Demographic Research a free, expedited, online journal
of peer-reviewed research and commentary
in the population sciences published by the
Max Planck Institute for Demographic Research
Konrad-Zuse Str. 1, D-18057 Rostock · GERMANY
www.demographic-research.org

DEMOGRAPHIC RESEARCH

VOLUME 19, ARTICLE 32, PAGES 1217-1248
PUBLISHED 15 JULY 2008
http://www.demographic-research.org/Volumes/Vol19/32/
DOI: 10.4054/DemRes.2008.19.32

Research Article

Hotspots and Coldspots: Household and village-level variation in orphanhood prevalence in rural Malawi

Alexander Weinreb

Patrick Gerland

Peter Fleming

This publication is part of the proposed Special Collection "HIV/AIDS in sub-Saharan Africa", edited by Susan Watkins, Jere Behrman, Hans-Peter Kohler, and Simona Bignami-Van Assche.

Table of Contents

1	Introduction	1218
2	Setting and data	1219
3	Analysis	1220
3.1	Expectations	1222
3.2	Spatial distribution	1224
3.3	Characteristics of Hotspots and Coldspots	1230
3.3.1	Models	1232
3.3.2	Main results	1233
3.3.3	More spatial results	1237
3.3.4	Complex variation	1239
4	Discussion and conclusion	1240
5	Acknowledgements	1243
	References	1244

Demographic Research: Volume 19, Article 32
research article

Hotspots and Coldspots: Household and village-level variation in orphanhood prevalence in rural Malawi

Alexander Weinreb[1]

Patrick Gerland[2]

Peter Fleming[3]

Abstract

We explore the characteristics of households and villages in which orphans are resident in two areas of Malawi. We first review pertinent themes in qualitative data collected in our research sites. Then, using spatial analysis, we show how positive and negative clusters of orphans – which we term orphanhood "hotspots" and "coldspots" – can be found at the village and sub-village levels. In the third and longest section of the paper, and using multilevel analyses with both simple and complex variance structures, we evaluate the relationship between the presence of orphans and a range of individual, household and village-level characteristics, including households' spatial relationship to each other and to other local sites of significance. This series of analyses shows that the most important covariates of orphan presence are household size, wealth, and religious characteristics, with all measured simultaneously at both household and village-level. In addition, most of these have heterogenous effects across villages. We conclude by reviewing some difficulties in explaining causal mechanisms underlying these observed relationships, and discuss conceptual, theoretical and programmatic implications.

[1] Department of Sociology and Anthropology, Hebrew University. Jerusalem 91905, Israel.
E-mail: awein@mscc.huji.ac.il.
[2] United Nations Population Division. Estimates and Projection Section. Room DC2-1906 - 2 UN Plaza. New York, NY 10017, USA. E-mail: gerland@un.org.
[3] University of Pennsylvania. Population Studies Center. 3718 Locust Walk. Philadelphia, PA 19104, USA.
E-mail: pfleming@pop.upenn.edu.

1. Introduction

Since the emergence of the AIDS pandemic, the fate of children of HIV+ adults has been on the academic and public policy radar, triggering research across a number of disciplines. Within demography and related fields, this research has predominantly been one of three types. The first has been concerned with estimating the scale of the orphanhood phenomenon (e.g., UNAIDS et al. 2002; Grassly et al. 2004). The second has focused on comparative outcomes of "AIDS orphans" and children whose parents are HIV-free, focusing in particular on morbidity, mortality, and schooling (e.g., Ainsworth and Semali 2000, Ainsworth and Filmer 2002, Bicego et al. 2003, Crampin et al. 2003, Case et al. 2004, Monasch and Boerma 2004, Mishra et al 2005, Sarker et al. 2005, Beegle et al. 2005, Ford and Hosegood 2005, Sharma 2005, Zaba et al. 2005, Andrews, Skinner and Zuma 2006). The third type of literature has looked more at the socio-cultural context within which AIDS-orphans' outcomes, like those of other types of orphans, are determined. In sub-Saharan Africa, this refers primarily to intrafamilial or communal systems of child-fostering (Seeley et al. 1993, Foster et al. 1997, Madhavan 2004, Nyamukapa and Gregson 2005).

Recently, this literature has also begun to benefit from the contribution of the geographical sciences. In particular, taking advantage of developments in Geographical Information System (GIS) capabilities, researchers have been laying the groundwork for mapping child-related outcomes, including mortality and orphanhood, whether as an end in itself, or as a means to more effectively direct health and educational infrastructure or other types of services (e.g., Balk et al. 2003, Abebe 2005, Erskine and Wilson 2005). In this paper we build on that emerging literature. We use a variety of data sources from two rural areas in Malawi in order to accomplish two specific aims. First, we describe the prevalence of orphanhood across the research area and, using geocoded household survey data, describe the spatial distribution of orphans at the household level in order to identify high- and low-prevalence orphanhood clusters. Second, in order to identify the characteristics of high- and low-orphan prevalence areas, we explore the association between orphan prevalence, changes in orphan prevalence, and several types of village- and household-level characteristics. Both aims, we suggest, address a missing empirical area in orphan-related research that, in turn, make it difficult to fully describe types of orphan trajectories or, more generally, to address the medium- and long-term consequences of orphanhood.

2. Setting and data

Malawi is an appropriate setting for this study for a number of reasons. First, it is broadly representative of high HIV prevalence countries in the southern and eastern regions of sub-Saharan Africa (SSA) on two key dimensions. (*i*) It has a mature, high-prevalence epidemic – 14% of adults are HIV positive – which, as of 2001, had led to a 3-fold increase in adult mortality over levels observed during the 1980s intercensal period (Doctor and Weinreb 2005), about two-thirds of which was symptomatically associated with AIDS (Doctor and Weinreb 2003). (*ii*) According to 2003 estimates, this level of mortality had generated about 500,000 orphaned children aged 0-17 nationwide (UNAIDS 2004), representing about 16% of all children in those age-groups, and 4% of the total national population (Ainsworth and Filmer 2002, PRB 2005).

Second, unusually rich data on rural Malawi – home to 85% of the national population – are available from the Malawi Diffusion and Ideational Change Project (MDICP). Since 1998, the MDICP has collected an array of data from approximately 2,500 married adults living in 1,500 rural households in three rural districts – Rumphi, Mchinji and Balaka. These are, respectively, in Malawi's Northern, Central and Southern regions (see http://malawi.pop.upenn.edu for specific sampling details).

Our analysis focuses on the MDICP's Northern and Southern region research areas. These are the most structurally dissimilar of our research sites, allowing for a useful comparison of orphan distribution patterns. Specifically, the northern areas are the most geographically isolated from the political and commercial heart of the country. They are dominated by the Tumbuka ethnic group, historically patrilineal and patrilocal, almost wholly Christian, and said to be disproportionately represented among Malawi's professional classes. In contrast, the southern areas are close to Malawi's commercial capital, and to commercial networks linking Malawi with Zimbabwe and South Africa. There is also a very different ethno-religious mosaic in the south. The dominant ethnic group in our research sites is the Yao, a predominantly Muslim group that has historically privileged descent through matrilineages and favored matrilocal postmarital residence. On the other hand, because the southern sites also include some non-Muslim (and non-Yao) villages, there is some interpenetration of region with religion, allowing us to differentiate any Christian-Muslim differences – in addition to differences between Christian denominations – from those of region.

We use a few different types and sources of data in our analysis. All data on orphanhood and household characteristics were collected in 2004. The orphanhood data were collected in a household roster administered before the main survey to *all households* in sampled villages. This included 1,163 households (41 villages) in Rumphi and 2,679 households (17 villages) in Balaka.

The main MDICP survey instrument was then administered to roughly one quarter of the women of reproductive age living in these households. These are the primary source of data on household characteristics. In addition, 90 percent of these women consented to an HIV test.

Other data include the spatial position ("geocode") of every household in all sampled villages – collected at the same time as the household rosters – and the geocode of all local infrastructure (markets, shops, wells, boreholes, roads, and so on) and primary institutions (schools, churches and mosques, village leader's home, and so on). These were collected using handheld Global Position System (GPS) units.

