Agricultural productivity and mortality: evidence from Kagera, Tanzania

Ikegami, Munenobu

30 April 2009

Agricultural Productivity and Mortality: Evidence from Kagera, Tanzania☆

Munenobu Ikegami[a]

[a] *International Livestock Research Institute, Old Naivasha Road, P.O. Box 30709, Nairobi 00100 Kenya.*

Abstract

We ask whether prime-age adult mortality due to HIV/AIDS decreases the endowment of knowledge for agricultural production in Kagera, Tanzania, reducing total factor productivity. We also quantify how much this negative effect contributes to the decrease in long-term household agricultural output growth compared to the contribution of decreased accumulation of productive assets; household members, land, and livestock. We find that prime-age adult mortality decreases the accumulation of knowledge stock as total factor productivity and the contribution of this negative effect to the decrease in agricultural output growth is larger than the contribution of decreased accumulation of each productive asset.

Keywords: mortality, human capital, HIV/AIDS, agriculture, total factor productivity, Tanzania

JEL classification: D9, I1, O12, Q12

☆The author thanks Michael R. Carter, Jean-Paul Chavas, Salvador Navarro, Laura Schecter, Christopher Udry, James R. Walker and seminar participants at "Agriculture for Development in Sub-Saharan Africa" organized by African Economic Research Consortium and University of California Berkeley in May 2009, "Taking Action for the World's Poor and Hungry People: Doctoral Students Workshop" organized by International Food Policy Research Institute and Beijing Normal University in October 2007, and several workshops at University of Wisconsin-Madison. The author thanks the World Bank for making publicly available the data analyzed in this paper. The author sorely remains responsible for any errors contained herein.

Email address: ikegami4@gmail.com (Munenobu Ikegami)

1. Introduction

In the regions affected by Human Immunodeficiency Virus / Acquired Immune Deficiency Syndrome (HIV/AIDS), prime-age adult mortality negatively affects household welfare by decreasing household income and consumption. Previous studies on the effects of prime-age adult mortality on household agricultural production show that the mortality decreases household size and productive assets such as land and livestock. In this study, we further ask whether prime-age adult mortality due to HIV/AIDS decreases the endowment of knowledge for agricultural production in Kagera, Tanzania, reducing total factor productivity(TFP). Equivalently, we ask whether prime-age adult mortality due to HIV/AIDS destroys household agricultural production by magnitude beyond the decreases in observed productive assets such as household members, land, and livestock. We also quantify how much decreased TFP growth contributes to the decrease in long-term household agricultural output growth compared to the decreased accumulation of each productive asset.

Kagera was estimated to be one of the regions in Tanzania most affected by the HIV/AIDS epidemic (World Bank (1992), Beegle (2005)). Kagera is also the region where AIDS cases were reported first in hospitals in Tanzania (Tibaijuka (1997)). In 1983, the first 3 AIDS cases were reported and the number of cases increased rapidly to 5,116 cases in 1994. On the other hand, the share of reported AIDS cases in Kagera to Tanzania decreased from 100% in 1983 to 10% in 1994. In 2003, the percentage of HIV positive in Kagera among age 15-49 is 3.7% while the figure in Tanzania is 7.0% (TACAIDS (2005)) and thus HIV/AIDS pandemic in Kagera has been alleviated compared to other regions in Tanzania. We use the Kagera Health and Development Survey (KHDS) which collects the detailed information on households in Kagera in 1991-94 and 2003. The survey samples households hit by prime-age adult mortality more than households without the mortality and the data allow us to study the long-term effects of prime-age adult mortality on agricultural production. In the data, 36.7% of prime-age adult mortality is considered to be due to HIV/AIDS by deceased individuals' families.

