



DISCUSSION PAPER SERIES

IZA DP No. 4759

The Analytical Returns to Measuring a Detailed Household Roster

Richard Akresh
Eric Edmonds

February 2010

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

The Analytical Returns to Measuring a Detailed Household Roster

Richard Akresh

*University of Illinois at Urbana-Champaign,
BREAD and IZA*

Eric Edmonds

*Dartmouth College,
NBER and IZA*

Discussion Paper No. 4759
February 2010

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

The Analytical Returns to Measuring a Detailed Household Roster*

Households are dynamic while most surveys only collect information on individuals who are present at a single point in time. We exploit a unique and thorough household membership enumeration in Burkina Faso to consider the analytical costs of the typical static household roster. We document that households are extremely fluid with 10 percent of individuals spending sometime away over a three year period, averaging 16 of the 36 months away. The residency status of persons age 16 to 24 is most in flux. A more complete enumeration offers substantial analytical richness that is especially important for the analysis of issues that are intertwined with who is present in the household, such as the measurement of income inequality and the nature of sibling interactions in education decisions. We find that evidence of sibling rivalry in Burkina Faso appears to owe to the correlation between the presence of sisters in a household and nonagricultural income. We argue for more detailed and thorough measurement of household composition in future multi-purpose household surveys.

JEL Classification: O15, J12, I21, J13

Keywords: sibling rivalry, Lorenz curves, household composition, schooling

Corresponding author:

Richard Akresh
University of Illinois at Urbana-Champaign
Department of Economics
1407 West Gregory Drive
David Kinley Hall, Room 411
Urbana, IL 61801
USA
E-mail: akresh@illinois.edu

* We thank David McKenzie, Robert Townsend, and the participants at the Conference on Survey Design and Measurement in Development Economics for helpful suggestions. Akresh acknowledges the collaboration between Yale University and l'Institut Supérieur des Sciences de la Population of the University of Ouagadougou that aided the data collection. The fieldwork was funded by the National Science Foundation (Grant No. 0082840), Social Science Research Council, J. William Fulbright Fellowship, National Security Education Program, Institute for the Study of World Politics, and Yale Center for International and Area Studies. Finally, Akresh thanks the members of the field research team, in particular the supervisors, Ouedraogo Touende Bertrand and Hubert Barka Traore.

1. Introduction

The conventional household roster of a multi-purpose household survey is an accounting of the individuals who are present at the time of sampling. The basis for this type of roster lies in the tradition of the census where the purpose of the survey was a full enumeration of a population present in a country at a specific point in time. However, a more complete enumeration of individuals associated with a household is potentially important when household membership is dynamic and responsive to the economic circumstances under study.

This study asks the question of who we miss in a conventional household roster. Transitory migrants and departing children are the most likely to be affected by different restrictions on the roster, and there are large, related literatures on who migrants (de Brauw and Giles 2008, McKenzie 2008), when and why children depart (Rosenzweig and Stark 1989, Parish and Willis 1993), and the partitioning of households (Foster 1993, Foster and Rosenzweig 2002). Issues related to movements in and out of households are especially substantive in panel surveys, and there is a corresponding literature on attrition bias (Thomas, Frankenberg, and Smith 2001, Maluccio 2004, Baird, Hamory, Miguel 2008, Foster and Rosenzweig 2008).¹ We consider how the answer to the question of who we are missing in a conventional household roster influences the assessment of living standards and the analysis of the impact of household composition.

The Burkina Faso Child Fostering Survey (BCFS), collected by Akresh in Bazega province in Burkina Faso, contains a novel household roster that we exploit to examine these issues. There are two principal methodological innovations in the BCFS household roster. First, the BCFS collects information on every biological child of the household head regardless of whether the child is currently resident at the time of the survey or not. Second, the survey collects information on all individuals (regardless of relationship to the head) who have lived

¹ The distinction between household migration and individual migration is important to consider and emphasized in some studies of panel attrition. For example, de Laet (2005) focuses on migrant husbands in Kenya and their effort to monitor their rural spouses. Chen (2006) using data from China finds that for households with absentee migrant husbands, there are significant differences between child outcomes that are easily monitored and those that are not.

within the sampled household at any point during the previous three years.² This detailed roster information allows us to examine turnover within the household over time (both movements in and out) and to consider the implications of this turnover for research questions that are dependent on the assignment of individuals to a given household.

Using these rich data, we document several points that have implications for economic analysis using household rosters. First, there is enough mobility in Burkina Faso that choices about how to define a household, especially about the appropriate period for membership, will have a substantive influence on analysis that is sensitive to household composition. This is especially true if the 16 to 24 age group is important for the analysis, as this group exhibits the highest mobility. Second, in and out movement is weakly correlated with household characteristics, so the economic analysis of household level issues will be less affected by mobility and limited rosters, except when those issues depend critically on who is present in the household. In contrast, individual characteristics such as gender, age, marital status, occupation, and education are strong predictors of mobility. Hence, investigations focusing on individual level variation or explaining individual attributes face particular difficulty with limited rosters.

We highlight two research questions where the broader roster detail changes inference. First, while mobility is weakly correlated with household characteristics, it can be substantive for the measurement of household living standards and income inequality through adjustments for household composition (including corrections for economies of scale and adult equivalency). We document the measurement of inequality is affected by the way household membership is defined. This is especially important when household economies of scale are thought to be minimal since scale economies diminish the impact of small changes in household membership.

² There were two reasons for selecting a three-year recall period. First, during extensive pre-testing of the survey instrument, households were able to reliably recall information about household membership, income, assets and transfers over a three-year time period, while the reliability and willingness to answer these questions about longer time periods was reduced. Second, this survey is part of a broader research program examining the impact of fostering on children's welfare and involved locating and interviewing the sending and receiving household for each child fostering exchange that occurred during this three-year period. Based on the pre-testing, expanding the time period beyond three years would have created logistical difficulties in finding these linked foster households.

Second, the 16 to 24 age group is the most mobile. These young adults often have younger siblings that would be the subject of schooling studies. BCFS's more detailed household roster substantively alters findings about the nature of sibling rivalry. Using siblings present, we document strong support for the sibling rivalry hypothesis that children are made better off by having more female siblings because of their relatively lower return on investment. When we replace resident siblings with all siblings (taking advantage of the detailed data on children of the household head), we find little support for sibling rivalry. This difference between results using present siblings and all siblings owes to a correlation between having females present and non-agricultural income. Higher non-agricultural income is associated with both keeping females at home and having children at home attend school.

The next section describes the differences between the BCFS and a more conventional multi-purpose household roster. Section 3 documents the extent of mobility in the population under study and thus the sensitivity of economic analysis to decisions made about the roster and the household membership definition. Section 4 explores the impact of the more detailed household roster on the distribution of income analysis. Section 5 measures how inferences about sibling rivalry are impacted when using the more detailed roster. Section 6 concludes.

2. Burkina Child Fostering Survey (BCFS) Household Roster

These data were collected by Akresh in 2001 in the province of Bazega in central Burkina Faso. They have been used in a previous study of the determinants of child fostering (Akresh 2009) and of the schooling effects of child fostering (Akresh 2008). Bazega province is located approximately 50 miles from the Burkina Faso capital, Ouagadougou. The area is poor with an average annual household income of \$158. The economy in Bazega is largely agricultural with sorghum and groundnuts the two principal crops.

There were two phases of the BCFS. Phase 1 (BCFS-1) was a baseline survey and are the data analyzed in this paper. The survey consists of interviews with 606 randomly selected households in 15 randomly selected villages in Bazega province. BCFS-1 took place between

February 15, 2001 and April 12, 2001. The baseline survey collected the necessary information to locate the host family for any child who had been fostered out and to locate the biological parents for any child who had been fostered in. Phase 2 (BCFS-2) consisted of finding the paired households that exchanged a foster child and interviewing them using the same survey instrument as Phase 1. BCFS-2 was conducted between May 6, 2001 and August 30, 2001. Because of the non-random nature of the tracked household sample collected in the second phase, those Phase 2 surveys are not used in this paper's analysis.

Two aspects of the household roster in BCFS are unique and substantive in the current analysis. First, for every household, all survey questions in the household roster were asked about every biological child of the household head, regardless of whether the child was currently resident at the time of the survey or not. Second, for every household, all survey questions in the household roster were asked for any individual who had lived within the sampled household for at least four months at any point during the three years prior to the survey. These survey questions include the age, gender, marital status, occupation, educational attainment, school enrollment status during the three years prior to the survey (1998, 1999, and 2000), and the number of months that the individual lived outside of the village during 1998, 1999, and 2000.

These two unique aspects (all biological children and anyone resident during the three years prior to the survey) of expanding the targets for the household roster differ from traditional, multipurpose household surveys in developing countries. Traditional household surveys will generally define membership in the household based on residency at the time of the survey. For example, the Demographic and Household Surveys define household membership as "persons who usually live in your household and guests ... who stayed here last night" (ICF Macro 2010, pp. HH-4). There is variation across surveys on how long the individual must have been resident, with durations ranging from one day to several months. The Indonesian Family Life Survey defines a household member as, "anyone who usually lives in the household, whether she/he is at home during the survey or is temporarily absent. A householder who has been away for 6 or

more months, and a householder who has been away for less than 6 months but plans to move out/be away for 6 or more months is not regarded as a householder.”

However, these official definitions to determine household membership are often more ambiguous and fluid when the survey is in the field. The Burkina Faso ICRISAT (International Crop Research Institute for the Semi-Arid Tropics) data collected from 1981 to 1985 is an example of how on-the-ground realities impacted defining household membership. Matlon (1988) in the ICRISAT data manual describes the issue:

“An entirely unambiguous, consistent, and universal definition of the ‘household’ for use in sampling, data collection, and analysis, proved to be elusive. This is due to:

- 1) strong underlying social and economic ties which link even apparently independent production/consumption units
- 2) the fragmentation of intra-unit decision-making by type of economic activity (e.g. production vs. consumption; agricultural vs. non-agricultural production; individual vs. communal or ‘household’ farming)
- 3) changes in the loci of decision-making and in group structures across seasons (in part as a function of seasonal flux in economic activities).”

It is clear that determining which individuals are members of a given surveyed household is complex and incorporates decisions about production, consumption, and residency. Recent research in Tanzania by Leone, Coast, and Randal (2009) used in-depth case study interviews to explore how household membership as defined by survey practitioners differed from the social units that people themselves defined as the relevant household. Household age and sex structure changed significantly across these definitions. Beaman and Dillon (2009) in Mali randomized four different household definitions (varying production and consumption membership restrictions) that resulted in changes in measures of household composition, particularly the number of male household members. One key difference in our paper is that we have the entire universe of ever born biological children of the household head and we are able to then explore how residency of these individuals interacts with substantive research questions.