Sample sizes somewhat vary across analyses. Spatial analyses use both the complete village samples (N=~3,850 households) from the household roster and data from 582 households in the southern and northern sites in which at least one adult woman was interviewed. Naturally, all regression analyses are also restricted to those same 582 MDICP households. Note that this is considerably less than the theoretical maximum of 1025 – the full sample of from these two sites – since we had surprising difficulty in matching the household rosters and survey data in the field. It is also important to note that these 582 households had marginally lower – but not statistically significant – orphan prevalence than the other households in MDICP villages. Consequently, we consider the reduced sample size unfortunate in terms of its effect on statistical power, but since the issue of matching household rosters from the village to survey data is unlikely to be selective on any orphan-related characteristic, we expect it to have minimal effect on the overall validity of the results, at least those related to identifying the characteristics of households in which orphans are found.[4]

3. Analysis

According to the household rosters, and as shown in Table 1, 14.2% of children aged less than fifteen from our northern and southern research areas had lost at least one

[4] We attempted to match as many survey and household roster records as possible using a number of methods. The primary one, conducted in MS-Access 2003, was a "Fuzzy Matching" technique based on four similarity algorithms: Dice coefficient, Levenshtein Edit Distance, Longest Common Subsequence, and Double Metaphone (Brown 2004; OpnSeason 2006). Specifically, at the village level we used a name matching procedure which involved comparing last name, first name and any aliases collected in the 1998, 2001 and 2004 waves of the MDICP survey with those collected in the 2004 village household rosters. Homonym cases were manually resolved by cross-checking household and individual characteristics and/or geographic coordinates. We then cleaned a relatively small number of unresolved cases in ArcView 3.3 and Stata 9.2. Overall, the 903 cases used in full analyses here have an estimated matching coefficient of 81%, with all cases falling in the 60-100% range. All final analyses reported below were also run using the matching coefficient as an analytic weight. There were *no substantive differences whatsoever* between these weighted models and those presented below.

parent. Of these, 8.5 percent were paternal orphans, 1.5 percent were maternal orphans, and 4.2 percent of those were double orphans. Across all three types, orphanhood was marginally higher in the south than in the north. Both these absolute levels, and the marginally higher southern region levels, are consistent with orphan estimates from the 2004 Integrated Household Survey (Republic of Malawi 2005).

Table 1: Village-level orphan prevalence, ages 0-14, 2004

Type of orphanhood	All	Regression Sample
Paternal orphans	8.5	7.7
Maternal orphans	1.5	0.5
Double orphans	4.2	4.0
All	14.2	12.2

These orphans were located in 12.5 percent of MDICP households. Our analysis has two specific aims related to these residential patterns. The first is to map the distribution of these orphaned children across our sites. In particular, we want to see whether, and to what extent, orphanhood is spatially clustered. The second is to identify individual, household and village-level characteristics which covary with that distribution. In addition, we seek to evaluate whether there is any heterogeneity across villages in those characteristics.

It should be noted that beyond their conceptual and programmatic implications, the answers to these questions are necessary first steps toward understanding more causal patterns related to orphan placement. Although we do not directly grapple with those causal patterns here – our data do not allow us to identify whether or not orphans are in their original home, how long they have been in the current location, and so on – we frame our analysis of orphan's location in soft causal tones. This is justified on the basis of our qualitative data. In particular, in our Malawian settings it is clear that double orphans move in almost all cases (only the oldest remain in their natal home); that maternal orphans almost always move (since legitimate carers are women), and that paternal orphans often move, though less immediately (since a widowed woman often remarries and will typically take only the youngest of her children to her new home, leaving the older ones with an older relative). In short, no matter what type of orphanhood, in the Malawian setting it is more than likely to lead to some type of physical move.

3.1 Expectations

Our *a priori* expectation was that we would find some spatial clustering. That is, rather than finding orphans randomly dispersed across a given sample of households, we would find areas with greater-than-expected concentrations of orphans: we refer to these as "hotspots." On the flipside, we also expected to find areas with less-than-expected concentrations: we refer to these as "coldspots." Over and above the relatively mechanistic, network-related characteristics of HIV infection – which we imagined independently affecting clustering of AIDS orphans – our expectation drew on two main sources.

The first was the extant Africa-focused literature on orphans (Barnett and Whiteside 2002, Guest 2003), extended families (Goode 1963, Adamchak et al. 1991, Weinreb 2002, 2006), and AIDS in general (Gregson et al. 1999, Garner 2000). One of the meta-messages uniting these literatures is that the strength of extended family norms varies from place to place. Along with that variation, it is reasonable to surmise, so would the willingness to open one's home to the children of a deceased family member, especially when one candidate could argue that another family member was better suited to the task.

The second source for our orphan-cluster expectations drew on qualitative data collected in our research sites in 2005. By qualitative data we primarily refer to open-ended conversations conducted by an experienced female interviewer in her 30s who had been long associated with the project. She sought out stories of local people who had died and left children, of others who had fostered children from within or outside the village, or of others, yet, who had refused to foster them – this turned out to be an extremely rare event (one known case across the three sites, described below). Based on these stories, we became much more acquainted with the process by which children are fostered and in particular with the negotiations – usually within-family and involving the orphans' older siblings, aunts, uncles, and grandparents, but sometimes also involving key local leaders like headmen and pastors – that often accompanied final decisions. Thus, it appeared that although, on some ideal level, orphans would be allocated by the extended family to a household that was most able and willing to take care of them, it was also clear that ability and willingness were not non-negotiable criteria in and of themselves. They could, for example, be variably portrayed as a combination of qualities including the potential hosting household's size, wealth, and physical placement, as well as characteristics of key individuals in the household, including their emotional attributes and gender. Consequently, negotiations about where to place orphans involved not only decisions about housing, but also about a much more general division of responsibilities with regard to other types of support in which all able-bodied members of the family would be expected to contribute something, whether monetary or in-kind. Similarly, negotiations also involved debates about how to balance

the acknowledged ideal that orphans be fostered with their siblings with the burden that this sudden influx would most likely cause.

These points are pertinent to the issue of orphan clusters – and also to the issue of heterogeneity in the effect of predictors across villages – since one of the key themes which emerged in the interviews is the extent to which sociocultural and, more recently, material mechanisms appeared to engender greater willingness to foster orphans. On the sociocultural front, for example, several informants described how some village and religious leaders would publicly laud those who had fostered. In one site, a couple of informants also recounted cautionary tales about someone who had refused to foster children of deceased relatives – the only case known across the three sites. Later, when that person was herself in need, leaders and others were said to have refused her assistance. Such reports, in short, seemed to imply that orphanhood clusters would emerge in settings where village or religious leaders were more actively engaged in promoting the fostering of orphans as the right thing to do, whether the justification drew on religious motifs or on a perceived tradition of care within extended families. In either case, the greater public recognition of fostering families in some settings than in others excited our sociological imagination since it raised the possibility that, in certain areas, there might be competition for foster children – since public recognition would turn foster children into a means of acquiring status. Or perhaps more realistically (and moderately), refusing to foster an orphan would be more costly in some settings than others.

Of course this desire for status, or the desire to avoid a loss of status, was not independent of certain financial factors, since fostering children entailed monetary expenses and other outlays. But here, too, we heard how material mechanisms might be increasingly creating the conditions for the creation of spatial clusters of orphan. In particular, the increasing visibility of special aid for AIDS orphans and "orphan-care projects" in our research sites and throughout Malawi – many established locally and funded through the Malawi Social Action Fund (MASAF) – led to a number of complaints by non-fosterers that orphans were receiving food assistance and special training that could more equitably be directed at impoverished children of all types: that is, those who had lost a parent, and those yet to lose one. Like other development aid (Walters et al. 1999; Mansuri and Rao 2004; Gould 2005), we wondered if an orphan cluster might arise where a particularly energetic or well-connected village headman would be able to establish one of these projects directly, be party to its establishment, or able to persuade its directors to include his village on the list of donor destinations. Doing any of these would likely make his village a more attractive destination for orphan-related decision-makers.

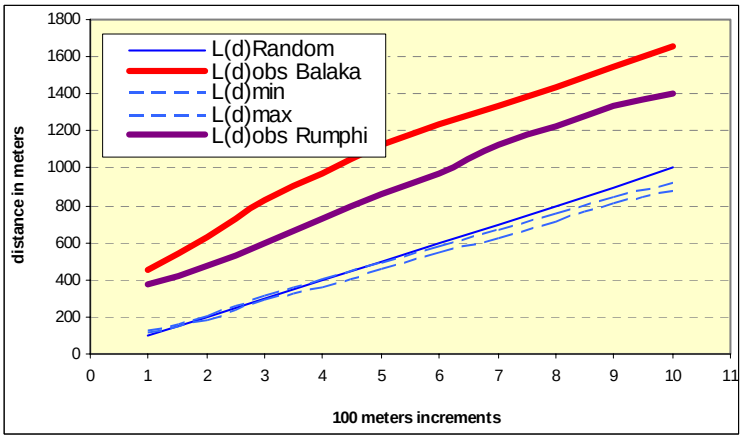
3.2 Spatial distribution

We used the geocoded household rosters in order to explore orphanhood prevalence in 2004 spatially. Specifically, using the Arcview Spatial Analyst (ESRI 2002) and SA extension (Scott 2001), we calculated two types of spatial statistics, each of which allowed us to evaluate the degree to which orphans cluster. Note that each of these spatial statistics was estimated in relation to both the absolute number of children in any given household who are orphans, and the proportion of children in the household that are orphans. That said, we consider the absolute number a more appropriate measure for two reasons. First, although it does not control for the overall density of children in a given household, it is likely to produce less biased estimates (the proportionate measure biases the estimates toward 100 percent where orphans are living with grandparents or other older kin whose own children live elsewhere). Second, the absolute number is numerically closer to our dependent variable in the regression analyses (section 3.3).

