

Migration and Economic Mobility in Tanzania:

Evidence from a Tracking Survey

Forthcoming in the Review of Economics and Statistics

Kathleen Beegle
The World Bank

Joachim De Weerd
EDI, Tanzania

Stefan Dercon
Oxford University, UK

March 2010

This study explores to what extent migration has contributed to improved living standards of individuals in Tanzania. Using a 13-year panel survey, the authors find that migration between 1991 and 2004 added 36 percentage points to consumption growth. Although moving out of agriculture resulted in much higher growth than staying in agriculture, growth was always greater in any sector if the individual physically moved. As to why more people do not move if returns to geographical mobility are so high, the analysis finds evidence consistent with models in which exit barriers are set by home communities that prevent the migration of some categories of people.

JEL classifications: O15, O12, R23

* We thank Karen Macours, David McKenzie, Peter Neary, two anonymous referees, the editor of this journal and seminar participants at the Massachusetts Avenue Development Seminar, Oxford University, and the World Bank for very useful comments. All views are those of the authors and do not necessarily reflect the views of the World Bank or its member countries.

1. Introduction

Finding routes out of poverty remains a key issue for households and policy makers alike. A long-term vision of development suggests that poverty reduction is associated with intergenerational mobility out of rural areas and agriculture, and into urban non-agricultural settings. Physical and economic mobility seem to go hand-in-hand. Standard economic theory has multiple narratives of how physical and economic mobility interact. The Lewis model offers a stylized description of rural transformation, with sector mobility of labor from agriculture into “modern” production processes. At least in its original specification, the model suggests an initial gap in earnings between rural and urban locations (Lewis, 1954).¹ The Harris-Todaro model emphasizes the migration process and the fact that relative individual earnings incentives matter, so that both pull and push factors drive migration. A gap between rural and expected urban earnings drives migration. Unemployment (or an informal sector offering low earnings) would nevertheless allow an actual gap between urban and rural wages to persist, with the premium a function of the unemployment rate (Harris and Todaro, 1970). Other work, such as the “new economics of migration” (Stark and Bloom, 1985), emphasizes that migration is part of a general livelihood strategy for the initial household as a whole. Migration is part of a welfare maximizing strategy with a clear role for overall household income growth, but also a role for risk sharing. For example, Rosenzweig and Stark (1989) find that migration patterns for marriage in rural India are consistent with the risk-sharing strategies of the initial household. Recent evidence has highlighted not just the role of networks in facilitating migration from home areas, but also how migration is closely linked to migrants’ access to social networks in destination areas (Munshi, 2003) and to community rates of out-migration (Kilic et al., 2009).

Although the emphasis on the process of migration in most recent empirical work has provided many insights, few of these studies convincingly address the question of whether migration leads to improved

living conditions. A major problem is having access to data that allow a careful and convincing assessment of the relative welfare of migrants and non-migrants, due to the standard evaluation problem: an individual cannot be observed to be both a migrant and a non-migrant. A few studies have access to experimental data, such as international migration lotteries (e.g., McKenzie, Gibson, and Stillman, 2009), but most studies have to work with non-experimental data. Without experimental data, the key concern, unobserved heterogeneity affecting both outcomes and the process of migration, persists. This leads to the quest for imaginative and convincing instruments for migration (see the review of the migration and poverty literature by McKenzie and Sasin, 2007, and the references therein). An additional hurdle is the need for panel data to study migration and economic mobility. The costs and difficulties of re-surveying means that attrition may be relatively high for this group and may also result in the loss of some of the most relevant households for the study of this process (Beegle, 2000; Rosenzweig, 2003).

This paper uses unique data from a region in Tanzania to address the question: What is the impact on poverty and wealth of physical movement out of the original community? Although we do not have experimental data, the nature of our data allows us to limit the potential sources of unobserved heterogeneity. Building on a detailed panel survey conducted in the early 1990s, we re-interviewed individuals in 2004, making a notable effort to track individuals who had moved.

The tracking of individuals to new locations proves crucially important for assessing welfare changes among the baseline sample. The average consumption change of individuals who migrated was more than four times greater than that of individuals who did not move. Those who had moved out of Kagera by 2004 experienced consumption growth that was 10 times greater compared with those who remained in their original community. These averages translate into very different patterns of poverty dynamics for the physically mobile and immobile. For those who stayed in the community, the poverty rate decreased by about 4 percentage points over the 13 years. For those who moved elsewhere within the region, the poverty rate decreased by about 12 percentage points; and for those who moved out of the region, the

poverty rate decreased by 23 percentage points. Had we not tracked and interviewed people who moved out of the community – a practice that is not carried out in many panel surveys – we would have seriously underestimated the extent to which poverty decreased during 1991 to 2004 in Kagera; we would have reported poverty reduction at about half its true value. Clemens and Pritchett (2008) raise similar concerns in the context of income growth and international migration. Furthermore, the tracking and re-interviewing enabled us to collect valuable information about pathways out of poverty.

Still, these statistics do not provide evidence that moving out of the community leads to higher income growth. As noted above, we cannot observe the counterfactual: what would income growth have been for migrants had they not migrated? We exploit some unique features of the data to address concerns about unobserved heterogeneity. First, individual fixed effects regressions for movers and stayers produce a difference-in-difference estimation of the impact of physical movement, controlling for any fixed individual factors that affect consumption. Second, we can control for initial household fixed effects in the growth rate of consumption because we observe baseline households in which some individuals migrate and others do not. This controls for observable and unobservable factors fixed to the family that can affect the growth rate of consumption. Thus, we identify the impact of migration on income using within-household variation in migration. Unlike most studies of migration, our identification does not rely on household shocks, distances to possible destinations, or the existence of family networks at the destination to identify the migration decision. Such variables are likely to impact the income of those migrating as well as those staying behind, and so the exclusion restriction will not be satisfied. In our study, we are able move beyond these approaches; in addition to using panel data on migrants and non-migrants, we compare siblings and other relatives who were living together at the baseline.

These estimations address many possible sources of heterogeneity, such as (genetic) health and ability endowments, risk aversion, wealth constraints, and market, risk and environmental circumstances. We find that movement out of the community results in 36 percent higher consumption relative to staying.

Comparison of the results with and without fixed effects suggests that migrants are more likely to be from families with greater potential for growth in earnings.

However, a weakness of this approach is the implicit assumption that *within* families migration is random, which is a strong assumption. For example, in view of the standard Harris-Todaro model of individual migration, earnings differentials drive migration, so those who are observed to have migrated from within a household tend to have had greater earning potential than those who stayed behind, implying that within-family migration may not be random.² We use two-stage least squares (2SLS) methods to deal with this potential endogeneity. We assert that opportunities to migrate depend on the interaction of household circumstances with the individual's status and position *within* the household at baseline. The 2SLS estimates show limited evidence of unobserved individual heterogeneity affecting consumption growth. In short, unobservables at the household level correlated with growth potential appear to matter, whereas individual heterogeneity does not.

We explore two additional avenues of interest. First, does migration to urban areas drive the results? Second, does migration capture changes in the sector of work that would explain the consumption growth we observe? We find suggestive evidence that physical mobility has an independent effect beyond its association with moving out of agriculture or moving to a more urban area. We use these results in conjunction with the literature on network externalities and poverty traps to explain why, if migration has such large pay-offs, more people do not move. We conclude that the findings are consistent with models in which exit barriers are set by home communities (through social and family norms), preventing migration of certain categories of people when windows of opportunity arise. Being willing and able to leave behind what you know appears to be a strong determinant of economic mobility. There is no evidence of financial constraints to migration.

In the next section, we provide the context of changes in economic fortunes in Tanzania in the past decade. Section 3 presents the data used in the analysis and Section 4 provides the basic indicators we use to assess economic and welfare changes. Section 5 briefly describes the method we use to assess the impact of migration, Section 6 presents the results, and Section 7 carries out some robustness checks. Section 8 builds a narrative around the regressions and aims to explain why more people do not migrate when the benefits of doing so are so high.

2. The Setting: Tanzania and Kagera, 1994-2004

During 1994 to 2004, Tanzania experienced a period of relatively rapid macroeconomic growth, attributed to liberalization, a renewed trade orientation, a stable political context, and a relatively positive business climate that helped to boost economic performance. Real GDP growth was on the order of 4.2 percent per year between 1994 and 2004, while annual population growth was around 3.2 percent (URT, 2004). There is also evidence that growth accelerated in the last few years of the period compared with the 1990s. However, growth was not sufficiently broad-based to result in rapid poverty reduction. On the basis of the available evidence, poverty rates declined only slightly and most of the progress in poverty reduction was in urban areas. According to the Household Budget Survey (HBS), between 1991 and 2000/01, poverty declined from 39 to 36 percent in mainland Tanzania. The decline in poverty was steep in Dar es Salaam (from 28 to 18 percent) but minimal in rural Tanzania (from 41 to 39 percent).

For the purposes of this study, it is useful to consider the Kagera region specifically. The region is far from the capital and the coast, bordering Lake Victoria, Rwanda, Burundi, and Uganda. It is overwhelmingly rural and primarily engaged in producing bananas and coffee in the north and rain-fed annual crops (maize, sorghum, and cotton) in the south. Relatively low-quality coffee exports and agricultural produce are the main sources of income. Mean per capita consumption was near the mean of mainland Tanzania in 2000. Likewise, the region appeared to mirror the rest of the country in terms of

growth and poverty reduction: real GDP growth was just over 4 percent per year between 1994 and 2004, while poverty in Kagera is estimated to have fallen from 31 to 29 percent between 1991 and 2000/01 (Demombynes and Hoogeveen, 2007).

The challenges of poverty reduction in Kagera seem to be representative for provincial Tanzania as a whole: some pockets, such as Dar es Salaam, have had substantial growth and poverty reduction, but this has not spread to other areas. This reflects the typical problem of land-locked, agriculture-based economies: how to deliver poverty reduction if the main engine of growth appears to be elsewhere (De Weerd, 2010).

3. The Data

The Kagera Health and Development Survey (KHDS) was originally conducted by the World Bank and Muhimbili University College of Health Sciences (MUCHS), and consisted of about 915 households interviewed up to four times from fall 1991 to January 1994 (at 6-7 month intervals) (see World Bank, 2004, and <http://www.worldbank.org/lsm/>). The KHDS 1991-1994 serves as the baseline data for this paper. Initially designed to assess the impact of the health crisis linked to the HIV-AIDS epidemic in the area, it used a stratified design to ensure relatively appropriate sampling of households with adult mortality. Comparisons with the 1991 HBS suggest that in terms of basic welfare and other indicators, the KHDS can be used as a representative sample for this period for Kagera (although not necessarily for the rest of Tanzania; the results are available upon request).

The objective of the KHDS 2004 survey was to re-interview all individuals who were household members in any round of the KHDS 1991-1994 and who were alive at the last interview (Beegle, De Weerd and Dercon, 2006). This effectively meant turning the original household survey into an individual longitudinal survey. Each household in which any of the panel individuals lived would be administered

the full household questionnaire. Because the set of household members at the baseline had subsequently moved, and usually not as a unit, the 2004 round had more than 2,700 household interviews (from the baseline sample of 912 households).

Although the KHDS is a panel of respondents and the concept of a “household” after 10-13 years is a vague notion, it is common in panel surveys to consider re-contact rates in terms of households. Excluding households in which all previous members were deceased (17 households with 27 people), the field team managed to re-contact 93 percent of the baseline households. This is an excellent rate of re-contact compared with panel surveys in low-income *and* high-income countries. The KHDS panel has an attrition rate that is much lower than that of other well-known panel surveys summarized in Alderman et al. (2001), in which the rates ranged from 17.5 percent attrition *per year* to the lowest rate of 1.5 percent per year, with most of these surveys covering considerably shorter time periods (two to five years).

Figure 1 charts the evolution of households from the baseline to 2004. Half of all households interviewed were tracking cases, meaning they did not reside in the baseline communities. Of those households tracked, only 38 percent were located nearby the baseline community. Overall, 32 percent of all households were neither located in nor relatively close to the baseline communities. While tracking is costly, it is an important exercise because migration and dissolution of households are often hypothesized to be important responses to hardship and a strategy for escaping poverty. Excluding these households in the sample raises obvious concerns regarding the selectivity of attrition. In particular, out-migration from the village, dissolution of households, and even marriage may be responses to changing economic or family circumstances. Tracking surveys provide a unique opportunity to study these responses: who uses them, their effects, and whether they get people out of poverty.

Turning to the re-contact rates of the sample of 6,352 respondents, Table 1 shows the status of the respondents by age group (based on their age at first interview in the 1991-1994 rounds). The surviving

older respondents were much more likely to be located, which is consistent with higher migration rates among the young adults in the sample. Among the youngest respondents, more than three quarters were successfully re-interviewed. Excluding people who died, 82 percent of all respondents were re-interviewed. Table 2 shows the location of the respondents. Without tracking, re-interview rates of surviving respondents would have fallen from 82 to 52 percent (2,780 of 5,394 survivors). Non-local migration is important: restricting the tracking to nearby villages would have resulted in 63 percent re-contact of survivors. Migration also proved to be an important factor in determining whether someone was re-contacted. Respondents who were not traced were much more likely to reside outside Kagera (43 percent) compared with their counterparts who were re-interviewed (8 percent).

