Family Networks and Orphan Caretaking in Tanzania*

Christopher Ksoll†

Yale University

January 21, 2007

Abstract

Most of the nearly 8 million orphaned children in Sub-Saharan Africa are taken care of by members of their extended family network. This paper studies how the documented negative effects of orphanhood vary with the characteristics of the family members taking care of them. To answer this question, I designed a survey and collected data on the extended family networks of orphans in Tanzania and combine these data with information on orphan outcomes. Using information on the potential and actual placement of orphans, I first examine the determinants of the placement of the orphan within the family network. The empirical results show that the occupations and relative wealth of potential caretakers influence where the orphan is placed. At the same time, cultural norms are also important, as paternal relatives are more likely to provide care than relatives of the mother.

*This version is preliminary and incomplete. Please do not cite. I am indebted to Joseph Altonji, Mark Rosenzweig, Paul Schultz and Christopher Udry for advice and support. I would like to thank Richard Akresh, Lori Beaman, Michael Boozer, Amit Khandelwal, Siddharth Sharma, Tavneet Suri, Harsha Thirumurthy, Ulrich Wagner for comments and help. I would also like to thank Edi-Africa, in particular Joachim DeWeerdt and Respicius D. Mitti, for their hospitality while preparing the field work. This research was funded by the Social Science Research Council through an International Dissertation Research Fellowship and a Risk and Development Grant, the National Science Foundation (Grant No. SES 0418096), the Yale Center for International and Area Studies and the Yale Economic Growth Center. Finally, I thank the field research team, in particular the supervisors, the late George Baisi, Henry Lugakingira and Elizabeth Massawe. All errors are my own.

†Economic Growth Center, Department of Economics, 27 Hillhouse Avenue, New Haven, CT 06520, *email:* christopher.ksoll@yale.edu, *website*: www.econ.yale.edu/∼cjk29

The systematic placement of the orphan within the family network leads to a placement bias when estimating how orphan welfare is affected by caretaker characteristics, as is shown analytically and empirically. To correct for the placement bias, the predicted probabilities from the placement estimation are used to set up a selection correction term. In the empirical analysis, I find that household wealth is an important determinant of orphans' education when correcting for the placement of the orphan. Neglecting the endogenous placement, however, leads to the erroneous conclusion that household wealth is insignificant.

Keywords: Orphans; Extended Family; Caregiving; Tanzania
JEL Classification: O15, D10, I3, J12

1 Introduction

Orphaned children in Africa have come into the spotlight as the HIV/AIDS pandemic takes its toll, adding to the already high death rate among parents from other causes. According to UNAIDS nearly 8 million children under the age of 15 have lost both parents (UNAIDS, 2004). There is some evidence that these children are worse off in terms of access to education and health outcomes because of the loss of their parents (Beegle et al., 2006; Case and Ardington, 2006; Evans and Miguel, 2005).

This paper examines how the effects of orphanhood vary with the characteristics of the caretakers in a context where most of the orphans are taken care of by close family members. Do richer caretakers provide better care than poorer caretakers? Or is the relationship between the orphan and the caretaker more important, such that grandparents might provide better care even if they are somewhat poorer? The answer to these questions is complicated by the fact that the characteristics of the caretakers are endogenous, because orphans are not randomly placed in the family network. Understanding the underlying mechanism of how the orphan is placed within the family network is thus crucial for understanding how the caretaker characteristics affect orphan welfare.

I first develop a model of orphan placement which shows that the systematic placement of the orphan leads to a bias when regressing orphan outcomes on caretaker characteristics. The intuition behind this is that observable and unobservable characteristics will influence both the caretaking decision as well as orphan outcomes: When a family member with very low levels of observable characteristics, such as wealth, takes care of an orphan, it is probably because she has some compensating unobserved characteristics, such as a high preference for the orphan. This leads to a negative correlation between the observed and the unobserved preferences. Thus, in a regression of orphan outcomes on the observed characteristics of the caretakers the estimated coefficients tend to be biased downward.[1] The theoretical model of orphan placement also suggests a method for correcting for this placement bias. Using information on actual as well as potential caretakers, a correction term can be constructed via the two-step procedure developed by Lee (1983). This involves estimating the determinants of the placement of the orphan in the network and predicting the probability that a network member is chosen as caretaker.

In order to estimate the model of orphan placement I designed a survey and collected primary data in Tanzania that contains information on the extended family networks of double orphans. I sampled both network members who took care of orphans and those who did not, allowing me to estimate the determinants of the placement of the orphan. In order to provide evidence on how orphan welfare is affected by caretaker characteristics, the data on the family networks of the orphans is linked back to a panel dataset. The sampling frame for my orphan survey is the Kagera Health and Development Survey (KHDS) which provides data on long-term health and education outcomes of respondents between 1991 and 2004. The orphan sample contains the children who lost both parents during this period. Thus, using the linked data I can analyze the long-term effects of orphanhood, while controlling for placement selection.

The empirical analysis proceeds in two steps. First the determinants of the placement decision are analyzed, yielding information on how households systematically place orphans. This step can also be thought of as a first stage for the correction of the placement bias. The second step is then to estimate how orphan welfare is affected by caretaker characteristics while controlling for placement bias.

I estimate the determinants of the placement of the orphan using a conditional logit model which accounts for stratified sampling from the family network (McFadden, 1978; Bierlaire,

[1] The estimates will be downward biased when observed and unobserved characteristics are unconditionally uncorrelated and both positively influence the outcomes of orphans and the placement of the orphan in the family network.

Bolduc and McFadden, 2006). A mixed logit specification allows for flexible substitution patterns between groups of network members. I find that economic variables are important. In particular, a one standard deviation increase in the value of land owned by the network member's household increases the probability of caretaking by 2.8 percentage points whereas owning business assets increases the probability of caretaking by 7.5 percentage points. These economic concerns are balanced with cultural norms: paternal relatives are much more likely to be caretakers than maternal relatives. Yet, the traditional caretakers, paternal uncles and aunts are much less likely to provide care than the grandparents or older siblings of the orphans. The mixed logit results show that there is evidence for heterogeneity in the effect of maternal grandmothers on the placement of the orphan. In examining how orphans are systematically placed within family networks this analysis contributes to understanding how households in Sub-Saharan Africa adapt in response to shocks.

The empirical analysis of how orphans' education is affected by caretaker characteristics reveals that the placement bias is substantial. Not accounting for the placement of the orphan biases the estimated coefficients downwards for variables that positively influence both the placement of the orphan as well as the education of the orphan. In particular, variables that capture household wealth seem much less important and, in part, insignificant when neglecting the endogenous placement of the orphan. Variables that capture the relationship between the orphan and the caretaker are also affected: Not accounting for the bias would lead us to conclude that paternal grandmothers (compared to paternal uncles and aunts) are detrimental for orphans' education while the corrected coefficients show that paternal grandmothers have a slightly positive but insignificant effect. A somewhat similar pattern occurs for paternal grandfathers. The coefficients that mistakenly seem small but positive become much larger and statistically significant when including the placement correction term.

These results provide evidence for the existence of a substantial and significant placement bias. Since the estimated coefficients from the placement regression and the outcome regression typically are of the same sign, this shows that family networks place orphans where they are better off in terms of education relative to placing them with a network member with average characteristics. Family networks in Tanzania are institutions in which the risks associated with parental mortality are at least partially mitigated.

The remainder of the paper is organized as follows. The following section reviews the literature and provides some background on orphans in Tanzania. Section 3 develops a model of caretaking. Section 4 outlines its econometric implementation. Section 5 describes the data and presents some descriptive statistics. Section 6 present results. Section 7 outlines how to estimate the welfare effects of orphanhood using the KHDS data. Section 8 offers a summary and an outlook.

2 Background and Literature Review

With a PPP adjusted per capita income of around $670, Tanzania is one of the poorest countries in the world. The life expectancy is around 46 years, the under 5 mortality rate is 126 per 1000 and the adult illiteracy rate is over 30% (World Bank, 2004b). The number of orphans is estimated to be 1.3 million of which 960,000 have lost their father, 525,000 have lost their mother and (counted in both categories) 165,000 double orphans who have lost both parents (Measure et al., 2001). While not all cases of orphanhood are due to HIV/AIDS, HIV/AIDS is seen as a major contributory factor since it affects mainly the reproductively active population. The Northwestern region of Kagera, where this research is situated, has the lowest regional per capita GDP (TZNBS 2003), which at $166 is about half of the unadjusted national average of 320 (World Bank, 2004b). Kagera was the first region within Tanzania to report an HIV/AIDS death in 1983. Kagera borders on Uganda, Rwanda and Burundi and has traditionally been an area with frequent population movement. While Kagera was originally seen as a region of high HIV infection rates, the current estimates for adults aged 15-49 using a representative sample are 3.9% for men and 3.5% for women (TACAIDS et al., 2005).[2] Kagera's HIV infection rates are thus quite low compared to the national average of 6.3 and 7.7 respectively. There is some evidence from antenatal clinics that, during the 1990s, Kagera saw a decrease in HIV infection rates, but it also seems plausible that the original infection estimates (ranging from 20-25%) were too high since they were based on non-representative samples, namely pregnant mothers at antenatal clinics. Orphan rates in Kagera are nonetheless comparable to the national average; 10.5% versus 10.8% respectively (TACAIDS et al., 2005). The rate of double orphans, however,

[2] Among both sexes the percentage of respondents who refused testing refusals was equally low at 2.3%. In addition, 1.0% of women and 4.1% of men were sampled, but not interviewed.

is 1.7% in Kagera which is substantially higher than 1% at the national level.

Studies that have analyzed the impact of orphanhood are of 3 types: cross-sectional studies, caretaking household fixed effects studies, and panel data studies. These studies contribute to answering different questions about orphan welfare, typically measured by enrollment as a measure of investment in education or body mass index as an indicator of health status. Cross-sectional studies typically regress these outcomes on orphan status and sometimes household controls (Ainsworth and Filmer, 2002; Case et al., 2004; Bicego et al., 2003). These studies provide information on how well orphans fare compared to the general child population and whether targeting orphans is an efficient strategy to reduce general poverty. The empirical findings from these studies have been mixed for Tanzania.

Household fixed effects studies use cross-sectional data and compare outcomes of orphans to those of non-orphans who reside in the same households (Case et al., 2004; Seck, 2006) . They find evidence that orphans receive less investment in education than the biological children of the household head. These findings can provide some justification for conditional transfers to households taking care of orphans. In general studies based on cross-sectional data answer some policy relevant questions, in particular on regional priorities, as they are usually based on large nationally representative samples. However, they do not identify the effects of orphanhood unless parents who die are a random sample of the population and the orphans are placed randomly within the family network. Both assumptions are unlikely to hold.

Much of the recent literature uses panel data in order to account for non-random parental death. These studies compare orphans' outcomes before and after the death of their parents and can thus correct for differences in parental background and often for individual fixed effects (Ainsworth et al., 2005; Ainsworth and Semali, 2000; Beegle et al., 2005; Case and Ardington, 2004; Evans and Miguel, 2005). They conclude that there is evidence of negative effects of orphanhood on the health and education of orphans. The effects typically vary with which parent died and whether the other parent (if still alive) was still taking care of the orphan. In general, maternal death seems to have more severe impacts than paternal death.

Ainsworth et al. (2005) study the effects of orphanhood on schooling participation and hours spent in school using household panel data from Kagera (the baseline of the data used in Beegle et al. (2005) and in this study). They find that maternal deaths delay schooling for younger children. Children already in school do not drop out of school, however, the hours spent in school are significantly reduced. Ainsworth and Semali (2000) use the same data to analyze the effects of orphanhood on health. Individual fixed effects estimates do not show any impacts of maternal or paternal orphanhood on childrens' height-for-age z-scores in their panel, possibly due to the small number of children, the short duration of the panel (two years) and the fact that households receive financial assistance following the adult death (Lundberg et al., 2000).

Case and Ardington (2004) find that maternal death negatively affects school participation using a 2-3 year panel survey from South Africa, whereas paternal deaths do not seem to have an effect. Since the mother's death does not seem to be associated with differences in socio-economic conditions in the household before her death, Case et al. interpret their results as causal. Evans and Miguel (2005) analyze the impact of orphanhood on school participation using a sample of Kenyan school children in 1998 who were re-interviewed in 2002. Using child-specific fixed effects, they find evidence of effects of maternal death on school participation. However, this is true also for the period of 1-2 years immediately before the mother's death. Paternal death does not seem to have any effect. Beegle et al. (2005) present evidence from Tanzania that orphanhood has negative long-run consequences on health and education for some orphans. Among other things, maternal orphans are 2 cm shorter and receive about 1 year of schooling less than they would have had they not been orphaned. By including orphan characteristics from the first wave of the study, they can control for parental background and individual endowments. Beegle et al. also note that the living arrangements of the orphan before the death of the parent mattered. An interesting feature of the survey data that Beegle et al. use is that the survey attempted to interview all individual respondents of a previous household survey even when they had left the household or the village. Beegle et al. (2005) find that the results change when they restrict the sample to those who remained in the village compared to the whole sample which underlines the importance of including orphans who moved out of the households and out of their villages. Households change their structure in response to shocks (Akresh, 2006; Taiwo, 2006) and not accounting for this endogenous choice can bias the estimated coefficients (via sample attrition in the case of Beegle et al, and endogenous placement in the current study). Where the orphan is placed in the family network may affect the welfare of the child. Whether orphans move out of their households and to where they move depends on the options available to them and is also influenced by cultural norms. In the following section I describe traditional caretaking norms in Kagera.