The first statistic we estimated was the Weighted K-function (Getis and Ord 1992, Aldstadt et al. 1998). Our aim here was to test for the spatial clustering of households with respect to the level of orphanhood within and between villages. This allowed us to determine whether the observed distribution of orphans across mapped households matches what the distribution would be if they were randomly distributed across those same households (HHs)? The L(d) Random values in Figure 1 – and the minimum and maximum lines representing the expected lower and upper bounds – represent the expected spatial pattern for randomly distributed households. Both the observed northern and southern distributions fall above this line, indicating significant spatial clustering between households with orphans (especially among households within a distance of 800 meters or more from the reference household). Note that similar results are observed when we estimate the Weighted K-function for proportion of children in the household who are orphans.

The second spatial statistic we calculated were Gi* statistics (Ord and Getis 1995, Aldstadt et al. 1998). Here our key aim was to test for statistically significant local spatial clustering of orphans in 2004 – the "hotspots" referred to above. We first computed the Gi* scores for the villages based on the absolute number of orphans per household, using a starting distance of 200 meter radius from a reference household and incrementally increasing that distance by 200 meters. In Figures 2 and 3 we present the results for a distance of 1200 meters between households, but the results are similar up to a distance of 1800 meters. Figure 2 presents results for southern district and Figure 3 for northern district. Each figure contains two panels. Panel (a) maps Gi* statistics for the full sample of ~3850 households, and panel (b) maps Gi* statistics for the MDICP household sample of 582 households. In each case, as indicated in the map legend, the color-coded Gi* statistics can be interpreted as z-scores.

Figure 1: K-function analysis (meters) with 100 meters increments



Panel (a) in both figures shows that the distribution of households with orphans in a small number of villages (around village numbers 3, 18 and 12 in the south, and 113 in the north) far exceeds what might exist completely at random (Gi*>1.96). There is also some indication of orphan clustering around villages 107 and 108 in the north (Gi* = 1.65 – 1.96). These are orphanhood hotspots. Equally interesting, panel (a) shows that there are areas within both research sites, particularly the south, in which orphan prevalence is much less than one might expect at random (i.e., villages 4, 7 , and 14 in the south and 117 in the north). These are orphanhood coldspots. In general, it appears that there is greater positive and negative clustering of orphans in the south than the north.

An initial glance at panels (a) across both Figures suggests that there is no obvious structural pattern to these orphanhood hotspots and coldspots. For example, in the southern area, one of the two primary orphan hotspots (around village 3) is close to a main (paved) road while the other, around village 18, is much further. Each is roughly equidistant from the area's main trading center (around village 9). In the north, the single hotspot is several kilometers from the main paved road, but it very close to both a non-paved but graded road, and also to the area's main trading center. We revisit these – and other – spatial patterns more systematically in the regression analyses.

Figure 2: Orphan clusters in MDICP villages, Southern Region (Balaka District), 2004:

(a) all households in the villages;

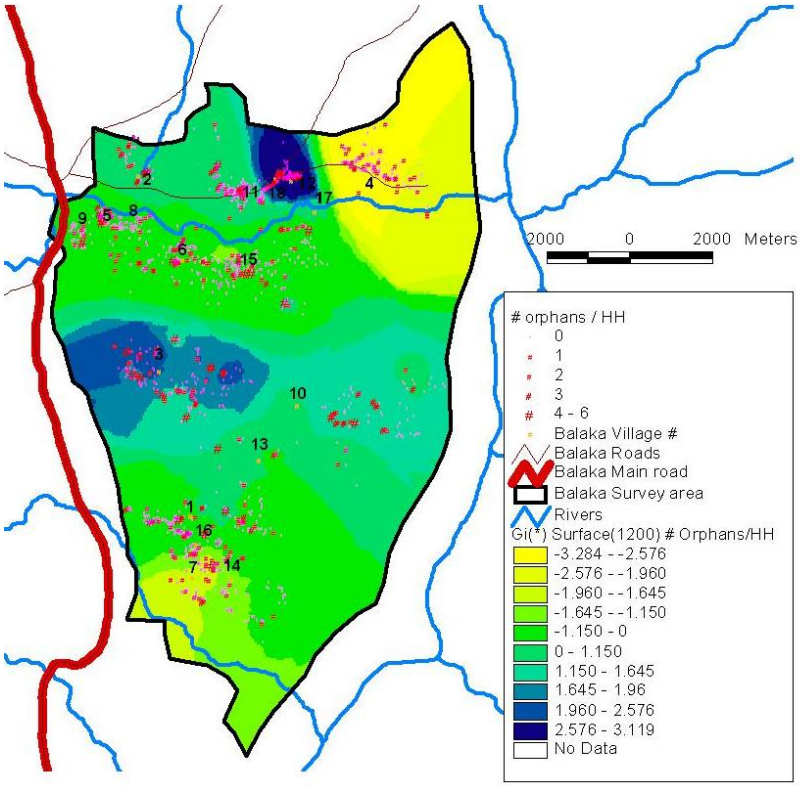


Figure 2: (continued)

(b) MDICP sample only

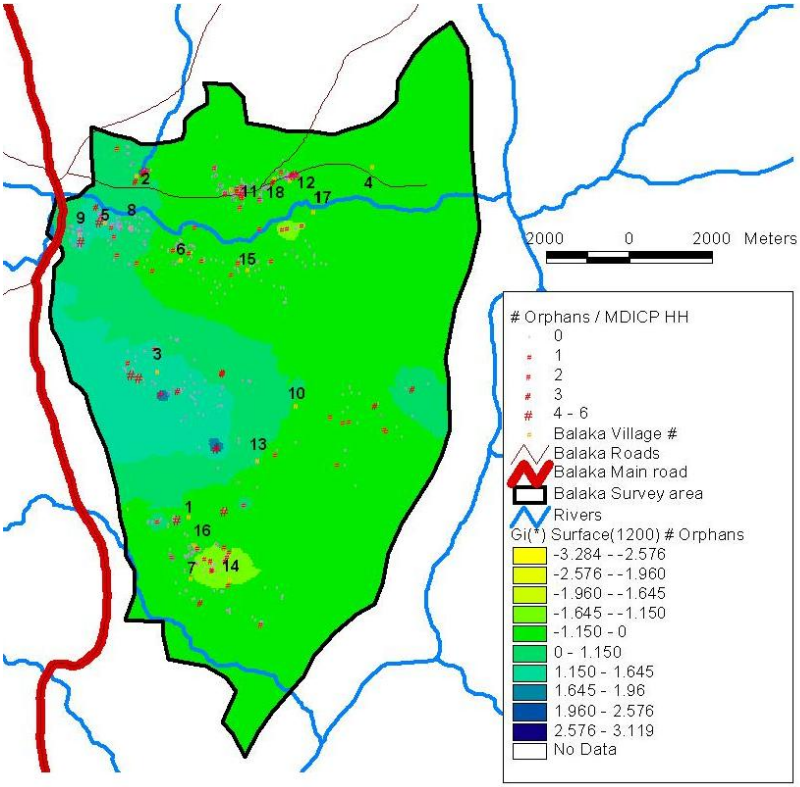


Figure 3: Orphan clusters in MDICP villages, Northern Region (Rumphi District*)*, 2004:

(a) all households in the villages;