The consumption data come from an extensive consumption module administered in 1991 and again in 2004. The consumption aggregate includes home produced and purchased food and non-food expenditure. The non-food component includes a range of non-food purchases, as well as utilities, expenditure on clothing/personal items, transfers out, and health expenditures. Funeral expenses and health expenses prior to the death of an ill person were excluded. Monetary levels were adjusted to account for spatial and temporal price differences, using price data collected in the Kagera survey in 1991 and 2004, and, for households outside Kagera, data from the National Household Budget Survey. Consumption is expressed in annual per capita terms. The poverty line is set at 109,663 Tanzanian shillings (TSh), calibrated to yield for our sample of respondents who remained in Kagera the same poverty rate as the 2000/01 National Household Budget Survey estimate for Kagera (29 percent). At the time of the survey one US dollar was worth around TSh 1,100.

4. Growth, Poverty, and Physical Mobility in Kagera

In this section, we discuss changes in living standards overall, and the changes for four mutually exclusive groups based on residence in 2004: (i) still residing in the baseline community, (ii) residing in a neighboring community, (iii) residing elsewhere in Kagera, and (iv) residing outside Kagera.

Table 3 shows that the basic needs poverty rate declined 8 percentage points in the full sample. This figure masks significant differences in changes between subgroups based on migration. For those found residing in the baseline community, poverty rates dropped by 3 percentage points, but rates dropped by 11, 12, and 23 percentage points for those who moved to neighboring communities, elsewhere in Kagera, and outside Kagera, respectively. A similar pattern is found for consumption per capita. Although mean consumption per capita grew by TSh 61,903 overall, or 38 percent, it grew by only 17 percent for those found in the same community and by 37, 56, and 161 percent for those who moved to neighboring communities, elsewhere in Kagera, and outside Kagera, respectively. Dividing consumption into food and non-food components gives the same result. The most basic assessment of welfare changes would have been wrong if we had focused only on individuals still residing in the community, a practice found in many panel data surveys. We would have underestimated the growth in consumption by half of its true increase.

For the groups in Table 3, the differences in consumption changes are statistically significant, as shown in Table 4. Excluding respondents who have relocated would omit those with greater rates of income growth and poverty reduction. Table 5 reports confidence intervals for the incremental samples (which are not mutually exclusive); it gives a more detailed picture of how inference on consumption growth and poverty reduction would have changed if we had not tracked movers. It is apparent that inference from a “simple” panel survey of respondents continuing to reside within the original communities would have produced underestimates of actual consumption growth and poverty reduction in this population.

These conclusions are robust across the distribution of consumption, as well as at the mean and the poverty line. Panel A in Figure 2 depicts the cumulative density function for consumption per capita for those people who remained in the same community. Panels B, C, and D show the cumulative density functions for respondents residing in neighboring communities, elsewhere in Kagera, and outside Kagera. For respondents who were located further from their location in 1991, the difference between the functions for 1991 and 2004 are more pronounced. For people who remained in the baseline community, the 1991 and 2004 distributions lie close to each other under the poverty line and diverge above it; for the other mobility categories, there is greater divergence.

Figures 3, 4, and 5 offer another cut of the data, comparing consumption of non-movers to movers in 1991 when both were living in the same community (Panel A) and in 2004 (Panel B). There is almost no difference between (future) non-movers and movers in 1991, but by 2004, we observe divergent income levels. The divergence is greater between those who stayed and those who moved further away (i.e., Figures 4 and 5).

What drives the association between migration and income growth? One plausible explanation is that migrants are relocating to less remote, less poor areas. By 1991, 68 percent of the sample was living in rural villages, of which a little more than half were categorized by the survey team as poorly connected in terms of infrastructure. The remainder of the sample were living in (or close to) the regional capital, Bukoba (17 percent) or other small urban centers in Kagera (14 percent). Table 6 investigates whether moving to a better connected center (e.g., from a poorly connected to a better connected village, or from a rural area to an urban center) is correlated with higher consumption growth.³ This is indeed the case: about 10 percent of the sample moved to a better-connected area and they experienced 90 percent consumption growth, on average. For those who moved to a similar area, consumption increased by 46

percent on average, while those who moved to a less urban or less-connected center experienced a lower increase at 28 percent. Clearly, it matters where people move, but moving seems in itself to matter too.⁴

Another plausible source of income growth for migrants is that they have moved to a different sector with respect to income. In Table 7, we explore whether migration is correlated with change in occupation or sector. Consumption growth was highest for those who moved into non-agriculture (67 percent), and there was considerable growth for those who started in non-agriculture. It is striking that the 10 percent who actually moved into agriculture from non-agriculture faced declining consumption, suggesting that this is a sign of hardship and possibly a means of coping with it. Table 8 reports consumption growth by both sector change and migration. A considerable number of people switched sectors without migrating but, within each category of sector status, migrants had much higher consumption growth than non-migrants. The main source of income matters for consumption growth, but it is strongly related to migration as well. For example, those who moved out of agriculture while also moving out of their original community in this period more than doubled their consumption levels, while those who switched into agriculture while staying within the community faced a 12 percent reduction in consumption.

5. Assessing the Impact of Migration on Consumption Outcomes

The correlations above do not resolve whether this consumption growth is in fact directly related to migration or whether it is spurious. To investigate this further, we explore several empirical approaches. First, we employ a difference-in-difference estimator, comparing the consumption growth of those who moved with those who stayed in their baseline community. We define $\ln C_{it}$ as the natural logarithm of consumption per capita for individual i in period t , and M_i as a dummy that is one if the individual was found to have physically moved out of the original community between t and $t+1$, and zero otherwise. The difference-in-difference specification is:

$$\Delta \ln C_{it+1,t} = \alpha + \beta M_i + \gamma X_{it} + \delta_{ih} + \varepsilon_{it} \quad (1)$$

in which $\Delta \ln C_{it+1,t}$ is $(\ln C_{it+1} - \ln C_{it})$, the growth rate of consumption per capita in the household in which i is residing in the two periods. This specification controls for individual fixed heterogeneity, which might impact the level of consumption in each period. This resolves a large number of possible sources of endogeneity, such as risk aversion or ability, which are likely to affect both migration and income outcomes. However, it does not address concerns about heterogeneity among families or individuals affecting growth in consumption and the migration decision. For example, current wealth may affect the ability to migrate as well as the potential to grow between t and $t+1$. McKenzie and Sasin (2007) discuss at length the issue of endogeneity with respect to measuring the impact of migration on poverty, stating that work that does not identify causal relations provides “rather weak grounds for policy recommendations.” McKenzie et al. (2009) find that ignoring selection led to overstating the gains from migration from Tonga to New Zealand.

Our data, while not experimental, still offer excellent opportunities to control for a wide set of factors in this respect. First, we have data on multiple individuals from the original household, which allows us to control for any initial household level heterogeneity (δ_{ih}) that may affect the growth of consumption by estimating (1) using “initial household” fixed effects (IHHFE). The result is that the impact of migration is identified using within initial household variation – differences between members of the same initial household, effectively controlling for initial growth paths. Second, we can control for a set of individual level factors that may affect consumption growth and possibly as well migration by including these as X_i in regression model (1). The variables used as individual conditioning variables for the growth of consumption from baseline are individual variables (sex, age, education relative to age-specific peer groups,⁵ and marital status) and family background variables (the number of biological children in the initial household at baseline interacted with the age-sex group of the children, the number of biological children living elsewhere interacted with the distance to the regional capital, and the years of education of

the biological mother and father). We also include a variable indicating whether the individual lost both parents between 1991 and 2004, allowing a separate effect if the individual was below age 15 at baseline. Quite a few of these variables – such as educational levels, marital status, parental death, or having children living elsewhere (offering opportunities for remittances) – are likely to affect migration, but also may have direct effects on consumption growth.

Despite controlling for fixed individual heterogeneity and both fixed and time-varying household-level heterogeneity (including initial growth paths) and the additional control variables, unobserved individual factors may still affect migration as well as consumption growth. We extend the analysis to 2SLS estimates, using three types of variables for instruments for the migration decision: pull-factors, push-factors, and variables reflecting social relationships.

The pull-factors include age and baseline location. Migration opportunities and incentives are typically stronger for young male adults, as employment in low skill and physically demanding activities is likely to be easier for them. Similarly, if a family were to decide on who should migrate to capture opportunities, then allowing a young male adult to go would seem sensible. Costs and information needs for migration may well be affected by how far the opportunities are located. We include an interaction term of the distance to the regional capital and whether the person is male and between 5 and 15 years old at the baseline (so between 18 and 28 in 2004) as a measure of the opportunities available.⁶

Individuals may also be pushed into migration (or families may decide to send someone) when shocks occur. We include a measure of economic shocks experienced by the household by including a measure of negative rainfall shock. Using data from 21 weather stations in Kagera from 1980-2004, each of the 51 baseline villages was mapped to the nearest station and 25-year average annual rainfall was computed. The largest deviation of rainfall between 1992 and 2002 from the long-run average was identified. This

rainfall shock variable was interacted with being in the 5-to-15 age group as a measure of this “push” factor (with higher values defined as high-deviation rainfall).

Finally, norms and social circumstances are likely to affect migration. In particular, within a household, who is able or expected to migrate is likely to be determined by the individual’s position in the household. We include indicators for being the head or spouse of the household head at the baseline. We expect these two positions in the household would make it less likely that the person would leave relative to others in the household. Age rank among those between 5 and 15 (with the oldest receiving the highest value) is also included. These indicators are unlikely to determine the consumption growth of the household, but may well affect whether a person is allowed, chosen, or chooses to migrate. Lastly, close family members, the closest relatives of the household head, sons and daughters, may have different probabilities of leaving the household’s community, compared with other residents, such as cousins or nephews. Local norms on marriage are patrilocal: girls are expected to move to the community of their husbands after marriage and husbands are expected to stay where their father was based. We include an indicator for being the son of the household head at baseline. Although both sons and daughters of the head may be expected to be more likely to stay in the community than other initial household members, patrilocality would make this probability higher for boys than for girls.

In sum, this means we are using a set of six instruments. Although we will show in the Appendix statistically convincing and close to identical results are obtained by using subsets of these instruments, we focus on the full set of instruments in the discussion of the results.

Although our main measure of migration (M_i) is an indicator for having moved, we also substitute this for the log of the distance moved (kilometers from the original community of the location in which the individual was found in 2004, “as the crow flies,” set to 0 for non-movers). We will also extend the

multivariate analysis to explore whether moving to a more urbanized area or changing employment sectors plays a role in increasing consumption growth.

6. Regression Results

Table 9 presents the basic results for the initial household fixed effects (IHHFE) and 2SLS estimates. (Appendix Table 1 presents the means and standard deviations for the covariates.) We estimate the regressions using an indicator for having moved and a measure of the distance of the move. The 2SLS estimates in columns 3 and 4 use the six instruments defined above. In Table 10, we present the first-stage results of regressions explaining migration or the distance travelled in migration.

Before turning to the variables of interest, we will briefly discuss the coefficients on the control variables. Recall that all effects are identified using *within* initial household variation. Those who are relatively better educated at baseline, relative to their peers and within the household, experienced much higher consumption growth – and the effect is strongly convex. Having an educated father has an additional effect on growth. The younger cohort did considerably better, as did males still unmarried at baseline.

Turning to the migration variables, we observe in the IHHFE regression that there is a larger and statistically significant impact of migration on consumption growth. Moving out of the community resulted in a 36-percentage-point increase in consumption growth over the 13-year period. As migrants move further from their baseline community, the impact is greater. These effects are large, with migration resulting in large divergence in income between people who initially lived together, usually parents, siblings, and other close relatives. Because this is the impact comparing within families, it nets out any transfers from migrants to non-movers. That is, if migrants sent remittances back to their origin households, then the estimates in Table 9 are a lower bound of the impact of moving (see also the results in next section on alternative definitions of the consumption aggregate, excluding transfers out). It also

seems counter to the theory that the migration decision is part of a household-level maximization strategy (although it cannot preclude that this is partly true).

For the first-stage results in Table 10, in terms of basic diagnostics, our set of excluded instruments appears strong and valid: the Cragg-Donald (F) test shows a value of 11.70 for the movement dummy and 9.07 for the distance regression. Especially in the former case, it is comfortably above the level of 10 often recommended for rejecting weak instruments (and in the latter case, still with relative limited bias in the Stock and Yogo, 2002, tables). The results are also robust to exclusion of any of the instruments (Appendix Table 2).