2.1 Caregiving Arrangements in Kagera

In Kagera, as in many parts of East Africa that are patrilineal, children belong to the paternal lineage. Traditionally, the caregiving duties thus lay with the paternal relatives, in particular paternal uncles and aunts (Foster, 2000). When husbands died, their wives' ability to remain on the land of their husband traditionally depended on whether they had children by the deceased and in general on their relations with their in-laws (Manji, 2000).[3]

When adult males die, members of the extended family and older men of the clan often come together and settle the inheritance and discuss the caregiving of the orphan children.[4] In many cases, dying parents have written a will concerning their inheritance and expressed a wish as to who should take care of the orphan. However, not all parents do and when they do, their will is not always respected.

The prevalence of paternal uncles and aunts in caregiving has gone down (Foster, 2000). More and more grandparents, maternal relatives and siblings have become caretakers, a fact that some observers attribute to a weakening of traditional safety nets. However, despite these changes, the "predominant caring unit" remains the extended family (Foster, 2000; Evans, 2004). Table 1 shows how many double orphans in my sample are taken care of by different categories of caregivers.[5] In my data collection effort and analysis, I defined the *family network* to be the

[3]Wives gained usufructuary rights over land through their husbands. In most parts of Kagera, only male children could inherit, with the major parts going to the eldest and the youngest male child. In some parts of Kagera, female children could inherit land, though their children, being from a different clan could not (Muchunguzi, 2002). Two things have changed more recently. First of all, more and more land is held not as clan land, and non-clan land can be bequeathed without restriction. In addition, recent legislative changes have invalidated obstacles to female inheritance in traditional laws. This invalidation is not perfect, however. Beyond issues of social acceptance, the process for a woman to claim land that she has inherited involves going to court and there invalidating the traditional laws that bar her from inheriting. Appleton (2000) mentions a law from colonial times that specified a wife's right to remain on her husband's land upon his death, but this colonial law was never mentioned by any person the author talked to and even if it exists, it might not be implementable.

[4]Information that I obtained from focus group interviews.

[5]Often, grandparents reside with their eldest son, so that it is not clear whether a grandfather is listed as a caregiver because he is truly the primary caregiver or because he is considered the head of the household in which

aunts, uncles, grandparents and siblings of the orphan, as well as half-siblings of the orphan and half-siblings of his or her parents. Around 80% of double orphans are taken care of by members of the family network thus defined.[6] In the following section I lay out a model of the decision process of caregiving within the network.

3 Model

For expository convenience, the model incorporates two people in the network and one orphan. The model is extended easily to more than two network members, and the econometric implementation allows for multiple orphans. The network members get together at the death of the deceased parent to decide on the caretaking arrangement of the orphan. They make their decision by maximizing the network welfare function which is the sum of the utility of consumption of the two network members, v_i, (where $i \epsilon \{1, 2\}$) and of the utility of the orphan, v_o, weighted by δ_N. When deciding on the placement of the orphan, the network members cannot commit to specific levels of consumption for the orphan. There are also no transfers between network members.[7] Thus the consumption of the caretaker and the orphan is determined by the caretaker, who divides the available resources between herself and the orphan, so as to maximize an altruistic utility function where the orphan's utility from consumption is weighted by a factor $\widetilde{\delta}_j$. I assume that network members have perfect knowledge about each other's resources and preferences. The network members incorporate the optimal levels of consumption from the within household utility maximization into the decision of where to place the orphan. Note that in the model the only relevant outcome is consumption, but one can think of this more generally as any outcome that affects the orphan's utility such as health and education.

the orphan resides.

[6] This is in fact similar to what Deininger et al. (2003) report for a different area of Tanzania. They find that between 80 and 95 % of single parent and double orphans are taken care of by the extended family. The range is due to the fact that their definition of extended family includes cousins, whereas mine does not.

In the estimation sample, I have to drop a number of cases in which orphans were taken care of by caregivers to whom I could not administer questionnaires (see below). Thus, in the estimation sample, the percent taken care of by family members is lower at 78%.

[7] Lundberg et al. (2000) point out that the transfers to the household that experienced an adult death petered out within a year.

In the model i denotes the person and j the placement. In general, the consumption of a network member i when person j is the caretaker is c_{ij}. In the household of the caretaker j, the consumption of the network member is c_{jj} and the consumption of the orphan is c_{oj}.

The network maximizes the following network welfare function by choosing to place the orphan with network member $j \epsilon \{1, 2\}$

$$U_j = v_1(c_{1j}) + \delta_{NU} v_o(c_{oj}) + v_2(c_{2j}) \tag{1}$$

where

- $v_{ij}, \; v_{oj}, \;\; i \epsilon \{1, 2\}, \;\; j \epsilon \{1, 2\}$ are the utility of consumption of the adult network member i and the orphan respectively, with v continuous and twice differentiable, $v' > 0, v'' < 0$
- c_{ij}, c_{oj} are the levels of consumption of adult network member i and of the orphan respectively, when the orphan is taken care of by j (it is possible that $i = j$)
- δ_{NU} is the utility weight of the orphan in the network's maximization problem. δ_{NU} could be the sum of the weights the households attach to the orphan's utility, denoted by $(\widetilde{\delta}_i, \widetilde{\delta}_j)$, or 1. For simplicity of notation I set $\delta_{NU} = 1$ although this does not change the qualitative results.

I can rewrite the network maximization problem as

$$\begin{matrix} \max \; U \\ j \; \epsilon \; \{1, 2\} \end{matrix} = \begin{cases} v_1(c_{11}) + v_o(c_{o1}) + v_2(c_{21}) & if \;\; j = 1 \\ v_1(c_{12}) + v_o(c_{o2}) + v_2(c_{22}) & if \;\; j = 2 \end{cases} \tag{2}$$

The consumption of the orphan is determined by the caretaker he is assigned to. The caretaker maximizes the following altruistic utility function by optimally dividing resources between herself and the orphan

$$\max_{c_{oj}} \quad v_j(c_{jj}) + \widetilde{\delta}_j v_o(c_{oj}) \tag{3}$$

where $\widetilde{\delta}_j$ is the weight placed on the utility of the orphan by network member j.The caretaking network member j faces the budget constraint that her consumption and the consumption have to be less than her own ressources and the contribution of the orphan;

$$c_{jj} + c_{oj} \leq w_j + r_{oj}$$

where

wj : wealth of the network member

r_{oj} : contribution of the orphan to the wealth of the network member, such as the orphan's inheritance as well as the value of the orphan's labor.

The budget constraint for the non-caretaking network member is much simpler, specifying simply that her consumption must be less than her wealth;

$$c_{kj} \leq w_k$$

3.1 Utility Maximization in the Households

In the household of the caretaker, the maximization of the altruistic utility function $v_j(c_{jj}) + \widetilde{\delta}_j v_o(c_{oj})$ leads to the first order condition

$$v_j'(w_j + r_{oj} - c_{oj}) = \widetilde{\delta}_j v_o'(c_{oj}) \tag{4}$$

which yields the optimal levels of consumption for the network member and the orphan, as a function of the exogenous parameters;

$$\begin{aligned} c^*_{jj} &= c_{jj}(w_j, r_{oj}, \widetilde{\delta}_j) \\ c^*_{oj} &= c_{oj}(w_j, r_{oj}, \widetilde{\delta}_j) \end{aligned}$$

The non-caretaker consumes her budget;

$$c^*_{kj} = w_k$$

The maximization problem of the network can be rewritten subtracting the sum of utilities of the network members when the other network member is the caretaker, $v_1(c_{12}) + v_2(c_{21})$[8]

$$\max_{j \,\epsilon\, \{1,2\}} U = \begin{cases} v_1(c_{11}) + v_o(c_{o1}) + v_2(c_{21}) - (v_1(c_{12}) + v_2(c_{21})) & if \quad j = 1 \\ v_1(c_{12}) + v_o(c_{o2}) + v_2(c_{22}) - (v_1(c_{12}) + v_2(c_{21})) & if \quad j = 2 \end{cases} \tag{5}$$

This simplification leads to a formulation in which the utility of the network from placing the orphan with network member i is only dependent on the characteristics of network member i and

[8] The levels of consumption in this case are determined by the wealth of the network member, $c_{12} = w_1; c_{21} = w_2$.

the contribution of the orphan to the budget of network member i.[9] Formally, the maximization problem of the network including the constraints can then be written as follows:

$$\Leftrightarrow \begin{array}{c} \max\ U \\ j\ \epsilon\ \{1,2\} \\ s.t. \end{array} = \begin{cases} v_1(c^*_{11}) + v_o(c^*_{o1}) - v_1(w_1) & if\ \ j=1 \\ v_o(c^*_{o2}) + v_2(c^*_{22}) - v_2(w_2) & if\ \ j=2 \end{cases}$$

$$\begin{aligned} c^*_{o1} &= \arg\max \quad v_1(c_{11}) + \widetilde{\delta}_1 v_o(c_{o1}) \quad s.t. \quad c_{11} + c_{o1} \leq w_1 + r_{o1} \\ c^*_{11} &= w_1 + r_{o1} - c^*_{o1} \end{aligned}$$

$$\begin{aligned} c^*_{o2} &= \arg\max \quad v_2(c_{22}) + \widetilde{\delta}_2 v_o(c_{o2}) \quad s.t. \quad c_{22} + c_{o2} \leq w_2 + r_{o2} \\ c^*_{22} &= w_2 + r_{o2} - c^*_{o2} \end{aligned}$$

Define the indirect utility function of placing the orphan in household 1 as:

$$U_1(w_1, r_{o1}, \widetilde{\delta}_1) = v_1(c^*_{11}) + \widetilde{\delta}_1 v_o(c^*_{o1}) - v_1(w_1) \tag{6}$$

The network chooses to place the orphan with the network member whose caretaking yields highest network utility. That implies that network member 1 will be chosen as caretaker if

$$U_1(w_1, r_{1o}, \widetilde{\delta}_1) \geq U_2(w_2, ro_2, \widetilde{\delta}_2) \tag{7}$$

In the Appendix, part A, I derive predictions for how $(w_1, r_{o1}, \widetilde{\delta}_1, w_2, r_{o2}, \widetilde{\delta}_2)$ affect the utility of the network from placing the orphan with a specific network member, and thus how they influence the placement of the orphan.

The derivations yield the intuitive results that the weight of the orphan in the altruistic utility function of a network member increases the probability that the orphan is taken care of by a network member. Moreover, the relative contribution of the orphan to a specific household increases the utility of placing the orphan there. The results for wealth are ambiguous. When

[9] U_j can now be thought of as the change in network utility occurring by network member j taking care of the orphan and optimally dividing her wealth (and possibly the contribution of the orphan) between herself and the orphan vs. network member j consuming her own budget.

the contribution of the orphan to the resources of the household (in terms of wealth) is larger than the level of consumption that he would receive in that household in the absence of a contribution, the probability of placing the orphan there first decreases with increasing wealth. The intuition is the following: when the orphan comes with a large inheritance and there is a network member j with very low own wealth, the network might assign the orphan to j in order to provide *her* with consumption. As her own wealth increases, her marginal utility from consumption decreases, reducing the incentive for the network to place the orphan with her just to provide her with consumption. As w continues to increase, it will positively affect caretaking. Thus, I cannot theoretically sign how the probability of placing the orphan in a household is affected by the network member's wealth.[10]

The main result of these derivations is that observed and unobserved characteristics of the network members influence both the placement of the orphan as well as the consumption of the orphan. The following section shows that this leads to a bias in a regression of the outcome of the orphan on observable characteristics of the caretaker.

3.2 Unobservable Preferences and Selection Bias

Take the following linearized version of the network utility

$$U_j = \beta_u w_j + \gamma_u \widetilde{\delta_j} \tag{8}$$

where for simplicity $r_{oj} = 0$.[11]

The following inequality constraints (if allocation 1 is chosen) must hold for allocation 1 to be chosen:

$$\beta_u w_1 + \gamma_u \widetilde{\delta_1} \geq \beta_u w_2 + \gamma_u \widetilde{\delta_2} \tag{9}$$

A linearized version of the consumption of the orphan as a function of wealth and preferences of the caretaking households would be (again ignoring the contribution of the orphan to the

[10] Including interactions between the orphans inheritance and the network members wealth might in future research shed light on this issue.