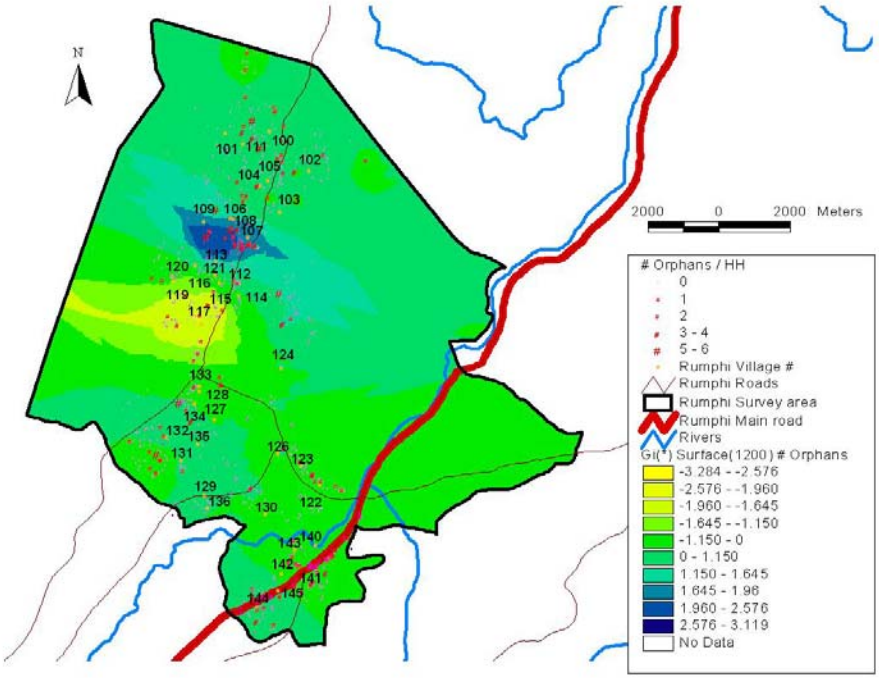
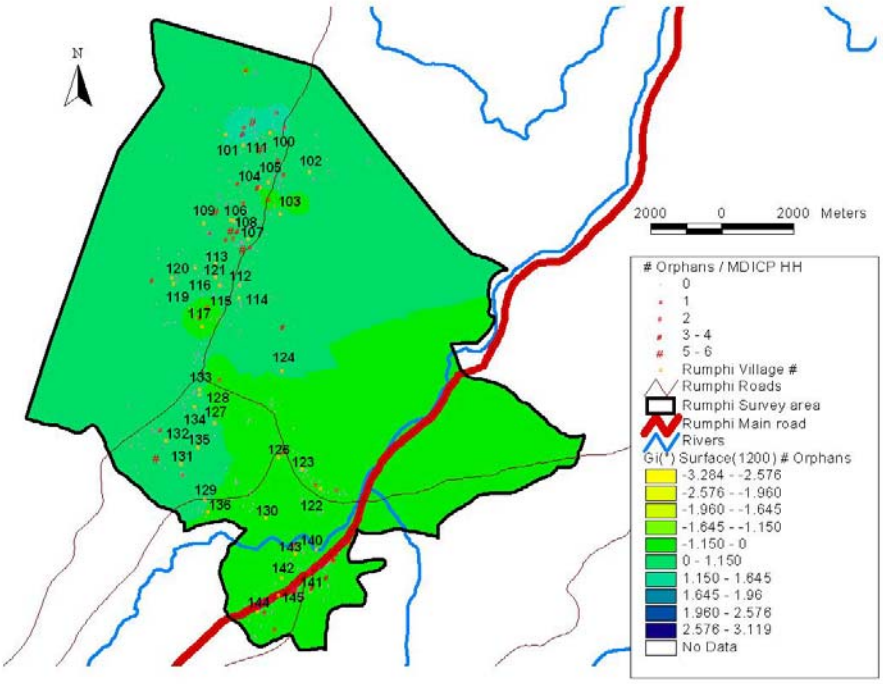


Figure 3: (continued)

(b) MDICP sample only



Panels (b) in both figures shows, not surprisingly, that restricting the sample to the MDICP households substantially reduces our ability to map geographic clusters with as much confidence as is possible with the complete roster. In Figure 2, for example, although we can still identify small areas of clustering in village 3 – though the clusters are no more than 200m in diameter – the extreme orphan hotspot in the northern villages (numbers 12 and 18) is no longer evident. Nor is the extreme orphan coldspot around village 4, although a smaller coldspot remains around villages 7 and 14 (Gi* = -1.96 - -1.65). The difference between the full roster and MDICP sample appears to be even more dramatic in Figure 3 (northern region sites). Here, no significant hotspot or coldspot whatsoever can be identified in the MDICP sample.

3.3 Characteristics of Hotspots and Coldspots

The residential patterns portrayed above, in particular in panels (a) of the figures, pose two types of questions. The first is related to place. Simply, what sort of places are orphanhood hotspots and coldspots? Do they vary in terms of observed aggregate characteristics like schooling, wealth, religion, access to market activities, HIV characteristics, and so on?

The second question indexes a different analytic level. Simply, these places are constituted by households. Consequently, what are the *households* like in which orphans are more or less likely to be located? This, rather than the first question, is the level at which much of the prior literature and our own qualitative data – both the semi-structured interviews mentioned above and a long series of informal conversations with people while conducting fieldwork in Malawi – operates. For example, fostering households were described as being relatively religious, as having access to market-related activities that would allow orphans to help support themselves, as having only middling levels of schooling and wealth (that is being neither poorest and least educated, nor wealthiest and most educated). Areas, in contrast, were not described this way.

On the other hand, it seemed to us that area-specific qualities might be reasonable factors in extended family decisions about where to send a particular orphan. For example, it would not be difficult to envisage a scenario where an extended family would send an orphan to an imperfect household which happened to be in a very suitable area with the thought that the advantages of the area would outweigh any disadvantages of the household.

In order to examine these potential household and village-level relationships we specified a range of explanatory variables in four key categories. These included:

i. Characteristics of the main survey respondent herself, including her age (actual age and an exponential term to capture non-linear effects), years of schooling, religious identity, HIV status, personal spending habits (for 3 months prior to the survey), and size of her AIDS- and religion-related conversational networks. Given normal patterns of within-household homogeneity, we assumed that some of these characteristics (e.g., schooling, religion, spending) would be correlated with those of other household members about whom we have no detailed information of the same type.

ii. Various household characteristics – these included household size (net of the number of listed orphans), number of deaths in the household in the last 3 years, and four indicators of household wealth, each designed to capture a slightly different dimension of the underlying parameter. The first indicator was a 0-6 scale indexing ownership of relatively simple durables, the second a 0-5 scale indexing ownership of more unusual and expensive items (e.g., cell-phones, televisions, and so on), the third a measure of the market value of all livestock holdings, and the fourth a measure of the number of crops grown by the household over the last agricultural cycle.

iii. Some village-level characteristics, including mean village-level aggregations of most of the variables described above.

iv. Indicators of the distance between an orphan's household and various types of infrastructure, local institutions and leaders, leisure sites, as well as other households. More specifically, by distance to *infrastructure* we refer to discrete measures of a household's distance to the closest market, shop, borehole, maize mill, and main road. Distance to *local institutions and leaders* refers to a household's distance to the closest primary school, church or mosque, village headman, and healer. Distance to *leisure sites* refers to a household's distance to a football pitch, bar, or dancing hall. And distance to other households refers to the mean and standard deviation of distance to the 10 closest households. In all cases, distances were estimated in ArcView 3.3, and are in kilometer-units with three decimal points (meaning measurement to the one meter level).[5]

[5] All distance variables are calculated using the standard Euclidean distance ("as the crow flies"), rather than being based on distance along footpaths, trails or dirt roads.. For some of the variables, especially distance to other households, this standard also provides a very reasonable estimate of distance along the footpath (since households in rural areas tend to be directly linked by footpaths and the distances between them tend to be short). More specific to the mean/variance measures for the 10 closest households: we calculated these using

3.3.1 Models

Models were estimated in a series of two-level regression analyses.[6] The baseline model was:

$$Y_{ij} = \gamma_{00} + \gamma_{10} x_{ij} + U_{0j} + R_{ij} \quad (1)$$

where Y represents a binomial variable indexing the presence of at least one orphan in a household i in village j (where n_j=53), x is a set of explanatory variables, γ_{00} the population grand mean, U_{0j} the specific effect of village j, and R_{ij} the residual effect for respondent/household i within village j. This structure allows us to examine the joint effect of individual, household, and village level factors on the likelihood of a household containing at least one orphan. By parsing the variance between individual and village-levels it also allows us to explore whether the hotspots and coldspots identified in Figures 2 and 3 can reasonably be mapped onto villages, or whether they involve sub-village clusters of households, or clusters which cross-village boundaries due, for example, to some shared religious affiliation.

An alternative estimation strategy was also used. Specifically, in order to explore possible heterogeneity in the effects of individual and village-level covariates across villages, we also specified a series of models with a more complex variance structure, as in:

$$Y_{ij} = \gamma_{00} + \gamma_{10} x_{1ij} + (U_{0j} + U_{2j} + \mathrm{cov}(U_{0j},U_{2j})) + R_{ij} \quad (2)$$

All terms in model [2] are equivalent to those in model [1], except in this second model we parsed the level-2 variance into three terms (leaving the variance at levels 1 as a "fixed" parameter). Thus, U_{0j} represents baseline level-2 variance, U_{2j} represents additional variance in the intercept associated with a given covariate of interest x_1, and the covariance term (U_{0j},U_{2j}) indexes differential slope of variance across level 2.