Some interesting patterns explaining migration emerge from Table 10. First, education offers strong and convex effects in leaving one's community. Being unmarried, especially being female and unmarried, is correlated with a higher probability of migration (consistent with patrilocality, whereby females move out of the paternal location at the time of marriage). When looking more specifically at the identifying instruments, we find significant effects, consistent with expectations: positional variables in the household matter, with the head and spouse less likely to leave, as are children of the head (relative to others belonging to the household). The effect is however considerably larger (more negative) for male children of the head – again consistent with patrilocality, as marriage norms make sons more likely to be expected to stay in the community than daughters. Older members among the children in the household are more likely to migrate, possibly reflecting some kind of pecking order, given the opportunities available. Rainfall shocks increase the probability of leaving. Finally, pull factors, like the interaction of being young, male, and residing close to the regional capital, increase the probability of leaving. The results are also consistent for the regressions with the dummy variable for migration and with the distance migrated variable. In short, although not aiming to obtain a structural model, we find suggestive correlates for the process of migration from within households. These include better income opportunities

(education and distance to the regional capital), norms of settlement and marriage, and other social factors.

The 2SLS results (IV with fixed effects) in columns 3 and 4 of Table 9 are almost identical to the IHHFE results. They are slightly less statistically significant (as can be expected from IV regressions given their lower efficiency), but still significant at 5 percent. Thus, there is no evidence that unobserved individual time-varying heterogeneity affects the non-instrumented results. For the distance variables, the results are marginally smaller (the coefficient is 0.10 compared with 0.12), suggesting limited evidence of a positive bias in the earlier results (i.e., migrants travelling longer distances are those with somewhat higher unobserved consumption growth potential, consistent with expectations). Still, the difference is remarkably small. The conclusion is strong: being able to move out of the village or community appears to be an important factor for consumption growth. If those who moved had stayed behind, our evidence suggests that they would not have done as well.

The fact that there is little change going from the IHHFE to the 2SLS results does not suggest that there is no selection process in the migration decision. For example, it could be expected that more able people migrate. There is some evidence that this occurs, yet this heterogeneity is not at the individual level but at the household level. Estimating the 2SLS without IHHFE increases the coefficient on migrant status by almost a third (from 0.37 to 0.57). This is consistent with the proposition of positive selection among households: individuals from households with *high* earning potential migrate. *Within* the household, there seems to be no unobserved heterogeneity in terms earning potential among those who do or do not migrate.

These results are not driven by the lack of a parsimonious set of instruments or relatively weak instruments. The results are similar when restricting the instrument set. When focusing only on the relational variables (head, spouse, son, daughter), the Cragg-Donald (F) statistics become 14.5 and 11.3,

the Sargan is not rejected, the coefficient on physical movement stays at 0.36, and the distance variable becomes 0.097, virtually identical to the results in Table 9.

The validity of our interacted instruments assumes that they do not capture different growth rates (e.g., because of different labor markets) across these groups within households. Although growth rates might be influenced by the distance to the regional capital, rainfall shocks, gender, age, etc. additively, there is no evidence to suggest that the interaction of these would capture different growth rates outside the migration effect. To explore this point, we exploit the fact that the 1991-1994 baseline data consist of four waves. The wave 1 data were used as the baseline for this paper because the consumption recall period was identical to the follow-up survey (wave 5). We use the three interim waves (2-4), which have similar recall periods to check the validity of our interacted instruments. Using a measure of annual consumption per capita growth for 1992-1993, we can check whether our instruments, appropriately defined for this period, jointly or individually explain the baseline consumption changes. We find that they do not, giving further confidence that the exclusion restriction is valid for our instruments: the instruments do not influence growth except through migration. Of course, this regression of baseline growth rates on our instruments can only be valid if migration can be plausibly omitted from it. We do find that 1992 was the year with the lowest and 1993 the third lowest migration rates of all the years between 1992 and 2004, suggesting that the omission of the migration variable from the regression should not lead to specification errors. As can be expected, the same exercise for the regressions in Table 9, with the endogenous moved variable replaced directly by the instruments, does yield jointly significant instruments (at 10 percent).

7. Robustness

We perform a variety of checks to verify the robustness of the findings. First, we use alternative definitions of the consumption aggregate, in particular excluding transfers out, which could be an

important driver of our results if remittances to one's origin village are large. We have data on transfers sent between the 2004 households of the same origin. The size of these remittances is on average only small a percentage of total consumption. Our findings are robust to excluding this component.

Second, we check the role of the configuration of the data. Our outcomes are household-level measures of consumption per capita in levels and growth, assigned to individuals. We re-structure the data to the 2004 household level in Appendix Tables 3 and 4 (using average characteristics as controls and appropriately defined household-level aggregated instruments). The results are similar and consistent irrespective of analyzing the data at the individual or household level.

Third, concerns may be raised that changes in household size and composition in new households in 2004 are driving the results. Appendix Table 5 shows that migrant households are smaller in terms of members or adult-equivalent members. Appendix Table 6 repeats the analysis using adult equivalent units rather than household size as the denominator and finds essentially similar results

Fourth, we investigate whether lack of common support drives the results. The coefficients in the IHHFE regressions are identified from the sample households that had "split up" from the baseline. Restricting the sample to the 2,940 individuals from at least two split-offs in 2004 yields identical results in both the IHHFE and 2SLS estimations. We further refine this by examining the sample of individuals from origin households that split off into at least one household that moved by 2004 (N=2,520) and the sample of individuals from origin households that had at least one split-off that remained in the village (N=2,777). These samples yield identical results, both for IHHFE and 2SLS. Restricting the sample further to baseline households that had at least one split-off that moved and one that remained in the village (N=2,357) yields identical IHHFE results, but has 2SLS estimates of 0.23 and 0.68 for the migration indicator variable and distance variables, significant at 10 percent and 7 percent, respectively, and with IV diagnostics that remain sound. Taken together, these sample restrictions do not cast doubt on the validity

of the results, although they suggest that the size of the effects may be slightly lower than indicated in Table 9.

As an alternative to the fixed effects model and the two-stage estimation, we investigate a number of matching models. Of course, the advantage of matching techniques is that they ensure comparison of like-with-like, with less restrictive functional form assumptions and omission of non-comparable observations. The disadvantage is that they ignore potential unobservables (such as ability) that drive selection into migration. Across an array of different matching techniques (Gaussian, nearest-neighbor, Epanechnikov), the main findings on the impact of migration on consumption gains are remarkably robust to the results in Table 10 (results not presented).

Finally, we examine the role of time-varying factors, specifically education. If migration itself is the result of the individual's efforts to increase his or her level of education, we might be capturing the gains to a migration-education bundle, rather than to migration *per se*. For secondary schooling, often in the form of boarding schools in Tanzania, this is a plausible concern. At the primary level, few people migrate for schooling opportunities and tertiary education is limited for this sample. Descriptively, we find that the likelihood of a move is not correlated with additional grades of schooling conditional on age, suggesting that the moves we observe are not specifically driven by demand for education. Exploring this further, we repeated the regressions in Table 9 but this time included the years of education gained between rounds, as well as the interaction of schooling gains with the migration variables. We have to be cautious in interpreting these results, as surely migrating between rounds and years of education gained between rounds are bound to be joint decisions. Nevertheless, the results can at least explore whether the observed premium is just driven by education gains. Table 11 shows the results (only reporting the migration and education variables, but the independent variables are otherwise identical to those in Table 9).⁷ Our findings are robust to including education gains. The results in columns 1 and 2 indicate that even controlling for educational gains, the premium remains high and virtually identical to the earlier results.

Including the interaction terms in columns 3 and 4 shows nevertheless that some of the gains may well work through education: the returns to moves are higher for those who added more years of schooling, although for those without additional schooling since the baseline, the returns to movement are still considerable.

8. Migration Incentives, Social Constraints, and Windows of Opportunity

The regressions in Tables 9 and 10 are suggestive about how the relatively traditional and tightly-knit society of Kagera reacted to growing economic opportunities in the past decade. There is substantial movement out of these communities and those moving capture a substantial premium when measured in consumption terms. At the same time, the high premium may indicate that there are opportunities that have not been taken.

In this section, we build a narrative around these results in four steps. First, we argue that there are windows of opportunity that arise over time and space in the region and people need to move in order to take advantage of these opportunities. Second, we further complete the discussion of the drivers of migration; the regressions in Table 10 used household fixed effects and, therefore, do not reveal correlates of the constraints to migration at the household and community level. Third, we discuss how social norms can prevent some people from moving. Finally, we argue why such welfare-reducing constraints may be imposed by society on its members, thus providing a potential answer to the question of why more people do not move if the pay-offs are so high. At the end of this section, we discuss what our results imply in terms of standard models of migration, and qualify this discussion by offering a few alternatives that cannot be rejected given the data available.

The economic landscape in the Kagera region, as in other regions in Africa, has been changing in the past two decades. Growth opportunities are continually being introduced and eliminated across time and

space, as the refugee crisis abates, links with war-ridden bordering countries change, and more localized negative and positive shocks manifest themselves with various degrees of severity. People need to be physically (geographically) mobile in order to respond to the opportunities. To elaborate on this point, we decompose the results of Table 9 further to examine the role of the location of a move (more or less remote areas) and moves associated with sector changes (i.e., out of agriculture into non-agricultural activities). Table 12 disaggregates the migration variable into three categories of migration (moving to a more or less connected or urbanized area). Even moving to a less-connected area is still correlated with higher growth compared with not moving, but moving to a more-connected area results in consumption growth that is 66 percentage points higher than not moving. The same result is found using the distance of the move in the second column. Although where individuals move matters for the magnitude of the effect, any movement has the potential to be welfare improving.

In Table 13, we interact migration with change in sector (out of agriculture). We pool people who moved out of agriculture and those who remained in non-agriculture; both groups had statistically indistinguishable findings in all regressions. The first column shows that moving out of agriculture is strongly linked to higher consumption growth (as noted above in the descriptive statistics). The next two columns show that there is a large and positive impact of moving, even after controlling sector shifts; there is also a strong interactive effect of this sector shift with physical movement out of the village. In other words, it is not just the move out of agriculture that accounts for the large growth differential, but migration as physical movement out of the village has strong additional and complementary effects.

Tables 12 and 13 thus show that movement in itself is important. A logical – for economists perhaps even tautological – consequence of this is that constraints to movement are impediments to growth for whomever they happen to constrain. To estimate the migration premium, the regressions in Tables 9 and 10 use initial household fixed effects. Any initial household and community characteristics are therefore a black box. Although this improves inference regarding the migration premium, it also offers an

incomplete narrative of why certain people migrate and therefore why the migration premium remains so high. To explore this, we replaced the initial household fixed effect in Table 10 by a set of household and community variables as measured at baseline. We include characteristics of the head of the household (age, sex, and years of education), household size, a dummy for whether the household is a farm household, and wealth characteristics (land for cultivation, consumption per capita, the value of the physical capital stock, and the flooring quality in the dwelling). The community variables included are the cluster (village) means of all the above characteristics, the distance to Bukoba (the regional capital), and whether the community is remote (defined as an area not closely connected to an urban center).

The results are shown in Table 14 (only the additional variables are reported as inference on the role of individual characteristics was superior in Table 10). The table offers the marginal effects from a probit regression on whether an individual moved and the coefficients from an OLS regression explaining the logarithm of the distance migrated.

We find suggestive evidence of some factors that matter for the migration decision. First, we find some weak evidence that migration is higher in communities where farming was still the most important activity for a higher number of families, at least with respect to the distance of moves (column 2). The median community in the sample has 83 percent of households mainly involved in agriculture. However, controlling for this, being a farmer decreases the probability of any move and the distance migrated. Those with more land do not migrate as far as those with less land. Taken together, this suggests that migration is to some extent driven by a move out of agricultural settings, but those not involved in agriculture and faced with land pressure are more likely to move away and move farther.

Second, the regression results reported in Table 10 show that education matters in explaining the migration of the individual, but Table 14 shows that the educational level of the family or community does not matter: there is no more or further migration from households or settings with more education.

Similarly, although living in better connected areas is positively associated with the probability of moving, it is not associated with distance, and being closer to the regional capital is not associated with either migration measure.

Finally, are there any wealth or credit constraint effects? Migration may be a costly activity, requiring a serious investment. It is also an indivisible investment. Thus, some households may not be able to afford any migration to take advantage of the high return or can only afford to have some members migrate. From Figures 3, 4, and 5, the samples of future movers and non-movers started off with relatively similar wealth distributions. The evidence from Table 14 is mixed. Among the household indicators, only the value of physical assets is associated negatively with moving. Neither consumption per capita nor good flooring is associated with migration. None of the community-mean wealth variables are statistically associated with migration. Overall, there is little evidence of credit or wealth constraints; if anything, there is a tendency for more migration from poorer households.

This still leaves open the question as to why more people do not migrate, given the high returns. The *individual* level variables in the regressions in Table 10 suggest that particular types of people within families can go and not others. Factors include individual education (with a convex effect), being unmarried and female (consistent with considerable migration for marriage by girls), being of a particular age group when rainfall shocks occur, and a series of positional variables in the household (including an age pecking order and gender effects). Distance to the regional capital, a migration pull factor, matters specifically for young males.