[11] This is not restrictive for the model, since if $r_{oj} = r_{ok}$, the terms drop out of the inequality constraints. Otherwise, if the contribution is observed, the variable behaves like w_j otherwise like $\widetilde{\delta_j}$.

household)[12]

$$c_{oj} = \beta_c w_j + \gamma_c \widetilde{\delta}_j \tag{11}$$

In the Appendix, part B, I derive the following results when assuming a uniform distribution for the observed and unobserved characteristics.[13] Denote the caretaker by j. Then

1. $E[\widetilde{\delta}_j] > E[\widetilde{\delta}]$

   In expectation, caretakers attach a higher weight to the utility of the orphans than non-caretakers

2. $E[w_j] > E[w]$

   In expectation, caretakers are richer than non-caretakers

3. $Cov[w_j, \widetilde{\delta}_j] < 0$

   Among caretakers there will be a negative relationship between observable and unobservable characteristics, even when there is none in the general population. Intuitively, a person with low levels of the observable variable wealth w_j must compensate for this with high values of the unobservable utility weight $\widetilde{\delta}_j$ to be chosen as caretaker, whereas for a rich person, the value of $\widetilde{\delta}_j$ can be quite low, and she might still be chosen as caretaker.

The typically available household data would contain information only on the caretaking household in the family network. A regression of the following form

$$c_{oj} = \beta_c w_j + e_{cj} \tag{12}$$

[12]The effect of orphanhood on orphans is the difference between the level of consumption in the network member's household and the level of consumption of the orphan before his parents died, where the latter is denoted by C_o^0. The previous equation then becomes

$$c_{oj} - c_o^0 = \beta_c w_j + \gamma_c \widetilde{\delta}_j - C_o^0 \tag{10}$$

To control for C_o^0 (and the possibility that it might be correlated with w_j or $\widetilde{\delta}_j$), I need information on the orphan before his parents died, i.e. panel data.

[13]The results are corroborated for other distributions using simulations.

would lead to biased results. In particular, the estimated coefficient on wealth will be biased *downward*

$$E[\widehat{\beta_c}|w_j] < \beta_c$$

because $Cov[w_j, \widetilde{\delta_j}] < 0$ and $\widetilde{\delta_j}$ is part of the unobserved characteristics in the caretaking household. I call this bias the *placement bias.*

3.2.1 Selection Correction

Many of the family networks in our data are quite large. Thus, the estimation of the placement of the orphans will be done with a conditional logit model. I then use the method of Lee (1983) and Trost and Lee (1984) is to correct for placement selection.[14] It involves estimating the probability of placement with a particular caretaker and then constructing a selection term for the selected caretaker;[15]

$$\lambda_1 = \frac{\phi(\Phi^{-1}(P(1|w_1, w_2))}{P(1|w_1, w_2)} \tag{13}$$

where $P(1|w_1, w_2)$ is the predicted probability of placement 1 from a logit model when the characteristics of the network members are w_1, w_2 and ϕ (Φ) is the standard normal pdf (cdf).

The estimation equation for the outcome of the orphan is

$$C_{oj} = \beta_c w_j + \varphi \lambda_j + e_{cj} \tag{14}$$

The selection correction term λ_j is identified through the use of w_k, the characteristics of the non-caretaker, in the probability of selection.

It may be noted that there will be a placement bias, even when a proxy for the utility weight is available, such as the degree of relatedness between the orphan and the caretaker. Let $\widetilde{\delta}_j = \delta_j + \epsilon_j$, where δ_j and ϵ_j are the observed and unobserved components of the utility weight respectively. ϵ_j still influences both the placement as well as the outcome and will thus lead to a placement bias.

[14] In this simple one variable setup a control function can be used that is constructed as the difference in levels of wealth. This term would be included with the wealth of the caretaker.

[15] This can be thought of as an inverse Mill's ratio, where the logistic error terms of the placement model are transformed into normally distributed error terms via $\Phi^{-1}(P(1|w_1, w_2))$. The denominator of the Mill's ratio is $\Phi[\Phi^{-1}(P(1|w_1, w_2))] = P(1|w_1, w_2)$.

4 Econometric Implementation of the Model

Options available to the orphan outside of his family network are not observed in my data. As a consequence, I conduct my analysis *conditional* on the orphan being taken care of by someone in the family network.

I first illustrate the econometric implementation of the model when the deceased parents only had one child. Let U_j be the network utility if person j were to take care of the orphan in a one-orphan network.[16] Following the setup of the previous model, this utility only depends on characteristics of network member j and characteristics of the orphan in that household. I denote the chosen allocation by $\widehat{j}$. By a revealed preference argument, it must be that the network valued caretaking by $\widehat{j}$ more than caretaking by k, that is: $U_{\widehat{j}} \geq U_k \;\; \forall\ k$.

This is illustrated for the case in which there is one orphan ($N_o = 1$), three family members ($N_{NTW} = 3$) who form the family network ($NTW = \{1, 2, 3\}$) and the first allocation was chosen ($\widehat{j} = 1$). Then the model implies the following inequality constraints: $U_1 \geq U_2$ and $U_1 \geq U_3$.

I assume that the value the network attaches to caretaking by j is a function of observable characteristics of j as well as unobserved characteristics: $U_j \equiv \overline{U_j} + \varepsilon_j$, where the component that is due to observable characteristics is a linear function of characteristics of the network member and interactions between the network member and orphan characteristics:

$$\overline{U_j} = X_j \beta_1 + X_j X^i \beta_2 \tag{15}$$

where

X_j : characteristics of the network member in allocation j

$X_j X^i$ network member-orphan interactions

$\beta = \left[\begin{smallmatrix}\beta_1 \\ \beta_2\end{smallmatrix}\right]$

The inequality constraints, when j is chosen are:

$$\overline{U_{\widehat{j}}} + \varepsilon_{\widehat{j}} \geq \overline{U_k} + \varepsilon_k \quad \forall\ k\ \neq\ \hat{\jmath} \tag{16}$$

[16] To denote the network welfare when multiple orphans are present, I use U_{jk} in the two orphan case and U_a in the general case. In the former, j is the caretaker for orphan 1 and k is the caretaker for orphan 2. In the general case, $a = \{j\}$ in the one orphan network, $a = \{j, k\}$ in the two orphan network and $a = \{j, k, l\}$ in the three orphan network.

I now extend the model to include family networks with two orphans (the setup is easily extended to more than two orphans) and lay out the assumptions needed for combining one and two orphan networks.

4.1 Two Orphan Case

Let U_{jk} be the network welfare in a network with two orphans when persons j and k are taking care of the first and second orphan respectively (and it is possible that $j = k$). The chosen allocation now consists of a chosen caretaker for the first orphan and another one for the second orphan: $\widehat{jk}$. The same revealed preference argument as above implies that

$$U_{\widehat{jk}} \geq U_{lm} \quad \forall\ lm$$

As an illustration, take the example in which there are two orphans ($N_o = 2$), the family consists of three network members ($N_{NTW} = 3$; $NTW = \{1,2,3\}$), and the allocation chosen is $\widehat{jk} = (1,1)$. The set of possible allocations in this case is $A = \{1,2,3\} \times \{1,2,3\}$. The model then implies the following inequality constraints:

$$\begin{aligned} U_{11} &\geq U_{12} \quad , \quad U_{11} \geq U_{21} \quad , \quad U_{11} \geq U_{13} \quad , \quad U_{11} \geq U_{31}, \\ U_{11} &\geq U_{23} \quad , \quad U_{11} \geq U_{32} \quad , \quad U_{11} \geq U_{33} \quad , \quad U_{11} \geq U_{22}. \end{aligned} \tag{17}$$

The number of allocations is $(N_{NTW})^{N_o} = 3^2 = 9$, and the number of inequalities is the number of allocations minus one, i.e. 8.

Define the predictable component of network utility from j taking care of orphan 1 imagining there was no second orphan as:

$$\overline{U_j^1} = X_j\beta_1 + X_jX^1\beta_2 \tag{18}$$

and the counterpart from person k taking care of orphan 2 as:

$$\overline{U_k^2} = X_k\beta_1 + X_kX^2\beta_2$$

The network utility from an allocation $\{j,k\}$ is assumed to be the sum of these utilities:

$$\begin{aligned} \overline{U_{jk}} &= \overline{U_j^1} + \overline{U_k^2} \\ &= X_j\beta_1 + X_jX^1\beta_2 + X_k\beta_1 + X_kX^2\beta_2 \\ &= (X_j + X_k)\beta_1 + (X_jX^1 + X_kX^2)\beta_2 \end{aligned} \tag{19}$$

This formulation specifies that the utility a person derives from caretaking of an orphan, say orphan 1, is the same when there is one orphan or when there are more orphans. I modify this assumption to allow for the possibility that the utility from taking care of a second orphan might not be the same.

4.1.1 Private Goods, Economies of Scale and Disutility from Separating Orphans

When taking care of multiple orphans, there might be economies of scale for household public goods. On the other hand, private goods become more scarce. Third, there could be a strong disutility from separating orphan siblings. I cannot distinguish among these three effects, but I want to allow for the utility to be affected when a caretaker is taking care of multiple orphans. Therefore, I let

$$\overline{U_{jk}} = \overline{U_j^1} + \overline{U_k^2} + \theta_2 ES \tag{20}$$

where ES is a dummy for whether there is more than one orphan in this allocation, i.e. $1\{j = k\}$. The coefficient θ_2 tries to capture the net effect of economies of scale, scarce resources and utility from keeping orphans together.[17]

4.1.2 Error Term

The inequality constraints, allowing for unobservables, are now given by:

$$\overline{U_{\widehat{jk}}} + \varepsilon_{\widehat{jk}} \geq \overline{U_{lm}} + \varepsilon_{lm} \quad \forall\ lm\ \neq \widehat{jk} \tag{21}$$

Using the inequality constraints, the parameters β can be estimated once a distribution of the error term has been assumed.

To combine the subsamples with one and two orphans, I need to make some assumptions on the parameters β and the variances of the error terms ε_k and ε_{lm}. Concerning the former, I assume that

$$\beta^1 = \beta^2 = \beta$$

where β^1 and β^2 denote the parameters in the one-orphan and two-orphan cases respectively. Concerning the latter, I assume that $var(\varepsilon_k) = var(\varepsilon_{lm})$

[17] A second approach would be to allow the effect of the second orphan to come through a multiplicative effect $U_{jk} = (U_j^1 + U_k^2)(1 + \theta_2 1\{j = k\})$.

4.2 Estimation with a Conditional Logit

Using McFadden's conditional logit model I estimate the determinants of placing the orphan with a particular family member, *conditional* on the orphan being placed in the family network.[18] In the usual conditional logit the choice probability of choosing allocation j (here for the one orphan case) is[19]

$$P_j = \frac{\exp(U_j + \alpha)}{\sum_{l \in A} \exp(U_l + \alpha)} = \frac{\exp(U_j)}{\sum_{l \in A} \exp(U_l)} \tag{22}$$

where the α allows for heterogeneity in the utility the networks derive from taking care of the orphan. The network fixed effect drops out when I condition on the orphan being placed in the network. I then maximize the likelihood function

$$L = \sum_{j \epsilon A} y_j \ln P_j(\beta) \tag{23}$$

where y_j denotes whether an allocation is chosen.

In the two orphan case, the placement probability of choosing allocation jk is (where for now I abstract from the possibility that a caretaker might take care of both orphans);

$$P_{jk} = \frac{\exp(U_j^1 + U_k^2)}{\sum_{l,p \in A} \exp(U_l^1 + U_p^2)} \tag{24}$$

The relative choice probability of P_{jk} relative to P_{lp} is

$$\frac{P_{jk}}{P_{lp}} = \frac{\exp(U_j^1 + U_k^2)}{\exp(U_l^1 + U_p^2)} \tag{25}$$

The relative choice probability of a choice jj relative to leaving the first orphan with j and placing the second orphan with m instead is:

$$\frac{P_{jj}}{P_{jm}} = \frac{\exp(U_j^1 + U_j^2)}{\exp(U_j^1 + U_m^2)} = \frac{\exp(U_j^2)}{\exp(U_m^2)} = \frac{P_j^2}{P_m^2} \tag{26}$$

Thus, the relative choice probabilities of placing the second orphan with j versus m is the same regardless of where the first orphan is placed. When I include the economies of scale variable, the relative choice probabilities become

$$\frac{P_{jj}}{P_{jm}} = \frac{\exp(U_j^1 + U_j^2 + \theta_2 ES)}{\exp(U_j^1 + U_m^2)} = \frac{\exp(U_j^2) * \exp(\theta_2 ES)}{\exp(U_m^2)} = \frac{P_j^2 * \exp(\theta_2 ES)}{P_m^2} \tag{27}$$

[18] The estimated coefficients on the determinants of the placement can be thought of as the parameters of the network's welfare function.

[19] In the following, I use U_j to denote the predictable component of utility, not $\overline{U_j}$.

