the child is to go to school consistently, independently of the poverty level. As a matter of fact, the closest relatives tend to make substantial commitments to ensure that their children attend school (UNICEF 2006).

2.3 Impact on health

The HIV/AIDS epidemic has had a strong impact on children's mortality. For the youngest age group (ages 0-3 years), the loss of a parent is significantly associated with the probability of survival. Zaba et al. (2005) estimate, from cohort studies in Uganda, Tanzania and Malawi, that the excess risk of mortality for children with an HIV-positive mother is 2.9 and lasts throughout childhood. The excess risk of mortality associated to maternal death is 3.9 in the 2-year period centered around the mother's death.

Children who become orphans are more vulnerable than non-orphans. Some of them become street children or prostitutes and are more likely to get infected with HIV (Richter and Swart-Kruger 1995).

Although there is not evidence of a general increase of morbidity and mortality in orphans, it is expected that the health of orphans, particularly those in the care of adolescents and elderly caregivers, is worse than the one of other children (Foster 1998). Orphans may be more malnourished than non-orphans, possibly because of reduced household resources or because parental illness or death interfere with child rearing. This may affect the incidence of morbidity in orphans.

Orphanhood has important consequences on psychological health, in addition to physical health. The stress and trauma of parental illness and death is amplified by stigmatization, dropping out of school, changes in friendships, increased workload, discrimination and social isolation (Foster and Williamson 2000). Sengendo and Nambi (1997) conducted a study on children in Uganda and found that most orphans were depressed, with lower expectations about the future than non-orphans. They observed that orphans that relocated from urban areas to rural areas were more depressed, implying that the failure to adapt to social change leads to psychological problems. They also noticed that depression was more likely in children living with a widowed father than in those living with a widowed mother, suggesting that the loss of the mother is more distressing than the loss of the father.

Both children and adults feel grief for the death of their parents. But children, unlike adults, may not immediately understand the finality of death and thus may not go through the grieving process that is fundamental to recover from the loss (Brodzinsky et al. 1986). Children may not have reached a stage of intellectual and emotional development that enable them to positively control negative emotions. They are more at risk of growing up with unresolved negative emotions which are often expressed with anger and depression. The support and encouragement from adults to express emotions are crucial to the psychological health of the orphans. This support may not be present in the context of a generalized HIV/AIDS epidemic, where basic material needs may not be met.

Children's behavior changes during parental illness. They often become sad, worried and stop playing to be near the parents. They are more likely to become solitary, to appear miserable, distressed, fearful of new situations and to develop low self-esteem (Foster and Williamson 2000).

AIDS-related orphans who lack social, economic and psychological support tend to become more vulnerable to HIV infection through early onset of sexual activity, commercial sex and sexual abuse. The lack of support therefore affects the future physical health of children also indirectly.

2.4 Coping mechanisms and the extended family

To a first approximation, a generalized HIV/AIDS epidemic strongly affects mortality and fertility rates, with important consequences on population age structure, sex ratio and probability of orphanhood. To a second degree of approximation, the epidemic affects household structure, movements in and out of the household, and the availability of kinship resources, both for young children and for the elderly (e.g. Wachter et al. 2002, 2003; Heuveline 2004).

In the sub-Saharan setting, the extended family is the predominant caring unit for orphans in communities that are severely affected by the HIV/AIDS epidemic (Ankrah 1993). The main mechanism that prevents families from falling into destitution operates through community members and the extended family, who provide material relief and economic support.

In this section, I discuss the coping mechanisms that are in place in the context of sub-Saharan Africa to deal with the orphanhood crisis. The most important safety net is provided by the extended family. Some nurturing roles are delegated to non-biological parents, through fostering practices. Social protection provided by governments is very limited or inexistent in most settings. When some forms of social protection are in place, indirect assistance to foster families or parents whose partner had died, may benefit only a small proportion of the population, namely members of the middle-class who are employed in the formal sector. Typically, especially for poor people, work is home-based and in the informal sector, without particular protections from the government. In a crisis situation, external help may come from religious organizations or NGOs, often funded by international organizations. More often, the primary source of assistance comes from the extended family.

A major issue is whether traditional fostering practices can adapt to the increasing stress imposed on them by the HIV/AIDS epidemic. It is important to identify who the care takers for orphans are and to what extent kinship members are involved in rearing orphans. I will discuss how the kinship role may change with the onset of the HIV/AIDS epidemic and I will conclude by pointing out the importance of a quantitative evaluation of the kinship resources available to orphans.

Who takes care of orphans? The question is of central importance. Providing an answer to it requires some discussion of fostering practices in sub-Saharan Africa. The coping mechanisms regarding orphans in sub-Saharan Africa are complex and vary across countries and social settings. However, a common element that distinguishes sub-Saharan Africa from western societies is that children are fostered, rather than 'adopted'. Child fostering consists of a culturally sanctioned arrangement such that children are reared by adults other than the biological parents. These arrangements are agreed upon by biological parents and other adults, often relatives. They contribute to strengthen ties across the community and they provide mutual benefits for both natal and fostering families. Although Institutional care exists, mostly in post-conflict countries, generally orphanages are not culturally and socially acceptable, in addition to being extremely expensive.

Fostering practices in sub-Saharan Africa can be categorized into two different classes: 'purposive' and 'crisis' fostering. There are several reasons to foster a child under voluntary circumstances. Isiugo-Abanihe (1985) reviews the motivations for purposive child fostering in West Africa. Most fostering in West Africa takes place within the kinship network and it is largely motivated by the need to reallocate resources within the extended family or clan, in order to maximize the survival probabilities of the kinship unit and to strengthen kinship ties. Altogether, fostering practices are strongly related to reasons such as kinship obligations, apprenticeship/training, alliance building, domestic labour and education. Purposive fostering relies on reciprocal advantages, responsibilities and rights. There are a set of social rules that determine the age structure of the exchange, which is intended to be symmetrical. Crisis fostering, on the other hand, involves some specific obligations. Esther Goody, in her classical book about fostering roles in West Africa, pointed out that kin members who have the right to the child in voluntary fostering are also obligated to foster the child in a period of crisis.

Orphanhood is not a new problem in sub-Saharan Africa, a region where mortality rates have been relatively high since before the onset of the HIV/AIDS epidemic. In the past, the combination of fostering practices, the obligation of relatives to take care of orphans, and the relative abundance of kinship resources alleviated the problem (e.g. Ntozi and Nakayiwa 1999). In addition to kinship obligations, the choice of the foster parent also depends on the reasons for fostering. If the objective is labour or simple companionship, grandmothers and childless couples tend to be an obvious destination (Goody 1982; Isiugo-Abanihe 1985).

With the increasing prevalence of HIV, how do foster practices for orphans change? First, with the onset of the HIV/AIDS epidemic, not only the overall number of orphans increases, but also the proportion of them who are double orphans increases, because of HIV transmission that occurs within couples. The increased number of double orphans, coupled with the higher mortality rate of adults, reduces the number of adult kin and increases the burden on grandparents. This situation raises questions on whether the logic of fostering, that enabled a sustainable distribution of obligations among kin in the past, may be overwhelmed by the rapid increase in AIDS-related deaths (Madhavan 2004).

Drew et al. (1998) observe that, although traditionally in Zimbabwe orphans have been incorporated into the extended family, the very high number of adult deaths has shifted the burden to elderly and adolescents. As a result, the phenomenon of grandparent-headed households and adolescent-headed households has been increasing. In the traditional Zimbabwean society, orphan children were cared for by members of the extended family. Geoff Foster (2000), in particular, describes how the caregiving functions of parents were usually taken on by paternal aunts and uncles. More recently, the safety net provided by the extended family has been weakened. According to Foster (2000), there are several reasons behind this process, such as changes in the economy, labor migration and formal education. Another important reason is the reduction of the number of available aunts and uncles, due to the HIV/AIDS epidemic, during a time when orphans have been increasing. With the increasing number of orphaned children and the unavailability of traditional caregivers, grandparents are recruited into childcare (e.g. Foster et al. 1996). Grandparents are often a last resort and agree to take orphans because other relatives are not available or refuse, generating, in some cases, situations of mutual support, where frail grandparents become recipients of care from grandchildren (Foster 2000).

Foster (2000) suggested to measure the strength of the extended family safety net by monitoring certain proxies. In particular, we may expect that where traditional values are maintained, such as in rural communities, the extended family safety net is better preserved. Analogously, the prevalence of purposive fostering within a community is an indicator of the strength of the extended family safety net. Where traditional widow inheritance is common, then orphan inheritance is likely to be prevalent too. The higher the frequency of regular contacts between relatives, the smaller the risk for orphans to be abandoned. Conversely, such risk is higher in situations where unions are established without the payment of a bride-price.

The role played by members of the extended family in providing care for orphans is strongly related to cultural practices that vary across time and geographical regions. The onset of the HIV/AIDS epidemic has influenced such practices through changes in behavior, for instance in terms of stigmatization of households affected by HIV/AIDS. The epidemic has also dramatically altered the demographic structure of the population, reducing the amount of kinship resources available to orphans and to the elderly. There have been several anthropological studies about the effect of the epidemic on the role of the extended family as a safety net in sub-Saharan Africa. Most of the quantitative research in this area has focused on a very specific set of kin members, that is 'parents'. There has not been a comprehensive quantitative evaluation of the effect of the epidemic on kinship structure in sub-Saharan Africa. I will review relevant quantitative studies on estimation of orphan numbers and orphanhood probabilities in the next section. I will then develop on some of the concepts used to estimate orphanhood probabilities in order to quantitatively evaluate the effect of the HIV/AIDS epidemic on kinship resources available to children and the elderly.

3 Background on estimation and prediction of the number of AIDS-related orphans

In order to adequately address the orphanhood problem in sub-Saharan Africa, a statistically sound approach to estimate the extent of the crisis is necessary.

In this section, I review the main methodologies that have been proposed in the literature to estimate and predict the number of orphans and the probability of being orphan at a specific age. I then provide some estimates of these key quantities. The review focuses on methods and estimates for the geographical setting of sub-Saharan Africa, in the context of the generalized HIV/AIDS epidemic.

Before going into the details of the methodologies, I would like to clarify the terminology that I will use. An orphan is a child (in the age range 0-15 years, unless specified) who has at least one biological parent dead. An AIDS-related orphan is a child who has at least one parent dead as a result of AIDS. A maternal (paternal) orphan is a child whose mother (father) has died. If the mother (father) has died of a cause associated with AIDS, then I refer to the child as AIDS-related maternal (paternal) orphan. A double orphan is a child whose parents have both died. A double AIDS-related orphan is a child whose parents have both died, at least one of them as a result of AIDS.

3.1 Survey estimation

The first and probably most obvious way to estimate the number of orphans is to 'count' them. For several countries, censuses are available and the number of orphans can be computed from census data. In some circumstances, census data may not have been collected, or they may not be reliable. In those situations sample surveys are a useful tool to estimate the number of orphans.

For a large set of developing countries and, in particular, for sub-Saharan countries, nationally representative household surveys that provide information on orphans are available. The most relevant surveys are part of the Demographic and Health Survey (DHS) program, sponsored by USAID. These surveys collect information on household members and they ask questions on survivorship of the mother and the father of the child. For instance, questions like "Is this child's mother alive?" and "Is this child's father alive?" for every child aged 0 to 14 years are included in the surveys. This is a primary source of information on orphanhood from which direct estimates of number and prevalence of orphans can be obtained. These estimates can also be disaggregated by geographical area and by socio-economic characteristics of the household.

A second relevant source of information that comes from DHS surveys is given by the sibling history portion of the individual women's and men's interviews. Information about fertility and mortality of the siblings of the respondent is collected. In particular, we know the age at death for the sister and the number of children she gave birth to during her life. With these data it is possible to indirectly estimate the number of maternal orphans.

The main concern with survey estimation is the reliability of the estimates. In situations where there may be economic help to households that take care of orphans, it is possible to have overreporting problems. Conversely, in certain cultural context, an orphan who moves into the household of a relative may not be considered an orphan anymore by the householder, thus leading to a problem of underreporting.

3.2 Methods based on demographic and epidemiological modeling

The HIV/AIDS epidemic affects the number of orphans mainly through increased mortality of children and adults, and reduced fertility of HIV-positive women.

Different methodologies, based on different sets of assumptions, can be used to estimate the number of AIDS-related orphans. For instance, Gregson et al. (1994) evaluate the potential impact of the HIV/AIDS epidemic on orphanhood in sub-Saharan Africa using a mathematical model that combines demographic, biological and behavioral parameters. Jones (2005) uses mathematical demography of kinship to evaluate the impact of the generalized HIV/AIDS epidemic in sub-Saharan Africa on maternal orphanhood probabilities.

Estimates of orphanhood probabilities have been produced by International Organizations with different methods. Recently, there has been some interest in harmonizing the methods across agencies that produce estimates of number of orphans. Here I present the methods that have been used by the US Census Bureau to produce estimates of the number of orphans related to AIDS causes and other causes of death. These methods have been adopted by the Joint United Nations Program on HIV/AIDS (UNAIDS), United Nations Children's Fund (UNICEF) and

US Agency for International Development (USAID) since 2002. The methodology is spelled out in Grassly et al. (2005).

The number of AIDS-related maternal orphans at age a is a quantity that depends on several demographic and economic factors. The HIV/AIDS epidemic affects the intensity of mortality rates for both children and adults. The stage s of HIV infection of the mother, when she gives birth, affects the probability of vertical transmission of the infection to the child (ξ_s). The probability of survival to age a for a child whose mother was in the stage s of HIV at childbearing is:

$$\xi_s l'_a + (1 - \xi_s) l_a$$

where l'_a and l_a are the survival probabilities to age a for an infected and an uninfected child, respectively.

If we know the proportion of women who were in the stage s of HIV when they gave birth and they die before their child reach age a, then we can compute the probability that a child survives to age a given that his or her mother died of an AIDS-related cause τ years ago:

$$\sum_{s=0}^{n} [Y_s(a - \tau)(\xi_s l'_a + (1 - \xi_s) l_a)]$$

where $Y_s(a - \tau)$ represents the proportion of adults who died of AIDS-related causes τ years before the reference date for the present and that were in stage s of the of HIV infection $(a - \tau)$ years before death, that is at the time of child's birth. n is the number of stages of HIV infection considered.

Two other important quantities enter into the calculation of the number of maternal orphans. First, the fertility history of a woman determines the probability that a woman who died of AIDS had a child at a particular age and time. Second, the number of women who died of AIDS-related causes at a specific time. By putting together this information with the previous elaborations and by summing up over age and stages of infection, Grassly et al. (2005) show that the number of maternal AIDS-related orphans of age a at time t, whose mother died τ years ago, can be expressed as:

$$\Omega'_{t,a,\tau} = \sum_{i=15}^{49} \{\mu'_{t-\tau,i+a-\tau} \sum_{s=0}^{n} [m_{i,s,t-a} Y_s(a - \tau)(\xi_s l'_a + (1 - \xi_s) l_a)]\} \tag{1}$$

where $\mu'_{t-\tau,i+a-\tau}$ is the number of women of age $i + a - \tau$ who died of AIDS-related causes at time $t - \tau$. $m_{i,s,t-a}$ is the fertility rate of women of age i in HIV stage s at time $t - a$.

The number of AIDS-related paternal orphans can be estimated using an approach analogous to the one described for calculating the number of AIDS-related maternal orphans, given that data on age-specific fertility profiles of men are available. The major difficulty in evaluating paternal orphanhood is that estimates of concordance of HIV status of the mother and the father are required in order to correctly determine probabilities of vertical transmission and the survivorship of fathers.

Several factors influence the concordance of HIV status of partners. Transmission between partners depends on a series of factors such as the length of the partnership, the stage of the man's HIV infection, the number and types of sexual acts and the prevalence of cofactors that enhance the transmission of HIV (Grassly et al. 2005). The prevalence of HIV among all women is also indicative of how likely it is for a woman to become infected in a different partnership and the level of risk associated to HIV-positive men who select high-risk women. Since for most countries there is not much data on all these quantities, Grassly et al. (2005) estimated the concordance of HIV status between partners using a logistic regression model of prevalence of HIV among the female partners of HIV-positive men against prevalence in the general population, as measured at ante-natal clinics (ANCs). Their regression equation relates the probability that a woman is HIV-positive, given that her partner is positive, to HIV prevalence measured at ANCs. They found that, when the epidemic is not generalized in the population, about 30% of women with a positive partner are predicted to be positive themselves as a result of transmission within the partnership. When the prevalence in the general population rises, then the fraction of HIV-positive women with an HIV-positive partner linearly increases, mainly due to the increased risk of preexisting infection. The prevalence of HIV infection among women with uninfected partners is significantly lower than in the general female population.