3. How Much Information is Gained With the More Detailed Roster?

The obvious first question in discussing the additional roster detail collected in the BCFS is whether it nets additional information. If households are stagnant and household members

infrequently move, then the expanded BCFS roster structure will yield little analytical gain while contributing to additional fieldwork complications. In our context of rural Burkina Faso, there appears to be a great deal of movement in and out of households.

We first plot the distribution of move-out rates for BCFS households. For each household present at the time of sampling in early 2001, the roster provides information about any individual present in 1998, 1999, or 2000. We compute if individuals present in 1998 were still present at any point in the final reference year, 2000. An individual is viewed as moving out if they were present at some point in 1998 and not present at any point in 2000. The household's move-out rate is the fraction of individuals present at some point in 1998 but not in 2000. The distribution of move-out rates across the 606 sampled households is in Figure 1.

Most individuals do not move between 1998 and 2000. Of those present at some point in 1998, 94 percent are present at some point in 2000.³ However, movers are spread over many households so that 41 percent of households experience some out movement between 1998 and 2000. This is evident in Figure 1 as the mass to the right of zero. At the household level, the move-out rate distribution ranges from 0 to 2/3 of household members present in 1998 not present in 2000. For households that experience departure, the modal move-out rate is 12.5 percent of 1998 members. Figure 1 and these calculations explicitly do not include deaths. If deaths are included, then 61 percent of households experience changes in the population that is still present in 2000.

An alternative measure of mobility is to look at the rate of new entrants. Figure 1 treats the 1998 (first period) population as the reference group and examines how many leave. To examine new entrants, we treat the 2000 (last period) population as the reference group and

³ The move-out rate from 1999 to 2000 is just over three percent. If this one-year and two-year move-out rate remained constant over a ten-year time period, then conducting a panel survey on this population would indicate that roughly 30 to 35 percent of individuals present in the baseline survey would no longer be living in the village at the time of a follow-up panel. Tracking these movers as was done in Kagera (Beegle, De Weerd, and Dercon 2008) would be critical for preventing biased analyses. In addition, this six percent overall move-out rate masks the much higher move-out rate of 23 percent for those individuals aged 16 to 24.

examine how many people present in 2000 are new entrants to the household over the two previous years. The distribution of new entrant rates is pictured in Figure 2. Births are not included in the figure.

Most individuals present in 2000 are not new entrants. Of those present in 2000, 93 percent were present two years earlier.⁴ At the household level, 35 percent of households experience some new entry other than births during the previous two years, and the range of the fraction of household members that are new entrants is 0 to 100 percent. A household that exists in 2000 but did not exist in 1998 would have 100 percent of its members as new entrants and would be missed entirely in Figure 1. The distribution of new entry rates across households is substantively more concentrated around zero than is the out movement rate. Fewer than 10 percent of households have had more than 20 percent of their members arrive within the last two years. When births are considered as new entrants, we observe substantively greater rates of new entry as fertility is high in the study area. When births are included, 25 percent of those present in 2000 are new entrants and 89 percent of households experience new entrants.

These two different mobility measures in Figures 1 and 2 are coarse in that many individuals can be both absent and present in a given year. To gauge how much time someone spends away from home, we calculate the fraction of the previous 36 months an individual on the roster lived outside of the village.⁵ The distribution of the share of time spent outside the village is shown in Figure 3. As in the previous figures, births and deaths are excluded.

Seventy-seven percent of individuals spent all 36 months present in the village. Thus, nearly 1/4 of individuals spend some time living outside their village. This contrasts with 94 percent of those present in 1998 were present in 2000, and 93 percent of those present in 2000 were present in 1998. It is obvious that a sizeable portion of the population spends some time

⁴ For those individuals aged 16 to 24, 15 percent of those present in 2000 were not present two years earlier.

⁵ The roster also collects information on all children of the household head, some of whom may not have been present. Hence, we tabulate the fraction of the last 36 months an individual was present only for individuals present at some point.

away, while returning for parts of the year. On average, individuals spent about 10 percent of the three-year period living away from home. For those individuals who spent a non-zero amount of time living outside the village, they spent on average 16 of the 36 months away. This group may easily be missed by conventional rosters. The relatively small overall share of total time spent away for this group suggests that they may appropriately be considered a part of the household. Hence, their exclusion is potentially a substantive cause for concern in the economic analysis of data using conventional household rosters.

Since most individuals do not move but a large number of households experience some movement, this implies the correlation between household attributes and out migration will be weak. That is, the distribution of movers will be spread across households. In fact, 74 percent of the variation in the share of the last 36 months spent away is within household and only 26 percent is between households. Household characteristics will be weak predictors of mobility. We therefore focus our discussion on the correlates of mobility foremost on individual attributes.

A more detailed examination of the frequency of mobility will help identify where household membership decisions are most likely to impact analysis. Table 1 tabulates by age and gender the distribution of the share of the last 36 months spent outside the village.⁶ It is already known that young adults are the most likely population to migrate. Table 1 supports this in that the 16 to 24 age group are those most likely to have left their residence in the previous 36 months. A majority of this age group spent some time away. The distribution of time away for this 16 to 24 age group looks fairly flat across the possible distribution of time away. Older adults (ages 25 to 49) and school age children (ages 8 to 15) are also more likely to spend time away than the oldest and youngest age groups, but their time away is more concentrated in a relatively brief period (under 6 months) when contrasted with the 16 to 24 age group.

⁶ There are 109 missing values for age for individuals that were alive and present at some point between 1998 and 2000. They are included in Figures 1 to 3, but are excluded from Table 1.

Panels B and C in Table 1 present a tabulation by age of time spent outside the village for males and females. It is striking that the distribution of time away and the frequency of time away looks similar for both genders in the 16 to 24 age group. It is substantively different for older age groups, especially ages 25 to 49, where women are far more likely to have been present for all 36 months, whereas men ages 25 to 49 are more likely to be away for 1 to 12 months.

The significant individual mobility documented in Table 1 has clear implications for a number of household characteristics. In particular, measures of household size will vary based on whether individuals who are temporarily living outside of the village are included as members of the household. In Table 2, we document the magnitude of the changes in household size, dependency ratios, and household per capita income using alternative residence requirements for inclusion in the household. Column 1 is the most restrictive, indicating an individual is considered a member only if they have never lived outside the village during the previous three years. Column 6 presents the most inclusive definition, which treats any individual that has been resident in the household at any point during the previous three years as a member. Columns 2 to 5 present intermediate definitions based on the number of months a person lived outside the village. Household size varies from 5.8 to 7.5 members, a 29 percent increase due to expanding the residency restrictions on membership. Consistent with Table 1, the largest change in household size is due to adding individuals aged 16 to 24, and moving from column 1 to 6 represents more than a doubling in members of this age group for both males and females.

Related to changes in household size are the subsequent changes in dependency ratios. We calculate a total dependency ratio as the number of children age 0 to 15 plus the number of adults over age 65 divided by the number of individuals age 16 to 64 multiplied by 100. Dependency ratios are quite high in rural Africa, partly due to high fertility, and we find ratios between 134 and 155 depending on which individuals are considered members of the household. Moving from the most exclusive definition in column 1 to the most inclusive and given the large movement of adults age 16 to 24, the dependency ratio drops about 11 percent as additional

individuals are considered included in the household. Total household per capita income, per capita agricultural income, and per capita non-agricultural income drops by 24 to 30 percent as we move from the most exclusive household membership definition to the most inclusive.

The share of the previous 36 months away is only one of many possible measures of mobility and it confounds those who exit, with those who enter, and those who are intermittently away. We can separate out these different concepts of mobility with the detailed roster in the BCFS. In Table 3a, we examine for all individuals the correlation between individual attributes and these different mobility concepts. Each column in Table 3a reflects a different dependent variable: an indicator for whether an individual has lived outside of the sampled village, an indicator that an individual moves out during the 36 months of reference and does not return, an indicator that an individual moves out and returns, and the fraction of the 36 month period spent away. We employ linear regression to compute conditional means of the relationship between individual characteristics and the mobility concept. These conditional means in Table 3 are not causal relationships. Of course, this simple cross-tabulation of correlation between migration and individual characteristics can mask important interactions (McKenzie and Rapoport 2007b).

Young adults 16 to 24 are more likely to live outside of the village, are more likely to move out and not return, and spend a larger share of the last 36 months away. Thus, some of the mobility of young adults may reflect children growing up and establishing independent lives, although the difference in mobility for young adults persists even controlling for marital status and education in Table 3a. However, young adults are also more likely to move out for a temporary period and return (column 3) than any other age group.

In general, the correlates of moving outside the village (column 1 of Table 3a) and the fraction of the 36 months spent away (column 4 of Table 3a) are similar. Individuals involved in non-agricultural occupations are more likely to move and spend more time away. The widowed, divorced, and never married individuals are less likely to move. In contrast to biological children

of the household head (the reference population), wives and heads of households are less likely to migrate while other relations are more likely to migrate.

Comparing individuals who move out and do not return to individuals who move out and return (columns 2 and 3) is informative of what attributes are associated with permanent versus temporary migration. Of course, moving out and not returning (column 2) is a poor proxy for permanent migration as we are limited in information, and some of those who left without returning during the last 36 months will return in the future. Nonetheless, compared to the married group, those who have never married are less likely to move out and never return and more likely to move out and return. Compared to biological children of the household head, wives are more likely to leave and return as are other relations. The more educated are also more likely to move out and return and are less likely to move out without returning. This might reflect that some of the movement in this high education group is for schooling as secondary school access is far from universal in our study population.

Men are less likely to migrate than women given their other characteristics in Table 3. We observed a similar pattern in Table 1 (without regression controls), and in Table 1 it appeared to be driven by young girls spending more time away than boys. This may reflect the types of activities girls are involved in, their greater involvement in domestic work, and their lower rates of school enrollment. In Table 3b we reproduce the conditional means of Table 3a bifurcating the sample by gender. The most noticeable differences are that males age 25 to 49 appear to have similar migration probabilities as those aged 16 to 24, whereas older women have lower probabilities of living outside of the village than young adult women. Never married women are much less likely to live outside of the village compared to never married men. Finally, females with secondary or higher education are much more likely to live outside of the village while education does not appear highly correlated with male migration.