Shifting the second orphan from person j to person m now involves trading off the characteristics of j and m as well as the utility from keeping orphans together

4.3 The Structure of the Data

The variables for the estimation via conditional logit are created as follows: I create $(N_{NTW})^{N_o}$ allocations for each network, where N_{NTW} is the number of network members and N_o is the number of orphans. For each allocation, I add up the variables for the N_{NTW} network members in that allocation. For the one orphan case, the structure of the dataset looks as follows

Figure 1: One Orphan Case

Caretaking Possibilities			Allocations	DATA		
1	2	3		y	Network Member	Netw. Memb.*Orphan
a	0	0	(1)	0	X_1	$X_1 * X^a$
0	**a**	**0**	**(2)**	1	X_2	$X_2 * X^a$
0	0	a	(3)	0	X_3	$X_3 * X^a$

The left-hand side displays the three possible allocations. The right-hand side shows the structure of the data where every row represents an allocation (or in the one orphan case, a network member).

The structure of the data for the two orphan case I use is displayed in the following figure, where ES denotes the economies of scale variable.

Figure 2: Two Orphan Case

Caretaking Possibilities 1	2	3	Allocations	DATA y	Netw. Memb.	Netw. Memb.*Orphan	ES
a,b	0	0	(1,1)	0	$2X_1$	$X_1 * X^a + X_1 * X^b$	1
0	a,b	0	(2,2)	0	$2X_2$	$X_2 * X^a + X_2 * X^b$	1
0	0	a,b	(3,3)	0	$2X_3$	$X_3 * X^a + X_3 * X^b$	1
a	b	0	(1,2)	0	$X_1 + X_2$	$X_1 * X^a + X_2 * X^b$	0
a	**0**	**b**	**(1,3)**	**1**	$X_1 + X_3$	$X_1 * X^a + X_3 * X^b$	**0**
0	a	b	(2,3)	0	$X_2 + X_3$	$X_2 * X^a + X_3 * X^b$	0
b	a	0	(2,1)	0	$X_2 + X_1$	$X_2 * X^a + X_1 * X^b$	0
b	0	a	(3,1)	0	$X_3 + X_1$	$X_3 * X^a + X_1 * X^b$	0
0	b	a	(3,2)	0	$X_3 + X_2$	$X_3 * X^a + X_2 * X^b$	0

With 3 network members and 2 orphans there are $3^2 = 9$ possible allocations. Most of the data falls into the top third, where one caretaker takes care of all orphans. However, in the actual data, some orphans are split up and one of the other six remaining allocations is chosen. The allocation $\{1, 3\}$ (in bold) is chosen. Network member 1 takes care of orphan a and network member 3 takes care of orphan b. Thus the allocation that is chosen is $(1, 3)$. The utility from this allocation is $U_1^a + U_3^b$ and the table on the right shows how the regression variables are formed.

4.4 Adjustment for Sampling from the Choice Set

The sample of network members collected for the analysis is a stratified sample from the family network.

McFadden (1978) derives the formula for a consistent estimator of β when the sample includes the chosen alternative and a stratified random sample of the non-chosen alternatives. In my case, this corresponds to sampling from the network of the orphan.[20]

$$P(j|B) = \frac{\exp(U_j + \log(csp_j))}{\sum_{k\in A}\exp(U_k + \log(csp_k))} = \frac{\exp(U_j) * csp_j}{\sum_{k\in A}\{\exp(U_k) * csp_k\}} \tag{28}$$

where csp_j is defined as the probability that the observed subset of network members (B) is observed conditional on j being the chosen alternative.[21] With random sampling from the choice set the correction terms fall out due to the fact that $P(B|j) = P(B|k)$ or $csp_j = csp_k$.

However with stratified random sampling from the choice set, the choice set probability is different depending on whether I condition on j or k, i.e. $csp_j \neq csp_k$. Since this is how my sample was constructed I need to adjust for the stratified random sampling from the choice set. As I show in the empirical results, this adjustment has substantial consequences for the estimated coefficients for various groups of network members. The estimated coefficients for the non-relationship variables are however consistent without this correction (Bierlaire, Bolduc and McFadden, 2006).

4.5 Flexible Substitution Patterns

I allow for flexible substitution patterns between different categories of caretakers, such as paternal and maternal grandparents, aunts and uncles as well as siblings. The parameters β_g on these dummies might be heterogeneous in different networks, so I allow them to be random variables from a specified distribution.[22] I estimate this model with a mixed logit specification. Mixed logit models incorporate a mixture of a logit error term and another error term. Conditional on the non-logit error term, the choice probabilities follow the usual logit formula. The mixed logit model has the advantage of allowing a flexible error structure while keeping the computational advantage of a closed form solution for the choice probabilities. Moreover, McFadden and Train (2000) have shown that appropriately defined mixed logit models can approximate any random

[20] The correction is similar to corrections for choice-based sampling (Manski and McFadden, 1981; Cosslett, 1981) but McFadden (1978) and Bierlaire, Bolduc and McFadden (2006) also explicitly deal with sampling from the choice set. The semi-parametric estimator of Fox (2005) which does pairwise comparisons is outside the scope of our sample. In addition, the selection correction term requires the predicted probabilities in the case of sampling from the choice set.

[21] In McFadden (1978) and Bierlaire et al.(2006) B is called the *observed choice set.*

[22] Train (2003) points out that using random coefficients in this way is similar to specifying a nested logit model. We use the mixed logit specification, because the multiple orphan case leads to overlapping nests.

utility model. To begin, I describe the mixed logit model in the case where there is one orphan in the network.[23] I assume the category dummies follow a known distribution, where the mean and variance are unknown. The following equation displays the network welfare function from placing orphan 1 from network n with person j, where I drop the interaction terms for simplicity;

$$\begin{aligned} U_j^n &= \overline{U_j^n} + v_j^n + \eta_j^n \\ &= X_j^n \beta + v_j^n + \eta_j^n \end{aligned} \tag{29}$$

where η_j^n is a logistic error term and $v_j^n = D_j^n * \xi^n$, $\xi^n \sim F(0, \Omega)$, where D_j^n is the set of dummies for the different groups and Ω a variance-covariance matrix.[24] The distribution F is assumed to be a triangular distribution.

Noting again that D_j^n is contained in X_j^n, I can rewrite this as follows (where ξ^n now contains zeros for all X_j^n columns that are not category-specific dummies)

$$\begin{aligned} U_j^n &= X_j^n \beta + X_j^n * \xi^n + \eta_j^n \\ &= X_j^n (\beta + \xi^n) + \eta_j^n \\ &= X_j^n \widetilde{\beta} + \eta_j^n \end{aligned} \tag{30}$$

where $\widetilde{\beta} = \beta + \xi^n \sim F(\beta, \Omega)$.

If the values of $\widetilde{\beta}$ are known, i.e. conditional on values of $\overline{\widetilde{\beta}}$, the choice probabilities follow the usual logit formula.

$$P_{lm}(v) = \frac{\exp(X_j^n \overline{\widetilde{\beta}})}{\sum_{k \in A} \exp(X_k^n \overline{\widetilde{\beta}})} \tag{31}$$

The researcher does not know $\overline{\widetilde{\beta}}$. However, I can get at the unconditional random coefficients choice probabilities by integrating over the $\overline{\widetilde{\beta}}$. That is, if I denote the density of $\widetilde{\beta}$ by $f(\widetilde{\beta}|\beta, \Omega)$, the unconditional random coefficients choice probabilities can be written as:

[23] This exposition follows Train (2003).

[24] If there are K categories, we can only estimate $K - 1$ variances for the coefficients (Bolduc, Ben-Akiva and Walker; 2001). I set the variance term with the lowest estimated standard deviation to zero, when the variances are estimated one at a time. The matrix Ω is assumed to be diagonal, so that the random coefficients are not correlated.

$$P_j(\beta,\Omega) = \int .. \int P_j(\widetilde{\beta}) f(\widetilde{\beta}|\beta,\Omega) d\widetilde{\beta} \tag{32}$$

From this formula it is easy to see that the independence of irrelevant alternatives property will not hold.[25] The extension to multiple orphans is simple: in equation 31 the X_j^n would now be the sum of the X_j^n for the different caretakers.

The likelihood function is given by:

$$L(\beta,\Omega) = \sum_{j \epsilon A} y_j \ln P_j(\beta,\Omega) \tag{33}$$

where $y_j = 1$ if the allocation j is chosen. The likelihood function cannot be estimated directly with maximum likelihood, since the $P_j(\theta)$ involves an integral. The mean β and the variance covariance matrix Ω of $\widetilde{\beta}$ can be estimated through simulation (Train, 1998; McFadden and Train, 2000).[26] Given parameters θ I simulate draws from $f(v|\theta)$ and calculate $P_j(\beta^r, \Omega^r)$.[27] The average of these draws is the simulated probability:

$$P_j\widehat{(\beta,\Omega)} = \frac{1}{R} \sum_{r=1}^{R} P_j(\beta^r, \Omega^r) \tag{34}$$

The Simulated Loglikelihood is

$$SL = \sum_{j \epsilon A} y_j \ln \widehat{P_{lm}}(\beta,\Omega) \tag{35}$$

In the empirical analysis, I regress a dummy whether an allocation of orphans was chosen on characteristics of the network members in that allocation as well as network member-orphan interactions.[28]

The following section describes how my data was collected and presents descriptive statistics on the variables I use in the empirical specification.

[25] As Train (2003) points out, since the denominators of the logit formula are inside the integral, they do not cancel out.

[26] In the estimation, I assume the covariance matrix is diagonal.

[27] I use Halton sequences (Train, 2000).

[28] The characteristics of the orphan alone are not included as in the conditional logit and mixed logit specifications used they drop out since they are the same across alternatives.

5 Data and Empirical Setting

I designed and collected a survey on the placement of orphans within their family network as well as on the characteristics of the caretakers and other members of the family network. The orphan children were identified from a panel dataset, the Kagera Health and Development Survey (KHDS) collected by the World Bank. The KHDS is a survey comprised of two parts, a panel dataset with four rounds conducted from September 1991 to January 1994 as well as a re-interview in 2004. While the first part of KHDS (henceforth KHDS-1) was a panel of households, the re-interview phase (KHDS-2) attempted to interview all individuals ever interviewed in KHDS-1. Of the 895 households interviewed in KHDS-1, the resurvey contacted members of 832 households who lived in over 2700 households in 2004. The long-term panel dimension of the KHDS is an extraordinary feature of a dataset in a developing country context. In addition, the KHDS-2 attempted to interview respondents even when they had moved outside the village and outside the region (and there are even some households interviewed in the neighboring country of Uganda). The mortality of household members was documented in a mortality questionnaire administered to all original households.[29] During the 14 years spanned by the KHDS, around 18 percent of household members died.

I sampled from KHDS-1 members who died between 1991 and 2004 and who had children under 15 at the time of their death. Information on whether the adults had children came from three sources: first, whether the parents had lived with their children in the interviewed households in 1991-1994; then, whether KHDS-1 listed these adults as having children who were living elsewhere, and third, whether respondents in 2004 mentioned the deceased as their

[29] The original sampling strategy of KHDS-1 involved stratified sampling at two levels: the first involved the agro-climatic zones of Kagera and within the zones regions of high and low adult mortality. The second level of stratification was within the selected clusters where households split into two groups according to a perceived level of risk of having a member die. The high risk group comprised all households in which there was a sick adult member or in which an adult death had occurred very recently. Because of the sampling scheme adult mortality in the KHDS-1 survey was twice as high as in the general population. The questionnaire of the KHDS-1 (and that of the KHDS-2) was similar to that of the Living Standard Measurement Surveys of the World Bank, adapted for the panel structure. For a more detailed description of the data, see Beegle et al. (2005) and World Bank (2004a).

parent.[30] The sample of deceased parents was then constructed as follows: I kept all parents who left behind double orphans (i.e. both parents died), all households in which there were two separate adult deaths and all cases in which at least one child of the parents died after their death. In total, of the 215 parents who died during the period from 1991 to 2004, I sampled 141 households. Of these, information could be found on 139 cases.[31] In the end, my sample comprised 67 cases in which both parents had died before the youngest child was 15 years of age, 40 cases in which there were orphans who lost their father but not their mother and 32 cases in which there were orphans who lost their mother but not their father.

I then constructed a family tree for every sampled case. The family tree consists of the following people: the deceased member and his or her "spouse(s)" who were defined as all persons with whom the deceased had children and all persons to whom the deceased was considered to be "married";[32] the children of the deceased and of his or her spouse; the biological parents of the deceased, the biological parents of the spouse(s) as well as siblings and half-siblings of the deceased and his or her spouse(s). Viewed from the perspective of the orphaned children, the family tree lists their parents, step-parents, aunts, uncles and grandparents and various step-relatives. The members listed on this family tree form the family network or extended family of the orphan. Since some of these networks were very large, respondents were then sampled from the network according to the following rules: the grandparents of the orphans and surviving

[30] My survey is thus not comprehensive: First and foremostI, we do not capture cases in which KHDS respondents did not have children by 1994, then had children between 1994 and 2004, died between 1994 and 2004, and their children were not living with any other household member from KHDS-1 in 2004. Unfortunately, I cannot estimate how frequent that is. Second, the information on children living elsewhere in KHDS-1 was not always very good, as I found out in our survey, especially for polygamous men, some of whose spouses were not known to (or mentioned by) KHDS-1 respondents. Nonwithstanding the shortcomings of my sample, the advantage of the present strategy was the integration within a 13-year panel. However, the multiple layers of sampling as well as the above limitation of the survey make it difficult to view this sample as truly "representative" of the population in Kagera.