Arcview extension "nearfeat.avx" (see http://www.jennessent.com/arcview/nearest_features.htm) (Jenness 2004).

[6] Our analysis draws heavily on Snijders and Boskers (1999). All multilevel analyses described in this paper were implemented in Stata 9 using the "xtmixed" command (do-files available from the authors upon request). Two details about model specification are worth noting here. First, no assumptions were made about the structure of the covariance matrix. Rather, all variances and covariances were distinctly estimated. Second, all models were estimated using restricted maximum likelihood (REML) over maximum likelihood (ML) since the latter is more sensitive to loss of degrees of freedom when dealing with a small number of groups (see Snijders and Boskers 1999: 56). As level-2 specification is set to the village level and there are 53 villages, this is of marginal concern.

3.3.2 Main results

Results of all bivariate estimates and models run in accordance with equation (1) are presented in Table 2. Specifically, Models 1 and 2 present estimated bivariate and multivariate relations, respectively, across the full range of explanatory variables (for brevity, Table 2 does not include estimates from other models evaluating the relationship between orphan presence and various measures of religiosity, or characteristics of conversational networks, since none of these models were statistically significant). Model 3 then presents the best-fitting model using a subset of these variables.

Table 2: Estimated relationship between orphan presence and selected variables (N=582)

	Bivariate Model 1	Multivariate Model 2[(1)]	Multivariate Model 3[(1)]
Respondent's own characteristics			
Age	.001 (.001)	.003 (.002)*	.004 (.001)**
Age * Age		-.0001 (.0001)	-.0001 (.0001)
Years of schooling	-.003 (.004)	-.002 (.006)	
Religion			
Catholic	.017 (.047)	.082 (.050)	.071 (.047)
CCAP, Anglican, Baptist	.038 (.039)	.100 (.045)*	.094 (.044)*
Other Christian	.041 (.038)	.053 (.043)	.055 (.042)
Muslim, Other	*reference*	*reference*	*reference*
Network size			
AIDS conversational net. Partners	-.001 (.002)	-.0007 (.002)	
Religion conversational net. Partners	-.001 (.001)	-.001 (.001)	
Total personal expenditures	.0001 (.001)	0.0001(0.0001)	
HIV+[(2)]	-.0002 (.021)		
General household characteristics			
Household size	-.032 (.005)***	-.040 (.006)***	-.041 (.006)***
Number of deaths in HH in last 3 yrs	.033 (.014)*	.036 (.014)**	.037 (.013)**
Wealth 1: Basic durables	-.025 (.010)**	-.030 (.011)**	-.026 (.010)*
Wealth 2: Luxury durables	-.036 (.054)	-.060 (.052)	
Wealth 3: Value of livestock	.0003 (.0003)	.0006 (.0003)$^{+}$	.0001 (.0000)*
Number of crops grown	.003 (.009)	.005 (.011)	

Table 2: (continued)

	Bivariate Model 1	Multivariate Model 2[(1)]	Multivariate Model 3[(1)]
Household's spatial characteristics			
Distance from closest 10 HHs (mean)	.086 (.068)	.063 (.011)	
Distance from closest 10 HHs (SD)	.307 (.266)	.332 (.336)	
Village characteristics			
Mean:			
Years of schooling	-.002 (.007)	-.005 (.017)	
Total personal expenditures	.0003 (.0004)	.0003 (.0005)	
Household size	.006 (.015)	.046 (.024)*	.037 (.020)$^{+}$
Number of deaths in HH in last 3 yrs	-.014 (.033)	-.019 (.042)	
Number died from AIDS in last year	-.011 (.019)	-.065 (.037)$^{+}$	-.062 (.029)*
Wealth 1: Basic durables	.029 (.025)	.163 (.052)**	.149 (.046)***
Wealth 2: Luxury durables	-.180 (.214)*	-.401 (.256)	
Wealth 3: Value of livestock	.0001 (.0008)	-.0001 (.0001)	-.0001(.0001)
Number of crops grown	.0003 (.018)	-.024 (.034)	
Mean HIV prevalence	-.036 (.147)	-.133 (.159)	
Modal religion is:			
CCAP, Anglican, Baptist	-.033 (.043)	-.152 (.065)*	-.142 (.055)**
Other Christian	-.010 (.046)	-.043 (.059)	-.045 (.055)**
Constant	n/a	-.334 (.148)*	-.330 (.122)**
Variance estimates			
Individual/household	n/a	.087 (.005)***	.087 (.005)***
Village	n/a	.006 (.003)*	.005 (.003)$^{+}$
-Log likelihood		-232.74	-186.26

Notes:
(1) Two-level model with simple variance structure (equivalent to model (1) in text).
(2) Not used in final models since only 522 of these 582 women consented to the test and it contributes nothing to the explanatory power of the model.
Significance levels: *** = .01 percent; ** = 1 percent; * = 5 percent; $^{+}$ = 10 percent.

A number of important results can be seen in Table 2. First, there are relatively few bivariate relationships between the presence of an orphan and any of these variables. The most pronounced effect is related to household size. The larger it is – that is, not including orphans – the less likely one is to find an orphan in it. In addition, there

is a negative association between orphan presence and at least two measures of wealth: the ownership of basic durables at the household level, and luxury durables at the village level. Similarly, orphan presence is positively associated with the reported number of deaths in the household in the last 3 years.

Aside from these, orphan presence is associated with very little at the bivariate level: not with the respondent's age, religion, religiosity (not shown), network size, network characteristics (not shown), HIV status, nor with village-level aggregates of these variables.

In the multivariate model 2, additional relationships emerge. While the previous relationships with household size, number of deaths in the household in the last 3 years, and ownership of basic durables, remain – and are somewhat augmented – we see a borderline positive relationship between the presence of an orphan and household livestock holdings, and also positive relationships with woman's age and self-identification as CCAP, Anglican, or Baptist (in relation to Muslim and a small category of "other"). Likewise we observe positive relationships with village-level (mean) measures of household size and wealth (basic durables), as well as a negative relation with the village's modal religion being CCAP, Anglican, or Baptist. All these relationships remain significant in final, sparser model 3.

We can summarize these results as follows. First, there is a strong negative relationship between presence of an orphan and household size at the individual level in both bivariate and multivariate estimates. In contrast, the significant positive relationship to household size at the village level only emerges with controls for individual level. Taken together, these suggest that orphans are more likely to be located in smaller households in general – that is, smaller net of their own presence – but these households are more likely to be in places where the average household size is greater. We do not know why these households are greater. It may be that they already include orphans, or that a culture of fostering has developed, driven by the types of orphan-care discourses described above. Alternatively, perhaps there is simply greater tolerance for larger households, making smaller ones on their midst natural targets for family decision-makers.

Second, in relation to households in which the respondent claimed to be Muslim (the reference group), households in which the respondent claimed to be from a Christian denomination – in particular CCAP, Anglican, or Baptist – were also more likely to have an orphan. On the other hand, this result was moderated by the broader religious context in the village. In particular, orphanhood prevalence tended to be lower in villages in which the modal religion was CCAP, Anglican, Baptist, or any of the myriad other Christian denominations to be found in our research sites. It is beyond the scope of this article to explain this difference between household and village-level coefficients. But one intriguing possibility is that it is stems from being a religious

minority. Specifically, where someone is in a religious minority s/he may feel more duty bound to present a good and caring face of his/her religious identity. This would both justify the stubborn affiliation to the non-majority faith, but also – important given the strongly evangelical nature of religion in Africa – serve to promote the faith, particularly since, throughout sub-Saharan Africa, care of orphans is emerging as one of ways by which leaders' legitimacy is judged.

Third, there are also marked wealth effects on orphans' presence. As in the case of household size, these point to somewhat different relationships at the household and village level. In particular, where our indicator of wealth is ownership of basic durables, orphans are simultaneously more likely to be found in poorer households, but also in wealthier villages. Where our indicator of wealth is livestock holdings – a very different dimension of wealth in this rural sample – we find orphans are more likely to be in wealthier households. We revisit these results below.

Fourth, in neither the bivariate nor multivariate specification does there appear to be any significant relationship between the presence of an orphan and their households' distinct spatial characteristics, meaning its mean distance from its 10 closest neighbors, nor the standard deviation around that mean distance. In other words, there is no evidence that orphans are more or less likely to be found in isolated homesteads or in households clustered among others. We expand on orphans' geo-spatial characteristics in the next section.