The prevalence of maternal and paternal orphanhood is strongly affected by the HIV-AIDS epidemic, through its consequences on mortality and concordance of HIV status between partners. The correlation in HIV status between partners has a relevant effect on the prevalence of double orphanhood. If the risk of dying of a child's mother and father were independent, the expected proportion of double orphans among children of age a would be the product of the proportion of children of age a who lost their mother and the proportion of children of the same age who lost their father. Due to the concordance of HIV status between partners, the probability of being a double orphan is higher than the one expected assuming that the deaths of parents are independent. Grassly et al. (2005) used data from Demographic and Health Surveys conducted in 25 countries to predict double orphanhood as the product of its expected prevalence, under the assumption of independent mortality of parents, and the observed excess risk of double orphanhood relative to its expected risk. They estimated that, for the countries that they considered, the risk of being double orphan for children aged 0 to 14 years is between 2 and 5.7 times the expected risk that would be obtained assuming that the mortality of mother and father were independent.

3.3 Estimates

In this subsection, I provide a quantitative picture of the extent of the orphanhood crisis in sub-Saharan Africa. The data presentation is based on estimates released by international agencies, such as UNICEF and UNAIDS, and obtained using the methods described in the previous section. I focus on southern Africa, which is the region where the extent of the crisis is the largest.

According to UNICEF (2006), at the end of 2005, about 25 million people in sub-Saharan Africa were living with HIV, representing two thirds of the world's population infected with HIV. It is estimated that 2 million of these people were children under age 15.

In 2005, sub-Saharan Africa was home to over 48 million orphans younger than 18 years old. 12 million of them were AIDS-related orphans and about 9 million were double orphans. In other regions of the world, the total number of orphans has been declining over time. Conversely, in sub-Saharan Africa the total number of orphans has been increasing. In 1990, about 31 million orphans were estimated to be in the region, 1% of them orphaned by AIDS. It is expected that in 2010 there will be over 53 million orphans in the region, 30% of them orphaned by AIDS (UNICEF 2006).

Children are suffering the greatest parental loss in southern Africa, where the HIV prevalence rates are the highest, and where in most countries 15% or more of all children are orphans. Table 1 gives estimated summary statistics for orphans in southern African countries for 2005 (UNICEF 2006). It is striking to observe that in countries like Botswana and Zimbabwe the HIV/AIDS epidemic is responsible for the generation of more than 75% of the orphans in the age range 0-17. The highest proportions of children who are orphans in 2005 are observed in

Country	Total number of orphans	% of children who are orphans	Children orphaned by AIDS as % of all orphans	% of children aged 0-5 who are orphans	% of children aged 6-11 who are orphans	% of children aged 12-17 who are orphans
Angola	1,200,000	14	13	6	16	24
Botswana	150,000	19	76	8	22	27
Lesotho	150,000	17	64	8	20	25
Malawi	950,000	15	57	6	17	24
Mauritius	23,000	6	-	2	6	11
Mozambique	1,500,000	15	34	7	16	24
Namibia	140,000	14	62	6	15	19
South Africa	2,500,000	13	49	6	14	19
Swaziland	95,000	17	66	9	20	24
Zambia	1,200,000	20	57	9	23	30
Zimbabwe	1,400,000	21	77	9	24	30

Table 1: Estimated summary statistics for orphans in southern African countries, 2005, released by UNICEF (2006).

Botswana, Zambia and Zimbabwe, where the values are around 20%. We can also notice a steep increase by age in the proportion of orphans. For instance, in Zimbabwe in 2005, 9% of children aged 0-5 are estimated to be orphans. This percentage quickly rises to 30% for children aged 12-17.

Figure 1 shows the relationship between predicted percentage of children expected to be orphans in 2010 and adult HIV prevalence in 2005, for sub-Saharan countries. The dashed line is obtained using Friedman's 'super smoother' (Friedman 1984). It is interesting to observe how well the adult HIV prevalence rate at a particular time is predictive of the percentage of orphans in the country 5 years later.

4 Data

The availability of reliable and comparable data sources is an important limitation for research in sub-Saharan Africa. Even when data and estimates for key demographic quantities are available, it is often hard to assess their uncertainty and reliability. In this section, I give an overview of the data sources that I use for this project that focuses on Zimbabwe. I describe the main data sources that are available for the Zimbabwean setting and the methods that I use to indirectly extract specific sets of rates from available data.

For Zimbabwe, we have two main data sources that provide aggregate demographic rates to inform our models: United Nations (UN) population statistics and the Demographic and Health Surveys (DHS). Relevant demographic data from the United Nations can be found in the *2006 Revision of the World Population Prospects*. These data come in the form of a CD-ROM which contains essential demographic data such as estimates and projections of total births, total deaths, population counts, mortality and fertility indicators by five-year age groups and sex for the period 1950-2050. A complementary data source from the United Nations Population Division is the World Fertility and Marriage Database 2003.

The Demographic and Health Surveys provide rich sample surveys collected in Zimbabwe during the years 1988, 1994, 1999 and 2005/2006. The 2005/2006 DHS for Zimbabwe includes a module on HIV seroprevalence based on HIV testing administered to the respondents.

4.1 Mortality rates

The United Nations publishes estimates and forecasts of number of deaths in Zimbabwe, in presence and absence of AIDS, for age groups of 5 years and time intervals of 5 years, for the period 1980-2050. For the same period, the United Nations provide data on population counts for age groups of 5 years and time intervals of 5 years. These estimates come from different

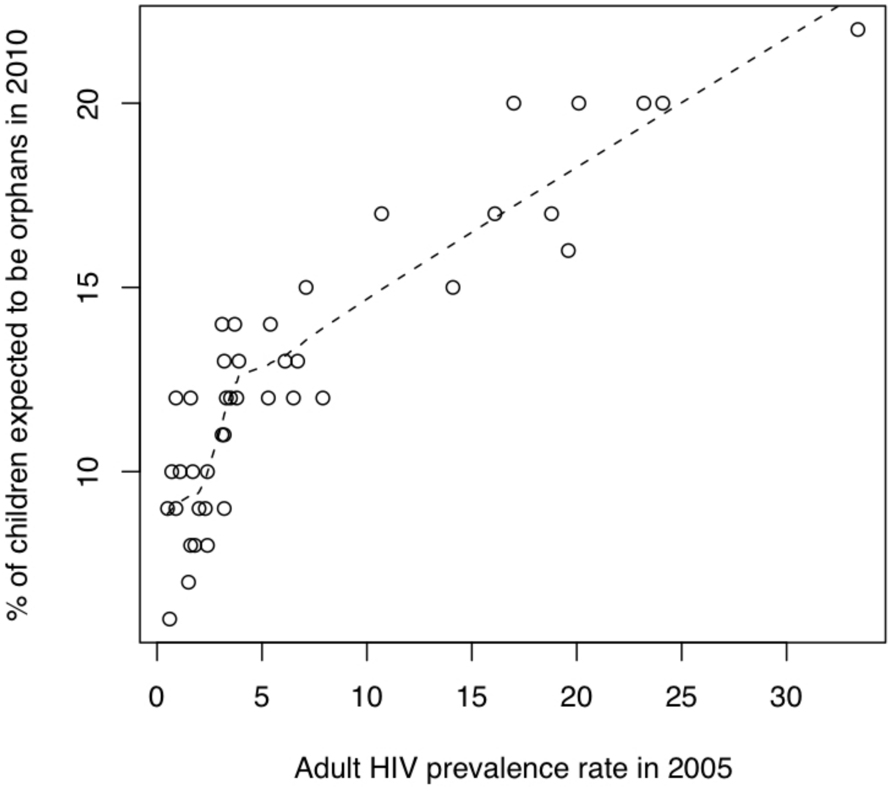

Figure 1: Relationship between predicted percentage of children expected to be orphans in 2010 and adult HIV prevalence in 2005 for sub-Saharan countries. Dots are observations, whereas the dashed line represents an interpolation obtained with a nonparametric smoother. Data source: UNICEF (2006).

data sources, such as census data and DHS. Given these data, it is straightforward to compute mortality rates and thus produce a life table.

Both the model based on aggregate demographic rates and the micro-simulation require information at a finer level, namely for age groups and periods of 1 year. I thus estimated these quantities based on the available data. First, I used a cubic spline interpolation method to estimate the number of deaths and population counts for single age groups over time. In practice, I estimated, for single age groups, the number of deaths and population counts for the mid-period year considered. For example, I obtained the mid-period number of deaths by dividing the whole number of deaths over the period by the length of the period (in years). Once I had estimates for the mid-period years, I then used a cubic spline smoother to predict the values for the years between the mid-period points. From these data I could then obtain the mortality rates for the 5-year age groups considered for periods of single years. Second, I estimated probabilities of death for single-year age groups. From the 5-year age groups mortality rates, I obtained 5-year probabilities of death (${}_4q_x$) using standard life table techniques. By then assuming that within each 5-year age group, the probabilities of death are constant, I could convert such probabilities of death into single years probabilities of death using standard demographic techniques (see Wachter 2007, for details). A special treatment was necessary for the youngest 5-year age group, since mortality is much more likely during the first year of life than in the following 4 years. For the years after 1995, the United Nations provide an estimate of the probability of survival to age 1: I thus used such probabilities for the ${}_nq_x$ conversion method. For years before 1995, I assumed that the relative probability of survival to age 1, with respect to age 5, is equal to the one observed in 1995.

4.2 Fertility rates and marriage statistics

The UN 2006 Revision of the World Population Prospects provides estimates and projections of age-specific fertility rates in Zimbabwe for the period 1995-2050. Rates are available for women in the age range 15-49 years, in age groups of 5 years and for periods of 5 years. For periods before 1995, estimates of total fertility rates for Zimbabwe are provided by the United Nations. I estimated the age-specific fertility rates for Zimbabwe for the period 1980-1994 by multiplying the age-specific profile of fertility of 1995 by a scalar, in order to match the UN estimate of total fertility rate for the respective period.

In order to obtain age-specific fertility rates for age groups and periods of 1 year, I used a cubic spline interpolation technique, with a procedure analogous to the one used for mortality rates.

Data on marital status by age mainly come from the World Fertility and Marriage Database 2003 and the DHS.

4.3 Estimates of new HIV infections by age

Rates of HIV infection are important inputs for the microsimulation. Estimates of new HIV cases are not published by the United Nations, but they can be obtained indirectly from published data. In this section, I present the strategy that I propose to estimate age-specific numbers of new HIV infections.

Using UN data we can compute a life table for all causes of death, including AIDS, and one for causes of death other than AIDS. I use these life tables to compute the survivorship from AIDS related causes only:

$$(1 - {}_nq_x^{ALL}) = (1 - {}_nq_x^{AIDS}) \times (1 - {}_nq_x^{OTHER})$$

Once I have a life table for AIDS-only, then I would like to have a picture of AIDS-related probability of death for a cohort. I thus follow the probability of survival over time and age to have a representation of AIDS mortality for a cohort. For instance, to follow the cohort of people born in 1980, I choose ${}_nl_0^{AIDS}$ for 1980, ${}_nl_1^{AIDS}$ for 1981, ${}_nl_2^{AIDS}$ for 1982, etc. Ultimately, when I observe an increase in the ${}_nl_x$ function, I consider the ${}_nl_x$ before the increase as the ${}_nl_x^{Ultimate}$ and from there on I assume that ${}_nl_{x+1} = 0.999{}_nl_x$.

Using the estimated 'cohort AIDS life table' and the number of births for specific years, we can estimate the number of survivors from AIDS at each age. For a specific cohort, I multiply the number of births by the survival probabilities to get the number of survivors, by age. The difference between consecutive numbers of survivors by age gives the number of AIDS-related deaths for each age group.

Now, let D^{AIDS} be a column vector containing the number of AIDS-related deaths by age for adults:

$$\begin{bmatrix} {}_1d_{15}^{AIDS} \\ {}_1d_{16}^{AIDS} \\ \cdot \\ \cdot \\ {}_1d_{50}^{AIDS} \end{bmatrix} \tag{2}$$

Let P be a matrix containing the probabilities of surviving exactly x years after getting infected with the HIV (p_x):

$$\begin{bmatrix} p_1 & 0 & 0 & 0 & \cdot \\ p_2 & p_1 & 0 & 0 & \cdot \\ p_3 & p_2 & p_1 & 0 & \cdot \\ p_4 & p_3 & p_2 & p_1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \tag{3}$$

Let N^{HIV} be a column vector with new cases of HIV by age:

$$\begin{bmatrix} {}_1n_{15}^{HIV} \\ {}_1n_{16}^{HIV} \\ \cdot \\ \cdot \\ {}_1n_{50}^{HIV} \end{bmatrix} \tag{4}$$

Then, for a cohort of individuals, it holds:

$$D^{AIDS} = P \times N^{HIV} \tag{5}$$

We have estimates for D^{AIDS} and P and we would like to estimate N^{HIV}. In order to estimate N^{HIV} we can perform a linear regression of D^{AIDS} on P. The estimated coefficients are N^{HIV}. Since the least-squares technique may be fairly unstable in this context, the L-1 norm regression is preferable.

5 Insights from the formal demography of kinship

This section is dedicated to the analysis of the orphanhood crisis in Zimbabwe using the tools of formal demography. I discuss the insights that we can get in the context of stable populations and populations with changing demographic rates over time. I also discuss ways to incorporate heterogeneity in models based on aggregate demographic rates. Methods based on the formal demography of kinship require only a basic set of demographic rates and are widely applicable to lots of different countries for which UN data are published. They are thus particularly interesting for comparative purposes and to get the main features of the effect of the epidemic on orphanhood probabilities.

5.1 Limiting probabilities of maternal orphanhood

A good starting point for understanding the process of orphans generation is the stable population theory. Keyfitz and Caswell (2005) discuss the stable theory of kinship with reference to the work of Lotka (1931), Burch (1970), Coale (1965), Goodman, Keyfitz and Pullum (1974) and Le Bras (1973). An important result that they present is the analytic representation of the probability that a girl aged a has a living mother $M_1(a)$, under a given regime of mortality and fertility:

$$M_1(a) = \int_\alpha^\beta \frac{l_{x+a}}{l_x} e^{-rx} l_x f_x f_{fab} dx \tag{6}$$

where l_x is the probability of survival to age x, f_x is the fertility rate at age x, f_{fab} is the fraction of females at birth, and r is the Lotka's intrinsic growth rate.

Using estimates and projections of vital rates for Zimbabwe for the period 1980-2050, we can compute the limiting stable probability of maternal orphanhood at age a $(1 - M_1(a))$ under different regimes of mortality and fertility. This means that, for each period of one year, we can evaluate the probabilities of orphanhood that result from the persistence of current mortality and fertility rates. In practice, this entails computing the Lotka growth rate for each year, and plugging the associated demographic rates into equation 6 (see Wachter 2007 for details about numerically finding the value of the Lotka r).

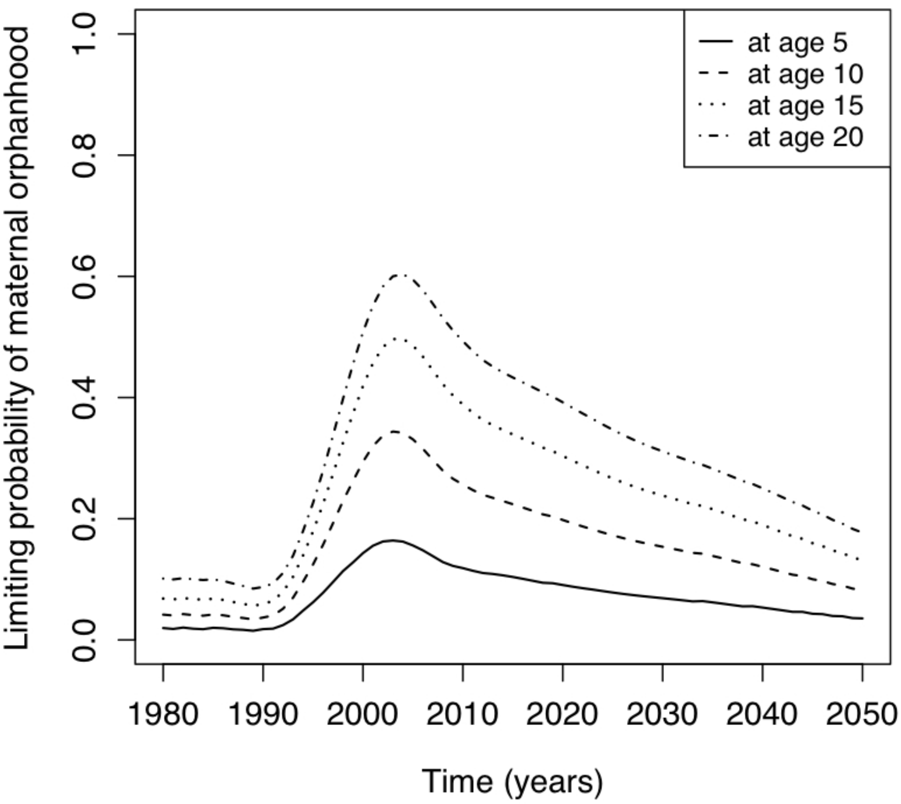

Figure 2: Estimates and projections of limiting stable probabilities of maternal orphanhood implied by mortality and fertility rates for selected years and ages in Zimbabwe. Data source: own elaborations on UN World Population Prospects, 2006 Revision.