[31] Two cases turned out to be incorrect because of discrepancies in parental information in KHDS-1. The 139 cases come from 122 KHDS-1 households, since some households had multiple KHDS-1 deaths.

[32] The concept of being married is very loosely defined in Tanzania, in particular because living together without at least claiming to be married is frowned upon. A relationship can be made official through three channels: a traditional marriage involving a bride price paid to the bride's relatives, a church or mosque wedding or a state marriage in front of the court. The last one is very uncommon.

spouses were sampled. In monogamous relationships, up to four siblings of the father and up to four siblings of the mother were sampled. In polygamous relationships, spouses whose children were not orphans (because they were 15 years or older at the time of death of the deceased) were not sampled. Among households that were to be included in the sample but that were living outside of the Kagera, I dropped those living outside the two largest towns of Tanzania, Dar es Salaam and Mwanza. This was (partially) justified by the fact that of households who moved outside of Kagera two thirds moved to these major towns. In these two urban centers the interview rate was much lower.

The families that were sampled from the family network were administered a household questionnaire. In addition, I collected information on the circumstances of death of the deceased, information on where the children lived after the death of their parents as well as what happened to the assets of the deceased. In total, the field teams collected slightly over 1100 questionnaires for 139 cases of adult deaths. Since some of the networks overlap (due to the fact that in some KHDS-1 households multiple parents died), this means that on average 9 households were interviewed for each of the 122 networks sampled.

5.1 Double Orphan Sample

The 67 cases of parental death in which there were double orphans lead to 120 children who were under the age of 15 when their parents died. Around 22 % are not taken care of by family members and I exclude them from the subsequent analysis. The largest number is taken care of by older siblings (about a third), followed by paternal grandparents. Maternal grandparents, paternal uncles and aunts and maternal uncles and aunts play a much smaller role in caregiving. The 120 orphans analyzed come from 58 networks and Table 2 shows how many orphans are in these networks. Table 3 then leaves out non-cared-for orphans and shows how many caretakers take care of the orphans in a network. In about one third of the networks with more than one orphan, the orphans are split up.

Differences between caregivers and non-caregivers are presented in Table 4, which displays average characteristics of caregivers and non-caregivers. Grandparents make up a much larger percentage of the caregivers, whereas aunts, uncles and stepmothers are much less likely to be caregivers. The prevalence of grandparents in caregiving also explains the higher average age of caregivers. That caregivers are more likely to be male could be due to the fact that a male is more often considered to be the head of the household, and that the head of a household might be reported as caretaker. Network members who are caretakers are more likely to be traders, to have a good house flooring and to have more business assets.

6 Determinants of Orphan Placement

6.1 Regression Results

In the empirical analysis, I regress a dummy whether an allocation was chosen on the characteristics of the network member(s) in that allocation as well as interactions of the network member(s) with the orphan(s). Using the assumptions on the error term of a conditional logit model (and a mixed logit model) I estimate the following equation;

$$\begin{aligned} y_{an} &= RD_{an}\beta_1 + DM_{an}\beta_2 + W_{an}\beta_3 + OCC_{an}\beta_4 + \\ &+ COI_{an}\beta_5 + ES_{an}\theta + \alpha_n + \varepsilon_{an} \end{aligned} \tag{36}$$

where

- $y_{an}\epsilon\{0,1\}$: Dummy whether allocation a in network n is chosen
- RD : Dummies to capture the biological relationship to the orphan
- DM : Demographic characteristics of the network member(s)
- W : Wealth variables
- OCC : Occupation of network member(s)
- COI : Network member(s)-orphan interactions
- ES : "Economies of scale" variable
- α_n : Network specific fixed effect
- ε_{an} : Logit error, or composite error

In table 6, I estimate the determinants of the caretaking decision, accounting for the joint determination of where orphans go. The chosen allocation is regressed on the characteristics of the caretakers in that allocation.

Column 1 presents results from a conditional logit model on the relationship variables alone. The omitted category is paternal uncles and aunts, the caretakers according to traditional norms. I find that paternal grandparents and older siblings of the orphans are significantly more likely to be caretakers than paternal uncles and aunts. Yet these coefficients understate the true effects since the aunts and uncles observed are a sample of all aunts and uncles, whereas grandparents are not. Column 2 and all following columns present results from a conditional logit model using McFadden's correction for sampling from the choice set (McFadden, 1978; Bierlaire et al., 2006).

Column 3 shows that demographic variables are fairly unimportant; in particular age and education are not significantly correlated with caretaking. In fact whether a person is considered the head of the household is the only demographic variable that is significantly correlated with caretaking. This is consistent with the hypothesis that being the head of the household entails a stronger bargaining position in the intrahousehold bargaining over whether to take care of the orphan.

The fourth column adds variables that reflect the economic situation of the network member. Richer households are more likely to be caretakers as measured by the value of land the household owns as well as whether the household owns business assets. The occupation of network members also matters. If the network member reports being a farmer or a trader he or she is significantly more likely to be a caretaker.[33] I interpret the coefficients to mean that the cost of caregiving matters. Traders in Kagera typically operate from a small stall, selling produce or various other small amenities, and it is common to see children around. For farmers the presence of children does not affect their productivity as much.

I cannot answer whether the labor the orphan can provide is an important determinant of caretaking. The result that households with business assets and more land are more likely to be caretakers could point to child labor being demanded in these households. Interactions of the

[33] The omitted category includes various other occupations: skilled laborers, professionals, construction workers, fishermen as well as people reporting other or no job. These other variables are not significantly correlated with caretaking individually or in other combinations.

age of the orphan and business assets however point to households with businesses being more likely to take care of younger children.

In column 5 I find that there are significant economies of scale or a large disutility from separating orphans. This specification accounts for the joint determination of the placement of multiple orphans.

In table 7 I allow for flexible substitution patterns between different groups of caretakers. I find that there is significant heterogeneity only for maternal grandmothers. According to traditional norms the grandmothers should not be caretakers. The significant heterogeneity might point to the fact that these traditional norms on caretaking are less observed in some places than in others.[34]

6.2 Probability of Caretaking

In this section I report results from a simple robustness check. I keep only one orphan per caretaker and drop interaction effects. The inequality constraints from the model imply that the network welfare from assigning one (or more than one orphan) to a particular caretaker is higher compared with from assigning them to the non-caretakers. I now only consider whether a caretaker was chosen over the other members of the family network and do not incorporate the number of orphans in the decision. The probability that network member j is chosen as a caretaker of at least one orphan is

$$\Pr(j \ \ is \ \ caregiver) = \frac{\exp(\overline{U_j})}{\sum_l \exp(\overline{U_l})} \tag{37}$$

where j is any member of the family network. Comparing the results in table 8 with columns 4 and 5 of table 6, I find that this simpler methodology yields similar results, though some effects are not as precisely estimated. In particular, the coefficients on the network member being a sibling of the orphan or being a household head lose their significance and become (marginally) insignificant.

To present some indication of the magnitude of the effects, I calculated informal "marginal effects" which indicate how much more likely a network member is to become caretaker, if the

[34] In fact this might be linked to different tribes or a rural-urban divide, an issue I will explore further in the future.

right hand side variable changes. These are not the true marginal effects, because they neglect the possibility that orphans previously not taken care of by network members could, due to this change, be taken care of now. Conditional on being in the network in which orphans are currently being taken care of, the probability that a network member becomes the caretaker, for example, increases by 2.8 percentage points when his or her landholdings increase by one standard deviation and by 7.5 percent when he or she owns business assets.

7 Estimating the Impact of Caretaker Characteristics on Orphan Welfare

In this section I analyze the impact of the characteristics of their caretakers on the years of completed education of orphan children. The regression equation is the following

$$educ_{it} = orphan_{it} * (\alpha + X_{ij}\beta + \gamma\lambda_{ij}) + X_{i0}\delta_1 + age_{it}\delta_2 + age_{it}^2\delta_3 + \varepsilon_{it}$$

where

- X_{ij} : Characteristics of caretaker j
- X_{i0} : Characteristics of the households the orphans were living in before the death of their parents
- age_{it} : Polynomial in age
- λ_{ij} : Lee selection correction term
- $orphan_{i1}$: Dummy for whether the child was orphaned in time t
- ε_{it} : error term

The X_{i0} control for the level of education the orphan would have recieved in the households they were living in before the death of their parents. Alternatively, I also use a fixed effects specification, where there is an individual fixed effect.

Education in t=0 (1991):

$$educ_{i0} = \alpha_i + age_{i0}\delta_2 + age_{i1}^2\delta_3 + \varepsilon_{i0}$$

Education in t=1 (2004); $orphan_{i1} = 1$

$$educ_{i1} = \alpha + X_{ij}\beta + \gamma\lambda_{ij} + \alpha_i + age_{i1}\delta_2 + age_{i1}^2\delta_3 + \varepsilon_{i1}$$

This leads to the following first difference specification: (education in 2004 - education in 1991)

$$\begin{aligned} \Delta educ_i &= \alpha + X_{ij}\beta + \gamma\lambda_{ij} + \Delta age_{i1}\delta_2 + \Delta age_{i1}^2\delta_3 + \Delta\varepsilon_i \\ &= \alpha' + X_{ij}\beta + \gamma\lambda_{ij} + age_{i1}\delta_3 + \Delta\varepsilon_i \end{aligned}$$

The change in years of education for orphan i who is being taken care of in household j is assumed to be a linear function of the characteristics of the caretakers, X_{ij}, age variables, age_{ij}, an intercept α' which captures the average change in orphans' education, and an error term, ε_{ij}. λ_{ij} is a selection correction term for orphan i being in household j that is constructed as follows:

$$\lambda_{ij} = \frac{\phi(\Phi^{-1}(P_i(j|X_{ij}, X_{i,k\neq j}))}{P_i(j|X_{ij}, X_{i,k\neq j})}$$

where $P_i(j|X_{ij}, X_{i,k\neq j})$ is the estimated probability of j being selected as the caretaker for orphan i using the methodology outlined above, given the characteristics of the other network members, $X_{i,k\neq j}$.

Table 9 presents preliminary results from a regression of the change in years of education on the characteristics of the caretakers for orphan children[35]. The first column presents results without correcting for the placement of the orphan. It appears that paternal grandfathers do not have a precisely measured impact on the education of orphans. Moreover, compared to the left-out category, namely paternal uncles and aunts, maternal grandmothers seem to be particularly

[35] The results are preliminary, because the standard errors have not been corrected for the fact that the probabilities of selection are based on the estimated coefficients from the placement equation. Also, they do not account for the sampling from the family network.

bad for the educational prospects of the orphans. Household wealth does not seem to matter. The second column corrects for the endogenous placement using Lee's methodology (Lee, 1983). First of all, the correction term is significant and positive. This provides evidence that the unobserved determinants of the placement of the orphan are also important determinants of the education of the orphan. The second result is that the estimated coefficients are substantially affected. The coefficient on the paternal grandfather being the caretaker now becomes much larger and significant. The coefficient on the maternal grandmother, which is negative and significant in column 1, is, in fact, insignificant and slightly positive, in column 2. Third, the coefficient on whether a household owns business assets nearly doubles in size when accounting for selection. It is significantly correlated with orphans' education, whereas it does not seem to be in column 1.

In Table 10, I present results a regression of the levels of education on characteristics of the caretaking households, the selection correction term, a time trend, as well as characteristics of the original households, where the last are meant to control for parental background. The results follow the same patterns as those for the change in years of education.

It noteworthy that these strong results are based on a mere 56 observations. The fact that the empirical results are able to detect the bias with this sample size highlights how strong the placement bias is and how important it is to account for it.

The coefficients in the outcome regression and the coefficients in the placement regression are typically of the same sign. This implies that orphans are allocated within the network in a way that provides them with better education outcomes than a random allocation.

8 Conclusion

This paper examines the determinants of the placement of orphans within the family network and analyzes how orphan welfare is impacted by caretaker characteristics. I find that orphans are systematically placed within the family network according to the characteristics of potential caretakers. Economic factors such as household landholdings and whether the household owns business assets are positively correlated with caretaking; the probability of caretaking increase by 2.8 percentage points when the landholdings of the potential caretaker increases by one standard deviation and by 7.5 percent when he or she owns business assets. Members who are traders or farmers are more likely to be caretakers, presumably because of lower costs of caretaking and more opportunities for the orphan to be productive. Finally, the biological relationship between the network member and the orphan is important. Through documenting the systematic placement of the orphan within the family network this analysis also contributes to our understanding of how household structures adjust to mortality shocks in Sub-Saharan Africa.