4. Measuring Inequality

Most measures of economic well-being are calculated using household level data. There are three substantive issues in how the assignment of individuals to households influences measurement of the standard of living and the distribution of income. First, when an individual is an earner, his presence or exclusion from a household will influence the total income and expenditures for the household. Second, most welfare measures adjust the computed income or expenditure measure for the number of individuals that share in that income or expenditure. Third, there may be economies of scale in the household that cause living standards measurement to vary based on how individuals are assigned to households, even in the case where the distribution of income is uniform across individuals (which mitigates the import of the first two issues). Hence, the ways in which a roster assigns individuals to households will have substantive implications for welfare calculations and inequality measurement.

Computing a welfare measure includes assumptions about household economies of scale and the relative consumption needs of different demographic groups. Deaton and Paxson (1998) review this literature and the importance of different assumptions in the computation of adult equivalence. We define an adult equivalent as, $E = (A + \beta K)^\theta$, where A is the number of adults assigned to the household, K is the number of children, β adjusts for the fact children may have reduced consumption needs compared to adults. θ is the economy of scale parameter.⁷

Assumptions about economies of scale can be important in the measurement of the distribution of income and tracking welfare. They have received an enormous amount of study (see Logan (2008) for estimates of changes in economies of scale in U.S. history). The larger the economy of scale, the less substantive is how individuals are assigned to households in our study setting. If we had a large scale national survey, then economies of scale could be more

⁷ Often adult equivalence calculations treat one-adult and multiple-adult households differently to allow for fixed start-up costs. With an average household size between 5.8 and 7.5 depending on how individuals are assigned to households, the addition or subtraction of individuals from households based on different membership rules will not move many households between one and two members. Thus, we ignore this adjustment in our discussion.

substantive if we were shifting individuals from single person to large households with varying assumptions on how to define the household, but sampling is in one area of Burkina Faso. There is little risk we are observing the same migrant in multiple sampled households. Given the high mobility of the 16 to 24 age group (51 percent of these individuals spent some time outside of the village during 1998 to 2000), the important margin for differences in household composition based on how a household is defined will be in a population that is usually not considered in K . Hence, issues associated with β are potentially smaller, but there is still significant mobility among children (22 percent of children age 8 to 15 spent some time during the three year period outside of the village), and we consider the distribution of income for $\beta = 1$ and $\beta = 0.5$.

Our discussion of the impact of household composition on the measurement of living standards is driven by variation in counts of the number of individuals present, even though it is plausible that household income varies with how individuals are assigned to the household. While the BCFS instrument collects atypical detail in the roster, it is difficult to assign income directly to individuals in the context of our study area, and the instrument is not designed to capture consumption-based measures of well-being. Thus, the survey limits our discussion to treating income as fixed for an individual household, and we focus on agricultural income, specifically because it cannot be easily assigned to one individual earner in our data. We define agricultural income per adult equivalence as, $Income_E = \frac{Income}{(A + \beta K)^\theta}$, where $Income$ measures

total household agricultural income.

To illustrate the importance of assumptions about when an individual is classified as a household member, we focus on the most dramatic juxtaposition. We compare the case where adult equivalence is computed by restricting the household to individuals who are present for all 12 months of 2000 to the case where the household is defined as individuals who are present for any of 2000. Neither extreme is likely to be implemented in practice in household surveys, but by considering these bounds, it is easier to see that these assumptions about when an individual

is considered a member of the household have substance. We refer to the household population that has been present for all 12 months of 2000 as the exclusive population. We refer to the household population that has been present at any point in 2000 as the inclusive population.

We compare the distribution of income between the exclusive and inclusive populations by plotting the respective Lorenz curves. Figure 4 contains the Lorenz curves for the exclusive and inclusive populations when a child is treated as equivalent to an adult ($\beta = 1$). Figure 5 contains the Lorenz curves when a child is treated as half an adult ($\beta = 0.5$). This parameter range is chosen to represent the extremes in the literature on adult equivalence (Deaton 1997).

The most inclusive definition of the household's population implies less inequality than the exclusive definition. This is visible in both Figures 4 and 5 as the Lorenz curve for the inclusive population is everywhere to the interior of the exclusive population curve. Based on Figure 4, the corresponding Gini coefficient for the exclusive population is 0.557 and for the inclusive population is 0.519. Differences between the Lorenz curves for the two household definitions are not visible in the poorest 40 percent of the population, but it is the middle of the distribution where the differences becomes more apparent. One possible explanation for this is that the distribution of agricultural income is not strongly correlated with total household size (this comes up again in Table 5 below). Hence, many of the households measured as poorest have more members. The marginal effect on income per adult equivalent of assigning an additional individual to the household in a large family is less than in a small family.

Differences in Lorenz curves between inclusive and exclusive household definitions are more apparent when children are counted as less than adults. This is evident in Figure 5 that treats a child as 0.5 adults for welfare calculation purposes. Compared to Figure 4, the difference between the exclusive and inclusive Lorenz curves are slightly larger and the corresponding Gini coefficients are 0.565 and 0.518, respectively. The difference between the two curves is also measureable in a poorer and smaller segment of the population. We think this is because $\beta = 0.5$

reduces average household size, and therefore the marginal effect of adding a person to the household on income per adult equivalent is larger in Figure 5 than when $\beta = 1$ in Figure 4.

While there are measurable differences in inequality depending on how individuals are assigned to households, greater economies of scale attenuate the importance of how individuals are assigned to households. This is a mechanical point. The marginal effect on income per adult

equivalent of adding or subtracting an adult is approximately $\theta \frac{Income}{(A + \beta K)^{\theta-1}}$. Smaller θ implies

greater economies of scale and a smaller effect on income per adult equivalent of changing the number of adults. This is apparent in Figure 6. Figure 6 mimics Figures 4 and 5 in construction but with the economies of scale assumption that $\theta=0.5$. $\theta=0.5$ is the upper bound on the possible economies of scale considered in Deaton and Paxson (1998). Differences in the Lorenz curves between the exclusive and inclusive household definitions are barely observable in Figure 6.

The challenge with this discussion is that the BCFS allows us to illustrate that measures of inequality will be influenced by how individuals are assigned to households, but it does not provide us with a clear answer to what the ‘true’ response is to the question of who belongs in the household. One possible approach when income measurement is the focus is to also collect information to identify which of the family members are involved in the production of that income, but we cannot identify this in the BCFS survey. Also, the observation that a more inclusive household definition reduces inequality is, of course, only one aspect of how migrants influence inequality. While their assignment to households has an impact on measured welfare, migration itself has welfare consequence that may be welfare improving (de Brauw and Harigaya 2007) and inequality reducing (McKenzie and Rapoport's 2007a). Nonetheless, the illustration in this section shows how substantive household definitions can be in measuring welfare, especially in cases where there are limited economies of scale.

5. Sibling Rivalry

Most economic models of the household have the implication that household decisions depend on who is present. There are many reasons for a connection between household composition and household decisions ranging from the household decision-making process to shadow prices and budget constraints. Any empirical study of the interaction of household composition and outcomes of household decisions can be influenced by how composition is measured.

We focus on sibling interactions and sibling rivalry in particular. Sibling interactions are an obvious topic choice to examine, because many of the young adults who appear most mobile and transitory in our data have siblings in the household.

A number of studies have examined the interaction of siblings in order to understand why girls often receive less education than their brothers. Butcher and Case (1994) is an important example. They use data from the United States that contains explicit information on an individual's completed education and about their brothers and sisters. They emphasize three explanations for why sibling sex composition might influence education decisions: resource competition between siblings, sex-typing of tasks, and peer effects. Resource competition occurs if boys and girls have different relative prices for investments in education or returns to that investment. Sex-typing stems from parents sending messages to children describing appropriate behaviors and goals, while peer effects come from children developing traits that depend on how they interact with their siblings. Butcher and Case find that women with only brothers receive significantly more education on average than women with any sisters.

The Butcher and Case finding is different than typically found in low income country studies. Parish and Willis (1993) examine how sibling sex composition influences the education of girls in Taiwan. They emphasize that cultural traditions favoring male descent can cause parents to manipulate daughters for the benefit of their sons. Garg and Morduch (1997) show that observationally similar behaviors can follow in a model where parents are altruistic to all their children. Instead, Garg and Morduch emphasize sibling rivalry. Credit constrained households

have their decisions about their children's education influenced by both the number of children they have, resource dilution, and the sex composition of their children, sibling rivalry. Resource dilution occurs because more children imply fewer resources per child and credit constraints limit the family's ability to borrow against future returns. Sibling rivalry occurs because all children benefit from having fewer siblings with comparatively higher returns on investment.

Many studies document resource dilution and sibling rivalry in educational investments in poor countries. In addition to the Ghana and Taiwan studies (Garg and Morduch 1997, Parish and Willis 1993), the list includes: Binder (1998) for Mexico, Morduch (2000) for Tanzania, Edmonds (2007) for Nepal, Ota and Moffatt (2007) for India, and Dammert (2010) for Guatemala and Nicaragua. While the emphasis on the sibling rivalry cause in Garg and Morduch is credit constraints and differences in relative returns to education, Edmonds (2007) emphasizes that comparative advantage in home production can lead to similar implications when girls have comparative advantage and it is not possible to hire labor for home production. Both Edmonds (2007) and Dammert (2010) find evidence consistent with this sibling rivalry interpretation.

Studies of sibling rivalry in education typically regress a measure of schooling (attendance, enrollment, attainment) on counts of the number of siblings a child has and counts of the number of sisters a child has. That is:

$$(1) \quad e_{ih} = \alpha_0 X_i + \alpha_1 Z_h + \beta_0 S_i + \beta_1 F_i + \varepsilon_{ih}$$

where e_{ih} is the educational outcome for child i in household h , X_i is a vector of individual characteristics, Z_h is a vector of household characteristics, S_i is a count of the number of siblings the child has, and F_i is a count of the number of female siblings the child has. The interpretation of β_0 is the change in e_{ih} associated with an additional male sibling. The interpretation of β_1 is the change in e_{ih} associated with the thought experiment of converting a sibling from a male to a female. $\beta_0 + \beta_1$ is then the change in e_{ih} associated with adding an additional female sibling.