Figures 2 shows estimates and projection of limiting maternal orphanhood probabilities for Zimbabwe. The estimates for each year in time are generated assuming that the age-specific mortality and fertility rates for the year considered will persist unchanged in the future. The figure shows the considerable impact of the HIV/AIDS epidemic on the probability of maternal orphanhood. For instance, the estimated probability of maternal orphanhood at age 10 is 0.04 in 1990. With the rapid increase in the number of AIDS-related deaths, such probability rapidly grows to 0.12 in 1995, 0.29 in 2000, up to 0.34 in 2004, before slowly decreasing, in accordance with the projected demographic rates. The same pattern characterizes the probability of maternal orphanhood at other ages.

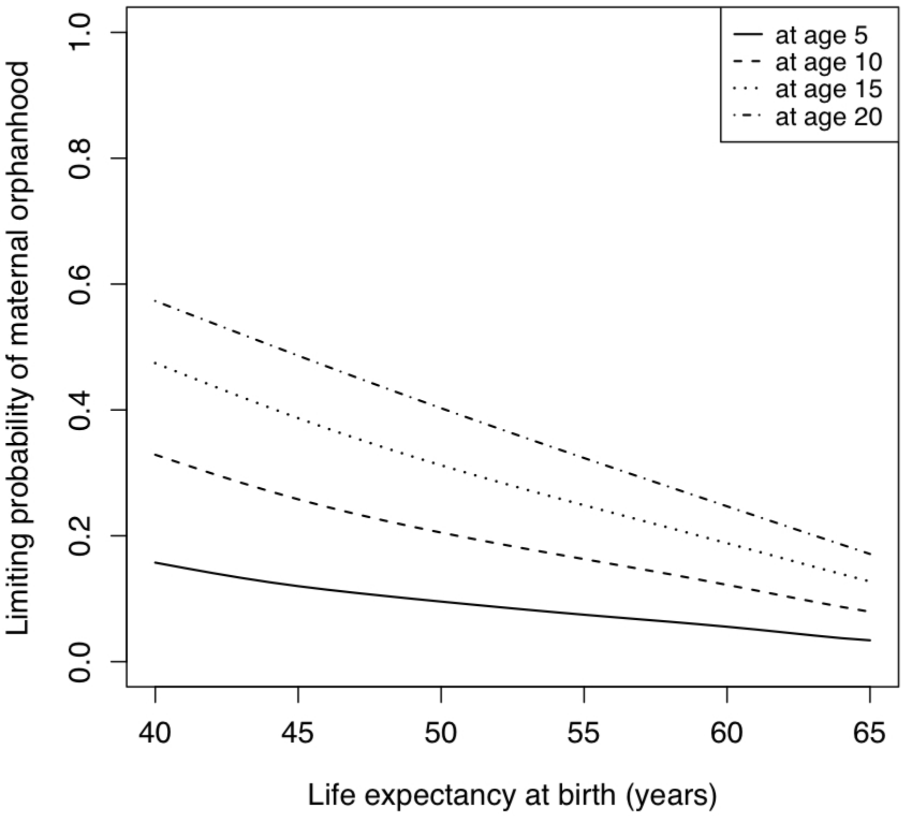

Figure 3: Estimates of limiting stable probabilities of maternal orphanhood associated to different levels of AIDS-related life expectancy at birth in Zimbabwe. Data source: own elaborations on UN World Population Prospects, 2006 Revision.

Figures 2 gives a representation of the impact of the HIV/AIDS epidemic over time and age. It is also interesting to understand the impact of AIDS-related changes in life expectancy on the probability of maternal orphanhood. If we choose some index of the level of mortality, then we can determine the effect on $M_1(a)$ of a change in mortality as the first derivative of $M_1(a)$ with respect to the index of mortality. For example, we can fit a model life table, such as the Brass relational logit model (Brass, 1974), to the estimates and projections of mortality rates in Zimbabwe for the period 1980-2050. We can choose the life table for the year at the beginning of the period under consideration and make it our standard for the Brass relational model. Then we estimate the shape and level parameters of the model over time and we associate them to different levels of life expectancy. We can thus relate couples of estimated parameters (α, β) to an index of mortality i, such that $\frac{\partial e_0}{\partial i} = 1$, and estimate the impact of a change in the index i on the probability of maternal orphanhood at age a. Analytically, this entails computing $\frac{\partial M_1(a)}{\partial i}$, a quantity that varies according to the initial level of life expectancy and the age groups that are mostly affected by the change in mortality patterns (under a generalized HIV/AIDS epidemic, a reduction in life expectancy is driven by an increase in adult mortality, rather than infant or old age mortality).

The approach that I just described is appealing, but it has a major limitation. The Brass model does not fit mortality profiles typical of the HIV/AIDS epidemic in a satisfactory way. I thus chose a different method to evaluate the effect of changing life expectancy on probabilities of maternal orphanhood. I looked at the probabilities of maternal orphanhood associated to different levels of life expectancy for the period 1990-2050 and I used a cubic spline smoother to interpolate such probabilities for missing values of life expectancy. Figure 3 shows the estimated probabilities of maternal orphanhood at age 5, 10, 15 and 20, respectively, associated to values of life expectancy between 40 and 65. As expected, higher levels of life expectancy are associated to lower orphanhood probabilities. It is also interesting to note that the slope of the lines is higher when evaluated at lower levels of life expectancy. This means that an improvement in life expectancy of one year, from a starting level of 40 years, generates a larger reduction in probabilities of orphanhood than an analogous improvement from a starting level of 60 years of life expectancy.

5.2 Approximation of probabilities of maternal orphanhood in the context of a generalized HIV/AIDS epidemic

If the maternity function is relatively concentrated around the mean age at childbearing, then the probability that a girl who is a years old has a living mother should be mainly determined by the probability of surviving a years past the mean age at childbearing. This intuition is behind an interesting approximation of $M_1(a)$, the probability that a girl aged a has a living mother under a given regime of mortality and fertility (e.g. Keyfitz and Caswell 2005). In this section, I will discuss such approximation in the context of a generalized HIV/AIDS epidemic. I will also consider the insights that we can get regarding the role that key demographic quantities have in shaping the probability of maternal orphanhood.

Keyfitz and Caswell (2005) showed that $M_1(a)$ can be approximated using a Taylor expansion of l_{x+a}/l_x around κ, the mean age at childbearing:

$$M_1(a) \approx \frac{l_{\kappa+a}}{l_\kappa} + \frac{\sigma^2}{2}(\frac{l_{\kappa+a}}{l_\kappa})'' \tag{7}$$

where σ^2 is the variance of ages at childbearing, and $(l_{\kappa+a}/l_\kappa)''$ is the second derivative of l_{x+a}/l_x evaluated at $x = \kappa$.

Let us now consider the probability that a girl aged a does not have a living mother. Following equation 7, an approximation to such quantity can be written as:

$$1 - M_1(a) \approx 1 - \frac{l_{\kappa+a}}{l_\kappa} - \frac{\sigma^2}{2}(\frac{l_{\kappa+a}}{l_\kappa})'' \tag{8}$$

Equation 8 shows that, to a linear extent, $l_{\kappa+a}/l_\kappa$ is a good approximation to $M_1(a)$. The two main factors affecting the accuracy of such approximation are the variance of ages of mothers and the concavity of the survivorship function.

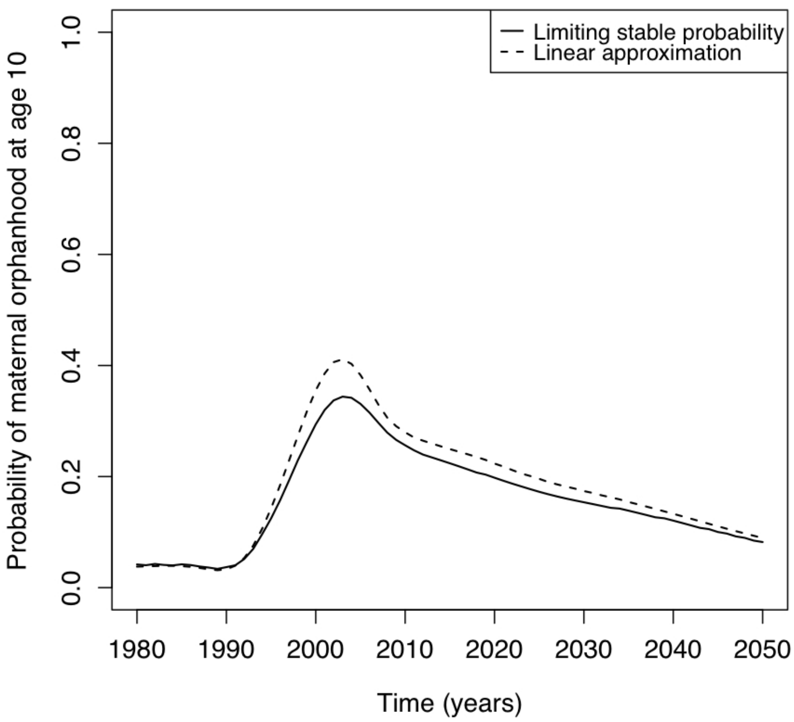

Figure 4: Estimates and projections of limiting stable probabilities of maternal orphanhood at age 10 in Zimbabwe and the respective approximation based on Taylor expansion. Data source: own elaborations on UN World Population Prospects, 2006 Revision.

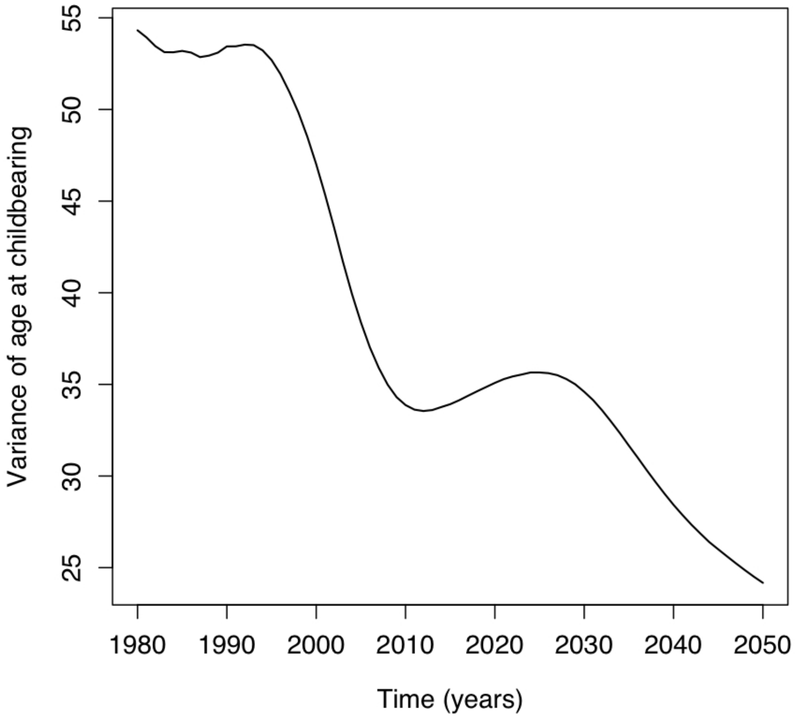

Figure 5: Estimates and projections for the variance of the age at childbearing in Zimbabwe from 1980 to 2050. Data source: own elaborations on UN World Population Prospects, 2006 Revision.

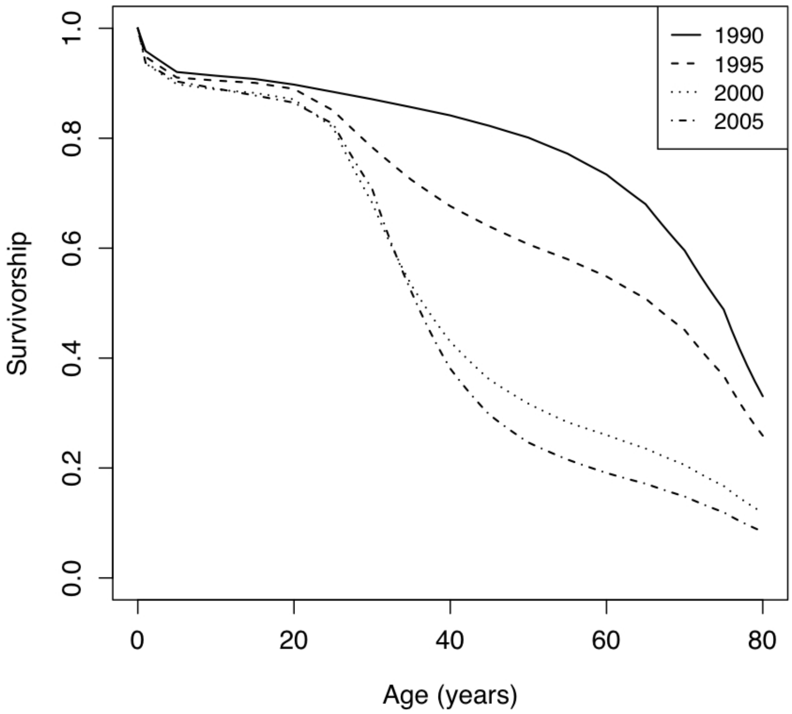

Figure 6: Estimates of female survival probabilities for selected years in Zimbabwe, in presence of HIV/AIDS. Data source: own elaborations on UN World Population Prospects, 2006 Revision.

Figure 4 shows the evolution over time of the probability of maternal orphanhood at age 10. Both the stable limiting probability and its linear approximation $(1 - \frac{l_{\kappa+10}}{l_\kappa})$ are presented. It is interesting to observe that the approximation is fairly accurate for the period from 1980 to the early 1990s. Then the approximation tends to systematically overestimate the exact value for the probability of orphanhood. This is related to the two factors mentioned earlier: the variance of the age at childbearing and the concavity of the survivorship function. Figure 5 shows the evolution over time of the variance of the age at childbearing. Figure 6 shows the female survival probabilities for selected years in Zimbabwe. Until the end of the 1980s, the survival function is concave at adult ages and its second derivative is thus negative. This makes the product $\frac{\sigma^2}{2}(\frac{l_{\kappa+a}}{l_\kappa})''$ negative and thus the approximation underestimates the exact value for the probability of orphanhood. Starting from the mid-1990s, we observe that the variance of the age at childbearing substantially decreases, except for a hump in the 2020s. We would expect the approximation to become increasingly more accurate. That is true until the mid-1990s, but it is not the case afterwards. The main reason is that the HIV/AIDS epidemic has completely transformed the shape of the survivorship function. The absolute value of the second derivative of the survivorship function, evaluated at the mean age at childbearing, decreases over time until the mid-1990s, when there is a change in sign of the second derivative, and then it rapidly increases. After the mid-1990s, the survivorship function becomes convex, thus explaining the overestimation of the exact value of the probability. In addition to that, the rapid increase in the value of the second derivative of the survivorship function more than counteracts the reduction in the variance of the age at childbearing, thus making the approximation less accurate.

5.3 Insights from the approximation of probabilities of maternal orphanhood

So far, I have focused my attention on the use of the theory of stable population to get insights on the impact of the HIV/AIDS epidemic on limiting probabilities of maternal orphanhood. In the previous sections, I considered the computation of orphanhood probabilities based on demographic rates for Zimbabwe, and I discussed an approximation of these quantities. In this section, I would like to further develop on the intuition behind the approximation that orphanhood probabilities are related to survival past the mean age at childbearing. For illustrative purposes, I will discuss the insights that we can get from an extremely simplified representation of the process of orphan formation.

Let us a consider a female population of adults who get infected with HIV at some point in their lives. Two demographic quantities are crucial in the explanation of related orphanhood probabilities: the age at childbearing (a_c) and the age at death (a_d) for causes associated to HIV infection. The age at death mainly depends on the age at infection with HIV, the incubation period from the infection to the development of AIDS, and the period from development of AIDS to death. For simplicity, and for practical purposes, I consider the incubation period and the period between development of AIDS and death as one period. I define this quantity inf, the period of infection before death. We have that $a_d = a_{hiv} + inf$, where a_{hiv} is the age at infection with HIV.

Let M be the difference between the age at death and the age at childbearing ($M = a_d - a_c = a_{hiv} + inf - a_c$). We can think of M as a random variable representing the number of years a hypothetical child spends with a living mother, given that the mother acquires HIV infection sometimes during her life. If M is negative, the woman had died before having the child. Such a death implies a reduction in the number of newborns, who are also potential orphans.