There is scope for further interpretation of these findings within the local social context. The regressions suggest that an individual needs to be in a position to move in order to take advantage of geographic and time-specific economic opportunities, while at the same time there are a number of crucial social constraints in place that may prevent an individual from doing so. Social and family norms interacting

with pull (nearby towns) and push (shocks) factors are determinants of who may be allowed (or chosen) to move.

There appear to be windows of opportunity – being in the right place at the right time – that certain categories of people can take advantage of: not having social and family constraints in a window of time when physical mobility has large pay-offs. Missing these windows implies remaining trapped in a low-return environment. But this still begs the question of why we do not see more migration given these high returns, and why barriers remain in place if they are welfare-reducing.

Our results are consistent with the literature that links network externalities to poverty traps, and so endogenizes exit barriers in the village. In Hoff and Sen (2006), the kinship group decides how high to set the exit barrier for its members. They start from the observation that kin who have moved and remain loyal to their kinship group at home will sometimes need to undertake actions with negative consequences for their employers (securing jobs for kin) or landlords (sharing housing), etc. This creates an entry-barrier for anyone with obvious, strong kinship ties to their home village. In order to overcome such entry barriers, an individual may have to sever ties with his or her kinship group, implying the loss of a productive element (from the kinship group's point of view). To avoid this *ex ante*, the kinship group may decide to manipulate exit barriers – raising them through social norms about migration in order not to lose productive members. Hoff and Sen's model finds that it may be in the interest of the kinship group to prevent some of its members from taking advantage of economic opportunities.

Our results offer an empirical qualification of this basic result and suggest that exit barriers are not equal over time because they depend on interactions among gender, age, age-rank, and the degree of connectedness to the household head. Furthermore, our results suggest that exit barriers are binding constraints only when geographic and time-specific push or pull factors offer a window for economic advancement through migration.

Are our results consistent with standard models of migration? As in Harris and Todaro (1970), higher benefits appear to drive migration. In their model, the assumption of unemployment in the urban sector allows the persistence of a *wage* premium between urban and rural areas, linked to urban unemployment in a context of imperfect labor markets. In equilibrium, expected urban earnings (across workers and the unemployed) are equal to rural earnings (Harris and Todaro, 1970, p.129). As a result, observing living standards in urban areas to be above those in rural areas would suggest that equilibrium has not yet been attained, with continuing streams of migrants, so our results are inconsistent with the migration equilibrium, provided that expectations reflect true conditions.⁸ We cannot discount that we may be observing the migration process in a state of disequilibrium as part of a dynamic adjustment process to a long-run equilibrium with equal returns in expectation. Nevertheless, the scale of the disequilibrium is not easily explained given the available data and the results in Tables 10 and 13.

Other interpretations can nevertheless be offered as to why migration may be limited despite high returns in consumption terms. Consumption may be a poor measure of the overall net welfare benefit of migration. People may find the alienation from their original home environment costly in subjective terms. For example, a recent re-survey of the ICRISAT households in six villages in Maharashtra and Andhra Pradesh in India used a similar tracking methodology to the current paper (Dercon et al., 2009). Tracking all the people living in the original 240 ICRISAT households in 1975-84, the study found that by 2004, those who had migrated had a premium of about 20 percent in consumption, controlling for initial household fixed effects. However, again controlling for initial household fixed effects, there was a negative premium on being a migrant in regressions with subjective well-being or subjective assessment of overall wealth as the dependent variable. In short, migrants had higher consumption in real terms but lower subjective well-being compared with those from the same original households who did not migrate, possibly as if a premium in terms of the former is required to compensate for the latter. In such

circumstances, there is no reason why the consumption of migrants would ever equate to the consumption of non-migrants; a gap would remain.

9. Conclusion

This paper explores the impact of migration on poverty and living standards in Tanzania. We use a unique 13-year panel data set, offering information on split-off households and migrants. Assessing the impact of migration on living standards is particularly difficult, because we cannot observe someone to be a migrant and remain in the original community at the same time. A relatively simple difference-in-difference model is used to assess the impact of migration on consumption levels, thereby controlling for fixed individual heterogeneity in determining the level of consumption. Furthermore, we can identify the impact of migration on the growth rate of consumption using within-household variation in the subsequent migration of individual members. This initial household fixed effects estimator controls for the unobserved heterogeneity in the growth rate of consumption that is common among baseline household members. Finally, a number of specific individual factors are added as controls, and IV estimates are also presented. We avoid identifying the migration decision based on the household circumstances (such as shocks, distance to potential areas of destination, and existence of family networks) used in most studies of migration impacts, but that are unlikely to satisfy the exclusion restriction.

The identified effects are remarkably large and robust: migrants experienced 36 percentage points higher consumption growth compared with those who stayed behind. We also find that transfers from migrants to non-migrating household members are relatively limited. We find suggestive evidence that it matters where one moves. Moving to more connected areas has substantially higher returns, but even moving to a less connected area is correlated with higher growth. Moving into non-agricultural activities contributes to

consumption growth, but this growth is always greater in any sector or activity if the individual physically moved.

Relative to the theory, the paper offers evidence consistent with features in the original Lewis model, with persistent wage differentials, at least at this stage of the rural transformation.⁹ Relative to the “new economics of migration,” the evidence is less conclusive. The robust difference in welfare levels between those who migrated and those who did not is not consistent with either complete altruism or risk-sharing, although it is possible that the results reflect partial risk-sharing and some transfers. We offer suggestive evidence that some transfers indeed occur, but they are relatively limited from those who migrated long distances.

Because we have not constructed a structural model of migration (only a first stage in a 2SLS procedure), our evidence does not shed full light on the migration process. However, we provide suggestive evidence that within-family social structures matter for who gets the opportunity to migrate, how far they go, and who, therefore, can move up economically in Tanzania. This evidence also helps us to understand better how our results relate to the predictions of the Harris-Todaro model. Unlike the equilibrium conditions in Harris-Todaro, on average the premium on migration is positive and rather large: in expectation earnings do not appear to be equalized, suggesting that equilibrium has not been obtained and returns to migration remain high. There appear to be barriers to physical movement, so that potential returns are unexploited. Our evidence suggests that, just as in Hoff and Sen (2006), some of these barriers may be exit barriers that result in less than efficient levels of migration. Alternatively, other welfare costs related to migration may also limit migration.

Figure 1: KHDS 2004: Re-contacting Respondents after 10+ years*

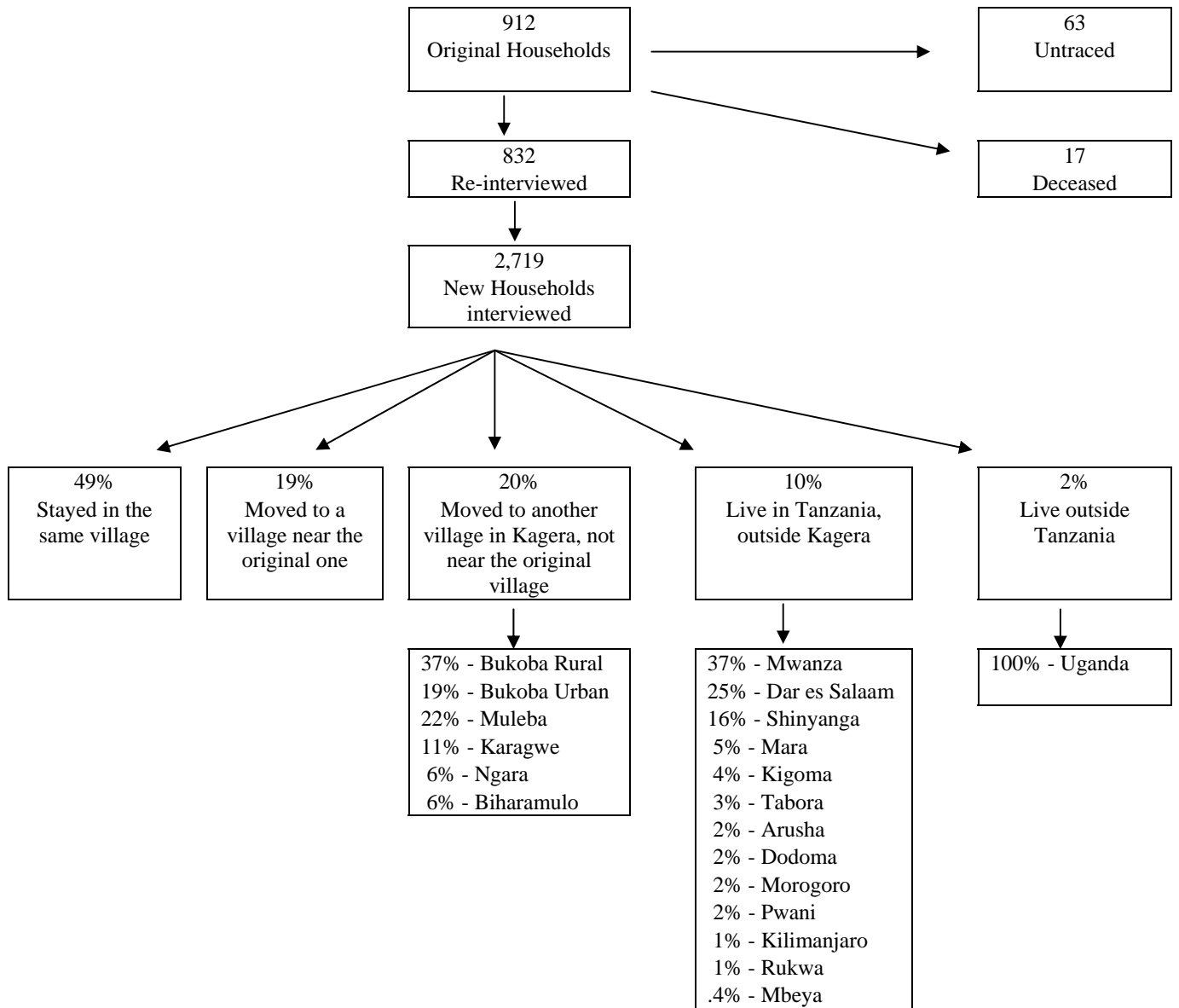
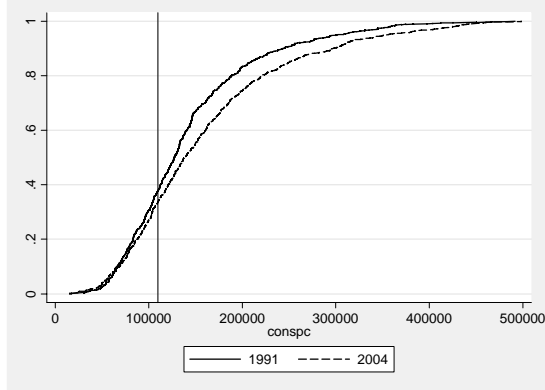
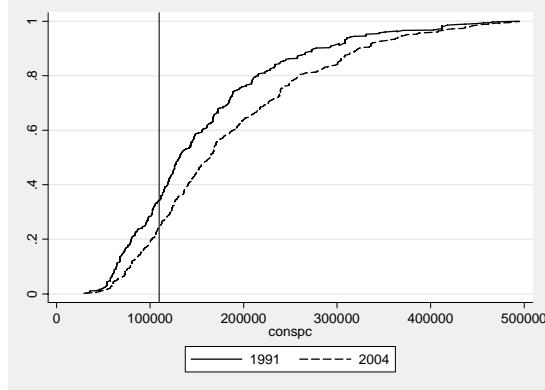


Figure 2: Cumulative Density Functions of Consumption per Capita (truncated at TSh 500,000)

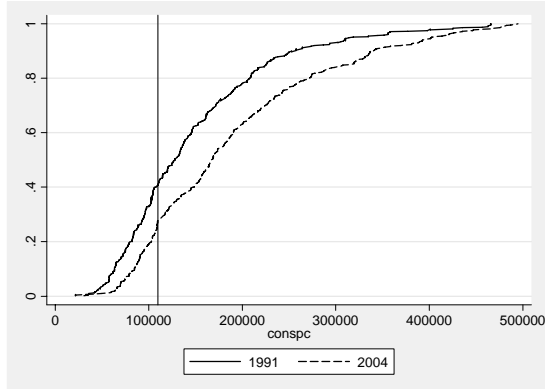
PANEL A: Within Community ($N=2447$)



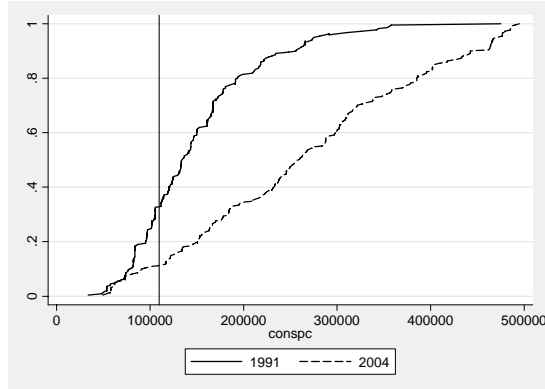
PANEL B: Nearby Community ($N=509$)



PANEL C: Elsewhere in Kagera ($N=489$)



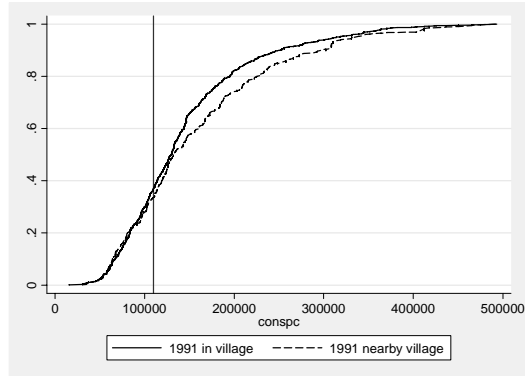
PANEL D: Outside Kagera ($N=221$)



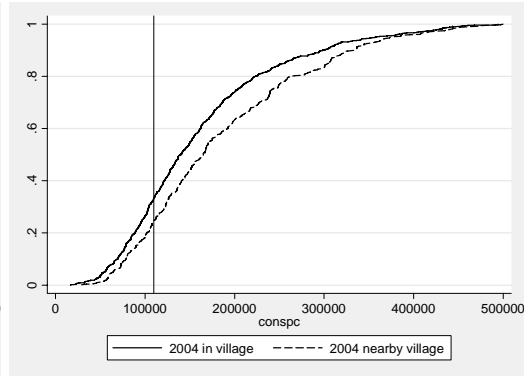
Note: The vertical line is the basic needs poverty line (TSh 109,663).