Conventional household rosters create two problems for researchers. First, it is difficult to identify siblings in settings where extended family members are present. Rosters typically define

relationships relative to the household head. Hence, a study must choose between analyzing the correlation between education and household membership or the correlation between education and siblings only for children of the household head. Restricting the study to children of the household head is typical in the relevant literature and is what is done in this paper. Second, it is not possible to accurately identify sibling counts when all siblings of the household head do not meet the criteria for being included in the household. This might just introduce classical measurement error if errors in counting siblings are random, but that seems unlikely. We already observed a correlation between non-agricultural occupations and migration in our discussion of Table 3a. We expect the measurement error in the number of siblings to be correlated with other factors including the siblings' age, the activities available to siblings (and thereby gender, local labor market conditions, and migration opportunities), wealth, and preferences.

The BCFS roster collects detailed information for all ever born children of the household head regardless of their residence status. Hence, our analysis is restricted to children of the household head as well, but we are able to gauge how important missing, non-resident siblings are for the study of sibling rivalry. Of course one source of difficulty in interpretation that is common to all studies of sibling interactions is the endogenous nature of fertility decisions. This problem plagues all studies of sibling rivalry and the present study is no exception.

There appears to be sibling rivalry in Burkina Faso when we estimate a conventional sibling rivalry regression as in equation 1 that is based on *resident* children of the household head *only*. This finding is documented in Table 4. In the first column, we regress whether a child age 8 to 15 of the household head attends school on the number of resident siblings and the number of resident sisters as well as dummy variables for child age and gender. Robust standard errors are clustered at the village level because of survey design. Each additional male sibling reduces the probability the child attends school by 2 percentage points or 8 percent (26 percent of children are enrolled in school). This is the resource dilution effect. Each additional female sibling has no impact on whether the child attends school. Hence the thought experiment of

switching a sibling's gender from boy to girl raises schooling by 2 percentage points. This is sibling rivalry. Column 2 adds village fixed effects to our regressions and results appear similar.⁸

Resource dilution and sibling rivalry effects only appear substantive for boys. Focusing on the village fixed effects specification in column 2 of the table, an additional boy lowers schooling for boys by 3 percentage points or 10 percent (30 percent of boys aged 8 to 15 are enrolled in school). The sibling rivalry thought experiment of converting a sibling from boy to girl is associated with school enrollment that is higher by 4 percentage points. Girls have lower school enrollment overall. Twenty-one percent of girls aged 8 to 15 attend school, and their schooling appears unrelated to the number of siblings or number of sisters in the household. We focus our remaining discussion on boys and the related results in Panel B.

When we look at the relationship between schooling and the number of *actual siblings* as opposed to co-resident siblings, we do not observe sibling rivalry. This is evident in columns 3 to 6 of Table 4. Counting the number of siblings, not just resident or living siblings, is feasible with the detailed BCFS household roster. Columns 3 and 4 use all siblings, including deceased ones, in calculating the sibling counts, while columns 5 and 6 only use living siblings to calculate the sibling count variables. For the regressions in these columns, the count of number of siblings and number of sisters is not restricted to siblings that are in residence in the surveyed household. While results are similar across the two sets of columns, we focus our discussion on columns 5 and 6, where only living children are included in the sibling counts.⁹ In the specification for boys without village fixed effects (column 5), the sibling rivalry coefficient is reduced by a factor of 10 compared to the baseline results in column 1. With village fixed effects (column 6), the sibling rivalry coefficient is reduced by 65 percent.

⁸ The advantage of village fixed effects is that they control for village characteristics that are apt to be correlated with school enrollment and household composition. The problem with including village effects is that some of these village characteristics are useful sources of variation in the presence of siblings and schooling choices.

⁹ A case can be made for including deceased children in sibling counts. There are a substantive number of deaths, with 20 percent of the household head's ever born children deceased at the time of enumeration. We focus our discussion on living siblings because we think the choice of current school enrollment is more likely to depend on current credit constraints or contemporaneous comparative advantage of siblings.

One possible reason for the disappearance of measurable sibling rivalry when all siblings are included is that it is not unusual to see a large age span across siblings in a household due to polygamy and re-marriage. Adding children from another time period might simply add measurement error to the measure of siblings and female siblings. This does not appear to be the case. In columns 7 and 8, we restrict sibling counts to siblings under the age of 25. Compared with the resident sibling specification (columns 1 and 2), counts of siblings under 25 provide much weaker evidence of sibling rivalry in the data although slightly larger than the associations observed in columns 5 and 6.

Another possible explanation for the difference between estimates of sibling rivalry with all siblings and all resident siblings is that residency is correlated with latent factors that will also be associated with schooling. Edmonds, Mammen, and Miller (2005) show that in South Africa non-labor income is associated with more young women present in the household. Obviously the South African study setting differs from ours in innumerable ways, but if higher incomes allow families to keep their girls at home, then we might observe a correlation between the presence of females and schooling that reflects the impact of income on each rather than sibling rivalry.

Two pieces of evidence suggest scope for the correlation between female residence and other factors that are also correlates of education. First, there is clearly selection in the gender mix that is present at any point in time. In Table 4, there are 1.08 male children of the household head aged 8 to 15 present for every girl in the same age range. However, there are 0.99 living male children 8 to 15 of the household head for every living female. Second, non-agricultural income also appears to be correlated with the presence of young women in the Burkina data.

Non-agricultural income is positively correlated with the probability that a biological sister is present. This is evident in Table 5. Table 5 reports the results of regressing the number of female siblings present in the household on non-agricultural income, the number of female siblings that are alive, and the number of biological siblings present. Columns 1 and 2 are estimated on the full sample. Columns 3 and 4 restrict the sample to boys. A causal model is not

being estimated. We are seeing if the data suggest a positive correlation between income and the presence of females. Of course, there may be strong theoretical reasons to expect a non-linear relationship between income and the presence of females or migration (see McKenzie and Rapoport 2007a), but our sample sizes are sufficiently small that we do not feel that we can identify higher order relationships. Our focus on non-agricultural income is driven by the correlation between migration and non-agricultural occupations documented in Table 3.¹⁰ Columns 2 and 4 include village fixed effects; columns 1 and 3 do not. Because we control for the number of living female siblings and the number of living biological siblings, the interpretation of the coefficient on non-agricultural income is how the income measure increases the number of living females that appear in the household.¹¹

We observe a positive association between the presence of sisters and non-agricultural income. Non-agricultural income is associated with fewer living female siblings (not shown), but a higher share of those living who are present. This appears valid in the full sample and when we focus on the presence of sisters of boys who are resident (last 2 columns). This evidence is consistent with studies that document a link between migration and non-farm enterprise operation and productivity (Dustmann and Kirkchamp 2002, Woodruff and Zenteno 2007), but of course there is also the possibility of a link between farms and the presence of children (Foster and Rosenzweig 2008 is a review). On this point, it is worth noting our analysis treats fertility as exogenous, but there are good reasons to think there exists a connection between latent labor demand and fertility (Rosenzweig 1977, Rosenzweig and Evenson 1977).

These findings imply that controlling for non-agricultural income should attenuate the correlation between sibling rivalry and schooling when sibling rivalry is identified based on resident siblings only. This is apparent in Table 6 when we control for non-agricultural income.

¹⁰ In unreported regressions, we also controlled for agricultural income and the value of assets. They do not appear to substantively influence the magnitude of the coefficient on non-agriculture income and do not have significant nor substantive predictive power themselves.

¹¹ The number of observations decreases between Tables 4 and 5, because of missing income data for some households, although the regression results in Table 4 are qualitatively similar using the decreased sample size in Table 5.

Table 6 is comparable to columns 1 and 2 of Table 4, which run a regression of schooling on the number of co-resident siblings and the number of co-resident sisters.¹² In column 2 with village fixed effects, controlling for non-agricultural income reduces the coefficient on number of resident sisters by 35 percent. For the boys only sample with village fixed effects (column 4), controlling for non-agricultural income reduces the coefficient on number of sisters by 21 percent. Controlling for non-agricultural income does not reduce the sibling rivalry effects for resident sisters by the same magnitude as replacing resident sisters with actual sisters. Thus, the underlying selection process that leads there to be a positive correlation between sister co-residency and education is more complicated than just non-agricultural income.

It is obvious that our design does not speak to the causal mechanisms behind this correlation between sibling selection and education. Other studies have attempted to separate the impact of migration on the left behind household from the correlates of selection into migration (de Brauw and Giles 2008, Gibson, McKenzie, and Stillman 2009). Overall, there appears ample reason to conclude that the types of sibling information collected in conventional rosters can be misleading when used to study sibling interactions. We find little evidence that supports the sibling rivalry hypothesis when we take advantage of the additional information on non-resident siblings collected in the BCFS.

6. Conclusion

Most multi-purpose household survey rosters are restricted to collecting data on individuals who are present in a narrow window of time. This approach to measuring household membership has an important advantage in that it minimizes the extent to which individuals are double counted in the survey. This produces more precise national statistics. However, this snapshot approach to the household roster has substantive analytical costs which have been documented herein.

¹² The sample size in Table 6 is reduced compared to Table 4 because of missing income data as in Table 5. We find reduced estimates of the extent of sibling rivalry when we control for income, and this reduction does not owe to the different sample size as estimates in the reduced sample are similar in magnitude to those reported for the full sample in Table 4.

Young adults are the most mobile population in our study area. Persons 16 to 24 spend more time out of their village than persons of other ages. They are more likely to leave their village and not return, and they are more likely to leave their village and return to it than other age groups. Young adults not only leave to form their own households, but they also are more likely to leave their family temporarily. This means that young adults are most apt to be missed in rosters with narrow time windows, and they are the group most likely to be affected by decisions about how to define a household.

We examine two types of research questions where inference is apt to be affected by the structure of the household roster. Inequality measures depend on the assignment of individuals to households, and young adults are typically counted as full adults in the computation of adult equivalence. This means that their inclusion or exclusion in the household will affect measurement of inequality, especially when there are minimal economies of scale.

Many young adults have younger siblings present within their parent's household. Many studies are interested in understanding sibling interactions, and thus the treatment of young adults in the household roster will have important implications for the analysis of sibling interactions. We document what appears to be sibling rivalry in households when we restrict our household definition to the narrowest feasible in our data. However, when we repeat the analysis using all siblings, rather than just siblings present, we find little support for sibling rivalry. In our data, there appears to be a correlation between nonagricultural income and the presence of female siblings that causes there to be a false relationship between female siblings and schooling.