I show analytically that the systematic placement of the orphan leads to a placement bias when regressing orphan outcomes on the characteristics of the caretakers. The theoretical framework developed here motivates a method for dealing with this bias. The method involves using the predicted probabilities of caretaking from the estimation of where the orphan is placed in the family network. In the empirical analysis I find that the bias is substantial. Without the selection correction, household wealth seems to be unimportant for orphans' education. However, after controlling for the endogenous placement of the orphan, the wealth of the caretaking households does significantly impact the education of orphans. This highlights the necessity of accounting for the endogenous placement of the orphan. For policymaking, this points out the need to incorporate information on the family networks of orphan children when determining eligibility for orphan aid. The empirical results further provide evidence for the importance of family networks in mitigating impacts associated with parental death. Orphans are systematically placed within the family network and as a consequence receive more education than if they had been placed randomly within the family network.

Appendix A: Predictions of the Model

To derive predictions for the influence of particular characteristics on the placement, I analyze critical values around which a formerly non-chosen placement is chosen (or a formerly chosen alternative is replaced). This is important since for some ranges of parameters even large changes in one parameter will not affect the placement. Take for example a very poor and a very rich household. The rich household is taking care of the orphan and is providing the orphan with more consumption than the total wealth of the poor household. When I vary the utility weight of the orphan in the poor household, nothing will change. The analysis below therefore assumes constellations of parameters where the placement can be affected by a change in the network members characteristics and/or the orphans contribution to that household. Conditional on this restriction, it is sufficient to analyze the change in the network utility from placing the orphan with a particular member in the neighborhood of the critical value for the parameter of interest.

The effect of w_1 on U_1 is ambiguous, depending on the other characteristics of network member 1, $r_{o1}, \widetilde{\delta}_1$. The ambiguity is documented first. Then it is shown that there exists a level of $w_1^+(r_{o1}, \widetilde{\delta}_1)$ s.th.$\forall\ w_1 > w_1^+$, $\partial U_1(w_1, r_{o1}, \widetilde{\delta}_1)/\partial w_1 > 0$.

$$\partial U_1(w_1, r_{o1}, \widetilde{\delta}_1)/\partial w_1 = \partial v_1(c_{11}^*)/\partial w_1 + \partial v_o(c_{o1}^*)/\partial w_1 - \partial v_1(c_{12})/\partial w_1$$

where $(c_{11}^*, c_{o1}^*) = \arg\max \quad v_1(c_{11}) + \widetilde{\delta}_1 v_o(c_{o1}) \quad \text{s.t.} \quad c_{11} + c_{o1} < w_1 + r_{o1}$

With $r_{oj} = 0, \quad \partial U_1(.)/\partial w_1 > 0$ as

$$\begin{aligned}
\frac{\partial v_1(c_{11})}{\partial c_{11}}\frac{\partial c_{11}}{\partial w_1} + \widetilde{\delta}_1 \frac{\partial v_o(w_1 - c_{11})}{\partial c_{o1}}\frac{\partial c_{o1}}{\partial w_1} &> \frac{\partial v_1(c_{12})}{\partial c_{12}}\frac{\partial c_{12}}{\partial w_1} \\
\frac{\partial v_1(c_{11})}{\partial c_{11}}\frac{\partial c_{11}}{\partial w_1} + \frac{\partial v_1(c_{11})}{\partial c_{11}}\frac{\partial c_{o1}}{\partial w_1} &> \frac{\partial v_1(c_{12})}{\partial c_{12}} \\
\frac{\partial v_1(c_{11})}{\partial c_{11}}\left(\frac{\partial c_{11}}{\partial w_1} + \frac{\partial c_{o1}}{\partial w_1}\right) &> \frac{\partial v_1(c_{12})}{\partial c_{12}} \\
\frac{\partial v_1(c_{11})}{\partial c_{11}} &> \frac{\partial v_1(c_{12})}{\partial c_{12}}
\end{aligned}$$

where the steps follow from the first order condition of the network members utility maximization problem and the fact that all available resources are consumed. The last inequality is true because $c_{12} = w_1$ and $c_{11} = w_1 - c_{o1}$. The intuition is very simple, the marginal utility from consumption is higher when consumption is lower. When network member shares her wealth between herself and the orphan, the marginal utility of extra wealth is higher.

In the case of $r_{oj} > 0$

$$\partial v_1(w_1 + r_{o1} - c_{o1})/\partial w <> \partial v_1(w_1)/\partial w_1$$

which depends on whether the consumption of the orphan is greater or less than his contribution, $c_{o1} <> r_{o1}$. When the orphan contributes more to the houshold than his share of consumption, then the marginal utility of consumption in that household is (at first) lower in the household with the network member alone. This can translate into the orphan being placed by the network with a network member to boost the consumption of that network member (and I can think of labor contribution or inheritance). For $c_{o1}(w_1, r_{o1}, \widetilde{\delta}_1) > r_{o1} : \partial U_1(.)/\partial w_1 > 0$

Define $w_1^+(r_{o1}, \widetilde{\delta}_1)$ as the level of c_{o1}, s.th. $c_{o1}(w_1, r_{o1}, \widetilde{\delta}_1) = r_{o1}$ then for $w_1 > w_1^+ \geq 0$: $\partial U_1(.)/\partial w_1 > 0$.

The network utility of placing the orphan with network member 2 is $U_2 = U_2(w_2, r_{o2}, \widetilde{\delta}_2) = \overline{U_2}$, which is unaffected by characteristics of network member 1, $(w_1, r_{o1}, \widetilde{\delta}_1)$.

It is possible that $U_1(w_1', r_{o1}, \widetilde{\delta}_1) > \overline{U_2} \ \ \forall \ \ w_1' \geq 0$. (This is the case where changing only one variable does not affect placement.) Otherwise, there exists a critical value of w_1, called $w_1^* \geq w_1^+$ where $U_1(w_1^*, r_{o1}, \widetilde{\delta}_1) = \overline{U_2}$ & $\partial U_1(w_1^*, r_{o1}, \widetilde{\delta}_1)/\partial w_1 > 0$. This results from the $\partial U_1(.)/\partial w_1 > 0$ (for $w_1^* \geq w_1^+$) and that U is unbounded,

Derivations with respect to r_{o1} and $\widetilde{\delta}_1$

For r_{oj} and $\widetilde{\delta}_1$the predictions are unambiguous, since they affect the network utility only when the orphan is placed with network member 1.

$$\begin{aligned}
\frac{\partial U_1(w_1, r_{o1}, \widetilde{\delta}_1)}{\partial r_{o1}} &= \frac{\partial v_1(c_{11})}{\partial c_{11}} \frac{\partial c_{11}}{\partial r_{o1}} + \widetilde{\delta}_1 \frac{\partial v_o(w_1 - c_{11})}{\partial c_{o1}} \frac{\partial c_{o1}}{\partial r_{o1}} - \frac{\partial v_1(c_{12})}{\partial c_{12}} \frac{\partial c_{12}}{\partial r_{o1}} \\
&= \frac{\partial v_1(c_{11})}{\partial c_{11}} \frac{\partial c_{11}}{\partial r_{o1}} + \widetilde{\delta}_1 \frac{\partial v_o(w_1 - c_{11})}{\partial c_{o1}} \frac{\partial c_{o1}}{\partial r_{o1}} - 0 \\
&> 0
\end{aligned}$$

$$\begin{aligned}
\frac{\partial U_1(w_1, r_{o1}, \widetilde{\delta}_1)}{\partial \widetilde{\delta}_1} &= \frac{\partial v_1(c_{11})}{\partial c_{11}} \frac{\partial c_{11}}{\partial \widetilde{\delta}_1} + \widetilde{\delta}_1 \frac{\partial v_o(w_1 - c_{11})}{\partial c_{o1}} \frac{\partial c_{o1}}{\partial \widetilde{\delta}_1} - \frac{\partial v_1(c_{12})}{\partial c_{12}} \frac{\partial c_{12}}{\partial \widetilde{\delta}_1} \\
&= \frac{\partial v_1(c_{11})}{\partial c_{11}} \frac{\partial c_{11}}{\partial \widetilde{\delta}_1} + \widetilde{\delta}_1 \frac{\partial v_o(w_1 - c_{11})}{\partial c_{o1}} \frac{\partial c_{o1}}{\partial \widetilde{\delta}_1} - 0 \\
&> 0
\end{aligned}$$

There is however a difference between how r_{o1} and $\widetilde{\delta}_1$ affect placement. It can easily be shown that there exists a critical value of r_{o1}, called $r^*_{o1} \geq 0$ such that $\forall \;\; r'_{o1} \geq r^*_{o1} \;\; U_1(w_1, r'_{o1}, \widetilde{\delta}_1) > \overline{U_2}$.

For $\widetilde{\delta}_1$ this is not the case. Intuitively, if a network member puts a lot of weight on the orphan and thus the share of the orphan is large, if the wealth of that network member is small, then giving the orphan an ever larger share of it will not necessarily make the orphans consumption larger than in the other household. Note that the parameter $\widetilde{\delta}_1 < \widetilde{\delta}_N$ because otherwise it could be that the network utility falls when $\widetilde{\delta}_1$ rises.

Appendix B: Signing the Bias

This section shows that the observable and unobservable characteristics of the network members will be correlated, conditional on caretaking, even if the variables are unconditionally uncorrelated.

Assume there are two network members, 1 and 2, who both have observable characteristics, w, as well as unobservable characteristics, δ. Assume further that these variables are all independently uniformly distributed, i.e. $(w_1, \delta_1, w_2, \delta_2) \sim iid\ U[0, 1]$. The uniform distribution is chosen because the conditional distributions and covariances can be calculated. Then

$E[w_j] = E[\delta_j] = \frac{1}{2}$

$cov(w_j, \delta_j) = 0$

Define $x_j = w_j + \delta_j, \quad a_j = w_k + \delta_k - \delta_j$

The first allocation is chosen, i.e. $j = 1$, if any of the following four equivalent formulations holds

$$\begin{aligned} & w_1 + \delta_1 > w_2 + \delta_2 \quad \Leftrightarrow \;\; x_1 > x_2 \\ \Leftrightarrow \quad & w_1 > w_2 + \delta_2 - \delta_1 \quad \Leftrightarrow \;\; w_1 > a_1 \end{aligned}$$

The conditional expectation of wealth, conditional on allocation 1 being chosen, is

$$E[w_1 | w_1 > w_2 + \delta_2 - \delta_1] \equiv E[w_1 | j = 1] = \int w_1 f(w_1 | j = 1)\; dw_j$$

and the conditional density of wealth is

$$f(w_1 | j = 1) = \frac{f(w_1, w_1 > a)}{\Pr(j = 1)} = \frac{f(w_1) F_a(w_1)}{\Pr(j = 1)}$$

$F_a(A)$, the cdf of $a_1 = w_2 + \delta_2 - \delta_1$ is defined over $A\epsilon[-1, 2]$; however since w_1 is only defined on $[0, 1]$, this is the relevant range for $F_a(A)$. When $A\epsilon[-1, 0]$ the cdf $F_a(A)$ is

$$F_a(A) = \int_{\delta_1=-A}^{1} \int_{w_2=0}^{A+\delta_1} \int_{\delta_2=0}^{A+\delta_1-w_2} d\delta_2 dw_2 d\delta_1 \quad = \quad \frac{1}{6}(1+A)^3$$

where I use the fact that $f(\delta_1) = f(\delta_2) = f(w_2) = 1$. Thus $F_a(0) = \frac{1}{6}$.

When $A\epsilon[0, 1]$,

$$\begin{aligned}
F_a(A) \quad = \quad & \int_{\delta_1=0}^{1-A} \int_{w_2=0}^{\delta_1} \int_{\delta_2=\delta_1-w_2}^{A+\delta_1-w_2} d\delta_2 dw_2 d\delta_1 \\
& + \int_{\delta_1=0}^{1-A} \int_{w_2=\delta_1}^{A+\delta_1} \int_{\delta_2=0}^{A+\delta_1-w_2} d\delta_2 dw_2 d\delta_1 \\
& + \int_{\delta_1=1-A}^{1} \int_{w_2=0}^{A+\delta_1-1} \int_{\delta_2=\delta_1-w_2}^{1} d\delta_2 dw_2 d\delta_1 \\
& + \int_{\delta_1=1-A}^{1} \int_{w_2=A+\delta_1-1}^{\delta_1} \int_{\delta_2=\delta_1-w_2}^{A+\delta_1-w_2} d\delta_2 dw_2 d\delta_1 \\
& + \int_{\delta_1=0}^{1-A} \int_{w_2=\delta_1}^{1} \int_{\delta_2=0}^{A+\delta_1-w_2} d\delta_2 dw_2 d\delta_1 \\
& + F_a(0) \\
= \quad & \frac{1}{2}(1-A)^2 A + \frac{1}{2}(1-A)A^2 + \frac{1}{3}A^3 + (A^2 - A^3) + \frac{1}{3}A^3 + F_a(0) \\
= \quad & \frac{1}{2}A + \frac{1}{2}A^2 - \frac{1}{3}A^3 \quad + \frac{1}{6}
\end{aligned}$$

The ex-ante probability that a particular allocation j will be chosen, $\Pr(j = 1)$, is 1/2 because all four variables are independent random variables from the same distribution, so that the probability that the sum of two is larger than the sum of the other two is 1/2 (this can easily be shown, as $\Pr(j = 1) = \int_0^1 f(w_1)F_a(w_1)\, dw_1$).