Figure 3: Cumulative Density Functions of Consumption per Capita within Community vs. Nearby Community (truncated at TSh 500,000)

PANEL A: 1991



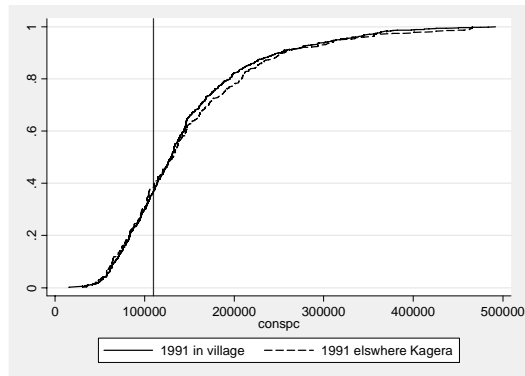
PANEL B: 2004



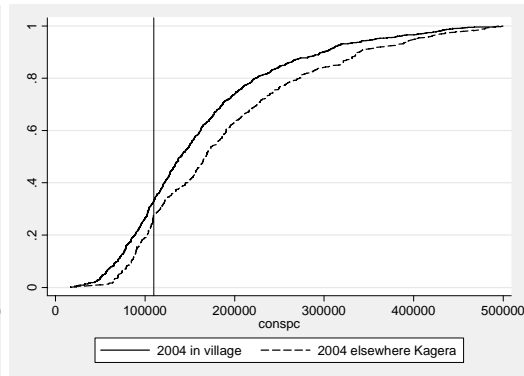
Note: The vertical line is the basic needs poverty line (TSh 109,663).

Figure 4: Cumulative Density Functions of Consumption per Capita within Community vs. Elsewhere in Kagera (truncated at TSh 500,000)

PANEL A: 1991



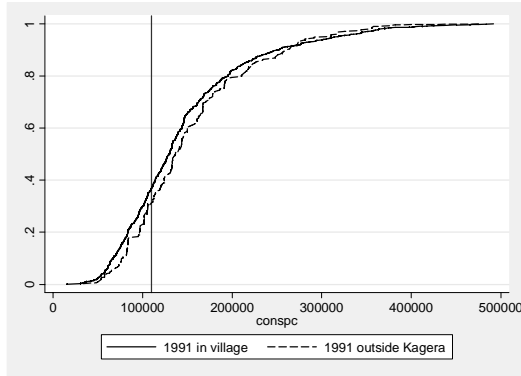
PANEL B: 2004



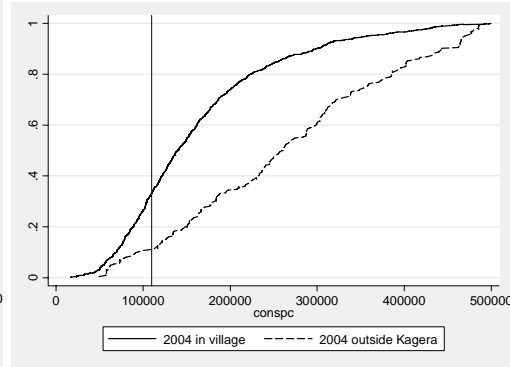
Note: The vertical line is the basic needs poverty line (TSh 109,663).

Figure 5: Cumulative Density Functions of Consumption per Capita within Community vs. outside Kagera (truncated at TSh 500,000)

PANEL A: 1991



PANEL B: 2004



Note: The vertical line is the basic needs poverty line (TSh 109,663).

Table 1: KHDS Individuals, by Age

Age at baseline 1991-1994	Re-contacted	Deceased	Untraced	Re-interview rate among survivors
<10 years	1,604 (77.1%)	160 (7.7%)	317 (15.2%)	83.5%
10-19 years	1,406 (73.2%)	104 (5.4%)	412 (21.4%)	77.3%
20-39 years	823 (63.3%)	285 (22.1%)	190 (14.6%)	81.2%
40-59 years	436 (70.6%)	147 (23.9%)	34 (5.5%)	92.8%
60+ years	163 (37.6%)	262 (60.4%)	9 (2.1%)	94.8%
Overall	4,432 (69.7%)	958 (15.1%)	962 (15.1%)	82.2%

Notes: Sample of individuals ever interviewed in KHDS 1991-1994 and alive at last interview. Age categories are based on age at first interview.

Table 2: KHDS Re-interview Rates by Location

	Number	Location	%
Baseline sample	6,352		
Re-interviewed	4,432		
		Same community	63.1
		Nearby community	14.1
		Elsewhere in Kagera	14.4
		Other region	7.1
		Other country	1.3
Untraced	962		
		Kagera	56.6
		Dar es Salaam	12.3
		Mwanza	10.4
		Other region	7.9
		Other country	5.5
		Don't know	7.3
Deceased	958		

Notes: Location for untraced respondents is reported by other household members from the baseline survey who were successfully located, interviewed, and able to provide location information on the respondent. In some cases, this information comes from other relatives or neighbors residing in the baseline communities.

**Table 3: Average Consumption Movements of Panel Respondents,
by 2004 Location**

	Mean 1991	Mean 2004	Difference in means	N
Consumption Poverty Headcount (%)				
Full Sample	0.34	0.27	-0.07***	4,116
Within community	0.35	0.31	-0.03***	2,620
Nearby community	0.33	0.21	-0.11***	577
Elsewhere in Kagera	0.36	0.24	-0.12***	595
Out of Kagera	0.30	0.07	-0.23***	324
Consumption per capita (TSh)				
Full Sample	164,434	226,337	61,903***	4,116
Within community	159,959	186,474	26,515***	2,620
Nearby community	171,493	234,973	63,480***	577
Elsewhere in Kagera	167,597	260,749	93,152***	595
Out of Kagera	180,707	472,474	291,767***	324
Food consumption per capita (TSh)				
Full Sample	106,805	146,701	39,896***	4,116
Within community	104,184	121,725	17,541***	2,620
Nearby community	111,207	152,624	41,417***	577
Elsewhere in Kagera	108,763	166,379	57,616***	595
Out of Kagera	115,704	303,453	187,749***	324
Non-Food consumption per capita (TSh)				
Full Sample	57,629	79,636	22,007***	4,116
Within community	55,775	64,748	8,973***	2,620
Nearby community	60,286	82,348	22,062***	577
Elsewhere in Kagera	58,834	94,369	35,535***	595
Out of Kagera	65,003	169,021	107,018***	324

Notes: Significance of the difference with the 1991 value using a paired *t*-test. *=10% **=5% ***=1%.

Table 4: Differences in Consumption and Poverty Head Count Changes by Mobility Categories

	N	Average change	t-test for equality change between both subgroups
Consumption per Capita (TSh)			
Stayed in community	2,620	25,940	$t=13.93$
Moved elsewhere	1,496	120,534	$p=0.0000$
Stayed in same or neighboring community	3,197	31,432	$t=16.67$
Moved elsewhere	919	160,820	$p=0.0000$
Stayed in Kagera	3,792	41,460	$t=20.25$
Moved elsewhere	324	281,064	$p=0.000$
Poverty Head Count (%)			
Stayed in community	2,620	-0.034	$t=5.41$
Moved elsewhere	1,496	-0.140	$p=0.000$
Stayed in same or neighboring community	3,197	-0.047	$t=5.11$
Moved elsewhere	919	-0.162	$p=0.000$
Stayed in Kagera	3,792	-0.059	$t=4.94$
Moved elsewhere	324	-0.231	$p=0.000$

Table 5: Sample Size, Mean, Standard Error, and 95% Confidence Interval for Incremental Samples

	N	mean	SE	95% CI	
Change in Consumption per Capita (TSh)					
(1) Only those who remained in community	2,620	25,940	3,057	19,945	31,935
(2) (1) + those who moved to neighboring communities	3,197	31,432	2,878	25,790	37,074
(3) (2) + those who moved elsewhere within Kagera	3,792	41,460	2,985	35,609	47,312
(4) (3) + those who moved outside Kagera Region (=full sample)	4,061	56,392	3,259	50,003	62,782
Change in Poverty Head Count (%)					
(1) Only those who remained in community	2,620	-0.034	0.012	-0.058	-0.010
(2) (1) + those who moved to neighboring communities	3,197	-0.047	0.011	-0.068	-0.025
(3) (2) + those who moved elsewhere within Kagera	3,792	-0.059	0.010	-0.078	-0.039
(4) (3) + those who moved outside Kagera (=full sample)	4,061	-0.068	0.009	-0.087	-0.049

**Table 6: Mean and Median Consumption Growth
by Move to More or Less Remote Area 1991-2004**

	<i>Mean</i>	<i>Median</i>	<i>N</i>
Did not move	0.13	0.16	2,147
Move out of community	0.53	0.50	1,080
Out of those that moved out of community:			
Move to more remote area	0.28	0.21	380
Move to similar area	0.46	0.45	378
Move to less remote area	0.90	0.86	322

Notes: Remoteness is based on the changes in classification among 6 possibilities, in order of remoteness, island in Lake Victoria, remote village, connected village, urban center, district capital, and regional capital.

Table 7: Mean and Median Consumption Growth by Sector Allocation Change, 1991-2004

	<i>Mean</i>	<i>Median</i>	<i>N</i>
Stay in Agriculture	0.21	0.22	1,721
Move out of Agriculture into Non-Agriculture	0.69	0.67	408
Stay in Non-Agriculture	0.43	0.43	172
Move into Agriculture from Non-Agriculture	-0.05	-0.03	245
Total	0.28	0.27	2,546

Table 8: Mean Consumption Growth by Sector Allocation and Physical Movement, 1991-2004

	<i>Stayed in Community</i>	<i>Moved out of Community</i>	<i>All</i>
Stay in Agriculture	0.18 (1,248)	0.29 (473)	0.22 (1,721)
Move out of Agriculture into Non-Agriculture	0.42 (201)	1.04 (207)	0.67 (408)
Stay in Non-Agriculture	0.11 (88)	0.88 (84)	0.44 (172)
Move into Agriculture from Non-Agriculture	-0.12 (157)	-0.00 (88)	-0.03 (245)
Total	0.18 (1,694)	0.49 (852)	0.27 (2,546)