Assuming independence between the components of the age at death and age at childbearing, we have that the expected value of M is given by:

$$E[M] = E[a_{hiv}] + E[inf] - E[a_c] \tag{9}$$

The variance of M is:

$$Var[M] = Var[a_{hiv}] + Var[inf] + Var[a_c] \tag{10}$$

The expected value of M tells us approximately for how many years on average a child has a living mother. The variance for M gives us an idea of the uncertainty about the number of years a chid is expected to have a living mother. In absence of any other information, we can use the Chebishev inequality to evaluate the probability that M is less than zero or that M is bigger than 15. If M is less than zero, the woman has died before having the child. If M is bigger than 15, the child has lived at least the first 15 years of his/her life with a living mother.

The Chebishev inequality can be used to obtain information on one side of the distribution of the random variable, if the distribution is symmetric. Given our hypotheses on the random variables considered, we cannot necessarily make such assumption and the use of the Chebishev inequality is mostly for illustrative purposes. What is interesting to observe is that with the increase in the variance of M, which is ultimately related to an increase in the variance of its components, we have an increase in the probability of 'extreme' events, which implies a reduction in births of potential orphans and a reduction of those children who have a living mother for less 15 years.

The same intuitive conclusion can be obtained using a normal approximation. M is composed of three random variables. Each of these variables can be reasonably modeled with a gamma distribution. In the three cases, the distributions are characterized by shape parameters which are fairly large. Therefore, these gamma distributions can be well approximated by normal distributions. M can thus be seen as a sum of normal distributions with mean and variance given by the expressions in equations 9 and 10, respectively. We can thus compute the probability that M is less than zero or bigger than 15 using the normal distribution.

The very simple analytical scheme depicted in this section provides some important insights with regard to orphanhood probabilities related to the onset of the HIV/AIDS epidemic. First, the higher the expected value of the age at infection and the incubation period, relatively to the expected value of the age at childbearing, the higher the chances that a child would live throughout his childhood with a living mother. [†]

The second important observation is that the variance of the age at infection, the age at childbearing and the incubation period play an important role in shaping the probabilities of orphanhood. The larger the sum of these variances, the less important the values of their means are in determining the probabilities of orphanhood. If both age at childbearing and age at death from AIDS have very small variances, then most children will likely have a living mother for a number of years equal to the difference between the two random variables. If the variance of the age at childbearing is large, then we would observe some children living throughout childhood with a living mother and some others losing their mother while very young.

The reasoning developed in this section derives from the assumption that the random variables under consideration can be modeled with a Gamma distribution. Although this could be a good approximation, variables such as the age at childbearing are bounded between certain values. For a more realistic representation of the demographic processes involved, generalizations of Beta distributions could be used instead. That would make the illustrative example more rigorous, without adding much in terms of the intuition.

[†] This reasoning would not hold if people got infected very early in life and the incubation periods were so short that virtually everyone would die before reaching childbearing age.

5.4 Probabilities of maternal orphanhood when vital rates evolve over time

Keyfitz and Caswell (2005) show that the problem of evaluating the probability of orphanhood at a given age may be approached using a renewal equation. In particular, the number of living mothers for girls aged a at time t is given by:

$$\int_{\alpha}^{\beta} B(t-a-x) l_x f_x f_{fab} \frac{l_{x+a}}{l_x} l_a dx \tag{11}$$

where $B(t-a-x)$ stands for female births at time $t-a-x$.

In equation 11, it is implicitly assumed that mortality and fertility rates are constant over time. If instead we assume that vital rates change over time, and we have estimates and projections of these rates for a rather long period of time, we can use different sets of fertility and survivorship schedules in order to account for the fact that mothers and daughters at different periods of time experience different fertility and mortality conditions. In such a case, the number of living mothers to girls aged a at time t is:

$$\int_{\alpha}^{\beta} B(t-a-x) l_x(t-a-x) f_x(t-a-x) f_{fab} \frac{l_{x+a}(t-a-x)}{l_x(t-a-x)} l_a(t-a) dx \tag{12}$$

The elements between parentheses in expression 12 are to be intended as the year of birth of the cohort members to which the rates apply. Expression 12 is appealing, but there are two main limitations to its practical application. The first one is that data series that extend far back in the past and future are required. For instance, we need to know the number of births of prospective mothers in the past, their survival probabilities and fertility schedules, and the survival probabilities of recently born children up to the age of interest. The second limitation is that the approach relies on cohort data, that is a longitudinal representation of fertility histories and hazards of mortality.

The two limitations can be overcome by re-writing expression 12. As a matter of fact, orphanhood probabilities can be evaluated using estimates and projections of demographic rates over time provided by the United Nations. With regard to the first limitation, it is important to observe that we do not need to go as far back in time to keep track of the number of births for the cohorts of prospective mothers. The relevant information that we need to know is the probability of the age of a woman who gave birth to a child a years before the time of reference. This probability depends on the age structure of the population and the age-specific fertility rates in place a years before the time of reference. Let A_b be the age at giving birth for a woman. Then the probability that the age at giving birth for a woman a years before the time of reference is equal to x, $P(A_b = x, t-a)$, is given by the proportion of births to women of age x, a years ago:

$$P(A_b = x, t-a) = \frac{{}_1f_x(t-a) \times {}_1K_x^f(t-a)}{\sum_{x=0}^{100} {}_1f_x(t-a) \times {}_1K_x^f(t-a)} \tag{13}$$

Equation 13 is expressed in discrete terms. Although a continuous representation of the equation is more compact and elegant, a discrete one is used to be consistent with the empirical analysis which is carried out in discrete terms. Figure 7 shows the evolution over time of the probability of mother's age at childbirth in Zimbabwe. We observe that with the onset of the HIV/AIDS epidemic, the probability of giving birth at an earlier age increases. Projections for the next decades show that the age at childbearing will increase. This is consistent with the estimated mean ages at childbearing over time, as they appear in figure 8

With regard to the second limitation, that is the absence of cohort data, we can estimate approximate cohort histories from period data. For instance, using demographic rates provided by the United Nations, we can generate a rectangular array of probabilities of surviving n

years past age x, $(\frac{l_{x+n}}{l_x})$. Such array will typically have rows representing age and columns representing years. By looking at the array over diagonals, we consider rates that apply to the same cohort of individuals and we can thus have a picture of cohort histories.

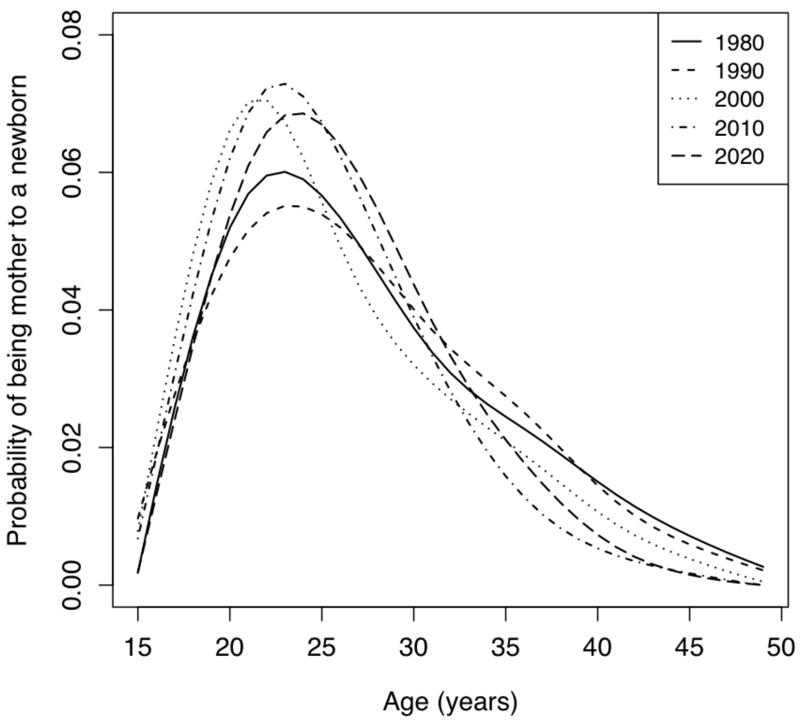

Figure 7: Estimates and projections of probabilities of mother's age at childbirth in Zimbabwe. Data source: own elaborations on UN World Population Prospects, 2006 Revision.

The probability that at time t a child aged a randomly selected from the population has lost her mother is given by the probability that his/her mother has not survived a years past giving birth to the child considered:

$$[1 - (\sum_{x=0}^{100} \frac{l_{x+a}(t-x-a)}{l_x(t-x-a)} \times P(A_b = x, t-a))] \tag{14}$$

Expression 14 gives the probability that a child standing in front of us has lost her mother, when we do not account for the survival probabilities of the child. Expression 14 is a weighted average of the probabilities of surviving a years past childbearing, where the weights are the probabilities of giving birth at a specific age.

Figure 9 shows estimates and projections of maternal orphanhood probabilities over time for a randomly selected child of age 5, 10, 15 and 20, independently of children survival.

If we want to account for the fact that some children will not survive long enough to become orphans and we thus want to compute the probability that a random newborn will be orphan at age a, then we need to take into account the survival probabilities of children. The probability

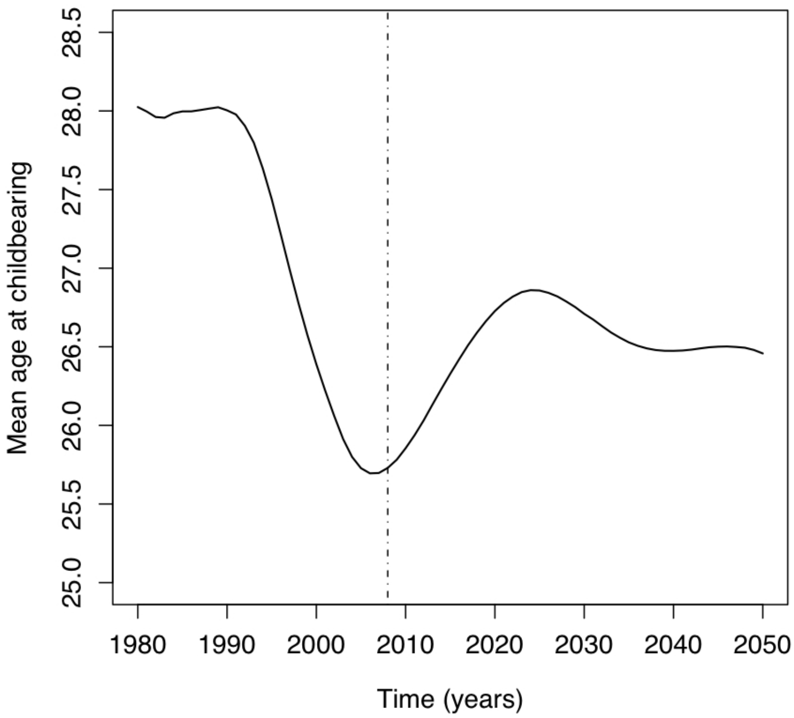

Figure 8: Estimates and Projections of mean age at childbearing for Zimbabwe. Data source: own elaborations on UN World Population Prospects, 2006 Revision.

that at time t a child of age a is alive, whereas his/her mother is not becomes:

$$l_a(t-a) \times [1 - (\sum_{x=0}^{100} \frac{l_{x+a}(t-x-a)}{l_x(t-x-a)} \times P(A_b = x, t-a))] \quad (15)$$

Figure 10 shows estimates and projections of maternal orphanhood probabilities over time for a child of age 5, 10, 15 and 20. Such estimates account for the survival probability of the child.

The results showed in figures 9 and 10, although different in scale, provide the same general picture for the trend in probabilities of orphanhood for the future. It is interesting to note how the generation of orphans is a process with a lag with respect to the generation HIV cases. In the previous section, we observed the probabilities of orphanhood implied by the persistence of specific conditions over time. We saw that the highest probabilities of orphanhood are implied by the persistence of the conditions at the peak of the HIV epidemic. Here we see how the consequences of a peak in the HIV epidemic on generation of orphans persist for a rather long period of time past the reduction in HIV prevalence. After reaching the peak in the epidemic, the probability of being orphan at age 5 decreases more quickly than the probability of being orphan at age 15. With the reduction in AIDS-related deaths, the most recently born cohorts of children are less likely to become orphaned. For the older children, the probability of being orphan is more related to the conditions in the past, as the probability of being orphan at a specific age is the cumulative result of the probabilities of being orphan at previous ages. These

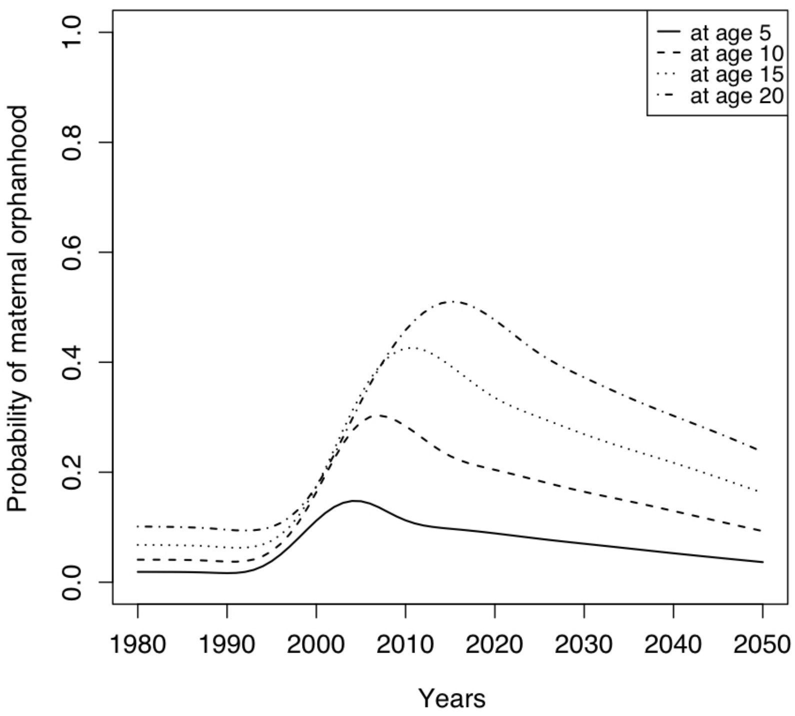

Figure 9: Estimates and projections of maternal orphanhood probabilities for children, at age 5, 10, 15, 20 in Zimbabwe. Data source: own elaborations on UN World Population Prospects, 2006 Revision.

results indicate that with the containment of the epidemic, the scale of the orphanhood problem may become smaller for young children, but may continue increasing for teenagers.

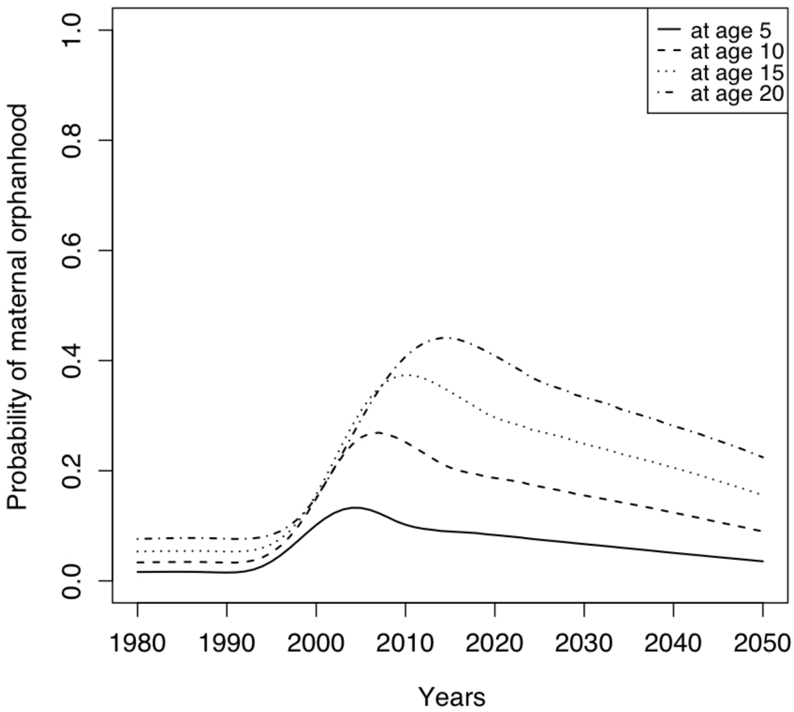

Figure 10: Estimates and projections of maternal orphanhood probabilities for children, accounting for survival probabilities of children, at age 5, 10, 15, 20 in Zimbabwe. Data source: own elaborations on UN World Population Prospects, 2006 Revision.

5.5 Estimation and projection of the number of maternal orphans

We previously considered the problem of estimating probabilities of orphanhood at different ages. It is also important to quantify the scale of the orphanhood problem and to provide estimates and projections of absolute numbers of orphans. In this section, I develop on the ideas that I discussed earlier and I suggest an approach to the estimation of the number of maternal orphans.