Fifth, a number of notable non-results also stand out. There is no relationship between presence of an orphan and the respondent's years of schooling (including those exploring non-linear effects – not shown here). Similarly, there is also no relationship between the presence of an orphan and a woman's conversational network size (expected given that more networked women may be more sociable or have more sources of support), reported expenditures on herself over the last 3 months (expected since much of the current literature implies that the presence of orphans should reduce spending on all other things), and HIV status (somewhat less surprising given that many HIV positive do not know their status, but nonetheless informative).

Finally, underlying these estimated relations and non-relations is a final observation regarding the analytic level. The relative size of the individual/household and village-level variance terms in models 2 and 3 shows that individuals account for 93.5 and 94.6 percent of the total variance across the two models. Given the strength of the orphanhood clusters observed in Figures 2 and 3, and at all distances tested (Figure 1), this suggests that the spatial clustering is not isomorphic with village boundaries. Nor, we think, can it reasonably be mapped onto village boundaries.

3.3.3 More spatial results

Taking advantage of other geo-coded data available in the MDICP, we estimated an additional series of models looking at the effects of distance between households and various types of infrastructure on orphans' presence. Results from these models are presented in Table 3. The first column presents bivariate estimates, and the second estimates from a series of multivariate models in which each discrete distance variable was added to Table 2's model 3.

Table 3: Estimated relationship between orphan presence and distance of household (in 1000 meters) from various local sites, by type of model (N=582) [(1)]

HH distance from:	Bivariate (1)	Multivariate[(2)] (2)
Infrastructure		
Market	.023 (.018)	.014 (.019)
Shop	.077 (.023)**	.071 (.023)**
Borehole	.051 (.028)*	.067 (.028)*
Maize mill	.013 (.013)	.007 (.013)
Main road	.010 (.006)	.015 (.007)
Local institutions & leaders		
Primary school	.003 (.023)	-.010 (.022)
Church or mosque	-.005 (.015)	-.018 (.016)
Village headman	.025 (.028)	.019 (.027)
Healer	.026 (.016)+	.022 (.016)
Leisure		
Football pitch	.028 (.020)	.030 (.020)
Bar or beer hall	.007 (.008)	.013 (.010)
Dancing hall	.006 (.010)	.024 (.014)+

Notes
(1) All estimates are from two-level random effects models with simple variance structure, equivalent to equation [1] in text.
(2) Multivariate estimates are net of variables used in model 3 in Table 2.
Significance levels: *** = .01 percent; ** = 1 percent; * = 5 percent; + = 10 percent.

The overarching result of this series of analyses – perhaps not surprising given the general lack of spatial patterns observed in panel (b) of Figures 2 and 3 – is that, at least in these 53 villages, a household's distance from most types of infrastructure, local institutions, and leisure sites, does not predict much. Orphans are no more nor no less likely to be found closer to a market, maize mill, main road, primary school, place of worship, football pitch, bar, or the home of a village headman or healer.

In fact, proximity to only two types of sites is associated with the presence of orphans in both bivariate and multivariate specifications. Specifically, orphans tend to be further from the closest shop – this generally refers to a small informal outlet like a stall on a side road, not part of a larger and more organized collection of vendors that constitutes a market. Orphans also tend to be further from the closest borehole, this being the primary source of water in all MDICP villages. Again, it is not within the scope of this study to establish the causal mechanisms which underlie these patterns, but the nature of household chores in rural areas of Malawi mean that these are two frequently visited sites. Consequently, these results beg the question of whether family members most likely to find fostering attractive live in the worst placed households, or in those areas that allow for larger livestock holdings – as seen in Table 2. If so, then it may be that orphans represent a new, additional or alternative source of within-family labor to be directed at, among other things, fetching water from the closest borehole, running errands to the nearest shop, or taking care of livestock.

3.3.4 Complex variation

The final stage of our analysis deals with heterogeneity in the effects of individual and village-level covariates across villages, as described in relation to equation (2). This is an important analytic stage since it allows us to distinguish relations that are relatively stable across villages from those that vary. The implications of this variability for both fuller conceptual understanding of underlying phenomena and for more effective program design are self-evident.

We specified a series of 10 models identical to model 3 in Table 2, albeit with the addition of distance to closest shop and distance to closest borehole (as described in the last section). In each model we added two terms to the village-level random part of the model. The first was a random parameter for a particular explanatory variable. The second was a covariance term between that random parameter and the general level-2 variance.

Table 4: Variance estimates (standard errors) and measures of log likelihood (LL) and change in LL from models with complex variance, relative to baseline estimates from model with simple variance structure

		Individual / household (HH) characteristics							Village average		
	Baseline [1]	Respond. Age	HH size	Number of deaths	Wealth 1	Wealth 3	Distance to shop	Distance to borehole	HH size	Num. AIDS deaths	Wealth 1
Variance estimates											
Individual/HH level	.087 (.005)***	.087 (.005)***	.081 (.005)***	.087 (.005)***	.082 (.005)***	.084 (.005)***	.085 (.005)***	.085 (.005)**	.085 (.005)***	.087 (.005)***	.087 (.005)***
Village level (VIL)	.003 (.002)	.0025 (.0023)	.0023 (.0022)	.0022 (.0024)	.043 (.022)+	.0015 (.0017)	.0015 (.0024)	.0001 (.0007)	.119 (.076)	.0010 (.0031)	.064 (.051)
Explanatory variable (EV)	n/a	1.69e-06 (4.86e-06)	.0014 (.0006)*	.0001 (.0004)	.0047 (.0024)*	2.53e-08 (1.82e-08)	.0119 (.0076)	.0179 (.0163)	.0046 (.0028)	.0002 (.0007)	.0081 (.0060)
Covariance (VIL|EV)	n/a	.0001 (.0001)	-.0018 (.0010)+	.0005 (.0009)	-.0142 (.0072)*	6.13e-06 (3.39e-06)+	-.0042 (.0045)	-.0016 (.0046)	-.0232 (.0145)	.0004 (.0004)	-.0228 (.0173)
Measures of model fit											
Log likelihood (LL)	-186.65	-186.41	-178.69	-186.53	-181.18	-179.99	182.11	-183.6	-183.06	-186.53	-185.61
-2*(baseline LL – model LL)		0.48	15.92***	0.24	10.94**	13.32**	9.08*	6.10*	7.18*	0.24	2.08

Notes: [1] Equivalent to model (2) in text Significance levels for variance estimates (two-tailed test, z-distribution) and measures of model fit (chi-square test, 2 degrees of freedom) are: *** = .01 percent; ** = 1 percent; * = 5 percent; + = 10 percent.

Table 4 presents the four variance estimates – one specific to the individual-level, and three to the village-level – from this series of models. Each column represents a discrete model in which one explanatory term was allowed vary randomly. In addition, the bottom two rows of the table present the log likelihood from each of these estimated models and a test-for-difference between each of these and what we call the baseline model. This baseline has completely identical fixed components to the others – that is, it has exactly the same explanatory variables – but we restrict its random part to the simpler two-component structure, as per equation (1). This makes each pairing of baseline model plus new model nested, allowing us to evaluate whether or not complex variation improves model fit with a straightforward test-for-difference between the baseline and model log likelihood. Specifically, we compare -2*(baseline LL – model LL) to a chi-square distribution with two degrees of freedom.

Results confirm that there is some level of heterogeneity in the effect of predictors of orphan presence. In particular, the effects of individual/household-level measures of household size and ownership of basic durables vary significantly across the 53 villages.[7] By this we refer to both the village-specific intercept on this variable (the variance estimate on the explanatory variable) and the village-specific slope (the covariance term). In addition, the village-specific slope effect of livestock holding (the covariance term on the "wealth 3" measure) also varies significantly.

In relation to three other variables – distance to borehole, distance to shop, and village-level average for household size – we find no significant differences in random intercept or slope effects. However, the addition of the random parameters contributes to model fit, as indicated by the chi-squared test on the difference in log-likelihood.

Finally, on a number of other variables – respondent's age, number of deaths in the household in the last 3 years, and village averages for ownership of durables and estimated number of AIDS deaths – there is no heterogeneity in their effect on the presence of an orphan whatsoever, not it terms of village-specific intercepts, slopes, or their joint effect (where the latter is measured by the chi-squared test).

4. Discussion and conclusion

MDICP villages with relatively high orphan prevalence in 2004 look somewhat different from villages with low prevalence on two dimensions. Specifically, and as shown in Table 2, the high prevalence villages are more likely to be primarily Muslim

[7] Note that we use a more liberal alpha=.10 as the significance cut-off since these village-level effects are estimated across only 53 villages.

or Catholic. They also tend to be somewhat wealthier (in terms of ownership of durables), and to have somewhat larger households.