These two examples have been chosen as obvious places where the intermittent and permanent mobility of young adults are likely to be substantive for inference, but most research conditions on household composition. Hence, there are innumerable topics where limited information on all potential household members may be a source of bias. The easiest solution to these analytical problems is to collect more detailed rosters. In the Burkina Faso survey, data was collected on all individuals who have lived in the household over a three year period and all

children of the household head. With more detailed data, researchers can make their decisions about who is and who is not a household member, adapting this definition to match the appropriate research context. Concerns about double-counting and introducing biases into the computation of national statistics can easily be addressed in the same way, by imposing a definition of the household that minimizes this possibility. A broader roster will entail more data collection costs as it will take more time, but the analytical value of this additional information is likely to be large in most applications. Moreover, in places where there is not much mobility, where the issues raised herein are less substantive, the extra costs of the more detailed roster will also be minimal. Of course, a detailed household roster is not obviously the best way to collect information on migrants themselves when the purpose of the study is to study migrants as opposed to sampled households. When migrants are the study topic, a variety of other survey approaches are appropriate (Mistiaen and McKenzie 2007 is a review).

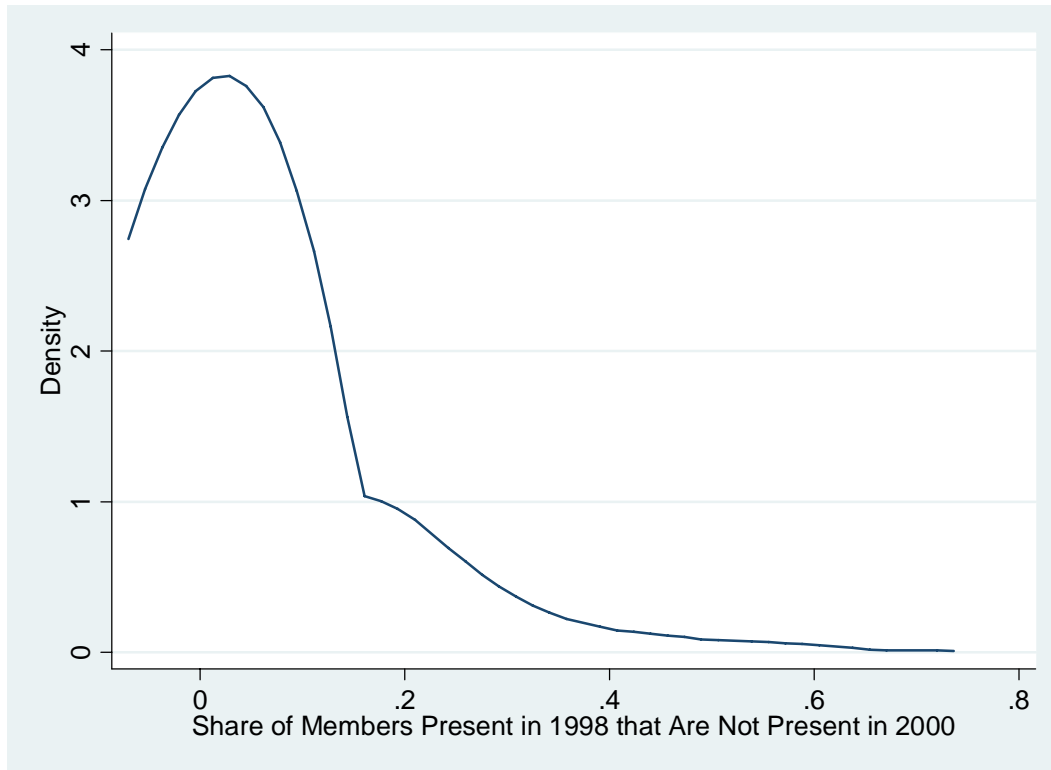
References

- Akresh, Richard. 2008. "School Enrollment Impacts of Non-traditional Household Structure." BREAD Working Paper No. 89, Cambridge MA.
- Akresh, Richard. 2009. "Flexibility of Household Structure: Child Fostering Decisions in Burkina Faso." *Journal of Human Resources*, 44(4), 976-997.
- Beegle, Kathleen, Joachim De Weerd, and Stefan Dercon. 2008. "Migration and Economic Mobility in Tanzania: Evidence from a Tracking Survey." World Bank Policy Research Working Paper 4798.
- Baird, Sarah, Joan Hamory, and Edward Miguel. 2008. "Tracking, Attrition and Data Quality in the Kenyan Life Panel Survey Round 1 (KLPS-1)." UC, Berkeley working paper.
- Beaman, Lori and Andrew Dillon. 2009. "Do Household Definitions Matter in Survey Design? Results from a Randomized Survey Experiment in Mali." Working paper.
- Binder, Melissa. 1998. "Family Background, Gender, and Schooling in Mexico." *Journal of Development Studies*, 35(2), 54-71.
- Butcher, Kristin and Anne Case. 1994. "The Effect of Sibling Sex Composition on Women's Education and Earnings." *Quarterly Journal of Economics*, 109(3), 531-563.
- Chen, Joyce. 2006. "Migration and Imperfect Monitoring: Implications for Intrahousehold Allocation." *American Economic Review Papers and Proceedings*, 96(2), 227-231.
- Dammert, Ana. 2010. "Siblings, Child Labor, and Schooling in Nicaragua and Guatemala." *Journal of Population Economics*, forthcoming.
- de Brauw, Alan and John Giles. 2008. "Migrant Labor Markets and the Welfare of Rural Households in the Developing World: Evidence from China." World Bank Policy Research Working Paper 4585.
- de Brauw, Alan and Tomoko Hiragaya. 2007. "Seasonal Migration and Improving Living Standards in Vietnam." *American Journal of Agricultural Economics*, 89(2), 430-447.
- de Laet, Joost. 2008. "Household Allocations and Endogenous Information." CIRPEE Working Paper 08-27.
- Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore, MD: Johns Hopkins.
- Deaton, Angus and Christina Paxson. 1998. "Economies of Scale, Household Size, and the Demand for Food." *Journal of Political Economy*, 106(5), 897-930.

- Dustmann, Christian and Oliver Kirkchamp. 2002. "The Optimal Migration Duration and Activity Choice After Re-Migration." *Journal of Development Economics*, 67(2), 351–372.
- Edmonds, Eric, Kristin Mammen, and Douglas Miller. 2005. "Rearranging the Family? Household Composition Responses to Large Pension Receipts." *Journal of Human Resources*, 40(1), 186-207.
- Edmonds, Eric. 2007. "Understanding Sibling Differences in Child Labor." *Journal of Population Economics*, 19(4), 795-821.
- Foster, Andrew. 1993. "Household Partition in Rural Bangladesh." *Population Studies*, 47(1), 97-114.
- Foster, Andrew and Mark Rosenzweig. 2002. "Household Division and Rural Economic Growth." *Review of Economic Studies*, 69(4), 839-69.
- Foster, Andrew and Mark Rosenzweig. 2008. "Economic Development and the Decline of Agricultural Employment." In *Handbook of Development Economics*, Volume 4, ed. T. Paul Schultz and John Strauss. Amsterdam: Elsevier.
- Garg, Ashish and Jonathan Morduch. 1998. "Sibling Rivalry and the Gender Gap: Evidence from Child Health Outcomes in Ghana." *Journal of Population Economics*, 11(4), 471-493.
- Gibson, John, David McKenzie, and Steven Stillman. 2009. "The Impacts of International Migration on Remaining Household Members: Omnibus Results From a Migration Lottery Program." World Bank Policy Research Working Paper 4956.
- ICF Macro. 2010. DHS Model Questionnaire - Phase 6 (2008-2013) QRE 11 Jan 2010. Calverton, MD: Measure DHS.
- Leone, Tiziana, Ernestina Coast, Sara Randal. 2009. "Did You Sleep Here Last Night? The Impact of the Household Definition in Sample Surveys: A Tanzanian Case Study." Presentation at British Society for Population Studies Annual Conference, Brighton, September 2009.
- Logan, Trevon. 2008. "Economies of Scale in the Household: Puzzles and Patterns from the American Past." NBER Working Paper #13869.
- Maluccio, John. 2004. "Using Quality of Interview Information to Assess Nonrandom Attrition Bias in Developing Country Panel Data." *Review of Development Economics*, 8(1), 91-100.
- Matlon, Peter J. 1988. "The ICRISAT Burkina Faso Farm-Level Studies: Survey Methods and Data Files." Village Level Studies and Miscellaneous Paper Series, ICRISAT.

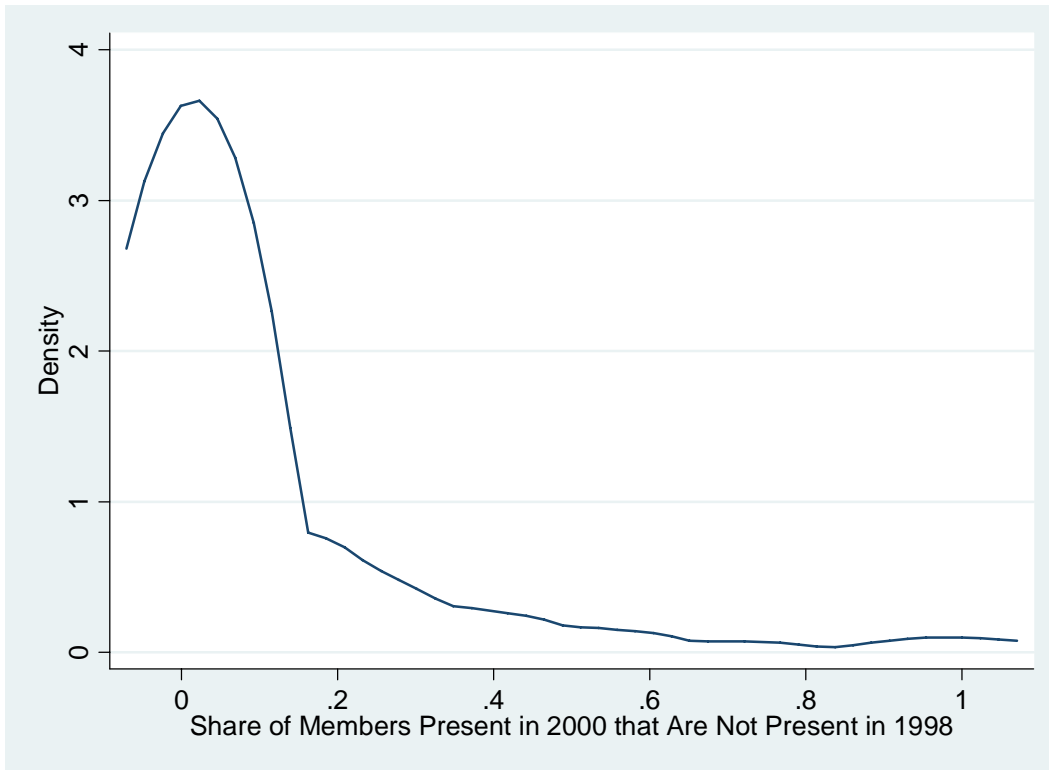
- McKenzie, David. 2008. "A Profile of the World's Young Developing Country International Migrants." *Population and Development Review*, 34(1), 115-135.
- McKenzie, David and Hillel Rapoport. 2007a. "Network Effects and the Dynamics of Migration and Inequality: Theory and Evidence from Mexico." *Journal of Development Economics*, 84(1), 1-24.
- McKenzie, David and Hillel Rapoport. 2007b. "Self-Selection Patterns in Mexico-US Migration: The Role of Migration Networks." World Bank Policy Research Working Paper 4118.
- Mistiaen, Johan Andre and David McKenzie. 2007. "Surveying Migrant Households: A Comparison of Census-Based, Snowball, and Intercept Point Surveys." World Bank Policy Research Working Paper 4419.
- Morduch, Jonathan. 2000. "Sibling Rivalry in Africa." *American Economic Association Papers and Proceedings*, 90(2), 405-409.
- Ota, Masako and Peter Moffatt. 2007. "The Within-household Schooling Decision: A Study of Children in Rural Andhra Pradesh." *Journal of Population Economics*, 20(1), 223-239.
- Parish, William and Robert Willis. 1993. "Daughters, Education, and Family Budgets: Taiwan Experiences." *Journal of Human Resources*, 28(4), 863-898.
- Rosenzweig, Mark. 1977. "The Demand for Children in Farm Households." *Journal of Political Economy*, 85(1), 123-46.
- Rosenzweig, Mark and Oded Stark. 1989. "Consumption Smoothing, Migration, and Marriage: Evidence from Rural India." *Journal of Political Economy*, 97(4), 905-926.
- Rosenzweig, Mark and Robert Evenson. 1977. "Fertility, Schooling, and the Economic Contribution of Children in Rural India: An Econometric Analysis." *Econometrica*, 45(5), 1065-1079.
- Thomas, Duncan, Elizabeth Frankenberg, and James P. Smith. 2001. "Lost But Not Forgotten: Attrition in the Indonesia Family Life Survey." *Journal of Human Resources*, 36(3), 556-592.
- Woodruff, Christopher and Rene Zenteno. 2007. "Migration Networks and Microenterprises in Mexico." *Journal of Development Economics*, 82(2), 509-528.