The United Nations World Population Prospects provide demographic rates for virtually all countries in the world in a standardized format. Here I propose a method to use such data set to estimate and project the number of maternal orphans over time. I will provide estimates for Zimbabwe, which are informative for the specific geographic setting. The relevance of the method is that can be easily applied to virtually all countries in the world and requires only demographic rates that are readily available. The main drawback of the method is that estimates may be less accurate than the ones obtained using several different data sources and epidemiological models.

The United Nations provides estimates and projections of population counts by age and sex, and age-specific fertility rates. Based on that information, it is straightforward to estimate the number of births to mothers of age x at time t ($B_x(t)$). Orphans of age a at time t are children who were born a years before time t, who survived a years and whose mothers have not survived a years past giving birth to them. The number of maternal orphans of age a at time t (MO_a^t) can be expressed as a weighted average of the number of births a years before time t, where the weights are the survival probabilities of children and their mothers:

$$MO_a^t = (1 - {}_aq_0(t-a)) \sum_{x=15}^{49} B_x(t-a) \times {}_aq_x(t-x-a) \tag{16}$$

The expressions between parentheses refer to the year of birth of the members of the cohort to which the demographic quantities apply. The idea behind this estimation procedure is to follow the same group of people over time and age in order to have demographic rates that give a longitudinal representation of life histories.

In some cases, it is interesting to know the overall number of maternal orphans in a specific age group at time t. For instance, UNICEF and other international organizations report estimates of orphans in the age group 0-17 years. Such a quantity, MO_{0-17}^t can be easily computed as:

$$MO_{0-17}^t = \sum_{a=0}^{18} MO_a^t \tag{17}$$

I applied equations 16 and 17 to data for Zimbabwe. Figure 11 shows estimates and projections of the number of orphans in the age group 0-17 years for the period 2000-2050. We observe a rapid increase in the number of maternal orphans between 2000 and 2010. Maternal orphans in Zimbabwe will continue to increase and will be more than 1 million between 2010 and 2020. Although HIV incidence is expected to steadily decline after a peak between 2000 and 2010, the consequences of the epidemic on orphanhood will be felt for decades. The number of maternal orphans are expected to slowly decrease only after 2030.

UNICEF does not provide estimates and projections of maternal orphans over a long period of time. In the 2006 report on children affected by AIDS, UNICEF reported for Zimbabwe an estimated number of orphans (both maternal and paternal) of 1.4 million in 2005 and a projected number of 1.3 million for 2010 (UNICEF 2006). The estimated number of maternal orphans for Zimbabwe for 2005 reported by UNICEF is 1.1 million, whereas paternal orphans are estimated to be 0.92 million and double orphans 0.7 million. More recent estimates of UNICEF for total orphans in 2007 give a value of 1.3 million for Zimbabwe. Such value is

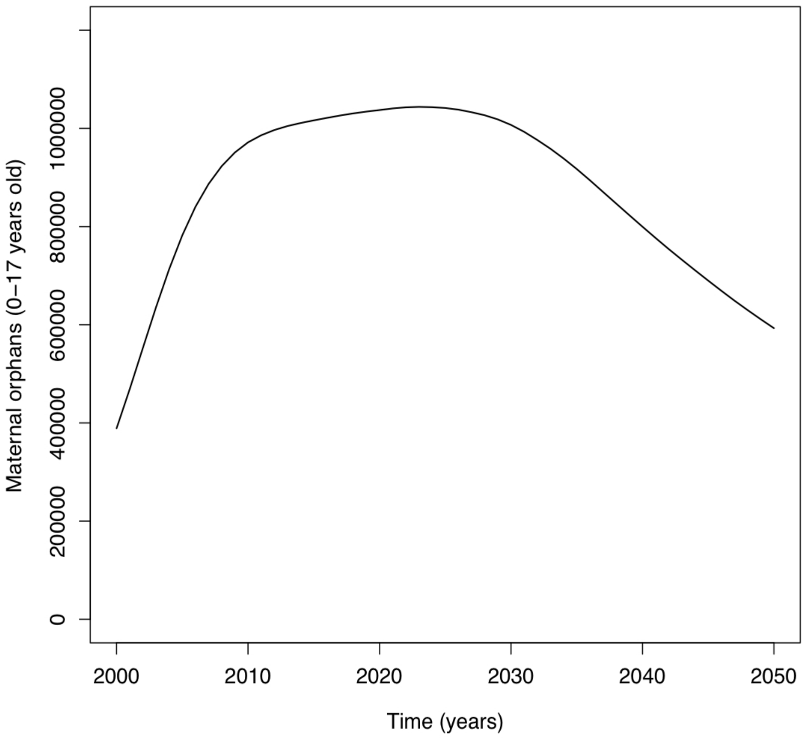

Figure 11: Estimates and projections of the number of maternal orphans who are between 0 and 17 years in Zimbabwe during the period 2000-2050. Data source: own elaborations on UN World Population Prospects, 2006 Revision.

consistent with a downward revision of the estimated number of orphans as a consequence of a downward revision in the estimates of HIV prevalence within the UN system. Altogether, estimates reported by UNICEF seem to be slightly higher than the ones obtained using the mathematical model based on demographic rates. Such differences may be due to different underlying assumptions about HIV prevalence, or may reflect randomness (UNICEF provides point estimates for maternal orphans only for one year). Independently of discrepancies that we may observe when we estimate the same quantities from different data sources (see, for instance, Robertson et al. 2008), the key insight of using the model that I suggested is that, even under the assumption of a fairly rapid reduction in HIV prevalence rates, the orphanhood problem in Zimbabwe will not decrease for at least a couple of decades.

5.6 The effect of heterogeneity on estimates of maternal orphanhood

In the previous sections, I discussed how the mathematical demography of kinship is of help for understanding the process of generation of maternal orphans. I then suggested an approach to estimate maternal orphans solely on the basis of demographic rates. The main advantage of estimating and forecasting orphans on the basis of aggregate demographic rates is that estimates and projections of those rates are widely available and allow for comparative analyses.

An important limitation of a procedure based on aggregate data is that the underlying population is implicitly assumed to be homogeneous, with the same set of rates that apply to everyone who belong to a specific age group. With the onset of the HIV epidemic, the assumption of homogeneity is strongly challenged. In the section, I deal with the issue of incorporating heterogeneity into mathematical models that rely on aggregate demographic quantities.

A convenient way of introducing heterogeneity into models for survival data is to allow for hazard rates to be different between individuals, according to a proportionality factor that is individual-specific. Such proportionality factors has become popular under the name of 'frailties', after Vaupel, Manton and Stallard (1979) formalized some ideas on using frailties for survival analysis, and to get insights on the dynamics of mortality.

In the previous section, with equation 16, I introduced a formula for estimating the number of maternal orphans of age a at time t, MO_a^t. Equation 16 can be written in terms of hazard rates:

$$MO_a^t = e^{-{}_aH_0(t-a)} \sum_{x=15}^{49} B_x(t-a) \times (1 - e^{-{}_aH_x(t-x-a)}) \tag{18}$$

where ${}_aH_x$ is the cumulative hazard rate from age x to age $x+a$. The expressions between parentheses refer to the year of birth of the members of the cohort to which the demographic quantities apply.

Assume that the hazard rate for each child and each mother in the population are proportional to the average hazard rate that we observe at the population level, for the respective groups, and that the coefficients of proportionality are positive random variables, z_c and z_m for children and mothers respectively. The expected value of the frailties z_c and z_m has to be equal to 1, so that the average hazard rates for the overall population are consistent with the aggregate ones. We also expect the frailties z_c and z_m to be correlated. Higher mortality risks for mothers, due to HIV infection or bad sanitary conditions, etc., are likely to be associated to high mortality risks for their children, since the mother and the child supposedly live in the same geographic area and there is a possibility of perinatal transmission of HIV.

By including frailties into our analysis, we can re-write equation 18 as:

$$MO_a^t = e^{-z_c {}_aH_0(t-a)} \sum_{x=15}^{49} B_x(t-a) \times (1 - e^{-z_m {}_aH_x(t-x-a)}) \tag{19}$$

or, equivalently:

$$MO_a^t = \sum_{x=15}^{49} (B_x(t-a) \times e^{-z_c {}_aH_0(t-a)}) - (B_x(t-a) \times e^{-z_c {}_aH_0(t-a) - z_m {}_aH_x(t-x-a)}) \tag{20}$$

To get some insights on the role of heterogeneity within a population, we can focus on a specific component of equation 20: $e^{-z_c {}_aH_0(t-a) - z_m {}_aH_x(t-x-a)}$. Such component gives the proportion of births that survive to age a with a living mother. If we derive a second order Taylor expansion for the component of interest, we get (omitting the terms between parenthesis for readability purposes):

$$\begin{aligned} e^{-z_c {}_aH_0 - z_m {}_aH_x} \approx\ & 1 - {}_aH_0 - {}_aH_x + \frac{1}{2}(({}_aH_0)^2 + {}_aH_0 {}_aH_x){z_c}^2 + \frac{1}{2}(({}_aH_x)^2 + {}_aH_0 {}_aH_x){z_m}^2 + \\ & + \frac{1}{2}(({}_aH_0)^2 + ({}_aH_x)^2 + 2{}_aH_0 {}_aH_x) z_c z_m \end{aligned} \tag{21}$$

We observe that the expected value of $e^{-z_c {}_aH_0 - z_m {}_aH_x}$ is related to $\mathrm{E}[{z_c}^2]$, $\mathrm{E}[{z_m}^2]$ and $\mathrm{E}[z_c z_m]$, which are proportional to, respectively, the variance of z_c, the variance of z_m and the covariance of z_c and z_m. In other words, the larger the heterogeneity in the population and the larger the covariance between the frailty of the mother and the one of her children, the higher the proportion of births that survive to age a with a living mother. When risks of mortality are correlated between mother and children, and having a specific disease, such as HIV/AIDS, highly increases the probability of death for a subgroup of the population and for their children, then it is more likely to observe that either both the mother and the child survive or they both die within a years from the birth of the child. The probability of maternal orphanhood (the child survives, but not the mother) is reduced in a context of a heterogeneous population, compared to a homogeneous one, but the reason for such reduction is a high level of mortality for a subgroup of the population.

5.7 Estimation of other kinship quantities

In the previous sections, I discussed and estimated probabilities of maternal orphanhood in Zimbabwe at different ages and for different periods of time. There are other relevant quantities of interest, such as the expected number of siblings or aunts available to children, or the probability that a specific number of siblings or aunts are available to children (either orphans or not). It is not the purpose of this section to discuss those quantities in detail, and to provide estimates based on an analytical model. A full quantitative description of the kinship structure and the kinship resources available to orphans will be obtained using a microsimulation model. In this section, I will give a brief introduction on how to approach the problem of estimating analytically the availability of kinship resources for orphans.

We can compute the probability that, at time t, a child of age a has k living older siblings (brothers or sisters), $P(S_a = k)$, using either data on age-specific parity or age-specific fertility rates. When data on parity are available (for instance from sample surveys such as the DHS), we can compute such probability as:

$$\begin{aligned} P(S_a(t) = k) &= l_a(t-a) \times \sum_{x=0}^{100} P(Parity_{age_{x-1}}(t-a-1) = k) \times \\ &\times \; P(A_b = x-1, t-a-1) \times E[\prod^{k} l_{j+a}(t-a-j)] \qquad (22) \end{aligned}$$

where js are random integer numbers drawn with replacement in the interval [1, x-15].

In equation 22, the probability that a child of age a has k living older siblings is given by the probability that the child survives until age a, multiplied by the probability that his/her mother had parity k before the birth of the child, multiplied by the expected probability of survival of the k children. The probability that the mother had parity k before the birth of the child under consideration depends on the age of the mother at the birth of the child that we consider, and the probability of giving birth at different ages. We sum over all ages of the mother to remove the conditional probability. Finally, the survival of the siblings is related to when they were born. Since we do not know such information, we assume that those children were evenly spaced during the fertile period of the mother.

If we do not have data on parity by age, but only age-specific fertility rates, then we can think of the number of children for a woman of age x as the sum of $(x-1)$ Poisson trials, where for each trial the probability of success (having '1' instead of '0', or in other words having a newborn) is equal to the age-specific fertility rates for the respective age groups. Let T_s be a random variable such that:

$$T_s = \begin{cases} 1 & \text{with probability equal to the age-specific fertility rate for age } s,\ {}_1f_s \\ 0 & \text{otherwise} \end{cases}$$

Then the probability that a child of age a at time t has k living older siblings (brothers or sisters), $P(S_a(t) = k)$ is:

$$\begin{aligned} P(S_a(t) = k) &= l_a(t-a) \times \sum_{x=0}^{100} P(\sum_{s=15}^{x-1} T_s(t-a-1) = k) \times \\ &\times \quad E[\prod_{j=1}^{k} l_{j+a}(t-a-j)] \times P(A_b = x-1, t-a-1) \end{aligned} \tag{23}$$

An analogous approach can be used to estimate the number of younger siblings. The formulas that I just presented can also be generalized in order to estimate the probability that at time t a child of age a has a certain number of maternal aunts or uncles who are alive. The idea behind such generalization is to consider maternal aunts and uncles as brothers and sisters of the mother and develop on the formulas for siblings.

So far I have outlined an approach to estimate the extent of maternal orphanhood. But critical quantities are also the probabilities of paternal orphanhood and dual orphanhood. If we ignore the dependencies between the survivorship of the mother and the father, then the only additional data requirement is either the age-specific fertility schedule for men or the probability that the sexual partner of a woman of age x is of age y. An approach based on fertility schedules of men is described in Grassly et al. (2005). Here I discuss a method that relies on probabilities of partnership by age.

We would like to know $P(F = y|M = x)$, the probability that the age of the father is y, given that the age of the mother is x. We can estimate this probability from sample surveys such as the DHS: we can look at households where both partners are alive and use data on their age (or age differences) to estimate the quantity $P(F = y|M = x)$. Alternatively, we can use data on the number of married people by age group and sex: these numbers for males and females can be used to form the marginals of a matrix. The single cells of the matrix can be estimated using techniques such as iterative proportional fitting (e.g., Deming and Stephan 1940). The probability that, at time t, a child of age a is alive, whereas her/his father is not, is equal to:

$$l_a(t-a) \times [1 - (\sum_{y=0}^{100} \frac{l_{y+a}(t-y-a)}{l_y(t-y-a)} \times \sum_{x=0}^{100} P(F = y|M = x) \times P(M = x, t-a))] \tag{24}$$

The idea behind equation 24 is that once we know the probability of the age of the father, given the age of the mother, then we can evaluate the probability of survival for fathers, given their age at the birth of the child.

The probability of being an orphan of both parents can be obtained by accounting for the survival probability of the mother, in addition to that of the father. If we assume that such probabilities are independent, then we can simply multiply them. However, it is more likely that such probabilities are positively correlated, since partners tend to be either both HIV-positive or both HIV-negative. To account for such deviation from a homogeneous population, we can use frailties in a way analogous to what we did for perinatal transmission in equation 20. The use of correlated frailties would show that the probability of double orphanhood is higher in a heterogeneous population compared to a population where the mortality risks of parents are independent. Grassly et al. (2005) empirically evaluated the effect of correlated parental probabilities of infection on double orphanhood using cross-sectional data for sub-Saharan countries.

6 The microsimulation program

6.1 The microsimulation core: SOCSIM

The individual-based model that I have built to perform my analysis relies on SOCSIM, a stochastic microsimulation program whose core was designed in the 1970s at the University of California, Berkeley. The infrastructure of the first version of SOCSIM was developed by Eugene Hammel and Kenneth Wachter, at the Department of Demography, UC Berkeley (e.g., Hammel, Mason and Wachter 1990). Marcia Feitel and Carl Mason are among the computer programmers who wrote the source code for the software.

SOCSIM has been used very successfully to model the dynamics of kinship structure in historical and contemporary populations (e.g., Wachter, Hammel and Laslett 1978; Wachter 1997; Wachter, Knodel and VanLandingham 1997). The core microsimulation package is very flexible and freely available to users who would like to customize it. It has been designed to model very detailed sub-groups of a population, and to address a wide range of research questions.

Each individual in the simulation is an observation in a rectangular data file, with records of demographic characteristics for the individual, and identification numbers for key kinship members. SOCSIM is efficiently written in C and takes full advantage of arrays of linked lists to keep track of kinship relationships and to store information. The simulator takes as input population files and demographic rates. It returns updated population files as output. The supervisory file (.sup) represents the interface between the user and the source code. The user provides information to the core simulator with regards to where input files are stored, where output files are to be placed, and where demographic rates are located. In addition to that, the supervisory file contains switches for specific features, such as fertility heterogeneity and birth spacing. For each segment of the simulation, the input populations are composed of two files, one that has records for individuals (.opop) and one that has records for marriages (.omar). The demographic rates consist of fertility, mortality, marriage, and group transition rates. They can vary with the age, sex, marital status and group affiliation of the individual. If the simulation has more than one segment, which is a typical situation when demographic rates change over time, the output population for a segment can be used as input for the next one.