That said, it is difficult to interpret any of these patterns with full confidence. It is not only that there is considerable heterogeneity across villages in some of these effects, as seen in relation to Table 4. It is also that we cannot identify the causal patterns that underlie these distributions. For example, are villages with high orphan prevalence more likely to be primarily Muslim or Catholic because religious leaders associated with those religions more actively support fostering? This seems unlikely given the positive coefficients on Christian denominations at the individual level, but it is possible if those positive coefficients reflect a type of minority-status effect in areas where they are not the modal or majority group. Alternatively, it seems equally plausible that there are more orphans in Muslim and Catholic areas because other denominations have lower HIV prevalence and thereby generate fewer orphans within their specific religious or denominational networks (on the assumption that extended families tend to be relatively homogamous with respect to religious identity).

Similar questions can be raised in relation to the effects of household size. For example, within villages with larger mean household size, are orphans more likely to be sent to smaller households from the start? Or do these households become smaller with time – perhaps as non-orphan residents leave, feeling crowded out by the presence of the orphans. Assuming that orphans are more likely to be sent to smaller households from the start, is it because smaller households are thought to serve orphans and their needs more effectively? Or are they smaller because of a combination of AIDS mortality clustering in family networks (due to correlated risk factors like labor migration, sexual habits, widow inheritance, and so on) and orphans remaining in family networks after their parent's death? Or do the smaller household sizes stem from the fact that the within-family clustering of HIV is also associated with lower fertility, partly due to reduced fecundability (Carpenter et al. 1997, Fylkesnes et al. 1998, Gray et al. 1998, Kigadye et al. 1993, Kilian et al. 1998).

Results on wealth-related variables also pose interpretive challenges. In particular, without knowing more about the orphans' specific trajectories – in particular, the original wealth status of their families of origin and foster families across the different dimensions of wealth – we cannot disentangle two quite different causes of the observed relation, each of which appears to us to be equally reasonable *a priori*: either orphans are placed in poorer households from the start; or orphans are placed in averagely wealthy households – or even wealthier-than-average households – that become more impoverished over time (leaving them, in these data, with better-than-average livestock holdings but not much else).

These results and the string of questions that build on them, have a number of implications. The first is oriented toward research. Moving beyond the descriptive

nature of this study in order to unravel the intricate causal loops which underlie the observed relationships will require somewhat different types of data than those used here. Based on these results we imagine those data being: (i) prospective, in particular, following children into orphanhood, perhaps across multiple households or locations (though the ethical implications of this type of research are problematic); (ii) multilevel, thereby facilitating the identification of interaction between different analytic levels, represented here by the household – including its spatial characteristics – and village; and (iii) and multidimensional, allowing researchers to dig into a range of explanatory black-boxes like religious/denominational identifiers, wealth, and types of family networks.

The second implication is at once conceptual and programmatic, and consists of two related points. First, the fact that orphanhood clusters exist, and can be identified at a variety of cluster-sizes, is profound. For it implies that behaviors and values related to orphan care vary significantly from place to place – whether in relation to relatively minor differences in care, or to major ones like explaining how orphan-care norms appear to have collapsed in some settings but strengthened in others. Standard regression approaches tend to ignore such effects – hence the value of our analyses of heterogeneity in the predictors of orphanhood presence.

Second, although we cannot fully capture it with the current distinction between household and village levels, the initial spatial analyses suggest that whatever it is that distinguishes orphanhood hotspots from coldspots is not as arbitrary as a village boundary. Rather, it is community in its more day-to-day interactional mode, maintained by a particular frequency of social contact and exchanges that inevitably result in the flow of information and ideologies, some of which are bound to be related to how orphans should be cared for, but all of which flow across administrative boundaries. This brings us back to a key theoretical point made by Cohen (1985) about the true nature of community. It is also referenced in past studies of network data in both Kenya and Malawi, where other types of sub-village or cross-boundary phenomena have been identified (Kohler, Behrman and Watkins 2000; Weinreb 2003).

More globally, all these results have some – albeit limited – programmatic relevance. Simply, in this Malawian setting at least, there appear to be no golden rules by which the organizers of orphan programs – a burgeoning sector within the local development and AIDS program portfolio – should rationally choose a specific village in which to place their program. This is because although we can see orphans clustering in space, we have less success in explaining why or how these clusters arise, even with data as rich and multidimensional as available in the MDICP. And that is without even touching on issues left unmentioned in our analyses in this paper. We refer in particular to the stability of these clusters over time, to the potentially different motivations

driving the placement of orphans who are girls from those who are boys, or those who are younger from those who are older, or paternal from maternal or double orphans.

In fact, the orphans described here appear to rather neatly reflect the generalized AIDS epidemic from which most of them have emerged. To some extent, we can objectify them with maps and models. But as in much of the rest of the AIDS literature, a lot of the underlying processes that have led them into our spotlight remain backstage in the explanatory shadows. Consequently, we can hazard guesses about the types of network- or spatially-related changes in orphan-related discourse that are shaping orphans' lives and aggregate patterns. But our empirical explorations are, throughout, limited to observed characteristics which, though rich and varied in comparison to other social surveys, do not furnish us with the appropriate data to explain orphan-related processes in causally robust ways. It is not ideal. But it is also not too terrible a subsistence. Descriptions and partial explanations, as we have attempted to show here, can be enlightening in their own modest way.

5. Acknowledgements

The collection of data used in this paper was made possible by support from the National Institute of Health and National Institute of Child Health and Human Development (NICHD) grants RO1-HD044228-01 and RO1-HD/MH41713-0. These studies were approved by the institutional review boards at the University of Malawi College of Medicine and the University of Pennsylvania. A prior version including the spatial analyses was presented at the Annual Meetings of the Population Association of America, Los Angeles, March 30, 2006. Comments from Shelley Clark and Susan Watkins on that version, and from two anonymous reviewers on more recent versions, are much appreciated. We are also indebted to Stawa Shaibu in Malawi for her excellent qualitative interviewing. Finally, please note that the views expressed in this article are those of the authors and do not necessarily reflect those of the authors' institutions (respectively, Hebrew University, the United Nations, and the University of Pennsylvania). The article's contents have not been formally edited and cleared by any of these institutions.

References

Abebe, Tatek (2005). Geographical dimensions of AIDS orphanhood in sub-Saharan Africa. *Norsk Geografisk Tidsskrift—Norwegian Journal of Geography* 59: 37-47.

Adamchak, D. J., Wilson, A. O., Nyanguru, A., and Hampson, J. (1991). Elderly support and intergenerational transfer in Zimbabwe: An analysis by gender, marital status, and place of residence. *Gerontologist* 31: 505–513.

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

Ainsworth, M. and Semali, I. (2000). *The Impact of Adult Deaths on Children's Health in Northwestern Tanzania*. Washington, D.C.: The World Bank, Development Research Group, Poverty and Human Resources.

Aldstadt, Jared, Chen, DongMei, and Getis, Arthur (1998). Point Pattern Analysis (PPA) online software (Version 1.0a): K-Function (second-order analysis) and Local Gi* statistic. San Diego State University, Department of Geography. http://www.nku.edu/~longa/cgi-bin/cgi-tcl-examples/generic/ppa/ppa.cgi

Andrews, Gail, Skinner, Donald, and Zuma, Khangelani (2006). Epidemiology of health and vulnerability among children orphaned and made vulnerable by HIV/AIDS in sub-Saharan Africa. *AIDS-Care* 18(3): 269-276.

Balk, Deborah, Pullum, Tom, Storeygard, Adam, Greenwell, Fern, and Neuman, Melissa (2003). *Spatial Analysis of Childhood Mortality in West Africa*. Calverton, Maryland, USA: ORC Macro and Center for International Earth Science Information Network (CIESIN), Columbia University.

Barnett, T. and A. Whiteside (2002). *AIDS in the Twenty-First Century: Disease and Globalization*. Palgrave Macmillan.

Beegle, K., De Weerdt, J., and Dercon, S. (2005). *Orphanhood and the long-run impact on children*. Paper presented in IUSSP XXV International Population Conference, Tours, France, July 18-23 2005.

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

Brown, George (2004). Fuzzy Matching Demo in Access. http://www.codeguru.com/vb/gen/vb_misc/tips/print.php/c13137

Carpenter, L.M., Nakiyingi, J.S., Ruberantwari, A., Malamba, S.S., Kamali, A., and Whitworth, J.A.G. (1997). Estimates of the Impact of HIV Infection on Fertility in a Rural Ugandan Population Cohort. *Health Transition Review* 4(Suppl.): 113-26.