Then, the conditional density of wealth is

$$f(w_1|j = 1) = \frac{f(w_1)F_a(w_1)}{1/2} = w_1 + w_1^2 - \frac{2}{3}w_1^3 + \frac{1}{3}$$

The conditional expectation of wealth

$$\begin{aligned}
E[w_1|w_1 \quad > \quad & w_2 + \delta_2 - \delta_1] \\
= \quad & \int_0^1 w_1 f(w_1|j = 1) dw_j \\
= \quad & \int_0^1 w_1(w_1 + w_1^2 - \frac{2}{3}w_1^3 + \frac{1}{3}) dw_j \\
= \quad & \frac{37}{60} > \frac{1}{2} = E[w_1]
\end{aligned}$$

is larger than the unconditional expectation, proving that on average caretakers have better observable characteristics than non-caretaking network members. The proof for δ_1 follows the same equations, replacing w_1 with δ_1 and redefining a_1 appropriately.

The covariance between δ_1 and w_1, conditional on caretaker 1 being chosen, is:

$$\begin{aligned} Cov(\delta_1, w_1|j=1) &= \frac{1}{2}\left[Var\left(w_1+\delta_1|j=1\right) - Var\left(w_1|j=1\right) - Var\left(\delta_1|j=1\right)\right] \\ &= \frac{1}{2}\left[Var\left(w_1+\delta_1|j=1\right) - 2Var\left(w_1|j=1\right)\right] \end{aligned}$$

The second step follows from the fact that w_1 and δ_1 are drawn from the same distribution.

$$\begin{aligned} Var\left(w_1+\delta_1|j=1\right) &= Var\left(x_1|j=1\right) = \int \left(x_1 - E\left(x_1|j=1\right)\right) f\left(x_1|j=1\right) dx_1 \\ Var\left(w_1|j=1\right) &= Var\left(\delta_1|j=1\right) = \int \left(w_1 - E\left(w_1|j=1\right)\right) f\left(w_1|j=1\right) dw_1 \end{aligned}$$

The conditional expectation and distribution function of w_1 were derived above. The conditional expectation of x_1 is the sum of the conditional expectations of w_1 and δ_1.

$$E\left(x_1|j=1\right) = E\left(\delta_1|j=1\right) + E\left(w_1|j=1\right) = 2 * \frac{37}{60} = \frac{37}{30}$$

The density of $x_1 = w_1 +\ \delta_1$, conditional on being the caretaker, is

$$f\left(x_1|j=1\right) = \frac{f(x_1, x_1 > x_2)}{\Pr(x_1 > x_2)} = \begin{cases} x_1^3 & x_1 \epsilon [0,1] \\ 10x_1 - 6x_1^2 + x_1^3 - 4 & x_1 \epsilon [1,2] \end{cases}$$

because

$$\begin{aligned} f\left(x_1, x_1 > x_2\right) &= \int_0^{x_1} f(x_1, x_2)\ \ dx_2 \\ &= \begin{cases} \int_0^{x_1} x_1 x_2\ dx_2 & x_1 \epsilon [0,1] \\ \int_0^1 (2-x_1) x_2\ dx_2 + \int_1^{x_1} (2-x_1)(2-x_2)\ dx_2 & x_1 \epsilon [1,2] \end{cases} \\ &= \begin{cases} \frac{1}{2} x_1^3 & x_1 \epsilon [0,1] \\ 5x_1 - 3x_1^2 + \frac{1}{2} x_1^3 - 2 & x_1 \epsilon [1,2] \end{cases} \end{aligned}$$

and

$$\Pr(x_1 > x_2) = \frac{1}{2}$$

The conditional variance of x_1 is

$$
\begin{aligned}
Var\,(x_1|j=1) &= \int_0^1 \left(x_1 - \frac{37}{30}\right)^2 x_1^3 \; dx_1 + \int_1^2 \left(x_1 - \frac{37}{30}\right)^2 \left(10x_1 - 6x_1^2 + x_1^3 - 4\right) \; dx_1 \\
&= \frac{101}{900}
\end{aligned}
$$

and the conditional variance of w_1 is

$$
Var\,(w_1|j=1) = \int_0^1 \left(w_1 - \frac{37}{60}\right)^2 \left(w_1 + w_1^2 - \frac{2}{3}w_1^3 + \frac{1}{3}\right) = \frac{251}{3600}
$$

The covariance between the observable and unobservable characteristics is

$$
\begin{aligned}
Cov(\delta_1, w_1|j &= 1) = \frac{1}{2}\left[Var\,(w_1 + \delta_1|j=1) - 2Var\,(w_1|j=1)\right] \\
&= \frac{1}{2}\left[\frac{101}{900} - 2 * \frac{251}{3600}\right] \\
&= -\frac{49}{3600}
\end{aligned}
$$

which is smaller than zero. This negative conditional correlation between w_1 and δ_1 leads to a downward bias on the coefficient on w_1 when regressing orphan outcomes on caretaker characteristics.

References

Ainsworth, M., K. Beegle, and G. Koda (2005). The impact of adult mortality on primary school enrollment in Northwestern Tanzania. *Journal of Development Studies 41(3)*, 412–439.

Ainsworth, M. and D. Filmer (2002, September). Poverty, AIDS and children's schooling: A targeting dilemma. World Bank Policy Research Paper 2885.

Ainsworth, M. and I. Semali (2000). The impact of adult deaths on children's health in Northwestern Tanzania. Mimeo, World Bank.

Akresh, R. (2006). Risk, network quality and famly structure: Child fostering decisions in Burkina Faso. Yale Economic Growth Center Working Paper No. 902.

Beegle, K., J. De Weerdt, and S. Dercon (2006). Orphanhood and the long-run impact on children. mimeo, World Bank, EDI and Oxford University.

Ben-Akiva, M., D. Bolduc, and J. Walker (2001). Specification, estimation and identification of the logit kernel (or continuous mixed logit) model. Working Paper, Department of Civil Engineering, MIT.

Bicego, G., S. Rutstein, and K. Johnson (2003). Dimensions of the emerging orphan crisis in Sub-Saharan Africa. *Social Science and Medicine 56*, 1235–1247.

Bierlaire, M., D. Bolduc, and D. McFadden (2006). The estimation of generalized extreme value models from choice-based samples. Mimeo.

Case, A. and C. Ardington (2006). The impact of parental death on school enrollment and achievement: Longitudinal evidence from South Africa. *Demography forthcoming*. mimeo, Princeton University.

Case, A., C. Paxson, and J. Ableidinger (2004). Orphans in Africa: Parental death, poverty, and school enrollment. *Demography 41(3)*, 483–508.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economic Studies 47(1)*, 225–238.

Cosslett, S. (1981). Efficient estimation of discrete-choice models. In C. Manski and D. McFadden (Eds.), *Structural Analysis of Discrete Data and Econometric Applications*, pp. 51–111. Cambridge: The MIT Press.

Dercon, S. and P. Krishnan (2000). In sickness and in health: Risk sharing within households in Ethiopia. *Journal of Political Economy 108*(4), 688–727.

Evans, D. and E. Miguel (2005). Orphans and schooling in Africa: A longitudinal analysis. mimeo, University of California.

Foster, G. (2000). The capacity of the extended family safety net for orphans in Africa. *Psychology, Health and Medicine 5*(1), 55–62.

Fox, J. (2006). Semiparametric estimation of multinomiral discrete choice models using a subset of choices. Mimeo, University of Chicago.

Garen, J. (1984). The returns to schooling: A selectivity bias approach with a continuous choice variable. *Econometrica 52*, 1199–1218.

Lee, L.-F. (1983). Generalized econometric models with selectivity. *Econometrica 51 (2)*, 507–512.

Lundberg, M., M. Over, and P. Mujinja (2000). Sources of financial assistance for households suffering an adult death in Kagera, Tanzania. World Bank Policy Research Working Paper No. 2508.

Manji, A. (2000). 'Her Name is Kamundage': Understanding the impact of AIDS on women's relations to land amongst the Haya of Tanzania. *Africa 70*, 262–295.

Manski, C. and D. McFadden (1981). Alternative estimators and sample designs for discrete choice analysis. In C. Manski and D. McFadden (Eds.), *Structural Analysis of Discrete Data and Econometric Applications*, pp. 1–49. Cambridge: The MIT Press.

McFadden, D. (1978). Modelling the choice of residential location. In A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull (Eds.), *Spatial Interaction Theory and Planning Models*, pp. 105–142. Amsterdam: North-Holland.

McFadden, D. and K. Train (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics 15*, 447–470.

Measure, National AIDS Control Programme, Tanzania, and Bureau of Statistics, Tanzania (2001). AIDS in africa during the nineties: Tanzania. A review and analysis of surveys and research studies.

Muchunguzi, J. K. (2002). HIV/Aids and women's land ownership rights in Kagera region - northwestern Tanzania. Workshop in HIV/Aids and Land Tenure held in Pretoria, South Africa June 2002.

Tanzania Commission for AIDS (TACAIDS), National Bureau of Statistics (NBS), and ORC Macro (2005). *Tanzania HIV/AIDS Indicator Survey 2003-04*. Calverton, Maryland, USA: TACAIDS, NBS, and ORC Macro.

Train, K. (2000). Halton sequences for mixed logit. Working Paper No. E00-278, Department of Economics, University of California, Berkeley.

Train, K. (2003). *Discrete Choice Models with Simulation*. Cambridge: Cambridge University Press.

Trost, R. and L.-F. Lee (1984). Technical training and earnings: A polychotomous choice model with selectivity. *The Review of Economics and Statistics 66 (1)*, 151–156.

TZNBS (2003). *Kagera Region Socio-Economic Profile, 2nd Edition*. Dar es Salaam, Tanzania: Tanzania National Bureau of Statistics.

Udry, C. (1994). Risk and insurance in a rural credit market: An empirical investigation in Northern Nigeria. *Review of Economic Studies 61*(3), 495–526.

World Bank (2004a). User's guide to the Kagera Health and Development Survey datasets. mimeo.

World Bank (2004b). World Development Indicators. http://devdata.worldbank.org/data-query/.

World Bank (2005). Kagera Health and Development Survey 2004 basic information document. mimeo.

Table 1: Caregiving Arrangements of Double Orphans

	Number	Percent
No one in network took care	26	22
Brother or sister caretaker	27	22
Paternal grandparent	26	22
Paternal uncle or aunt	17	14
Maternal grandparent	10	8
Maternal uncle or aunt	14	12
Total	120	100

Notes: Table displays the degree of relationship between orphans and their caretakers in a sample of orphans in Kagera, Tanzania.

Source: Author's survey.

Table 2: Number of Orphaned Siblings in a Network

Number of orphans	Networks
1	26
2	19
3	10
4	3
Total	58

Notes: Table displays the number of orphaned siblings per network. Only orphaned children taken care of by a relative are included.

Table 3: Caregiving Arrangements of Double Orphans II

	Number	Percent
Brother or sister caretaker	27	28.7
Paternal grandparent	26	27.7
Paternal uncle or aunt	17	18.1
Maternal grandparent	10	10.6
Maternal uncle or aunt	14	14.9
Total	94	100

Notes: Table displays the degree of relationship between orphans and their caretakers in the sample of orphans used in the analysis. Orphans not taken care of by members of the family network as defined by the above categories are excluded.
Source: Author's survey.

Table 4: Differences in Characteristics of Caretakers and Non-Caretakers

	Non-caretakers	Caretakers	SE of Diff
Relationship to Orphan			
Paternal Grandfather	0.02	0.14	0.02***
Paternal Grandmother	0.01	0.12	0.02***
Maternal Grandfather	0.06	0.07	0.03
Maternal Grandmother	0.03	0.09	0.03**
Paternal Uncle or Aunt	0.30	0.19	0.06*
Maternal Uncle or Aunt	0.11	0.05	0.04
Other Spouse of Deceased Parent	0.29	0.09	0.06***
HH Assets			
Value of HH Land (excl. orphan land)[a,b]	15.39	21.22	4.31
Members HH owns Business Assets	0.08	0.24	0.04***
Demographics of Network Member			
Sex of Member	0.54	0.66	0.07*
Age of Member at Second Death	38.74	47.29	2.37***
Years of Education of Member[b]	6.68	7.06	0.41
Member is Head of HH	0.69	0.93	0.06***
Occupation of Network Member			
Member is Farmer	0.70	0.83	0.06**
Member is Trader	0.07	0.14	0.04*
N	357	58	

Notes: Table displays summary statistics for network members who did not take care of orphans (Column 1) and those who did (Column 2). Column 3 reports standard errors of the difference; *** denotes significance at 1% level, ** 5% level and * 10% level

a in 100000 Tanzania Shilling, approximately 100 USD

b Summary statistics for network members with non-missing information

Source: Author's survey

Table 5: Description of Variables used in the Empirical Analysis

Variable	Description
Dependent variable: this allocation was chosen	Members of the family network were asked who had been the main caretaker of the orphan right after the death of the second parent.
Relationship Variables	Using the family tree of the orphan provided information on the relationship of the orphan to the network members. The category "Other Spouse of Deceased Parent" refers to cases in which a parent had other spouses apart from the second parent of the orphan, e.g. in a polygamous and/or sequential monogamous setting. The left out category are the paternal aunts and uncles of the orphan, the traditional caretakers.
Sex of Member	Sex of the network member
Age of Member	Age of the network member at the death of the second parent of the orphan.
Years of Education	Years of completed education of the member
Head of Household	The network member is considered to be the head of the household he or she is living in.
Value of HH land	Value of land of the household the network was living in 2004 minus the value of land owned by any sibling of the orphan who at the time of death was below the age of 21 (to avoid biases through joint movements to other households with somewhat older sibling) and who was not the head of the household or his or her spouse. In 100000 Tanzania Shilling which corresponds to approximately 100 USD.
HH owns Business Assets	Dummy indicating whether the household owns any business assets, after subtracting the inheritance in business assets that the orphan(s) brought with them. This information is from the inheritance questionnaire which attempted to list all assets the deceased owned at their death.
Occupation Variables	Main occupation in last 14 days or main occupation in last 12 months. Left out category includes people who report having no job, a job as skilled worker, as a professional, as a fisher or in construction.
More than one orphan in HH	Dummy variable that indicates a particular allocation involves more than one orphan being taken care of by a specific network member.