The individual is the unit of analysis of the simulator. Each person is subject to a set of rates, expressed as monthly probabilities of events, given certain demographic characteristics such as age, sex, marital status, etc. Every month, each individual faces the risk of a number of events including childbirth, death, marriage and migration. The selection of the event and the waiting time until the event occurs are determined stochastically, using a competing risk model. Some other constraints are included in the simulation program in order to draw events only for individuals that are eligible for such events (e.g., to allow for a minimum amount of time between birth intervals from the same mother, to avoid social taboos such as incest, etc.).

Each event for which the individual is at risk is modeled as a piecewise exponential distribution. The waiting time until each event occurs is randomly generated according to the associated demographic rates. The individual's next event is the one with the shortest waiting time. Marriage formation is a bit more sophisticated. SOCSIM is a closed simulator, in the sense that all partners must be drawn from within the existing population and cannot be externally generated. The computer program uses a two-stage process to pair eligible males and females from within the simulated population. When the next scheduled event for an individual is 'marriage', then the person is placed in a pool of eligible members to form an union. If a member of the opposite sex with appropriate demographic characteristics is available in the pool, then the two individuals are paired. Otherwise the person stays in the pool until an appropriate mate 'picks' him/her, based on a random process with probabilities dependent on demographic characteristics of the two potential spouses.

At the end of the simulation, two main files are created, the population file and the marriage file. These files contain a list of everyone who ever lived in the population and a list of every marriage that ever occurred. From these data, it is possible to determine the main demographic characteristics of the population and the entire kin network of any individual at any time.

For more details about SOCSIM, its history, computer routines and applications, see Hammel et al. (1976), Wachter et al. (1997) and the online documentation (www.demog.berkeley.edu/∼socsim).

6.2 Modelling the HIV/AIDS epidemic with SOCSIM

SOCSIM has not been specifically programmed to model the dynamics of a generalized HIV/AIDS epidemic and its demographic consequences. However, the microsimulator has been designed to be customized and modified to address a wide range of research questions. In this section, I will talk about the simulation strategy that I pursued and the modifications to the original source code that have been made to appropriately model the HIV/AIDS epidemic.

I model some of the dynamics of the HIV/AIDS epidemic by taking advantage of the flexible 'group structures' in SOCSIM. Each individual in the simulation belongs to a specific group, where the meaning of group depends on the context and the purpose of the simulation. For instance, generally speaking, groups can represent ethnicities, geographical residence, country of origin, allegiance to a soccer team or any other sort of membership. Groups are mutually exclusive: each individual can belong to only one group at a time. In the context of an HIV epidemic, I use group structures to represent HIV status. Each individual can be either HIV positive or HIV negative and is subject to mortality rates that are dependent on his/her HIV status. Adult agents in the microsimulation become HIV positive according to age-specific rates of transmission. Their life expectancy at the time they become HIV positive is modeled to be about 10 years. Newborns to HIV positive mothers can become HIV positive through perinatal transmission of the virus. This specific transmission mode of the virus is modeled through inheritance of group membership. If the mother is HIV negative at the time she gives birth, her child is born HIV negative. If the mother is HIV positive at the time she gives birth, her child is HIV positive with a probability of 0.35 at the age of one month. HIV positive children are expected to live, on average, 7 years.

For a married individual, the probability of becoming HIV positive may be associated to the HIV status of the spouse. Positive correlations in HIV status of partners increase the probability of double orphanhood, compared to a baseline scenario where the HIV status of spouses are uncorrelated. The original version of socsim is not designed to model these correlations. The source code of the microsimulator was thus modified, in order to allow for more flexibility in this regard. In the version of socsim that I used, the baseline age-specific risk of transition from HIV negative to HIV positive status can be multiplied by a user-defined factor, when the individual's spouse is HIV positive. The choice of the value for the multiplier is not obvious, since there is not a lot of empirical evidence in the literature on how the risk of becoming HIV positive varies according to the HIV status of the partner. There is also quite a bit of variability across countries and levels of adult HIV prevalence rates. Based on some results in the literature (e.g., Grassly, Phil and Timaeus 2005; Todd et al. 2006), I chose a value of 9 as risk factor for the simulation for Zimbabwe.

In summary, SOCSIM has not been originally designed to model an HIV/AIDS epidemic. However, careful use of group structures and minor modifications to the core source code make it possible to model key characteristics of the HIV/AIDS epidemic, which are relevant for the process of orphans generation. In the next section, I will talk more in details about the model parameterization and the structure of the microsimulation.

6.3 Parameterization

In the previous section, I discussed how the simulation program has been customized in order to model the dynamics of a population affected by the HIV/AIDS epidemic. Here I introduce some key aspects of the parameterization of the microsimulation that I propose.

The microsimulation for the population of Zimbabwe covers the period 1980-2050. A starting population that matches key demographic characteristics of the the actual population of Zimbabwe in 1980 has been created by letting a small unmarried initial population evolve over 100 years of time. The rates that were used for this first segment of the simulation were approximately the ones estimated for Zimbabwe in 1980, based on United Nations data sources. The initial simulated population for 1980 is composed of about 50,000 living individuals. The population size of living individuals at the end of the simulation, in 2050, is about 150,000.

The simulation is composed of 15 segments. For each segment, the computer program reads in as input a population file and demographic rates. It then produces as output a new population file. For instance, the initial simulated population for 1980, together with a set of average demographic rates for Zimbabwe for the period 1980-1984, are used as input for segment 2 of the simulation. The output for segment 2 is a simulated population file for Zimbabwe in 1985. Segment 3 of the simulation takes as input the simulated population file for 1985, together with a set of average demographic rates for Zimbabwe for the period 1985-1989, and returns a population file for 1990. The process is analogous for every segment of the simulation. With 15 segments, the last population file that is generated is the one for 2050.

For each time interval of five years, new sets of demographic rates are used as input. Such rates are estimates obtained either from the United Nations estimates and forecasts (medium scenario), or from the Demographic and Health Surveys for Zimbabwe. Details about these data sources, and how the rates were estimated, are given in the data section. The basic set of age-specific and sex-specific rates that are needed for the simulation are fertility, marriage and mortality rates. The baseline age-specific patterns of fertility are obtained from the 2006 Revision of the World Population Prospect (medium scenario). The baseline age-specific patterns of nuptiality come from the World Fertility and Marriage Database 2003, and the Demographic and Health Surveys. As for mortality rates, we need to distinguish between HIV positive and HIV negative individuals. For HIV negative individuals, I estimated age-specific patterns of mortality using a procedure based on cause-specific life tables derived from the 2006 Revision of the World Population Prospect. For HIV positive individuals, the pattern of mortality has been chosen to reflect a life expectancy at the time of infection of about 10 years for adults and 7 years for children. Age-specific HIV infection rates, or HIV incidence, are calculated using a back-calculation technique based on UN estimates and projections of numbers of AIDS-related deaths, and some assumptions on the progression rates from the first stage of HIV infection to AIDS and death.

The estimated rates, which are obtained from a variety of data sources, are used as a baseline. In order to provide more flexibility to the model, and to account for the fact that there is quite a bit of uncertainty about the level of certain demographic rates, I introduced a set of parameters that rescale age-specific profiles of fertility rates, marriage rates, and transition rates from HIV negative to HIV positive status. This means that, in the simulation model, there are three rescaling parameters per segment. Given that the whole simulation is composed of 15 segments overall, the total number of parameters is 45. The estimation procedure for the parameters of interest relies on a Bayesian approach that I am not discussing in detail. In this paper I focus on results obtained from the microsimulation. In the next sections, I will show my analysis of the evolution of kinship resources available to orphans and, more generally, children in Zimbabwe.

7 Kinship resources available to orphans

7.1 Maternal, paternal and double orphanhood

In this section, I show some estimates and projections of probabilities of maternal, paternal and double orphanhood in Zimbabwe for the period 1980-2050. Orphanhood prevalence by age and time is obtained from the analysis of the output population generated with the microsimulation program that I described in the previous sections. The results of this section complement the ones of the section on the formal demography of kinship, which mainly focused on maternal orphanhood.

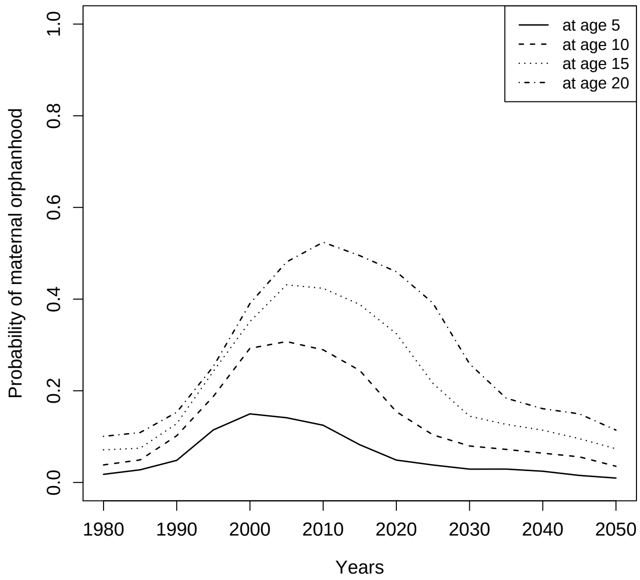

Figure 12: Estimated fractions of maternal orphans, by age, for the period 1980-2050 in Zimbabwe. Results are based on the output population file of the microsimulation.

Figure 12 shows estimates and projections of maternal orphanhood prevalence by age, and over time. The estimated pattern is consistent with the one obtained from the macrodemographic model in the section on the formal demography of kinship. The peak in orphanhood probabilities is between 2000 and 2010, depending on age, with younger ages reaching a peak earlier in time. Maternal orphanhood prevalence decreases to pre-epidemic levels within a couple of decades from the peak. It is relevant to observe that around 40% of children at age 15 are estimated to be maternal orphans in 2010, and that in 2030 such percentage is still about 30%. Compared to the results based only on demographic rates, the estimates from the microsimulation show a faster increase in orphanhood probabilities with the onset of the HIV/AIDS epidemic, and a faster rate of reduction in orphanhood probabilities after reaching a peak. These differences are related to both the assumptions about the evolution of the HIV/AIDS epidemic over time, and the effect of perinatal transmission of the HIV virus. In the simulation, the HIV/AIDS prevalence is expected to decrease fairly quickly after reaching a peak. That is consistent with UNAIDS estimates, and would potentially generate less deaths in the adult population than the ones that we would expect under the UN medium scenario. In addition, perinatal transmission of the HIV virus leads to increased child mortality, with a peak a few years after the one in adult HIV prevalence. Perinatal transmission increases the chances of death for those children who are also more likely to be or to become orphans. Therefore, it is reasonable to expect that the proportion of orphans decreases at a faster rate in the microsimulation, compared to the estimates based only on demographic rates, which do not account for perinatal transmission.

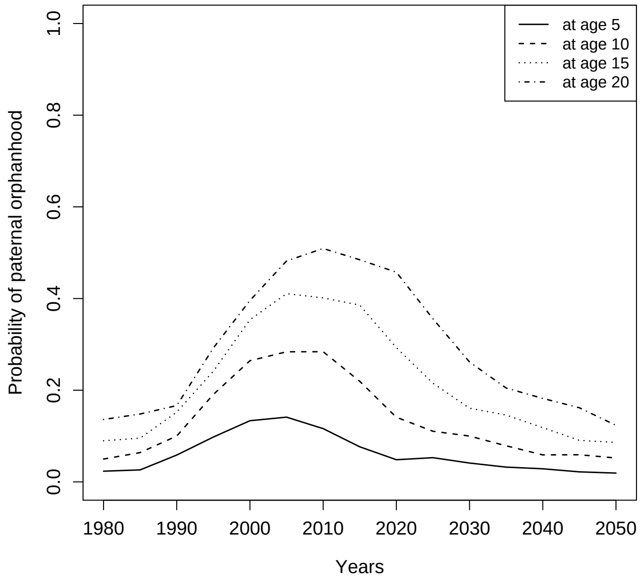

Figure 13: Estimated fractions of paternal orphans, by age, for the period 1980-2050 in Zimbabwe. Results are based on the output population file of the microsimulation.

Figure 13 shows the evolution of paternal orphanhood probabilities over time. The results based on the microsimulation complement the ones obtained in the section on the formal demography of kinship, where I pointed out the difficulties of estimating paternal and double orphanhood within an analytical framework. The pattern or paternal orphanhood is fairly similar to the one of maternal orphanhood. Most of the survey-based estimates significantly underestimate maternal orphanhood, compared to paternal orphanhood. For the specific context of Zimbabwe, it has been shown that the discrepancies between survey-based estimates and other demographic or epidemiological models is related to problems of misreporting of foster parents as natural parents, which is especially common among mothers (Robertson et al. 2008).

Figure 14 shows the evolution of probabilities of double orphanhood over time. The probability of having both parents dead is fairly low at age 5, but it rapidly increases with age. The probability of double orphanhood at age 10 reaches a peak at a level of around 15%. For teenagers of age 15, the peak is at about 20%. Figure 15 shows the estimated prevalence of double orphanhood over time for children in the age group 0-17 years old. We can observe a rapid increase in the probabilities of double orphanhood in the 1990s, with a peak around 2010 at a level of about 11%. Given the UNAIDS projections that the adult HIV prevalence rates will continue to decrease over time, the prevalence of double orphanhood will also decrease. I project the prevalence of double orphanhood to be about 6% in 2020 and 2% in 2030. Double orphanhood prevalence for 2005 is consistent with the estimates published by UNICEF (2006) which imply a prevalence in the age group 0-17 of about 11%. Projection of prevalence of

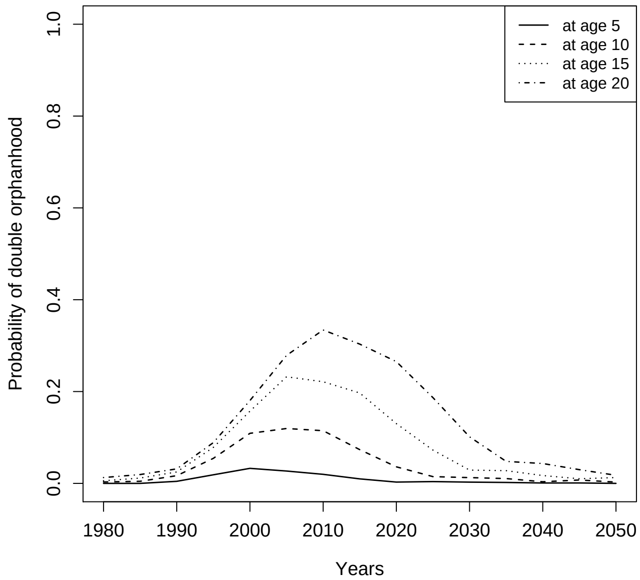

Figure 14: Estimated fractions of double orphans, by age, for the period 1980-2050 in Zimbabwe. Results are based on the output population file of the microsimulation.

double orphanhood for the future is a challenging task. As far as I know, the results of the microsimulation are the first set of projections which go beyond a 5-year horizon for Zimbabwe. These projections are novel and informative to address the lack of care.

The trend in prevalence of double orphans in Zimbabwe is of great concern. The double orphanhood condition has relevant negative consequences on the health, education and general well-being of the children who have lost their parents. The impact of double orphanhood is much more dramatic than the one of maternal or paternal orphanhood alone. A caregiver who is not a biological parent is needed for double orphans. In most cases, the caregiver who fosters the child is a member of the kinship group, either a grandparent or an uncle or aunt. The HIV epidemic has a strong impact on adult mortality and reduces the pool of kinship resources available to children, at a time when the number of orphans is very high. A quantitative evaluation of kinship resources available to orphans is missing in the literature. In the next sections, I will address this important problem.

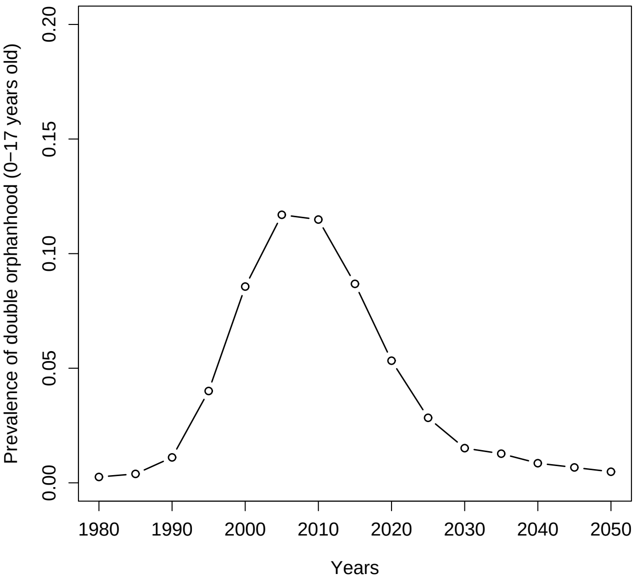

Figure 15: Estimated prevalence of double orphanhood in Zimbabwe for children in the age group 0-17 years old, for the period 1980-2050. Results are based on the output population file of the microsimulation.