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

Cohen, Anthony P. (1985). *The Symbolic Construction of Community*. London: Routledge.

Crampin, A. C., Glynn, J. R, Ngwira, B. M. M., Mwaungulu, F. D., Pönnighaus, J. M., Warndorff, D. K., and Fine, P. E. M. (2003). Trends and measurement of HIV prevalence in northern Malawi. *AIDS* 17(12), 1817–1825.

Doctor, H.V. and Weinreb, A.A. (2003). Estimation of AIDS adult mortality by verbal autopsy in rural Malawi. *AIDS* 17(17): 2509-2513.

Doctor, H.V. and Weinreb, A.A. (2005). Mortality among married men in rural Kenya and Malawi. *African Population Studies/Etude de la Population Africaine* 20(2): 165-177.

Erskine, Su and Wilson, Dan (2005). A spatial framework for the management and support to orphans and vulnerable children (OVC) in Kwazulu-Natal. Unpublished Powerpoint presentation.

ESRI (2002). *ArcView Spatial Analyst 2.0*. Environmental Research Systems Institute, Inc., Redlands, California. http://www.esri.com/software/arcview/extensions /spatialanalyst

Ford, Kathleen and Hosegood, Victoria (2005). AIDS mortality and the mobility of children in Kwazulu Natal, South Africa. *Demography* 42(4):757-68.

Foster, G., Makuta, C., Drew, R., and Kralovec, E. (1997). Factors leading to the establishment of child-headed households: The case of Zimbabawe. *Health Transition Review* Volume 7 (Supplement 2): 155–168.

Fylkesnes, K., Ndhlovu, Z., Kasumba, K., Musonda, R.M., and Sichone, M. (1998). Studying the Dynamics of the HIV Epidemic: Population-Based Data Compared With Sentinel Surveillance in Zambia. *AIDS* 12:1227-34.

Garner, Robert (2000). Safe Sects? Dynamic religion and AIDS in South Africa. *Journal of Modern African Studies* 38(1): 41-69.

Getis, A. and Ord, J.K. (1992). The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis* 24(3).

Goode, W. J. (1963). *World Revolution and Family Patterns*. New York: The Free Press.

Gould, Jeremy (2005). Timing, scale, and style: Capacity as governmentality in Tanzania, pp. 61-82 in Mosse, David and Lewis, David (eds.) *The Aid Effect: Giving and Governing in International Development*. London: Pluto Press.

Grassly, N. C., Lewis, J. J. C., Mahy, M., Walker, N., and Timæus, I. M. (2004). Comparison of household-survey estimates with projections of mortality and orphanhood in sub-Saharan Africa in the era of HIV/AIDS. *Population Studies* 58: 207-217.

Gray, R.H., Wawer, M.J., Serwadda, D., Sewankambo, N.K., Li, C., Wabwire-Mangen, F., Paxton, L., Kiwanuka, N., Kigozi, G., Konde-Lule, J., Quinn, T.C., and Gaydos, C.A. (1998). Population- Based Study of Fertility in Women With HIV-1 Infection in Uganda. *Lancet* 351(9096):98-103.

Gregson, Simon et al. (1999). Apostles and Zionists: The influence of religion on demographic change in rural Zimbabwe. *Population Studies* 53(2): 179-193.

Guest, E. (2003). *Children of AIDS: Africa's Orphan Crisis* (2nd edition). London: Pluto Press.

Jenness, J. (2004). Nearest Features, with Distances and Bearings (NearFeat.avx) extension for ArcView 3.x, v. 3.8. Jenness Enterprises. Available at: http://www.jennessent.com/arcview/nearest_features.htm

Kigadye, R.M., Klokke, A., Nicoll, A., Nyamuryekung'e, K.M., Borgdorff, M., Barongo, L., Laukamm-Josten, U., Lisekie, F., Grosskurth, H., and Kigadye, F. (1993). Sentinel Surveillance for HIV-1 Among Pregnant Women in a Developing Country: 3 Years' Experience and Comparison With A Population Serosurvey. *AIDS* 7:849-55.

Kilian, A.H.D., Gregson, S., Ndyanabangi, B., Walusaga, K., Kipp, W., Sahlmiiller, G., Garnett, G.P., Asiimwe-Okiror, G., Kabagambe, G., Weis, P., and von Sonnenburg, F. (1998). Reductions in Risk Behaviour Provide the Most Consistent Explanation of Declining HIV Prevalence in Uganda. *AIDS 13:391-98.*

Kohler, Hans-Peter, Behrman, Jere R., and Watkins, Susan C. (2000). Empirical Assessments of Social Networks, Fertility and Family Planning Programs: Nonlinearities and their Implications. *Demographic Research* 3(7).

Madhavan, Sangeetha (2004). Fosterage Patterns in the Age of AIDS: Continuity and Change. *Social Science and Medicine* 58:1443-54.

Mansuri, G. and Rao, V. (2004). Community-based and -driven development: A critical review. *World Bank Research Observer* 19(1): 1-39.

Mishra, Vinod, Arnold, Fred, Otieno, Fredrick, Cross, Anne, and Hong, Rathavuth (2005). Education and Nutritional Status of Orphans and Children of HIV-Infected Parents in Kenya. No.24, ORC Macro, Demographic and Health Research Division, Calverton, Maryland.

Monasch, R. and Boerma, J. T. (2004). Orphanhood and childcare patterns in sub-Saharan Africa: an analysis of national surveys from 40 countries. *AIDS* 18: 55-65.

Nyamukapa, C. and Gregson, S. (2005). Extended family's and women's roles in safeguarding orphans' education in AIDS-afflicted rural Zimbabwe. *Social Science and Medicine* **60**(10): 2155-67.

OpnSeason (2006). Approximate String Matching Engine: Fuzzy Matching Algorithm and Source Code - http://www.kdkeys.net/forums/thread/6450.aspx

Ord, J. K. and Getis, A. (1995). Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis*, Vol. 27: 286-306.

Population Reference Bureau (2005). 2004 World Population Data Sheet. Accessed at http://www.prb.org/pdf04/04WorldDataSheet_ENG.pdf on April 9, 2008.

Republic of Malawi. 2005. *Final Report: Integrated Household Survey 2004-2005*. Zomba: National Statistical Office.

Sarker, Malabika, Neckermann, Christina, and Müller, Olaf (2005). Assessing the health status of young AIDS and other orphans in Kampala, *Uganda Tropical Medicine & International Health* 10(3): 210–215.

Scott, Lauren (2001). *ArcView SA Extension for Gi* and K Function statistics*. San Diego State University and Center for Spatially Integrated Social Science, University of California, Santa Barbara.

Seeley, J., Kajura, E., Bachengana, C., Okongo, M., Wagner, U., and Mulder, D. (1993). The extended family and support for people with AIDS in a rural

population in southwest Uganda: A safety net with holes? *AIDS Care* 5: 117–122.

Sharma, Manohar (2005). Orphans in Malawi: Prevalence, Outcomes, and Targeting of Services. IFPRI. Washington, D.C. October 2005. 33 p. http://www.ifpri.org/themes/hiv/pdf/sharma2005.pdf

Snijders, T. A. B. and Bosker, R. J. (1999). *Multilevel analysis. An introduction to basic and advanced multilevel modelling*. London: Sage

UNAIDS, UNICEF, and USAID 2002. *Children on the Brink 2002: A Joint Report on Orphan Estimates and Program Strategies*. Washington, DC: UNAIDS, UNICEF, and USAID.

UNAIDS (2004). Report on the global HIV/AIDS epidemic (available online at www.unaids.org) New York: World Health Organization and UNAIDS.

Walters, B.B., Cadelina, A., Cardano, A., and Visitacion, E. (1999). Community history and rural development: why some farmers participate more readily than others. *Agricultural Systems* 59(2):193-214.

Weinreb, Alexander A. (2002). Lateral and vertical intergenerational exchange in rural Malawi. *Journal of Cross-Cultural Gerontology* 17(2): 101-138.

Weinreb, Alexander A. (2003) Change and instability: A multilevel analysis of AIDS-related conversational networks in Malawi. *Demographic Research* SC1: 373-396.

Weinreb, Alexander A. (2006). Substitution and substitutability: The effects of kin availability on intergenerational transfers in Malawi, pp. 13-38 in Anne H. Gauthier, Cyrus Chu, Shripad Tuljapurkar (eds.) *Allocating Public and Private Resources Across Generations: Riding the Age Waves – Volume 2*. Springer-Verlag Publishers.

Zaba, B. et al. (2005). HIV and mortality of mothers and children. Evidence from cohort studies in Uganda, Tanzania and Malawi. *Epidemiology* 16: 275-280.