Table 9: Explaining Consumption Change - IHHFE & 2SLS with IHHFE

	(1)	(2)	(3)	(4)
	IHHFE	IHHFE	2SLS with IHHFE	2SLS with IHHFE
Moved outside community	0.363*** (0.025)		0.378** (0.150)	
Kms moved (log of distance)		0.120*** (0.006)		0.104** (0.043)
<i>Individual characteristics at baseline</i>				
Deviation of years schooling from peers	0.013** (0.006)	0.009 (0.006)	0.013** (0.006)	0.010 (0.006)
Squared deviation of years schooling from peers	0.004*** (0.001)	0.004*** (0.001)	0.004*** (0.001)	0.004*** (0.001)
Male	-0.004 (0.038)	-0.009 (0.037)	-0.003 (0.038)	-0.010 (0.037)
Unmarried	-0.023 (0.056)	-0.020 (0.054)	-0.027 (0.064)	-0.011 (0.060)
Unmarried male	0.141*** (0.045)	0.131*** (0.044)	0.144*** (0.053)	0.123** (0.049)
Both parents died	-0.006 (0.084)	0.013 (0.081)	-0.006 (0.083)	0.010 (0.082)
Above 15 & both parents died	0.050 (0.100)	0.024 (0.098)	0.048 (0.101)	0.033 (0.100)
Years of education mother	-0.003 (0.006)	-0.004 (0.005)	-0.003 (0.006)	-0.003 (0.006)
Years of education father	0.008* (0.005)	0.007 (0.005)	0.008* (0.005)	0.007 (0.005)
<i>Biological children residing in HH at baseline</i>				
Male children 0-5	-0.028 (0.031)	-0.029 (0.030)	-0.028 (0.030)	-0.028 (0.030)
Female children 0-5	-0.027 (0.030)	-0.024 (0.029)	-0.027 (0.030)	-0.025 (0.029)
Male children 6-10	0.009 (0.035)	0.014 (0.034)	0.009 (0.035)	0.014 (0.034)
Female children 6-10	-0.045 (0.038)	-0.056 (0.037)	-0.046 (0.037)	-0.055 (0.037)
Male children 11-15	0.012 (0.036)	0.017 (0.035)	0.012 (0.036)	0.016 (0.035)
Female children 11-15	-0.000 (0.035)	-0.006 (0.034)	-0.000 (0.035)	-0.007 (0.034)
Male children 16-20	0.010 (0.041)	0.001 (0.040)	0.010 (0.041)	0.001 (0.040)
Female children 16-20	-0.085* (0.044)	-0.093** (0.043)	-0.085* (0.044)	-0.094** (0.043)
Male children 21+	0.033 (0.045)	0.026 (0.044)	0.033 (0.045)	0.028 (0.044)
Female children 21+	-0.073 (0.055)	-0.094* (0.054)	-0.072 (0.055)	-0.094* (0.054)

	(1)	(2)	(3)	(4)
	IHHFE	IHHFE	2SLS with IHHFE	2SLS with IHHFE
Number of children residing outside HH	-0.000 (0.011)	0.002 (0.011)	-0.000 (0.011)	0.001 (0.011)
Km from regional capital * number outside children	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
<i>Age at baseline (1991-1994)</i>				
5-15 years	0.143*** (0.030)	0.139*** (0.029)	0.140*** (0.043)	0.149*** (0.040)
16-25 years	0.059 (0.039)	0.059 (0.038)	0.056 (0.049)	0.069 (0.045)
26-35 years	0.108* (0.065)	0.105* (0.063)	0.107* (0.065)	0.108* (0.063)
36-45 years	0.132* (0.080)	0.130* (0.078)	0.130 (0.081)	0.135* (0.079)
46-55 years	0.149 (0.091)	0.163* (0.088)	0.148 (0.090)	0.164* (0.088)
56-65 years	0.118 (0.098)	0.123 (0.096)	0.118 (0.098)	0.124 (0.095)
66+ years	0.180 (0.121)	0.168 (0.118)	0.179 (0.120)	0.172 (0.118)
Constant	-0.023 (0.064)	-0.013 (0.063)		
Cragg-Donald			11.86	9.33
Sargan Statistic			6.26	7.28
Sargan p-value			0.28	0.20
Number of observations	3,227	3,227	3,227	3,227

Notes: Standard errors are in parentheses. *** indicates significance at 1%; ** at 5%; and, * at 10%.

Table 10: First-Stage Regressions of Table 9

	(1)	(2)
	Moved	Distance moved
<i>Baseline covariates: excluded instruments</i>		
Head or spouse	-0.218*** (0.038)	-0.634*** (0.147)
Child of head	-0.097*** (0.032)	-0.423*** (0.123)
Male child of head	-0.114*** (0.037)	-0.334** (0.144)
Age rank in HH * age 5-15	14.390* (8.003)	65.346* (30.884)
Km from reg. capital * male * age 5-15	-0.001*** (0.000)	-0.002** (0.001)
Rainfall shock * age 5-15	0.002** (0.001)	0.007** (0.003)
<i>Individual characteristics at baseline</i>		
Deviation of years schooling from peers	0.012** (0.005)	0.071*** (0.018)
Squared deviation of years schooling from peers	0.003** (0.001)	0.014*** (0.004)
Male	-0.017 (0.030)	-0.010 (0.116)
Unmarried	0.137*** (0.048)	0.464** (0.187)
Unmarried male	-0.105** (0.042)	-0.244 (0.164)
Both parents died	-0.029 (0.066)	-0.261 (0.253)
Above 15 & both parents died	0.113 (0.079)	0.562* (0.304)
Years of education mother	0.012*** (0.004)	0.040** (0.017)
Years of education father	-0.002 (0.004)	-0.000 (0.015)
<i>Biological children residing in HH at baseline</i>		
Male children 0-5	-0.001 (0.024)	0.008 (0.093)
Female children 0-5	-0.001 (0.024)	-0.010 (0.092)
Male children 6-10	-0.001 (0.028)	-0.059 (0.107)
Female children 6-10	-0.006 (0.030)	0.038 (0.116)
Male children 11-15	-0.011 (0.028)	-0.083 (0.110)
Female children 11-15	-0.035 (0.027)	-0.077 (0.105)

	(1)	(2)
	Moved	Distance moved
Male children 16-20	-0.022 (0.032)	-0.006 (0.125)
Female children 16-20	-0.031 (0.035)	-0.036 (0.134)
Male children 21+	0.020 (0.036)	0.127 (0.137)
Female children 21+	-0.016 (0.044)	0.127 (0.169)
Number of children residing outside HH	-0.008 (0.009)	-0.043 (0.033)
Km from regional capital * number outside children	0.000** (0.000)	0.001** (0.000)
<i>Age at baseline (1991-1994)</i>		
5-15 years	0.284*** (0.054)	0.886*** (0.210)
16-25 years	0.206*** (0.031)	0.603*** (0.118)
26-35 years	0.079 (0.051)	0.246 (0.198)
36-45 years	0.135** (0.063)	0.403* (0.243)
46-55 years	0.079 (0.071)	0.095 (0.276)
56-65 years	0.046 (0.078)	0.068 (0.300)
66+ years	0.056 (0.095)	0.246 (0.366)
Number of observations	3,227	3,227

Notes: Standard errors are in parentheses. *** indicates significance at 1%; ** at 5%; and, * at 10%. Linear probability model (column 1) and OLS (column 2) with household fixed effects.

**Table 11 Explaining Consumption Change - IHHFE,
with Change in Schooling and Interactions**

	(1) IHHFE coef/se	(2) IHHFE coef/se	(3) IHHFE coef/se	(4) IHHFE coef/se
Moved outside community	0.364*** (0.025)		0.262*** (0.033)	
Km moved		0.120*** (0.006)		0.099*** (0.009)
Gains in years of education	0.018*** (0.005)	0.018*** (0.005)	0.005 (0.006)	0.010* (0.005)
Gains in education * moved dummy			0.033*** (0.007)	
Gains in education * km moved				0.006*** (0.002)
Deviation of years schooling from peers	0.017*** (0.006)	0.013** (0.006)	0.018*** (0.006)	0.014** (0.006)
Squared dev. of years schooling from peers	0.005*** (0.001)	0.004*** (0.001)	0.005*** (0.001)	0.004*** (0.001)
Number of observations	3,226	3,226	3,226	3,226

note: *** p<0.01, ** p<0.05, * p<0.1. Number of observations is one less than Table 9 due to one excluded observation missing 2004 education. Other controls reported in Table 9 are included but not reported.

Table 12: Explaining Consumption Change – IHHFE, Characteristics of the Move

	(1) IHHFE	(2) IHHFE
<i>Characteristics of the move</i>		
Move to more remote area	0.176*** (0.036)	
Move to similar area	0.274*** (0.034)	
Move to more connected area	0.661*** (0.037)	
Km moved		0.073*** (0.011)
Distance moved if to similar area		0.032** (0.015)
Distance moved if to more connected area		0.070*** (0.013)
<i>Individual characteristics at baseline</i>		
Deviation of years schooling from peers	0.010* (0.006)	0.008 (0.006)
Squared deviation of years schooling from peers	0.004*** (0.001)	0.004*** (0.001)
Male	-0.004 (0.037)	-0.008 (0.037)
Unmarried	-0.008 (0.054)	-0.006 (0.054)
Unmarried male	0.127*** (0.044)	0.121*** (0.043)
Both parents died	0.005 (0.082)	0.025 (0.081)
Above 15 & both parents died	0.053 (0.098)	0.020 (0.097)
Years of education mother	-0.004 (0.005)	-0.004 (0.005)
Years of education father	0.006 (0.005)	0.006 (0.005)
<i>Biological children residing in HH at baseline</i>		
Male children 0-5	-0.021 (0.030)	-0.023 (0.030)
Female children 0-5	-0.026 (0.029)	-0.025 (0.029)
Male children 6-10	0.008 (0.034)	0.015 (0.034)
Female children 6-10	-0.048 (0.037)	-0.056 (0.036)
Male children 11-15	0.023 (0.035)	0.022 (0.035)
Female children 11-15	-0.010 (0.034)	-0.011 (0.033)
Male children 16-20	0.012	0.002

	(1)	(2)
	IHHFE	IHHFE
	(0.040)	(0.040)
Female children 16-20	-0.085*	-0.095**
	(0.043)	(0.043)
Male children 21+	0.023	0.020
	(0.044)	(0.044)
Female children 21+	-0.090*	-0.099*
	(0.054)	(0.054)
Number of children residing outside HH	-0.001	0.003
	(0.011)	(0.011)
Km from regional capital * number outside children	0.000	0.000
	(0.000)	(0.000)
<i>Age at baseline (1991-1994)</i>		
5-15 years	0.141***	0.143***
	(0.029)	(0.028)
16-25 years	0.063*	0.066*
	(0.038)	(0.038)
26-35 years	0.107*	0.102
	(0.063)	(0.063)
36-45 years	0.130*	0.131*
	(0.078)	(0.077)
46-55 years	0.164*	0.166*
	(0.088)	(0.088)
56-65 years	0.135	0.127
	(0.096)	(0.095)
66+ years	0.190	0.169
	(0.118)	(0.117)
Constant	-0.015	-0.007
	(0.063)	(0.062)
Number of observations	3,227	3,227

Notes: Standard errors are in parentheses. *** indicates significance at 1%; ** at 5%; and, * at 10%.

Table 13: Explaining Consumption Change – IHHFE, Moving out of Agriculture

	(1)	(2)	(3)
	IHHFE	IHHFE	IHHFE
<i>Characteristics of the move</i>			
Moved outside community		0.195*** (0.034)	
Kms moved (log of distance)			0.073*** (0.011)
Moved out of Agriculture	0.407*** (0.034)	0.126*** (0.044)	0.175*** (0.040)
Moved outside community & out of agriculture		0.449*** (0.059)	
Distance moved * moved out of agriculture			0.075*** (0.015)
<i>Individual characteristics at baseline</i>			
Deviation of years schooling from peers	0.013* (0.007)	0.011* (0.006)	0.010 (0.006)
Squared deviation of years schooling from peers	0.004** (0.002)	0.003** (0.001)	0.003** (0.001)
Male	-0.059 (0.042)	-0.020 (0.040)	-0.030 (0.040)
Unmarried	-0.005 (0.063)	-0.059 (0.061)	-0.049 (0.060)
Unmarried male	0.079 (0.051)	0.128** (0.050)	0.132*** (0.049)
Both parents died	-0.066 (0.113)	-0.045 (0.108)	-0.022 (0.108)
Above 15 & both parents died	0.110 (0.126)	0.076 (0.121)	0.045 (0.120)
Years of education mother	0.005 (0.007)	-0.002 (0.007)	-0.005 (0.007)
Years of education father	-0.003 (0.006)	0.001 (0.006)	0.000 (0.006)
<i>Biological children residing in HH at baseline</i>			
Male children 0-5	-0.048 (0.035)	-0.047 (0.034)	-0.044 (0.033)
Female children 0-5	-0.029 (0.034)	-0.018 (0.033)	-0.020 (0.033)
Male children 6-10	0.023 (0.039)	0.014 (0.038)	0.024 (0.037)
Female children 6-10	-0.057 (0.043)	-0.056 (0.041)	-0.067 (0.041)
Male children 11-15	0.018 (0.041)	0.010 (0.039)	0.023 (0.039)
Female children 11-15	0.004 (0.039)	0.010 (0.037)	0.008 (0.037)
Male children 16-20	-0.024 (0.046)	-0.009 (0.044)	-0.011 (0.044)

	(1)	(2)	(3)
	IHHFE	IHHFE	IHHFE
Female children 16-20	-0.100** (0.049)	-0.103** (0.047)	-0.111** (0.047)
Male children 21+	0.027 (0.053)	0.013 (0.051)	0.005 (0.050)
Female children 21+	-0.141** (0.067)	-0.103 (0.064)	-0.119* (0.064)
Number of children residing outside HH	0.006 (0.013)	0.004 (0.012)	0.004 (0.012)
Km from regional capital * number outside children	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
<i>Age at baseline (1991-1994)</i>			
5-15 years	0.177*** (0.049)	0.132*** (0.047)	0.135*** (0.047)
16-25 years	0.058 (0.057)	0.029 (0.055)	0.029 (0.054)
26-35 years	0.063 (0.083)	0.082 (0.079)	0.078 (0.079)
36-45 years	0.077 (0.098)	0.085 (0.094)	0.077 (0.093)
46-55 years	0.103 (0.110)	0.128 (0.106)	0.133 (0.105)
56-65 years	0.091 (0.119)	0.105 (0.114)	0.113 (0.114)
66+ years	0.195 (0.156)	0.246* (0.149)	0.233 (0.148)
Constant	0.084 (0.082)	0.054 (0.079)	0.057 (0.078)
Number of observations	2,546	2,546	2,546

Notes: Standard errors are in parentheses. *** indicates significance at 1%; ** at 5%; and, * at 10%.