Figure 1: Move Out Rates for Individuals Present in 1998



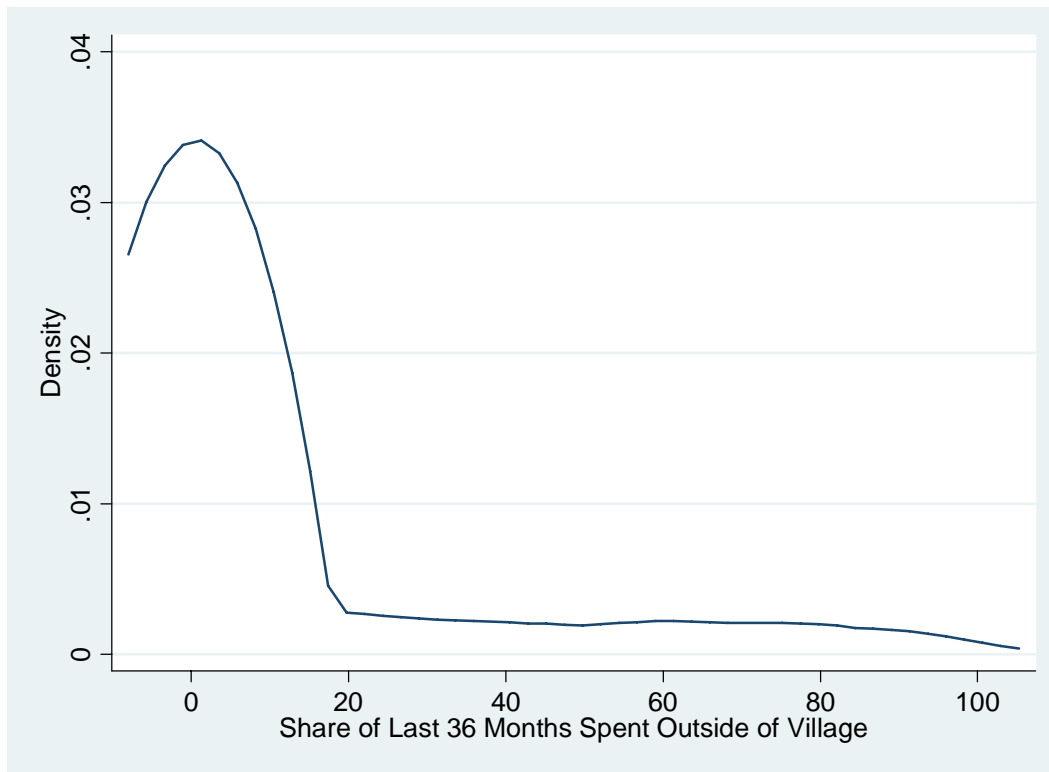
Notes: Kernel density estimation of the fraction of household members present in 1998 that were not present in 2000 (moved out) for a given household. Data source: Burkina Child Fostering Survey (BCFS).

Figure 2: New Entrants Rates for the 2000 Population



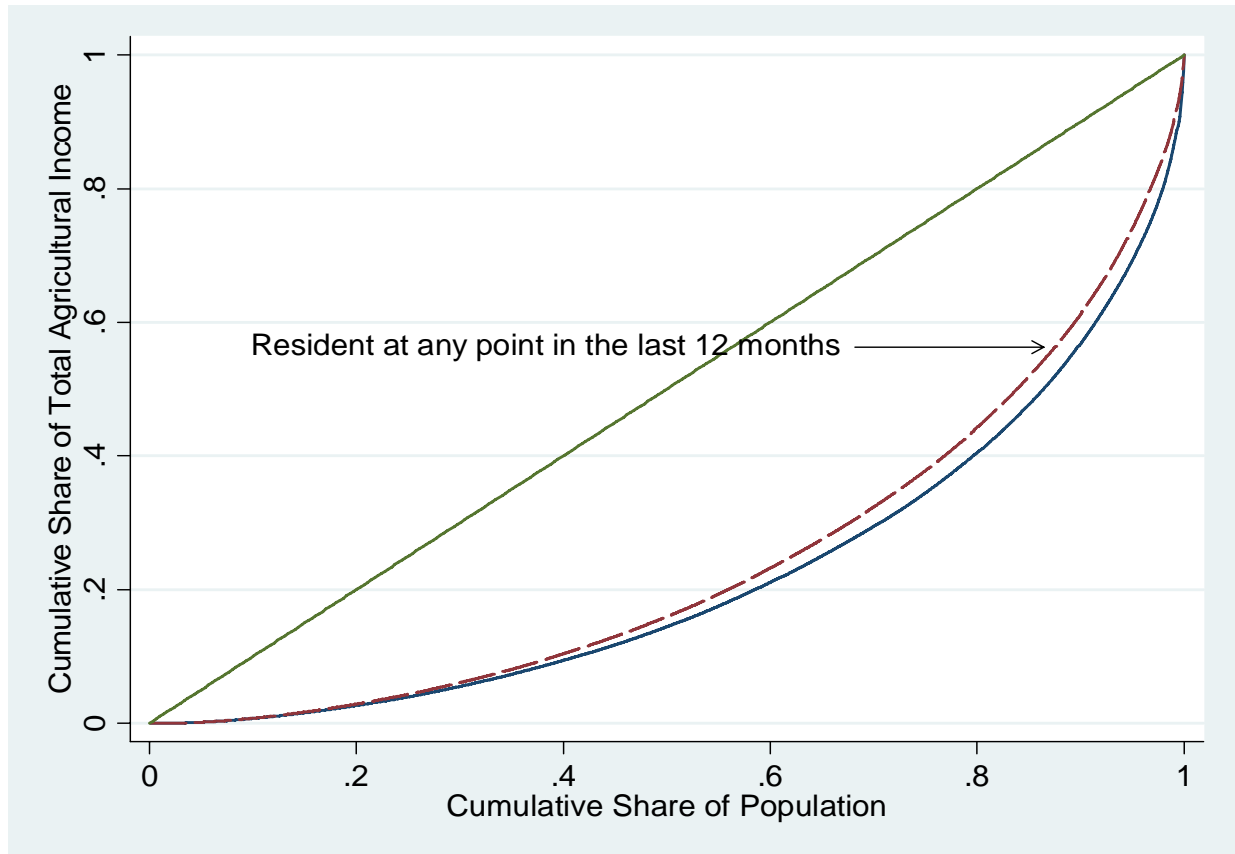
Notes: Kernel density estimation of fraction of household members present in 2000 that were not present in 1998 (new entrants) for a given household. Data source: Burkina Child Fostering Survey (BCFS).