Table 6: Determinants of the placement decision: conditional logit coefficients

Dependent Variable: Network member is caretaker	(1)	(2)	(3)	(4)	(5)
Paternal Grandfather	2.447 ***	2.740 ***	1.997 ***	4.145 ***	3.529 ***
	(0.471)	(0.516)	(1.015)	(1.161)	(0.980)
Paternal Grandmother	2.445 ***	2.696 ***	2.290 ***	3.141 ***	2.577 ***
	(0.666)	(0.648)	(1.169)	(1.190)	(0.877)
Maternal Grandfather	0.331	0.508	0.187	0.411	0.205
	(0.704)	(0.712)	(1.205)	(1.328)	(1.064)
Maternal Grandmother	1.112	1.257 *	1.463	1.915	1.517
	(0.746)	(0.724)	(1.265)	(1.283)	(0.936)
Maternal Uncle or Aunt	-0.739	-1.024	-0.833	-0.947	-0.778
	(0.687)	(0.774)	(0.940)	(0.869)	(0.772)
Sibling	1.459 **	1.364 **	2.114 **	1.358	1.164 *
	(0.697)	(0.675)	(1.049)	(0.860)	(0.700)
Other Spouse of Deceased Parent	-0.087	0.094	0.477	0.762	0.769
	(0.676)	(0.667)	(0.735)	(0.750)	(0.632)
Sex of Member			0.322	0.418	0.360
			(0.815)	(0.741)	(0.652)
Age of Member at Second Death			0.127	0.098	0.077
			(0.075)	(0.064)	(0.054)
Age of Member squared			-0.001	-0.001 *	-0.001
			(0.001)	(0.001)	(0.001)
Years of Education of Member			0.064	0.027	0.027
			(0.090)	(0.096)	(0.081)
Member is Head of HH			1.670 **	2.055 **	1.742 **
			(0.874)	(0.969)	(0.866)
Value of HH Land (excl. orphan land)				0.013 *	0.011 *
				(0.007)	(0.006)
Members HH owns Business Assets				1.364 **	1.117 **
				(0.655)	(0.564)
Member is Farmer				1.579 **	1.347 **
				(0.678)	(0.638)
Member is Trader				1.795 **	1.477 *
				(0.906)	(0.823)
Value of Land * Orphan under 5				0.015	0.012
				(0.012)	(0.012)
Memb. owns Bus. * Orphan under 5				3.790 **	3.255 **
				(1.626)	(1.486)
Member Farmer * Orphan under 5				0.534	0.575
				(1.013)	(0.922)
More than one orphan in HH					2.402 ***
					(0.738)
N	10643	10643	10643	10643	10643

a Coefficients from a conditional logit model estimating the determinants of whether an allocation of orphans was chosen.
b Robust standard errors in parenthesis, *** significant at 1% level,;** at 5% and * at 10%.
c Column 1 reports standard conditional logit results. All other columns correct for sampling from the network.
d Excluded category for relationship variables is paternal aunts and uncles; excluded category for occupation is a category comprising fishermen, construction workers, skilled workers, professionals, no job, other.
e Dummys for missing information on land and education included. Neither is significant.

Table 7: Determinants of the placement decision: mixed logit coefficients

Dependent Variable: Network member is caretaker	Coefficients on parameter	Variance Parameter
Paternal Grandfather	2.783 **	0
	(1.152)	(3.772)
Paternal Grandmother	0.799	3.147
	(1.993)	(3.687)
Maternal Grandfather	-0.031	0.31
	(0.242)	(6.982)
Maternal Grandmother	1.192	2.866 **
	(0.936)	(1.434)
Maternal Uncle or Aunt	-0.372	
	(0.699)	
Sibling	1.232 *	1.028
	(0.685)	(1.245)
Other Spouse of Deceased Parent	0.746	0.002
	(0.641)	(1.379)
Sex of Member	0.286	
	(0.632)	
Age of Member at Second Death	7.51	
	(5.067)	
Age of Member squared	-7.836	
	(5.236)	
Years of Education of Member	0.649	
	(0.801)	
Member is Head of HH	1.855 **	
	(0.847)	
Value of HH Land (excl. orphan land)	1.029	
	(0.692)	
Members HH owns Business Assets	0.999 *	
	(0.591)	
Member is Farmer	1.214 **	
	(0.601)	
Member is Trader	1.636 **	
	(0.793)	
Value of Land * Orphan under 5	0.114	
	(0.108)	
Memb. owns Bus. * Orphan under 5	0.212	
	(0.129)	
Member Farmer * Orphan under 5	0.022	
	(0.108)	
More than one orphan in HH	0.23 ***	
	(0.073)	
N	10643	

a Coefficients and variance parameters from a mixed logit model estimating the determinants of whether an allocation of orphans was chosen. For the variance parameter, the left out category was maternal uncles and aunts.
b Standard errors in parenthesis, *** significant at 1% level,;** at 5% and * at 10%.
c Excluded category for relationship variables is paternal aunts and uncles; excluded category for occupation is fishermen, construction workers, skilled workers, professionals, no job, other.
d Dummys for missing information on land and education included. Neither is significant

Table 8: Determinants of the placement decision (using one orphan per caretaker)

Dependent Variable: Network member is caretaker	(1)	(2)	(3)	(4)
Paternal Grandfather	2.375 ***	2.865 ***	2.010 **	3.235 ***
	(0.560)	(0.564)	(0.918)	(0.948)
Paternal Grandmother	2.391 ***	2.883 ***	2.387 **	2.792 **
	(0.611)	(0.610)	(1.111)	(1.064)
Maternal Grandfather	0.566	1.005	0.428	0.471
	(0.643)	(0.661)	(0.972)	(1.092)
Maternal Grandmother	1.606 **	1.994 ***	1.887	2.061 *
	(0.659)	(0.671)	(1.278)	(1.207)
Maternal Uncle or Aunt	-1.173 *	-1.412 *	-1.255	-1.361 *
	(0.663)	(0.751)	(0.933)	(0.797)
Sibling	0.952	0.889	1.739 *	0.827
	(0.604)	(0.594)	(0.895)	(0.767)
Other Spouse of Deceased Parent	-0.223	0.218	0.449	0.745
	(0.646)	(0.671)	(0.752)	(0.797)
Sex of Member			0.289	0.140
			(0.716)	(0.600)
Age of Member at Second Death			0.115	0.057
			(0.071)	(0.065)
Age of Member squared			-0.001	-0.001
			(0.001)	(0.001)
Years of Education of Member			0.044	0.032
			(0.079)	(0.089)
Member is Head of HH			0.994	1.329
			(0.882)	(0.844)
Value of HH Land (excl. orphan land)				0.010 *
				(0.006)
Members HH owns Business Assets				1.426 **
				(0.600)
Member is Farmer				1.370 **
				(0.646)
Member is Trader				1.671 **
				(0.670)
N	415	415	415	415

a Estimates from a conditional logit model estimating the determinants of whether a network member was chosen as caretaker. Only one orphan per caretaker is taken into account, observations on additional orphans were excluded.
b Robust standard errors in parenthesis, *** significant at 1% level,;** at 5% and * at 10%.
c Column 1 reports standard conditional logit results. All other columns correct for sampling from the network.
d Excluded category for relationship variables is paternal aunts and uncles; excluded category for occupation is fishermen, construction workers, skilled workers, professionals, no job, other.
e Dummys for missing information on land and education included. Neither is significant.

Table 9: Impact of caretaker characteristics on orphans' education

Dependent Variable: Change in the Years of Education between KHDS -1 and KHDS-2	(1)	(2)
Age in KHDS 1	1.095 ***	1.043 ***
	(0.288)	(0.269)
Age in KHDS 1 squared	-0.073 ***	-0.068 ***
	(0.024)	(0.022)
Orphans Sex	0.587	0.660
	(0.660)	(0.615)
Paternal Grandfather is caregiver	0.415	2.516 *
	(1.106)	(1.285)
Paternal Grandmother is caregiver	-2.555 *	0.075
	(1.512)	(1.705)
Maternal Grandfather is caregiver	-0.374	-1.126
	(1.887)	(1.780)
Maternal Grandmother is caregiver	-0.574	0.777
	(1.631)	(1.598)
Maternal Uncle or Aunt is caregiver	0.630	-1.105
	(1.315)	(1.380)
Sibling is caregiver	-1.460	-1.941 *
	(1.055)	(0.999)
Other Spouse of Deceased Parent is caregiver	0.991	0.461
	(2.447)	(2.289)
Members HH owns Business Assets	0.991	1.723 **
	(0.820)	(0.810)
Sex of Member	-0.384	0.105
	(1.002)	(0.951)
Lee selection correction term		2.362 ***
		(0.863)
Constant	2.310	-1.254
	(1.487)	(1.902)
Observations	56	56
R-squared	0.43	0.51

Notes: Standard deviations are in parenthesis. * denotes significance at the 10% level, ** at the 5% level, * at the 1% level.
Column 1 reports coefficients from an OLS regression of the change in orphans' education between KHDS-1 and KHDS-2 on characteristics of the caretakers. Column 2 reports coefficients from an OLS regression where a selection correction term controls for the placement of the orphan. Note that the standard errors are not corrected for the estimated selection correction term.

Table 10: Impact of caretaker characteristics on years of education

Dependent Variable: Years of Education of orphan children in KHDS-2		
Years of Education in KHDS1	-0.506	-0.335
	(0.606)	(0.584)
Enrolled in KHDS1	-0.321	0.452
	(1.561)	(1.535)
Age in KHDS 1	0.545 *	0.614 *
	(0.315)	(0.303)
Age in KHDS 1 squared	-0.001	-0.013
	(0.036)	(0.035)
Orphans Sex	-0.282	0.156
	(0.686)	(0.688)
Paternal Grandfather is caregiver	0.777	2.143
	(1.173)	(1.298)
Paternal Grandmother is caregiver	-2.170	0.210
	(1.700)	(1.986)
Maternal Grandfather is caregiver	-0.068	-0.942
	(2.033)	(1.985)
Maternal Grandmother is caregiver	1.715	1.952
	(1.967)	(1.881)
Maternal Uncle or Aunt is caregiver	1.889	-0.056
	(1.423)	(1.649)
Sibling is caregiver	-0.267	-1.066
	(1.053)	(1.076)
Other Spouse of Deceased Parent is caregiver	0.568	0.456
	(2.476)	(2.364)
Members HH owns Business Assets	1.762 *	2.273 **
	(0.924)	(0.915)
Sex of Member	0.042	0.231
	(1.295)	(1.239)
Lee selection correction term		2.041 **
		(0.982)
Age of KHDS1-hh-head	-0.048	-0.006
	(0.035)	(0.039)
Sex of KHDS1-hh-head	0.469	0.033
	(1.280)	(1.240)
Years of Educ of KHDS1-hh-head	0.161	0.139
	(0.169)	(0.161)
Residing w/ father in KHDS1	-1.124	-1.161
	(1.038)	(0.991)
Residing w/ mother in KHDS1	-0.511	-0.159
	(1.109)	(1.072)
Log consumption in KHDS1	-0.807	-0.420
	(0.760)	(0.749)
Dwelling in KHDS1	1.882	1.659
	(1.114)	(1.068)
Constant	13.454	4.193
	(9.083)	(9.747)
Observations	56	56
R-squared	0.66	0.70

Notes: Standard deviations are in parenthesis. * denotes significance at the 10% level, ** at the 5% level, * at the 1% level.

Column 1 reports coefficients from an OLS regression of the years of completed education of orphans in KHDS-2 on characteristics of the caretakers and characteristics of the original households. Column 2 reports coefficients from an OLS regression where a selection correction term controls for the placement of the orphan. Note that the standard errors are not corrected for the estimated selection correction term.