Table 14: Explaining Migration, Household, and Community Correlates

	(1) Probit: moved out of community	(2) OLS: km moved (log of distance)
Sex of hh head	0.012 (0.026)	-0.105 (0.093)
Age of hh head	0.001 (0.001)	0.009*** (0.002)
Educ of hh head	-0.005 (0.004)	0.000 (0.015)
household size	-0.000 (0.003)	-0.014 (0.012)
primary occupation is farming	-0.064** (0.031)	-0.355*** (0.099)
Acres of land cultivated	-0.003 (0.002)	-0.014* (0.008)
consumption pc (in millions of TSh)	-0.021 (0.124)	-0.674 (0.419)
Value of physical assets (in millions of TSh)	-0.008* (0.005)	-0.024* (0.014)
Good flooring in dwelling	-0.003 (0.031)	0.097 (0.105)
km from cluster to Bukoba	-0.000 (0.000)	-0.000 (0.001)
remote community	-0.047** (0.020)	-0.078 (0.068)
<i>Cluster mean of household characteristics</i>		
Sex of hh head	-0.207* (0.120)	0.068 (0.411)
Age of hh head	0.002 (0.002)	0.017** (0.008)
Educ of hh head	0.005 (0.013)	0.017 (0.043)
household size	0.007 (0.008)	0.049* (0.028)
primary occupation is farming	0.110 (0.087)	0.500* (0.289)
Acres of land cultivated	-0.001 (0.007)	-0.012 (0.023)
consumption pc (in millions of TSh)	0.256 (0.362)	1.545 (1.224)
Value of physical assets (in millions of TSh)	0.007 (0.015)	0.013 (0.053)
Good flooring in dwelling	-0.028 (0.084)	-0.152 (0.286)
Constant		-0.825 (0.755)
Number of observations	3,119	3,119

Notes: Standard errors are in parentheses. *** indicates significance at 1%; ** at 5%; and, * at 10%. Column 1 presents the marginal effects from a probit estimation. Individual characteristics included in Table 10 specifications are also included here, but are not presented in the table. The sample is slightly reduced (from 3,227) due to missing information on baseline value of physical stock for 108 households.

Appendix Table 1: Sample Means and Standard Deviations

	Mean	SD
Change in (logged) Consumption per Capita	0.26	(0.77)
Moved	0.33	(0.47)
Distance moved (kms)	35.02	(145.01)
Distance moved variable: log(kms+1)	1.06	(1.78)
<i>Baseline covariates: excluded instruments</i>		
Head or spouse	0.26	(0.44)
Child of head	0.49	(0.50)
Male child of head	0.25	(0.43)
Age rank in HH * age 5-15	0.00	(0.00)
Km from reg. capital * male * age 5-15	12.04	(38.43)
Rainfall shock (annual centimeters deviation) * age 5-15	-18.51	(28.71)
<i>Individual characteristics at baseline</i>		
Deviation of years schooling from peers	-0.25	(2.24)
Squared deviation of years schooling from peers	5.08	(9.50)
Male	0.47	(0.50)
Unmarried	0.69	(0.46)
Unmarried male	0.36	(0.48)
Both parents died	0.05	(0.22)
Above 15 & both parents died	0.02	(0.15)
Years of education mother	2.72	(3.02)
Years of education father	4.23	(3.32)
<i>Biological children residing in HH at baseline</i>		
Male children 0-5	0.15	(0.45)
Female children 0-5	0.14	(0.45)
Male children 6-10	0.10	(0.36)
Female children 6-10	0.09	(0.34)
Male children 11-15	0.10	(0.36)
Female children 11-15	0.11	(0.38)
Male children 16-20	0.06	(0.29)
Female children 16-20	0.06	(0.28)
Male children 21+	0.05	(0.26)
Female children 21+	0.04	(0.21)
Number of children residing outside HH	0.64	(1.83)
Km from regional capital * number outside children	44.62	(184.60)
<i>Age at baseline (1991-1994)</i>		
5-15 years	0.35	(0.48)
16-25 years	0.20	(0.40)
26-35 years	0.08	(0.27)
36-45 years	0.07	(0.26)
46-55 years	0.06	(0.23)
56-65 years	0.04	(0.20)
66+ years	0.02	(0.12)
Number of observations		3,227

Appendix Table 2: Alternative Sets of Instrumental Variables

	Excluded instrumental variable:						
All included	Head or spouse	Child of head	Male child of head	Age rank in HH * age 5-15	Km from reg. capital * male * age 5-15	Avg. Rainfall shock * age 5-15	
Moved outside community							
Coefficient	0.38	0.40	0.49	0.30	0.32	0.38	0.36
Standard error	0.15	0.20	0.16	0.16	0.15	0.16	0.16
Cragg-Donald	11.86	7.62	12.36	12.32	13.58	12.54	12.89
Kms moved (log of distance)							
Coefficient	0.10	0.10	0.16	0.08	0.08	0.10	0.10
Standard error	0.04	0.05	0.05	0.05	0.04	0.04	0.05
Cragg-Donald	9.33	7.45	8.81	10.11	10.29	10.23	9.87

Notes: Each coefficient is generated from a separate regression based on the 2SLS specification in Table 10.

**Appendix Table 3: Explaining Consumption Change - IHHFE and 2SLS,
Household-level Results**

	(1)	(2)	(3)	(4)
	IHHFE	IHHFE	2SLS	2SLS
Moved outside community	0.321*** (0.038)		0.520*** (0.154)	
Kms moved (log of distance)		0.112*** (0.009)		0.146*** (0.045)
<i>Individual characteristics at baseline</i>				
Deviation of years schooling from peers	0.013 (0.010)	0.007 (0.010)	0.011 (0.010)	0.005 (0.010)
Squared deviation of years schooling from peers	0.006*** (0.002)	0.006*** (0.002)	0.006*** (0.002)	0.005** (0.002)
Male	0.085 (0.096)	0.067 (0.093)	0.115 (0.098)	0.076 (0.093)
Unmarried	-0.114 (0.101)	-0.117 (0.098)	-0.144 (0.104)	-0.132 (0.100)
Unmarried male	0.179* (0.102)	0.182* (0.099)	0.208** (0.104)	0.197** (0.101)
Both parents died	-0.079 (0.136)	-0.042 (0.133)	-0.080 (0.136)	-0.032 (0.133)
Above 15 & both parents died	0.124 (0.175)	0.066 (0.171)	0.092 (0.177)	0.034 (0.175)
Years of education mother	-0.004 (0.010)	-0.006 (0.009)	-0.008 (0.010)	-0.009 (0.010)
Years of education father	0.016* (0.009)	0.015* (0.009)	0.017* (0.009)	0.015* (0.009)
<i>Biological children residing in HH at baseline</i>				
Male children 0-5	-0.049 (0.074)	-0.042 (0.072)	-0.049 (0.074)	-0.040 (0.072)
Female children 0-5	0.014 (0.078)	0.019 (0.076)	0.040 (0.081)	0.034 (0.078)
Male children 6-10	-0.112 (0.090)	-0.097 (0.087)	-0.101 (0.090)	-0.087 (0.088)
Female children 6-10	-0.181* (0.090)	-0.209** (0.088)	-0.179* (0.090)	-0.216** (0.088)
Male children 11-15	0.046 (0.076)	0.051 (0.074)	0.042 (0.076)	0.051 (0.073)
Female children 11-15	-0.046 (0.083)	-0.045 (0.081)	-0.018 (0.086)	-0.031 (0.082)
Male children 16-20	0.040 (0.088)	0.024 (0.086)	0.054 (0.089)	0.026 (0.085)
Female children 16-20	-0.200** (0.105)	-0.214** (0.102)	-0.174* (0.107)	-0.206** (0.102)
Male children 21+	0.046 (0.105)	0.041 (0.102)	0.035 (0.105)	0.035 (0.102)
Female children 21+	-0.204 (0.126)	-0.237* (0.122)	-0.166 (0.129)	-0.229* (0.122)
Number of children residing outside HH	-0.001	0.001	-0.004	0.000

	(1)	(2)	(3)	(4)
	IHHFE	IHHFE	2SLS	2SLS
Km from regional capital *	0.000	0.000	0.000	0.000
number outside children	(0.000)	(0.000)	(0.000)	(0.000)
<i>Age at baseline (1991-1994)</i>				
5-15 years	0.245***	0.246***	0.215***	0.229***
	(0.062)	(0.060)	(0.066)	(0.063)
16-25 years	0.060	0.066	0.047	0.058
	(0.071)	(0.069)	(0.072)	(0.069)
26-35 years	0.152	0.160	0.196	0.183
	(0.131)	(0.127)	(0.135)	(0.130)
36-45 years	0.141	0.167	0.155	0.161
	(0.168)	(0.163)	(0.168)	(0.163)
46-55 years	0.206	0.264	0.278	0.312*
	(0.187)	(0.182)	(0.194)	(0.193)
56-65 years	0.144	0.180	0.243	0.231
	(0.199)	(0.193)	(0.212)	(0.206)
66+ years	0.336	0.344	0.384*	0.357
	(0.236)	(0.229)	(0.238)	(0.230)
Constant	0.025	0.019		
	(0.129)	(0.125)		
Cragg-Donald			13.68	9.10
Sargan Statistic			5.67	7.76
Sargan p-value			0.34	0.17
Number of observations	1,909	1,909	1,909	1,909

Notes: Standard errors are in parentheses. *** indicates significance at 1%; ** at 5%; and, * at 10%.

Appendix Table 4: First-Stage Regressions of Appendix Table 3

	(1)	(2)
	Moved	Distance moved
<i>Baseline covariates: excluded instruments</i>		
Head or spouse	-0.179*** (0.063)	-0.479* (0.254)
Child of head	-0.058 (0.048)	-0.343* (0.193)
Male child of head	-0.157*** (0.056)	-0.455** (0.227)
Age rank in HH * age 5-15	-7.379 (12.036)	-1.269 (48.763)
Km from reg. capital * male * age 5-15	-0.001* (0.000)	-0.001 (0.001)
Rainfall shock * age 5-15	0.001*** (0.000)	0.002*** (0.000)
<i>Individual characteristics at baseline</i>		
Deviation of years schooling from peers	0.007 (0.007)	0.070** (0.030)
Squared deviation of years schooling from peers	0.003** (0.002)	0.018*** (0.006)
Male	-0.126* (0.070)	-0.233 (0.282)
Unmarried	0.095 (0.075)	0.364 (0.306)
Unmarried male	-0.022 (0.079)	-0.073 (0.322)
Both parents died	0.019 (0.098)	-0.276 (0.396)
Above 15 & both parents died	0.160 (0.125)	0.974* (0.507)
Years of education mother	0.015** (0.007)	0.060** (0.028)
Years of education father	-0.010 (0.007)	-0.019 (0.027)
<i>Biological children residing in HH at baseline</i>		
Male children 0-5	0.016 (0.053)	-0.011 (0.216)
Female children 0-5	-0.082 (0.056)	-0.246 (0.229)
Male children 6-10	-0.007 (0.065)	-0.162 (0.263)
Female children 6-10	0.005 (0.066)	0.221 (0.267)
Male children 11-15	0.037 (0.055)	0.045 (0.224)
Female children 11-15	-0.087 (0.061)	-0.274 (0.246)

	(1)	(2)
	Moved	Distance moved
Male children 16-20	-0.061 (0.063)	-0.040 (0.257)
Female children 16-20	-0.105 (0.075)	-0.169 (0.304)
Male children 21+	0.063 (0.075)	0.239 (0.305)
Female children 21+	-0.156* (0.091)	-0.168 (0.369)
Number of children residing outside HH	0.006 (0.016)	0.001 (0.064)
Km from regional capital * number outside children	0.000** (0.000)	0.001** (0.000)
<i>Age at baseline (1991-1994)</i>		
5-15 years	0.454*** (0.072)	1.351*** (0.292)
16-25 years	0.032 (0.052)	0.046 (0.209)
26-35 years	-0.227** (0.094)	-0.745* (0.383)
36-45 years	-0.089 (0.121)	-0.389 (0.491)
46-55 years	-0.340** (0.135)	-1.521*** (0.548)
56-65 years	-0.433*** (0.144)	-1.549*** (0.582)
66+ years	-0.239 (0.170)	-0.727 (0.688)
Number of observations	1,909	1,909

Notes: Standard errors are in parentheses. *** indicates significance at 1%; ** at 5%; and, * at 10%.