Figure 3: Fraction of Months Spent Outside the Village



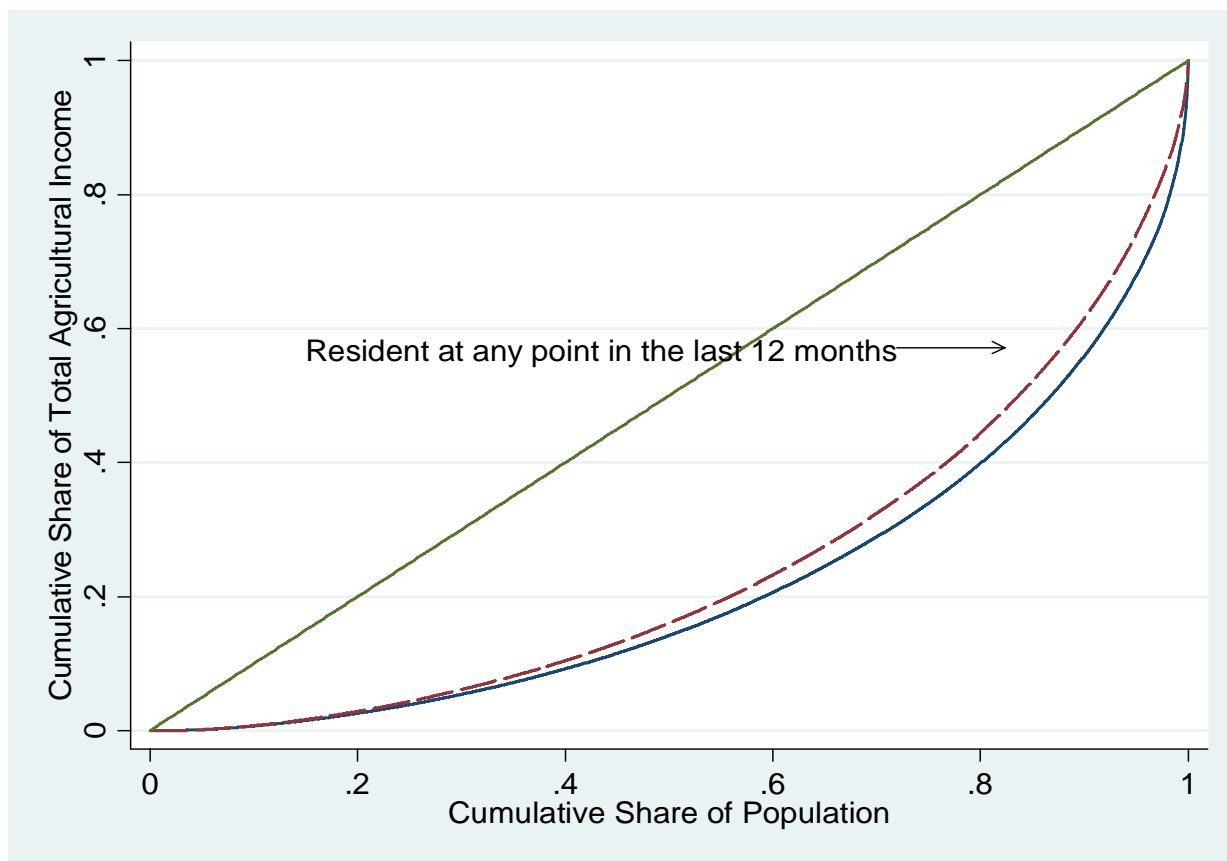
Notes: Kernel density estimation of fraction of months during the previous three years that the individual spent outside the village. Data source: Burkina Child Fostering Survey (BCFS).

Figure 4: Distribution of Income with the Most Inclusive and Exclusive Definitions of Household Membership (Children Treated as Equivalent to Adults)



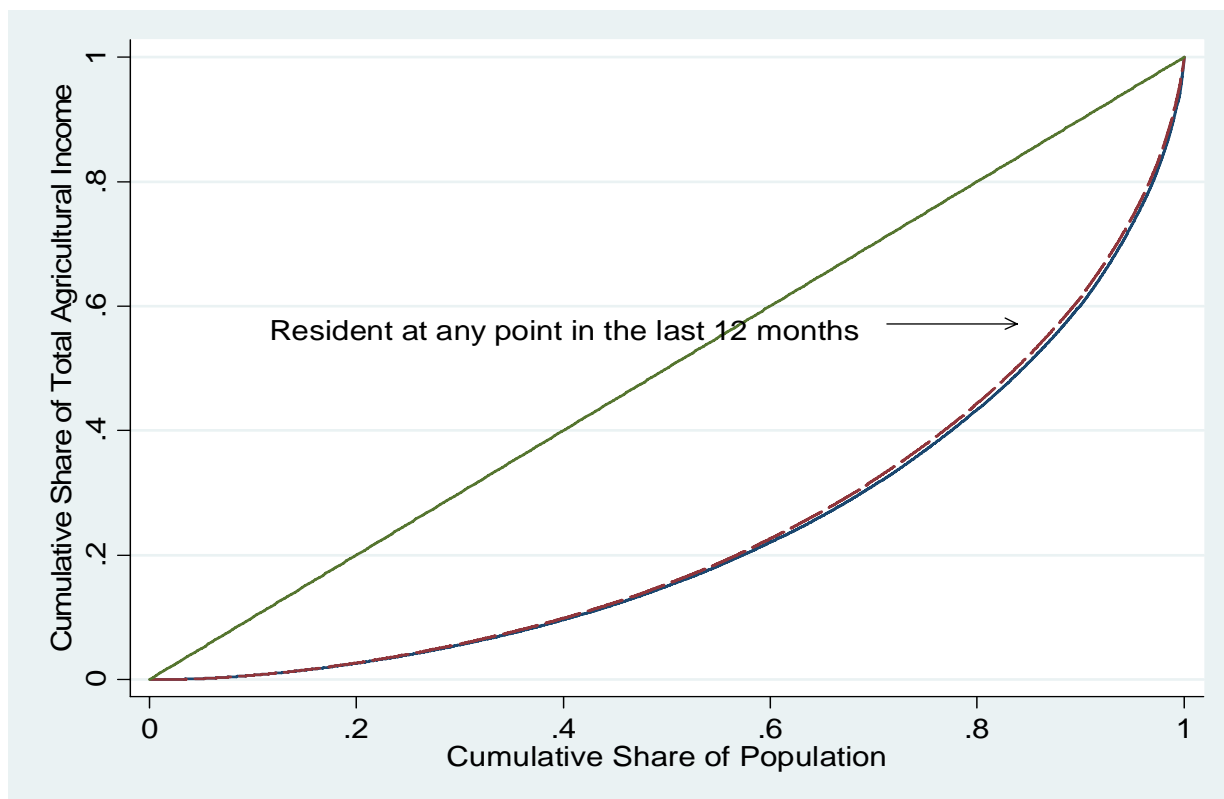
Notes: Lorenz curves graphing the cumulative share of total household agricultural income against the cumulative share of population. Children are treated equivalent to adults, $\beta = 1$. We assume the economies of scale parameter $\theta = 1$. Data source: Burkina Child Fostering Survey (BCFS).

Figure 5: Distribution of Income with the Most Inclusive and Exclusive Definitions of Household Membership (Children Treated as 0.5 Adults)



Notes: Lorenz curves graphing the cumulative share of total household agricultural income against the cumulative share of population. Children are treated as 0.5 adults, $\beta = 0.5$. We assume the economies of scale parameter $\theta = 1$. Data source: Burkina Child Fostering Survey (BCFS).

Figure 6: Distribution of Income with the Most Inclusive and Exclusive Definitions of Household Membership (Children Treated as Equivalent to Adults, Economies of Scale Parameter =0.5)



Notes: Lorenz curves graphing the cumulative share of total household agricultural income against the cumulative share of population. Children are treated equivalent to adults, $\beta = 1$. We assume the economies of scale parameter $\theta = 0.5$. Data source: Burkina Child Fostering Survey (BCFS).

Table 1: Frequency Distribution of Mobility

	Number of Individuals (1)	Mobility (Number of Months Outside Village)					24-35 Months (7)
		Never (2)	1-6 Months (3)	7-12 Months (4)	13-18 Months (5)	19-24 Months (6)	
Panel A: All Individuals							
Age 0-7	1165	86.7	3.9	2.5	1.6	1.9	3.4
Age 8-15	1170	78.0	6.5	3.9	3.1	3.9	4.6
Age 16-24	737	48.7	9.6	9.9	8.4	12.1	11.3
Age 25-49	956	79.3	6.9	4.0	2.3	3.3	4.2
Age 50+	473	93.6	3.6	0.9	0.9	0.4	0.6
Panel B: Male							
Age 0-7	555	90.4	2.9	2.2	0.7	1.6	2.2
Age 8-15	580	81.0	6.2	4.1	2.3	2.9	3.5
Age 16-24	322	48.8	12.4	10.2	8.4	9.6	10.6
Age 25-49	346	66.8	8.7	7.5	4.3	5.8	6.9
Age 50+	224	92.0	4.0	1.3	0.5	0.9	1.3
Panel C: Female							
Age 0-7	610	83.4	4.8	2.8	2.3	2.1	4.6
Age 8-15	590	75.1	6.8	3.7	3.9	4.7	5.8
Age 16-24	415	48.7	7.5	9.6	8.4	14.0	11.8
Age 25-49	610	86.4	5.9	2.0	1.1	2.0	2.6
Age 50+	249	95.2	3.2	0.4	1.2	0.0	0.0

Notes: Tabulations are restricted to individuals that were present at some point between 1998 and 2000. Numbers in columns 2 to 7 represent the fraction of individuals in that age group that spent the corresponding number of months outside the village

Table 2: Distribution of Household Characteristics by Mobility

	Mobility (Number of Months Outside Village)						Difference between Columns 1 and 6 (7)	Percentage Change from Column 1 to Column 6 (8)
	Never (1)	1-6 Months (2)	7-12 Months (3)	13-18 Months (4)	19-24 Months (5)	24-35 Months (6)		
Panel A: All Individuals								
Household Size	5.81	6.27	6.58	6.82	7.14	7.50	1.70	29.19
# individuals 0-7	1.69	1.76	1.81	1.84	1.88	1.94	0.26	15.23
# individuals 8-15	1.52	1.65	1.73	1.79	1.86	1.95	0.43	28.15
# individuals 16-24	0.60	0.72	0.84	0.94	1.09	1.23	0.63	105.29
# individuals 25-49	1.26	1.37	1.44	1.47	1.53	1.59	0.33	26.12
# individuals 50+	0.74	0.77	0.77	0.78	0.78	0.79	0.05	6.77
Panel B: Male								
Household Size	2.73	2.96	3.13	3.23	3.37	3.53	0.80	29.44
# individuals 0-7	0.87	0.90	0.92	0.93	0.95	0.97	0.09	10.56
# individuals 8-15	0.82	0.88	0.92	0.95	0.98	1.01	0.19	23.40
# individuals 16-24	0.27	0.34	0.40	0.45	0.50	0.56	0.29	105.10
# individuals 25-49	0.40	0.45	0.50	0.53	0.56	0.60	0.20	49.78
# individuals 50+	0.36	0.37	0.38	0.38	0.39	0.39	0.03	8.74
Panel C: Female								
Household Size	3.28	3.52	3.68	3.82	4.01	4.23	0.95	28.99
# individuals 0-7	0.87	0.92	0.95	0.97	0.99	1.04	0.17	19.84
# individuals 8-15	0.76	0.83	0.86	0.90	0.95	1.01	0.25	33.18
# individuals 16-24	0.35	0.40	0.47	0.53	0.63	0.71	0.36	105.45
# individuals 25-49	0.90	0.96	0.98	0.99	1.02	1.04	0.14	15.75
# individuals 50+	0.41	0.42	0.42	0.43	0.43	0.43	0.02	5.06
Panel D: Household Characteristics								
Total Dependency Ratio	155.14	151.80	149.24	145.19	139.61	134.74	-20.40	-13.15
Child Dependency Ratio	141.43	138.77	136.40	133.07	128.11	124.96	-16.47	-11.64
Total household per capita income	23726	21262	19429	19197	18253	17307	-6418	-27.05
Total household per capita agricultural income	14175	12395	11392	11005	10472	9952	-4223	-29.79
Total household per capita non- agricultural income	10907	10044	9110	9242	8774	8307	-2600	-23.84

Notes: Moving from column 1 to 6 adds in additional individuals that were present at some point during the 1998 to 2000 time period but were absent for the number of months noted at the top of each column. Total dependency ratio is calculated as the number of children 0 to 15 plus the number of adults over age 65 divided by the number of adults 16 to 64. Child dependency ratio is calculated as the number of children 0 to 15 divided by the number of adults 16 to 64.