7.2 Grandparents

In this section, I assess the evolution over time of grandparental resources available to orphans. The young population age structure of Zimbabwe, together with the relevant demographic impact of the HIV/AIDS epidemic, generate a strong imbalance between number of orphans and availability of grandparents as potential caregivers.

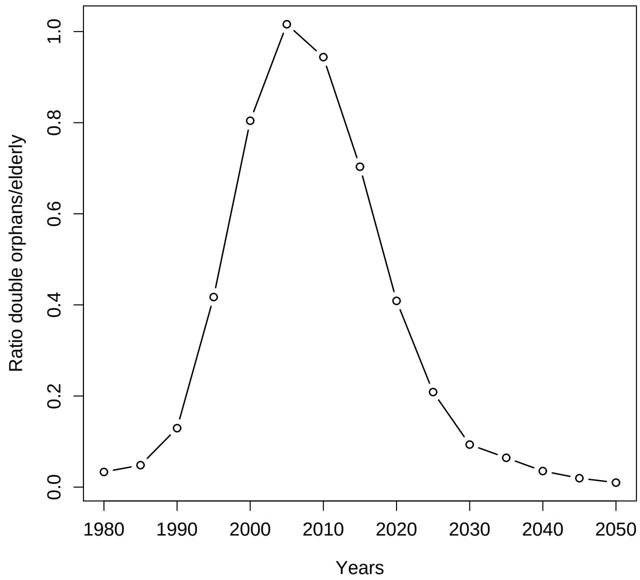

Figure 16: Estimated ratio of number of double orphans in the age group 0-17 years old, and number of elderly who are at least 60 years old, for the period 1980-2050 in Zimbabwe. Results are based on the output population file of the microsimulation.

Figure 16 shows the estimated evolution over time of the ratio of number of double orphans in the age group 0-17 years old, and number of elderly who are at least 60 years old. The ratio is very small in the early 1980s, but it then rapidly increases in the 1990s. In 2005, the number of double orphans is about equal to the number of elderly, meaning that the burden on the elderly is very high. Such ratio then slowly declines to pre-epidemic rates, as a consequence of population aging and the reduction in adult HIV prevalence rates.

The microsimulation output allows for a more detailed analysis than the one based on macrodemographic measures similar to dependency ratios. For instance, we can look at the average number of grandparents available specifically to double orphans. Figure 17 shows the evolution over time of such quantity for double orphans in the age group 0-17 years old. It is interesting to note that the lowest level of grandpaternal resources has yet to be reached. Based on the microsimulation results, I expect that the minimum value will be between 2020 and 2030. During that decade, double orphans will have, on average, only about one living grandparent to count on. Some of them will not have any grandparent at all. Figure 18 shows the proportion of double orphans whose grandparents are all dead. Between 2020 and 2030, we expect that about 35% of double orphans in the age group 0-17 years will not have any grandparent to rely on. These children will be particularly vulnerable. The problem of lack of care is very dramatic and needs to be addressed, especially in those situations where traditional kinship resources may not be enough. These results are therefore particularly informative for

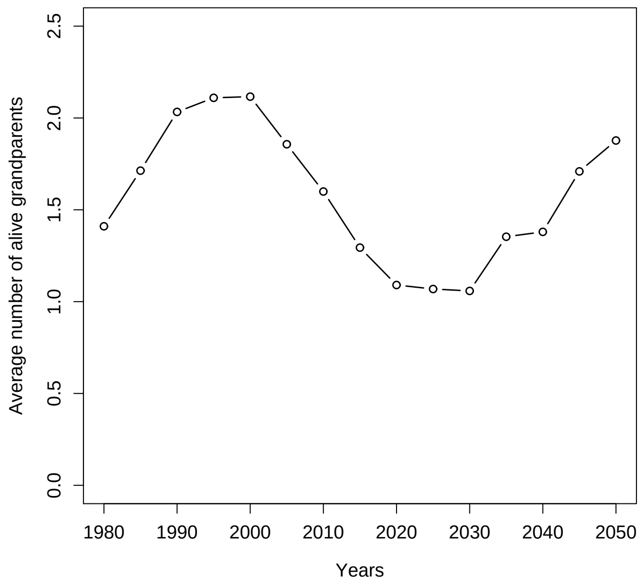

Figure 17: Average number of alive grandparents of double orphans 0-17 years old, for the period 1980-2050 in Zimbabwe. Results are based on the output population file of the microsimulation.

international organizations, such as UNICEF, whose mission is to protect children.

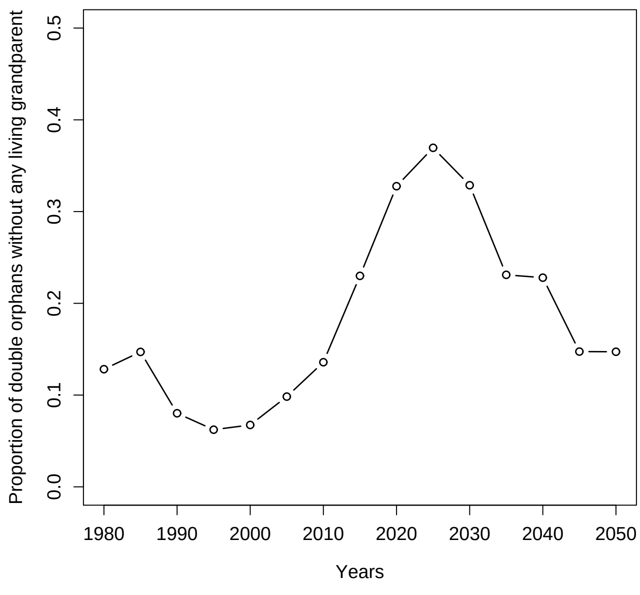

Figure 18: Estimated fraction of double orphans 0-17 years old whose grandparents are all dead, for the period 1980-2050 in Zimbabwe. Results are based on the output population file of the microsimulation.

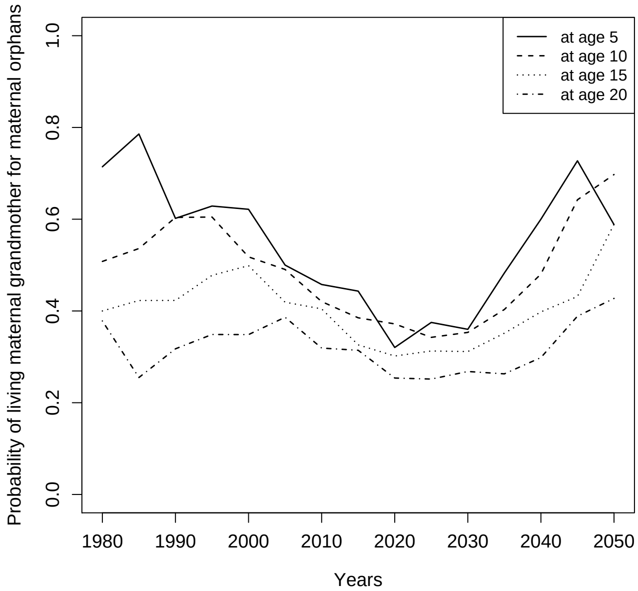

Figure 19: Estimated fraction of living maternal grandmothers for maternal orphans, by age, for the period 1980-2050 in Zimbabwe. Results are based on the output population file of the microsimulation.

A pattern analogous to the one that I have described so far emerges also from the analysis of probabilities of living grandparents by age of the orphan. Figure 19 shows the Estimated fraction of living maternal grandmothers for maternal orphans, by age, for the period 1980-2050 in Zimbabwe. I chose to show maternal orphans and maternal grandmothers because losing the mother has more adverse consequences on children's health and well-being than losing the father. Maternal grandmothers have an important role as caregiver when the mother of the child dies. We observe that the probabilities of having a living maternal grandmother are at their lowest levels in the decade 2020-2030. Younger orphans tend to have a slightly higher probability of having a living grandmother than older ones, since their grandmothers are younger, on average.

This section shows the relevant impact of the HIV/AIDS epidemic on grandpaternal resources available to orphans. The most vulnerable children are double orphans with no or very little grandpaternal resources. In some cases, uncles and aunts may step in as caregivers. In other cases, other foster families, international organizations or churches may be needed. The balance between these actors depends on how traditional forms of social relationships based on mutual support will be impacted by the epidemic. That is related to the amount of kinship resources that will be eroded as a consequence of AIDS-related deaths. In the next section, I will quantitatively evaluate the availability of uncles, aunts and siblings for orphans.

7.3 Uncles, aunts and siblings

The extended family is the traditional safety net for orphans in Zimbabwe. In this section, I will discuss the quantitative availability of uncles, aunts and siblings to orphans. In the patrilineal system of Zimbabwean communities, paternal uncles and aunts have an important role in raising children within the extended family, or in providing support to child-headed households (e.g., Foster, Makufa, Drew and Kralovec 1997).

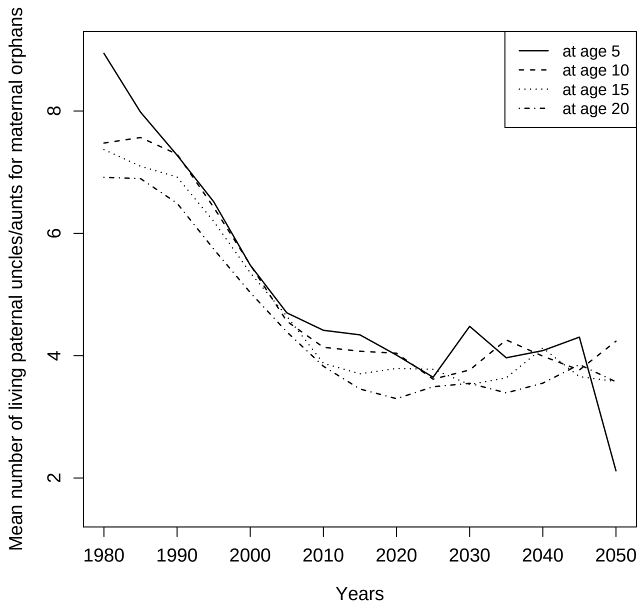

Figure 20: Estimated average number of living paternal uncles and aunts for a maternal orphan, by age, for the period 1980-2050 in Zimbabwe. Results are based on the output population file of the microsimulation.

Figure 20 shows the estimated average number of living paternal uncles and aunts for a maternal orphan, by age, for the period 1980-2050 in Zimbabwe. Maternal orphans are more vulnerable than paternal orphans. I decided to show paternal uncles/aunts, because of their important role as caregivers in the patrilineal system of Zimbabwe. Figure 20 shows a progressive decline of available aunts and uncles for orphans over time. There are some differences in the number of living aunts and uncles by age of the orphan: younger children tend to have younger uncles and aunts, who are subject to a different mortality risk than older individuals. The trend over time is similar, though. It is consistent with the observation that traditional forms of social relationships are under stress, and new members of the community are providing care to children. For instance, the mechanism of mutual help between members of the same patriline is less prevalent in urban areas, where it is becoming more common for maternal aunts and uncles to take care of orphans.

Figure 20 shows one side of the story: over time, children have a smaller number of uncles and aunts to rely on. The second important aspect of the problem is the burden on each of the uncles and aunts, which is related to the number of orphans that need to be taken care of. I will focus on double orphans, who are the children who need a caregiver the most. Figure 21 shows the ratio of number of living uncles/aunts and number of double orphans in the age group 0-17 years old, over time. The ratio has been computed by extracting the number of unique people

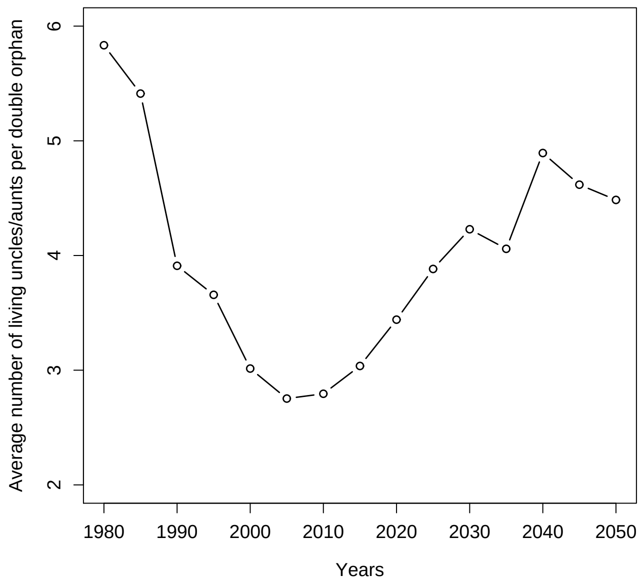

Figure 21: Estimated ratio of number of living uncles/aunts and number of double orphans (age group 0-17 years old) for the period 1980-2050, in Zimbabwe. Results are based on the output population file of the microsimulation.

who are uncles/aunts to double orphans in simulated population. Figure 21 shows that the burden on uncles and aunts is currently at its highest levels. In the next few decades, although orphans will continue on counting on approximately the same number of uncles/aunts, the overall number of orphans (and, more generally, children) will decrease and thus the per-capita burden on uncles and aunts will be alleviated.

Orphans are often a heavy burden on relatives, who may refuse to take care of them. Such refusal is a sign of the decline of traditional extended family practices. Foster et al. (1997) show that leading factors to the establishment of child-headed households are the death of the parents and the availability or relatives who provide support to the children, but do not accept them into their households. In other cases, relatives are nonexistent, or they are distant, sick or do not have the material means to provide for additional children. Households headed by adolescents are an additional coping mechanism in response to the impact of HIV/AIDS on communities. It is thus relevant to consider the availability of elder siblings to support their younger brothers and sisters. Figure 23 shows the trend over time of the fraction of double orphans younger than 10 years old who have at least one elder sibling who is older than 15 years old. There is some stochasticity after 2040, mostly related to the smaller number of expected double orphans at that time. However, ther trend is fairly clear. In the 1980s, before the HIV epidemic took off, half of the youngest double orphans (less than 10 years old) were expected to have at least one adolescent sibling older than 15 years old. By 2020, only about a third of the youngest orphans are expected to have a living adolescent sibling. Then the ratio is expected to improve over the course of the next decades. This result shows that at the time when the availability of grandparents, uncles and aunts is at its lowest levels, the resources provided by elder siblings are also thin and the additional coping mechanism provided by child-headed households may be seriously undermined by the age structure of siblings.

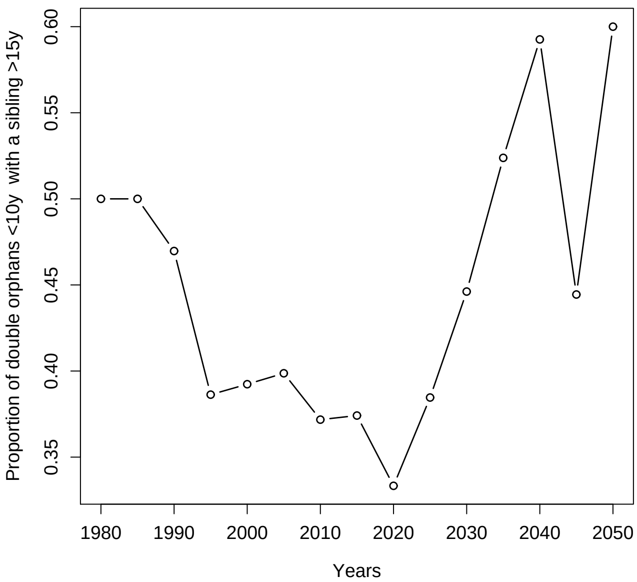

Figure 22: Estimated fraction of double orphans younger than 10 years old who have at least one elder sibling who is older than 15 years old, for the period 1980-2050 in Zimbabwe. Results are based on the output population file of the microsimulation.

7.4 An index of kinship resources

In the previous sections, I analyzed the effect of the HIV/AIDS epidemic on the generation of orphans and the availability of grandparents, uncles, aunts and siblings. In this section, I provide an index of overall kinship resources for young children.

I quantify the amount of kinship resources by weighting the availabilty of members of the kinship group by their relatedness to the child considered. The weights are obtained using the Hamilton's coefficient of relatedness, which is defined as the percentage of genes that two individuals share by common descent. A child inherits $\frac{1}{2}$ of his/her genome from a parent. The coefficient of relatedness for a child and one of his/her parents is thus $\frac{1}{2}$. The coefficient for a child and one of his/her grandparents is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. The coefficient of relatedness for a child and one of his/her uncles/aunts is $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$, and so on.