Appendix Table 5: Household Size at Baseline and Follow-up, by Mobility Categories, Mean (Median)

	Household size		Household size: adult equivalent		N
	1991	2004	1991	2004	
Same village	7.71 (7.0)	5.98 (6.0)	6.15 (5.7)	4.94 (4.6)	2,150
Neighboring community	8.20 (7.0)	4.93 (5.0)	6.59 (5.9)	3.87 (3.4)	400
Elsewhere in Kagera	7.65 (7.0)	4.47 (4.0)	6.17 (6.0)	3.55 (3.2)	437
Outside Kagera	8.45 (7.0)	4.45 (4.0)	6.74 (6.1)	3.69 (3.1)	251

Notes: Adult equivalence is defined following the National Bureau of Statistics with varying weights by age and sex.

**Appendix Table 6: Explaining Consumption Change - IHHFE and 2SLS
Adult Equivalent Consumption (rather than per capita)**

	(1)	(2)	(3)	(4)
	IHHFE	IHHFE	2SLS with IHHFE	2SLS with IHHFE
Moved outside community	0.363*** (0.024)		0.426*** (0.143)	
Kms moved (log of distance)		0.117*** (0.006)		0.123*** (0.041)
<i>Individual characteristics at baseline</i>				
Deviation of years schooling from peers	0.014** (0.006)	0.010* (0.006)	0.013** (0.006)	0.010 (0.006)
Squared deviation of years schooling from peers	0.004*** (0.001)	0.003** (0.001)	0.004*** (0.001)	0.003** (0.001)
Male	-0.010 (0.036)	-0.016 (0.035)	-0.008 (0.036)	-0.016 (0.035)
Unmarried	0.043 (0.053)	0.048 (0.051)	0.030 (0.060)	0.045 (0.057)
Unmarried male	0.087** (0.043)	0.076* (0.041)	0.099* (0.051)	0.079* (0.046)
Both parents died	0.007 (0.079)	0.026 (0.077)	0.009 (0.079)	0.027 (0.077)
Above 15 & both parents died	0.032 (0.095)	0.008 (0.093)	0.026 (0.096)	0.005 (0.095)
Years of education mother	-0.002 (0.005)	-0.003 (0.005)	-0.003 (0.006)	-0.003 (0.006)
Years of education father	0.008* (0.005)	0.007 (0.005)	0.008* (0.005)	0.007 (0.005)
<i>Biological children residing in HH at baseline</i>				
Male children 0-5	-0.041 (0.029)	-0.041 (0.028)	-0.041 (0.029)	-0.041 (0.028)
Female children 0-5	-0.027 (0.028)	-0.025 (0.028)	-0.027 (0.028)	-0.024 (0.028)
Male children 6-10	0.001 (0.033)	0.005 (0.032)	0.000 (0.033)	0.006 (0.032)
Female children 6-10	-0.042 (0.036)	-0.052 (0.035)	-0.042 (0.035)	-0.053 (0.035)
Male children 11-15	0.001 (0.034)	0.005 (0.033)	0.001 (0.034)	0.005 (0.033)
Female children 11-15	-0.014 (0.033)	-0.019 (0.032)	-0.011 (0.033)	-0.018 (0.032)
Male children 16-20	0.005 (0.039)	-0.003 (0.038)	0.006 (0.039)	-0.003 (0.038)
Female children 16-20	-0.056 (0.042)	-0.064 (0.041)	-0.054 (0.042)	-0.064 (0.041)
Male children 21+	0.013 (0.043)	0.006 (0.042)	0.011 (0.043)	0.005 (0.042)
Female children 21+	-0.066 (0.052)	-0.087* (0.051)	-0.063 (0.053)	-0.087* (0.051)

	(1)	(2)	(3)	(4)
	IHHFE	IHHFE	2SLS with IHHFE	2SLS with IHHFE
Number of children residing outside HH	0.002 (0.010)	0.004 (0.010)	0.003 (0.010)	0.004 (0.010)
Km from regional capital * number outside children	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
<i>Age at baseline (1991-1994)</i>				
5-15 years	0.200*** (0.028)	0.199*** (0.027)	0.187*** (0.041)	0.199*** (0.038)
16-25 years	0.175*** (0.037)	0.178*** (0.036)	0.163*** (0.046)	0.175*** (0.043)
26-35 years	0.192*** (0.061)	0.189*** (0.060)	0.188*** (0.061)	0.188*** (0.060)
36-45 years	0.197*** (0.075)	0.197*** (0.074)	0.190** (0.077)	0.195*** (0.075)
46-55 years	0.256*** (0.086)	0.270*** (0.084)	0.253*** (0.086)	0.270*** (0.083)
56-65 years	0.232** (0.093)	0.238*** (0.091)	0.231** (0.093)	0.238*** (0.090)
66+ years	0.313*** (0.114)	0.302*** (0.112)	0.310*** (0.114)	0.300*** (0.111)
Constant	-0.137** (0.061)	-0.125** (0.059)		
Cragg-Donald			11.86	9.33
Sargan Statistic			10.59	11.44
Sargan p-value			0.06	0.04
Number of observations	3,227	3,227	3,227	3,227

Notes: Standard errors are in parentheses. *** indicates significance at 1%; ** at 5%; and, * at 10%.

References

Alderman, Harold, Jere R. Berman, Hans-Peter Kohler, John Maluccio, and Susan Cotts Watkins. 2001. "Attrition in Longitudinal Household Survey Data: Some Tests for Three Developing-Country Samples." *Demographic Research* 5(4).

Beegle, Kathleen. 2000. "Economic Mobility in Indonesia and Vietnam: What Missing Data Can't Tell Us." mimeo.

Beegle, Kathleen, Joachim De Weerd, and Stefan Dercon. 2006. "Kagera Health and Development Survey (KHDS), 2004, Basic Information Document." mimeo.

Clemens, Michael and Lant Pritchett. 2008. "Income Per Natural: Measuring Development for People rather than Places." *Population and Development Review* 34 (3): 395-434.

De Weerd, Joachim. 2010. "Moving out of Poverty in Tanzania: Evidence from Kagera." *Journal of Development Studies* 46(2): 331-349.

Demombynes, Gabriel and Johannes Hoogeveen. 2007. "Growth, Inequality and Simulated Poverty Paths for Tanzania, 1992-2002." *Journal of African Economies* 16(4):596-628.

Dercon, Stefan, Pramila Krishnan and Sonya Krutikov. 2009. "Migration, Risk-sharing and Subjective Well-being." mimeo, University of Oxford.

Fields, Gary S. 2004. "Dualism in the Labor Market: a Perspective on the Lewis Model after Half a Century." *The Manchester School* 72(6): 724-735.

Harris, John and Michael Todaro. 1970. "Migration, Unemployment and Development: A Two-Sector Analysis." *American Economic Review* 60(1): 126-142.

Hoff, Karla and Arjit Sen. 2006. "The Kin System as a Poverty Trap?" in *Poverty Traps* Samuel Bowles, Steven Durlauf, and Karla Hoff (eds.), Princeton University Press: 91-115.

Kilic, Talip, Gero Carletto, Benjamin Davis and Alberto Zezza. 2009. "Investing Back Home: Return Migration and Business Ownership in Albania." *Economics of Transition* 17(3): 587-623.

Lewis, W. Arthur. 1954. "Economic Development with Unlimited Supplies of Labor." *The Manchester School* 22(2):139-91

McKenzie, David and Marcin J. Sasin. 2007. "Migration, Remittances and Human Capital: Conceptual and Empirical Challenges." World Bank Policy Research Working Paper 4272.

McKenzie, David, John Gibson, and Steven Stillman. 2009. "How Important is Selection? Experimental v. Non-Experimental Measures of the Income Gains from Migration." *Journal of the European Economic Association* forthcoming.

Munshi, Kaivan. 2003. "Networks in the Modern Economy: Mexican Migrants in the United States Labor Market." *Quarterly Journal of Economics* 118 (2): 549-597.

Rosenzweig, Mark. 2003. "Payoffs from Panels in Low-Income Countries: Economic Development and Economic Mobility." *American Economic Review* 93(2): 112-117.

Rosenzweig, Mark, and Oded Stark. 1989. "Consumption Smoothing, Migration, and Marriage: Evidence from Rural India." *Journal of Political Economy* 97(4): 905-926.

Stark, Oded, and David Bloom. 1985. "The New Economics of Labor Migration." *American Economic Review* 75(2): 173-178.

Stock, James and Motohiro Yogo. 2002. "Testing for Weak Instruments in Linear IV Regression." NBER Technical Working Paper 0284.

URT (United Republic of Tanzania). 2004. "The Economic Report, 2004." President's Office, Planning and Privatisations.

URT (United Republic of Tanzania). 2007. "Analytical Report for the Integrated Labor Force Survey." National Bureau of Statistics, Ministry of Planning, Economy and Empowerment.

World Bank. 2004. "User's Guide to the Kagera Health and Development Survey Datasets." mimeo.

¹ For example, Lewis (1954, p. 150) wrote: “Earnings in the subsistence sector set a floor to wages in the capitalist sector, but in practice wages have to be higher than this, and there is usually a gap of 30 per cent or more between capitalist wages and subsistence earnings.”

² This is correct, even if in equilibrium, when no further migration takes place, expected earnings are equal.

³ Tables 6 and onward are restricted to the sample in the main regressions (N=3,227). From the full sample of 4,432, we exclude, in this order: 715 people who were not interviewed in wave 1 (they were interviewed in waves 2, 3, and/or 4), 15 people in 1-person households, 267 people missing either wave 1 or wave 5 consumption expenditure, 120 people missing peer’s schooling, 2 people missing parental education, and 86 people with incomplete data in wave 1. Tables 7, 8, and 12 have 2,546 observations because of missing occupational data for 2004.

⁴ In order to investigate the clustering of migration patterns, all households were sorted into “tracking zones,” indicating the geographical area in which they resided in 2004. Tabulating, for each tracking zone, the village of origin of the households tracked in that zone did not reveal any discernable pattern of clustered migration. In each tracking zone, there was never any origin village that dominated, with the exception of villages that lie within or neighbor the tracking zone.

⁵ We used the variable “years of schooling completed relative to peers” rather than a straight “years of schooling completed” for two reasons. First, a substantial number of individuals in the sample were younger than 18 at the baseline and had therefore not necessarily completed their education. As such, years of schooling at the baseline might be less correlated with a move by 2004 than, say, eventual completed years of schooling. Second, akin to this concern, years of schooling is highly correlated with age for individuals of school-going age. The regressions also include a set of age variables, defined in broad age groups (for ease of interpretation and

discussion of results). One consequence could be that years of schooling at the baseline would pick up at least some age effect. To address this concern, rather than use education in years, we constructed a variable of education relative to peers: the absolute deviation of education levels compared with mean education of age-specific peers at the baseline for those younger than 18 and relative to other adults for the rest of the sample. This purges the education variable of an age effect it would otherwise pick up. All the regressions below were repeated using a straightforward “years of schooling” variable, rather than our “years of schooling relative to peers” variable. Neither the results nor their interpretation were affected.

⁶ The non-interacted variables are all included as determinants of consumption growth via X_i and δ_{ih} .

⁷ Only the IHHFE and not the 2SLS results are reported here. When estimating the first column using 2SLS with the same instruments as before, but also including the education gained in the first and second stage, gave close to identical results as the IHHFE regressions (not shown). In the first-stage regression, the variable additional years of education gained is not correlated with the likelihood of moving, conditional on age.

⁸ In principle, deviation of the estimated average difference between urban and rural living standards ex-post could also be possible in equilibrium in a Harris-Todaro-style model. For this to explain the apparent less-than-optimal migration levels, it would need to be the case that the gap in living standards between urban and rural areas was expected to be much lower than is now apparent. This could come about if expectations were not rational; if there was an unexpected higher level of urban wages by 2004; or, for example, if there was much higher employment in well-paid jobs. The premium seems to be too high for this to be a sufficient explanation, most likely because although urban wages may be high, unemployment levels for particular groups are also high. More than 20 percent of the urban population between ages 18 and 34 is unemployed, in the sense that they are looking for work and available for work. Unemployment is double this

figure using the national definition, which includes those with marginal or precarious job situations (URT, 2007).

⁹ This interpretation of Lewis (1954) is still debated, and not necessarily a feature of subsequent dual economy models. For a discussion, see Fields (2004).