Table 3a: Individual Correlates of Mobility

Dependent Variable:	All individuals			
	(1)	(2)	(3)	(4)
	Individual Lives Outside Village	Individual Moves Out and Does Not Return	Individual Moves Out and Returns	Fraction Last 36 Months Away
Age				
Age 0-7	-0.286*** [0.024]	-0.149*** [0.015]	-0.068*** [0.018]	-13.885*** [1.617]
Age 8-15	-0.235*** [0.028]	-0.107*** [0.016]	-0.092*** [0.022]	-12.517*** [1.678]
Age 25-49	-0.264*** [0.034]	-0.093*** [0.020]	-0.108*** [0.024]	-12.618*** [1.964]
Age 50+	-0.413*** [0.040]	-0.105*** [0.020]	-0.248*** [0.029]	-18.954*** [2.640]
Marital Status				
Widowed, Divorced	-0.276*** [0.025]	-0.158*** [0.017]	-0.103*** [0.030]	-17.625*** [1.368]
Never Married	-0.229*** [0.042]	-0.227*** [0.027]	0.028 [0.024]	-15.073*** [2.685]
Occupation				
Non-agricultural occupation	0.290*** [0.037]	0.181*** [0.030]	0.068 [0.051]	16.265*** [2.638]
Education				
Highest Education Level is Primary	0.008 [0.018]	0.001 [0.012]	0.002 [0.016]	0.448 [0.995]
Highest Education Level is Secondary	0.160** [0.068]	-0.090*** [0.023]	0.194*** [0.053]	10.061** [3.864]
Relationship to Head				
Wife	-0.274*** [0.038]	-0.342*** [0.029]	0.124*** [0.031]	-19.204*** [2.400]
Head of Household	-0.090** [0.035]	-0.270*** [0.021]	0.233*** [0.032]	-10.444*** [2.562]
Other Relationship to Head	0.192*** [0.035]	-0.077*** [0.008]	0.287*** [0.029]	9.690*** [1.612]
Gender				
Male	-0.052*** [0.013]	-0.035*** [0.006]	-0.013 [0.010]	-3.453*** [0.762]
Number of individuals	4462	4462	4462	4462

Notes: Robust standard errors corrected for village clustering in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Omitted variable categories are: ages 16 to 24, married, biological child, and no education.

Table 3b: Individual Correlates of Mobility, By Gender

Dependent Variable:	Male Only				Female Only			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Individual Lives Outside Village	Individual Moves Out and Does Not Return	Individual Moves Out and Returns	Fraction Last 36 Months Away	Individual Lives Outside Village	Individual Moves Out and Does Not Return	Individual Moves Out and Returns	Fraction Last 36 Months Away
Age								
Age 0-7	-0.315*** [0.031]	-0.120*** [0.017]	-0.114*** [0.019]	-14.049*** [1.345]	-0.160*** [0.045]	-0.138*** [0.031]	0.025 [0.035]	-9.046** [3.356]
Age 8-15	-0.261*** [0.026]	-0.098*** [0.019]	-0.110*** [0.021]	-12.434*** [1.644]	-0.115** [0.042]	-0.082*** [0.026]	-0.022 [0.037]	-8.216** [2.785]
Age 25-49	-0.086 [0.076]	-0.016 [0.043]	0.059 [0.063]	2.244 [3.790]	-0.320*** [0.040]	-0.093*** [0.017]	-0.186*** [0.037]	-16.738*** [2.737]
Age 50+	-0.269*** [0.063]	-0.006 [0.038]	-0.136** [0.053]	-4.67 [3.686]	-0.421*** [0.045]	-0.101*** [0.018]	-0.283*** [0.033]	-21.076*** [2.860]
Marital Status								
Widowed, Divorced	-0.296*** [0.059]	-0.087** [0.034]	-0.179*** [0.034]	-14.430*** [3.942]	-0.365*** [0.065]	-0.267*** [0.050]	-0.083* [0.047]	-26.892*** [4.330]
Never Married	-0.054 [0.085]	-0.095 [0.054]	0.079 [0.058]	-5.856 [6.149]	-0.407*** [0.038]	-0.317*** [0.032]	-0.038 [0.025]	-23.868*** [2.991]
Occupation								
Non-agricultural occupation	0.301*** [0.027]	0.238*** [0.043]	-0.0003 [0.041]	15.938*** [1.946]	0.202* [0.101]	0.042 [0.043]	0.174 [0.100]	13.498** [5.508]
Education								
Highest Education Level is Primary	0.01 [0.028]	0.009 [0.014]	-0.014 [0.021]	0.044 [1.216]	-0.001 [0.037]	-0.013 [0.018]	0.027 [0.034]	1.086 [1.863]
Highest Education Level is Secondary	0.072 [0.083]	-0.079** [0.036]	0.093 [0.089]	4.74 [4.858]	0.338*** [0.080]	-0.076*** [0.025]	0.350*** [0.060]	19.969*** [5.992]
Relationship to Head								
Wife					-0.297*** [0.044]	-0.410*** [0.043]	0.192*** [0.045]	-19.978*** [2.987]
Head of Household	-0.104 [0.069]	-0.215*** [0.050]	0.117** [0.047]	-16.650*** [5.572]	0.062 [0.064]	-0.131*** [0.027]	0.271*** [0.067]	6.959* [3.747]
Other Relationship to Head	0.199*** [0.034]	-0.032** [0.014]	0.219*** [0.038]	7.922*** [1.587]	0.198*** [0.052]	-0.093*** [0.016]	0.335*** [0.041]	12.883*** [3.156]
Number of individuals	1997	1997	1997	1997	2465	2465	2465	2465

Notes: Robust standard errors corrected for village clustering in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Omitted variable categories are: ages 16 to 24, married, no education, and biological child.

Table 4: School Enrollment and Sibling Rivlary
Restricted to Children 8-15 of the Household Head

	Resident Siblings		All Siblings, Including Deceased		All Living Siblings		All Living Siblings Under 25	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample limited to resident children	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Siblings limited to resident children	Yes	Yes	No	No	No	No	No	No
Siblings limited to living children	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Siblings limited to age under 25	No	No	No	No	No	No	Yes	Yes
Age & Gender Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Panel A: All Children								
# Siblings	-0.022**	-0.020**	-0.009	-0.012	-0.007	-0.010	-0.012	-0.013
	[0.009]	[0.009]	[0.008]	[0.007]	[0.010]	[0.008]	[0.009]	[0.008]
# Female Siblings	0.021	0.023*	0.010	0.017	0.001	0.008	0.007	0.012
	[0.013]	[0.013]	[0.014]	[0.011]	[0.015]	[0.011]	[0.014]	[0.011]
Panel B: Boys								
# Siblings	-0.029**	-0.032***	-0.008	-0.015*	-0.006	-0.014	-0.014	-0.019*
	[0.011]	[0.009]	[0.008]	[0.007]	[0.012]	[0.010]	[0.010]	[0.009]
# Female Siblings	0.039**	0.042**	0.012	0.023	0.003	0.015	0.015	0.022
	[0.019]	[0.019]	[0.015]	[0.014]	[0.018]	[0.016]	[0.017]	[0.018]
Panel C: Girls								
# Siblings	-0.013	-0.007	-0.008	-0.009	-0.006	-0.005	-0.007	-0.005
	[0.014]	[0.014]	[0.011]	[0.010]	[0.011]	[0.010]	[0.014]	[0.012]
# Female Siblings	0.000	0.001	0.005	0.012	-0.006	0.001	-0.006	-0.002
	[0.019]	[0.018]	[0.018]	[0.014]	[0.017]	[0.014]	[0.021]	[0.017]

Notes: Robust standard errors corrected for village clustering in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Regressions include 804 children in Panel A, 418 children in Panel B, and 386 children in Panel C. Dependent variable is current school enrollment status of the child.

Table 5: The Presence of Biological Sisters and Income*Sample restricted to children 8-15 of the household head*

	All children		Boys Only	
	(1)	(2)	(3)	(4)
Household Non-Agricultural Income	0.196* [0.112]	0.207* [0.114]	0.129 [0.083]	0.136 [0.098]
Number of Female Siblings	0.657*** [0.058]	0.681*** [0.052]	0.685*** [0.084]	0.742*** [0.079]
Number of Biological Siblings	-0.062 [0.042]	-0.073* [0.035]	-0.093 [0.062]	-0.123** [0.053]
Age and Gender Effects	Yes	Yes	Yes	Yes
Village Fixed Effects	No	Yes	No	Yes
Number of children	746	746	390	390

Notes: Dependent variable is the number of biological sisters resident in the household. Robust standard errors corrected for clustering at the village level. Number of observations is less than in Table 4 because of missing income data for some households. Table 4 results are similar using the child sample in Table 5.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: School Enrollment, Sibling Rivalry, and Income*Sample restricted to children 8-15 of the household head*

	All Children		Boys Only	
	(1)	(2)	(3)	(4)
Sample limited to resident children	Yes	Yes	Yes	Yes
Siblings limited to resident children	Yes	Yes	Yes	Yes
Siblings limited to living children	Yes	Yes	Yes	Yes
Siblings limited to age under 25	No	No	No	No
Age & Gender Fixed Effects	Yes	Yes	Yes	Yes
Village Fixed Effects	No	Yes	No	Yes
# Siblings	-0.015 [0.011]	-0.015 [0.010]	-0.021* [0.012]	-0.026** [0.011]
# Female Siblings	0.013 [0.015]	0.015 [0.015]	0.028 [0.021]	0.033 [0.022]
Household Non-Agricultural Income	0.019 [0.013]	-0.003 [0.020]	0.026 [0.021]	0.002 [0.021]
Number of Children	746	746	390	390

Notes: Robust standard errors corrected for village clustering in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Dependent variable is current school enrollment status of the child