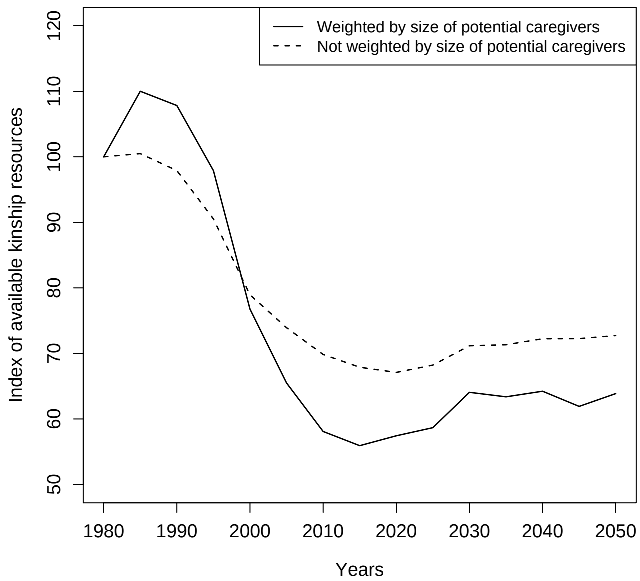

Figure 23: Two indexes of availability of kinship resources for children younger than 10 years old, for the period 1980-2050 in Zimbabwe. See main text for details on the indexes. Results are based on the output population file of the microsimulation.

Figure 23 shows two indexes of kinship resources for children younger than 10 years old in Zimbabwe, over time. The indexes are normalized so that the initial value in 1980 is equal to 100. The dashed line is the normalized average amount of kinship resources for a child less than 10 years old. For the calculation of the index, the kinship members considered are elder siblings, parents, grandparents, uncles and aunts. The members of the kinship group are weighted by their coefficient of relatedness with the child. The solid line is constructed in an analogous way, with the exception that the amount of kinship resources are also weighted by the number of unique individuals in the population that are potential caregivers. In other words, the index represented with a dashed line is multiplied by the number of unique members of the population who take up the role of siblings, parents, grandparents, uncles and aunts, to get the solid line index. The bigger the number of unique individuals, the less the average burden on each of them to take care of the children in the society.

The indexes in figure 23 show a transition from relatively high levels of kinship resources for children, to relatively low levels of kinship resources. The index corrected for the number of unique members of the potential caregivers group (solid line) shows a more rapid decline in kinship resources in the 1990s. Both indexes show a very slow increase in kinship resources for children after 2020.

These results open important questions on the future of traditional forms of social support based on reciprocal obligations. The demographic change associated to the HIV/AIDS epidemic, and the demographic transition, may reduce the amount of kinship resources available to children to a point that undermines the traditional role of the extended family as a safety net, and the existence of social mechanisms such as child fosterage.

8 Conclusion and Discussion

The focus of the paper is on the estimation of the evolution of kinship structure in Zimbabwe, one of the countries hit the hardest by the HIV/AIDS epidemic. I used both formal demography and microsimulation to quantify the extent of the orphanhood crisis and the kinship resources available to orphans. First, I provided some qualitative background information on the effect of the HIV/AIDS epidemic on children in sub-Saharan Africa. I discussed the coping mechanisms that are in place in the African context and the role played by the extended family in caregiving practices for orphans. I then discussed the relevant literature that deals with the quantitative evaluation of the extent of the orphanhood crisis in sub-Saharan Africa. I provided a description of the main ideas behind the methodologies that have been adopted by international agencies to estimate the number of orphans and I showed estimates of relevant quantities for countries in southern Africa.

I identified a gap in the literature. A comprehensive quantitative evaluation of the effect of the HIV/AIDS epidemic on kinship structure in the sub-Saharan context is missing. Such a quantitative assessment is necessary to evaluate the available kinship resources for orphans and the elderly. A quantitative evaluation of the effect of the HIV/AIDS epidemic on kinship resources for the elderly has been developed in the context of south-east Asia (Wachter et al. 2002, 2003), but there has not been a similar attempt for the geographical region of sub-Saharan Africa. A quantitative analysis of kinship resources for elderly and orphans contributes to the debate on the extent to which the safety net provided by the extended family is under stress in the context of the HIV/AIDS crisis.

Methodologically, I addressed the issue by using both formal demography and microsimulation. One way to analyze kinship resources is to use fairly general models that provide some intuition about the processes that regulate the formation of orphans and the availability of kin. Alternatively, more detailed parameterizations, based on microsimulation and tailored to a specific social and geographical context, can be pursued. The results of the analysis show that the orphanhood crisis may persist for decades after the peak in HIV prevalence. Together with increased number of orphans, predictions based on the microsimulation show a decrease in kinship resources for orphans. In particular, the number of available traditional caregivers, such grandparents and paternal uncles and aunts, are expected to sharply decline. This is likely to translate into an increased burden for living relatives and reduced resources for orphans.

The effort to extract relevant information from the microsimulation is in progress. In particular, I am working on the evaluation of uncertainty for estimates obtained from both the microsimulation and the formal analysis. I consider the application of the Bayesian melding (e.g. Poole and Raftery 2000; Sevcikova et al. 2007) to demographic microsimulation a promising area of research.

References

[1] Ankrah E.M. 1993. The Impact of HIV/AIDS on the Family and other Significant Relationships: The African Clan Revisited. *AIDS Care* 5:5-22.

[2] Bicego G., S. Rutstein and K. Johnson. 2003. Dimensions of the Emerging Orphan Crisis in Sub-Saharan Africa. *Social Science and Medicine* 56(6):1235-1247.

[3] Bor R. and J. Elford, eds. 1998. *The Family and HIV Today: Recent Researcha and Practice.* Continuum International Publishing Group (Edition 2).

[4] Brass W. 1974. Perspectives in Population Projection. *Journal of the Royal Statistical Society Series A*. 137:532-583.

[5] Brodzinsky D.M., A.V. Gormly and S.R. Ambron. 1986. *Lifespan Human Development* (3rd ed.). New York:CBS College Publishing.

[6] Burch T.K. 1970. Some Demographic Determinants of Average Household Size: An Analytic Approach. *Demography* 7:61-70.

[7] Clark S.J. 2001. An investigation into the impact of HIV on population dynamics in Africa. Unpublished PhD dissertation.

[8] Clark S.J. 2006. Demographic impacts of the HIV epidemic and consequences of population-wide treatment of HIV for the elderly: results from microsimulation. National Research Council.

[9] Coale A.J. 1965. Appendix: Estimates of Average Size of Household. In A.J. Coale et al., Aspects of the Analysis of Family Structure, pp. 64-69. Princeton University, Office of Population Research, Princeton, New Jersey, USA.

[10] Deming W.E. and F.F. Stephan. 1940. On Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known. *The Annals of Mathematical Statistics* 11(4):427-444.

[11] Drew R.S., G. Foster and J. Chitima. 1996. Cultural Practices Associated with Death in the North Nyanga District of Zimbabwe. *Journal of Social Development in Africa* 11:79-86.

[12] Drew R.S., C. Makufa and G. Foster. 1998. Strategies for Providing Care and Support to Children Orphaned by AIDS. *AIDS Care* 10(Suppl. 1):S9-S15.

[13] Foster G. 1998. Today's Children: Challenges to Child Health Promotion in Countries with Severe AIDS Epidemics. *AIDS Care* 10(1):S17-23.

[14] Foster G. 2000. The Capacity of the Extended Family Safety Net for Orphans in Africa. *Psychology, Health and Medicine* 5:55-62.

[15] Foster G., C. Makufa, R. Drew, S. Kambeu and K. Saurombe. 1996. Supporting Children in Need Through a Community-based Orphan Visiting Programme. *AIDS Care* 8(4)389-403.

[16] Foster G., C. Makufa, R. Drew, S. Mashumba and S. Kambeu. 1997. Perceptions of Children and Community Members Concerning the Circumstances of Orphans in Rural Zimbabwe. *AIDS Care* 9(4):391-405.

[17] Foster G., C. Makufa, R. Drew and E. Kralovec. 1997. Factors Leading to the Establishment of Child-headed Households: The Case of Zimbabwe. *Health Transition Review* 7(2):155-168.

[18] Foster G. and J. Williamson. 2000. A Review of Current Literature of the Impact of HIV/AIDS on Children in Sub-Saharan Africa. *AIDS* 14(3):275-284.

[19] Friedman J.H. 1984. *A Variable Span Scatterplot Smoother.* Laboratory for Computational Statistics, Stanford University Technical Report No.5.

[20] Garnett G.P. and R.M. Anderson. 1993. Factors Controlling the Spread of HIV in Heterosexual Communities in Developing Countries: Patterns of Mixing between Different Age and Sexual Activity Classes. *Phil Trans. R. Soc. Lond. B.* 342:137-159.

[21] Girosi F. and G. King. 2008. *Demographic Forecasting.* Princeton University Press.

[22] Goodman L.A., N. Keyfitz, and T.W. Pullum. 1974. Family Formation and the Frequency of Various Kinship Relationships. *Theoretical Population Biology* 5:1-27.

[23] Goody E. 1982. *Parenthood and Social Reproduction: Fostering and Occupational Roles in West Africa.* New York: Cambridge University Press.

[24] Grassly N.C., D. Phil and I.M. Timaeus. 2005 Methods to Estimate the Number of Orphans as a Result of AIDS and Other Causes in Sub-Saharan Africa. *Journal of Acquired Immune Deficiency Syndromes* 39(3):365-375.

[25] Gregson S., G.P. Garnett and R.M. Anderson. 1994. Assessing the Potential Impact of HIV-1 Epidemic on Orphanhood and the Demographic Structure of Populations in sub-Saharan Africa. *Population Studies* 48: 435-458.

[26] Hammel E.A., D. Hutchinson, K.W. Wachter, R. Lundy and R. Deuel. 1976. *The SOCSIM Demographic-Sociological Microsimulation Program.* Institute of International Studies, Berkeley, California.

[27] Hammel, E.A., C. Mason and K.W. Wachter. 1990. SOCSIM II, a Sociodemographic Microsimulation Program, rev. 1.0, Operating Manual: Graduate Group in Demography Working Paper No. 29. Berkeley, California, University of California, Institute of International Studies, Program in Population Research.

[28] Heuveline P. 2003. HIV and Population Dynamics: A General Model and Maximum-Likelihood Standards for East Africa. *Demography.* 40(2):217-245.

[29] Heuveline P. 2004. Impact of the HIV Epidemic on Population and Household Structure: The Dynamics and Evidence to Date. *AIDS* 18(2):45-53.

[30] Isiugo-Abanihe U.C. 1985. Child Fosterage in West Africa. *Population and Development Review* 11(1):53-73.

[31] Jones J.H. 2005. The Probability of Maternal Orphanhood Under a Generalized AIDS Epidemic. *Paper presented at the 2005 Population Association of America Meetings, Philadelphia, PA.*

[32] Keyfitz N. and H. Caswell. 2005. *Applied Mathematical Demography.* Springer, New York.

[33] Le Bras H. 1973. Parents, Grandparents, diaeresis bisaieux. *Population* 28:9-37.

[34] Lotka A.J. 1931. Orphanhood in Relation to Demographic Factors. Metron (Rome) 9:37-109.

[35] Madhavan S. 2004. Fosterage Patterns in the Age of AIDS: Continuity and Change. *Social Science and Medicine* 58:1443-1454.

[36] Merli M.G. and A. Palloni. 2006. The HIV/AIDS epidemic, kin relations, living arrangements, and the African elderly in South Africa. National Research Council. *Aging in sub-Saharan Africa: recommendations for furthering research.* Panel on policy research and data needs to meet the challenge of aging in Africa. Barney Cohen and Jane Menken Eds. Committee on population, Division of behavioral and social sciences and education. Washington, DC: The National Academies press.

[37] Monasch R. and J.T. Boerma. 2004. Orphanhood and Childcare Patterns in Sub-Saharan Africa: an Analysis of National Surveys from 40 Countries. *AIDS* 18:S55-S65.

[38] National AIDS Council. Ministry of Health and Child Welfare. The MEASURE Project. Center for Disease Control and Prevention (CDC/Zimbabwe). 2002. AIDS in Africa during the Nineties. Zimbabwe. A review and Analysis of survey and reasearch results.

[39] Ntozi J.P.M. and S. Nakayiwa. 1999. *AIDS in Uganda: How Has the Household Coped with the Epidemic?* In: The Continuing HIV/AIDS Epidemic in Africa: Responses and Coping Strategies, edited by I. O. Orubuloye, John C. Caldwell, and James P. M. Ntozi. Canberra, Australia, Australian National University, National Centre for Epidemiology and Population Health, Health Transition Centre, 1999:155-181.

[40] Palloni A. and L. Lamas. 1989. *The Palloni Approach: a Duration-dependent Model of the Spread of HIV/AIDS in Africa.* In: The AIDS Epidemic and its Demographic Consequences. Proceedings of the United Nations/ World Health Organization Workshop on Modelling the Demographic Impact of the AIDS Epidemic in Pattern II Countries.

[41] Palloni A. and M. Glicklich 1989. *Review of the Approaches to Modelling the Demographic Impact of the AIDS Epidemic.* In: The AIDS Epidemic and its Demographic Consequences. Proceedings of the United Nations/ World Health Organization Workshop on Modelling the Demographic Impact of the AIDS Epidemic in Pattern II Countries.

[42] Palloni A. and Y.J. Lee. 1992. Some Aspects of the Social Context of HIV and its Effects on Women, Children and Families. *Population Bulletin of the United Nations* 33:64-87.

[43] Poole D. and A.E. Raftery. 2000. Inference for Deterministic Simulation Models: the Bayesian Melding Approach. *Journal of the American Statistical Association* 95(452):1244-1255.

[44] Richter L.M. and J. Swart-Kruger. 1995. AIDS-risk among Street Children and Youth: Implications for Intervention. *South African Journal of Psychology* 25(1):31-38.

[45] Robson E. 2000. Invisible Carers: Young People in Zimbabwe's home-based health care. *Area* 32(1):59-69.

[46] Robertson L., S. Gregson, C. Madanhire, N. Walker, P. Mushati, G. Garnett and C. Nyamukapa. 2008. Discrepancies between UN Models and DHS Survey Estimates of Maternal Orphan Prevalence: Insights from Analyses of Survey Data from Zimbabwe. *Sexually Transmitted Infections* 84(Suppl I):i57-i62.

[47] Sengendo J. and J. Nambi. 1997. The Psychological Effect of Orphanhood: A Study of Orphans in Rakai District. *Health Transition Review (Supplement to)* 7:105-124.

[48] Sevcikova H., A.E. Raftery and P.A. Waddell. 2007. Assessing Uncertainty in Urban Simulations Using Bayesian Melding. *Transportation Research.* Part B 41:652-669.

[49] Todd J., H. Grosskurth, J. Changalucha, A. Obasi, F. Mosha et al. 2006. Risk Factors Influencing HIV Infection Incidence in a Rural African Population: A Nested Case-Control Study. *The Journal of Infection Diseases* 193:458-466.

[50] UNAIDS. 2002. Improved Methods and Assumptions for Estimation of the HIV/AIDS Epidemic and its Impact: Recommendations of the UNAIDS Reference Group on Estimates, Modelling and Projections. *AIDS* 16:W1-W14.

[51] UNICEF, USAID and UNAIDS. 2002. *Children on the Brink. A Joint Report on Orphan Estimates and Program Strategies.* New York, July 2002.

[52] UNICEF. 2006. Africa's Orphaned and Vulnerable Generations. Children Affected by AIDS.

[53] Vaupel J.W., K.G. Manton and E. Stallard. 1979. The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality. *Demography* 16:439-454.

[54] Wachter, K.W., E.A. Hammel and P. Laslett. 1978. *Statistical Studies of Historical Social Structure.* New York, Academic Press.

[55] Wachter K.W. 1997. Kinship resources for the elderly. *Philosophical transactions of the royal society of London - Series B: Biological sciences* 352(29).

[56] Wachter K.W., D. Blackwell and E.A. Hammel. 1997. Testing the Validity of Kinship Microsimulation. *Journal of Mathematical and Computer Modeling* 26:89-104.

[57] Wachter K.W., J.E. Knodel and M. VanLandingham. 2002. AIDS and the elderly of Thailand. *Demography* 39(1):25-41.

[58] Wachter K.W., J.E. Knodel and M. VanLandingham. 2003. Parental Bereavement: Heterogeneous Impacts of AIDS in Thailand. *Journal of Econometrics* 112:193-206.

[59] Wachter K.W. 2007. Essential Demographic Methods. *Original Manuscript.* Berkeley, California.

[60] World Bank. 1997. *Confronting AIDS: Public Priorities in a Global Epidemic.* New York: Oxford University Press.

[61] Zaba B., J. Whitworth, M. Martson, J. Nakiyingi, A. Ruberantwari et al. 2005. HIV and Mortality of Mothers and Children: Evidence from Cohort Studies in Uganda, Tanzania, and Malawi. *Epidemiology* 16(3):275-280.