We will focus on agricultural production among other income generating activities and we will study the effects of prime-age adult mortality on agricultural production in the region. Agriculture is the major income source in Kagera and also in Tanzania. In Kagera, 85% of household heads engage in agriculture in 2000/01 while 70% in Tanzania (JBIC (2006), Tanzania NBS (2002)). In Kagera, households engage in subsistence and traditional agriculture. Male adult members produce coffee and banana with or without cattle manure. Female adult members produce crops such as maize and yams mainly for own consumption. Prime-age adult mortality affects their production since their family business is labor-intensive. As shown below, households hit by prime-age adult mortality between 1990 and 2003 have less increase in household members by 1 person from 1991 and 2003 than households without the mortality. They also accumulate less other productive assets; land and livestock. As a consequence, their agricultural output growth is also smaller.

However, we do not find such clear differences in per capita asset accumulation and output growth between households with prime-age adult mortality and those without it. In order to explore the effects of the mortality on agricultural production more, we will study the difference in TFP growth. We study the hypothesis that a household hit by the mortality cannot increase TFP as much as a household not hit by it. We also decompose agricultural output growth into the contribution of the accumulation of each productive asset and TFP growth and compare the differences in those factors between households with and without the mortality.

The remainder of this paper is organized as follows. Section 2 reviews the previous studies on the effects of prime-age adult mortality on households' welfare based on household level micro data and the differences between the previous studies and this study. Section 3 outlines our conceptual model, hypothesis, and framework of empirical methods. Section 4 explains the characteristics of the original data, especially with respect to prime-age adult mortality, how we construct our data for the analysis from the original KHDS data, and discuss the relevancy of our specification of the model to study the data. Our empirical methods are explained in more details in Section 5 and the empirical results are shown and discussed in Section 6. Section 7 concludes this paper.

2. Previous studies

Whether and how much HIV/AIDS epidemic affects a household welfare is the important topic. We can categorize the literature of the effects of prime-age mortality due to HIV/AIDS on household welfare into consumption studies and production studies. Beegle et al. (2008) studies the effects of prime-age mortality on long-term consumption growth based on KHDS. Their regression equations have change in logarithm of per capita consumption from 1991 to 2003 as the dependent variable and dummy variables for deaths as explanatory variables. They use household fixed effects methods in order to control unobserved time-invariant characteristics and relax the endogeneity and self-selection problem of HIV/AIDS as other previous studies based on panel data do. They take into account which year each death occurred by using dummy variables for deaths in 1991-1995, 1996-1999, and 2000-2004. Their results show that the coefficients of dummy variables for deaths are negative but only dummy variables for deaths in 2000-2004 are statistically significantly different from zero. This characteristics of the results are robust in various specification of regression equations. Their results imply that there are negative effects of prime-age adult mortality on consumption growth but households may recover from the negative shock of the mortality after 5 years. They find that a prime-age adult death results in a 7% drop in consumption in the first 5 years after the death. Carter et al. (2007) use KwaZulu-Natal Income Study (KIDS), South Africa data and study the effects of prime-age mortality due to HIV/AIDS on long-term growth rate of per capita consumption and find the negative coefficients for dummy variables for deaths although they are not statistically significantly different from zero. They also find the large magnitude of the negative effects: a prime-age adult death lowers a household's 5 year growth rate by 21%.

Although the consumption studies above find the negative effects of prime-age adult mortality on household 5-year consumption growth, channels of the causality has not been made clear. Production studies analyze some potential channels of the causality. Beegle (2005) uses the first 4 waves of KHDS from 1991 to 1994 and studies the short-term effects of prime age adult mortality in a household on the household members' labor supplies. She constructs dummy variables for male and female deaths in future and past 0-6 months and 7-12 months and uses them as explanatory variables in regression equations. The dependent variables are the probabilities of (1) being in wage employment, (2) non-farm self-employment, (3) working on coffee production, (4) banana production or (5) maize, cassava, or beans production. She finds coefficients of some dummy variables for deaths are negative and statistically significantly different from zero in regression equation of (1) being in wage employment, working on (3) coffee production and (5) maize, cassava, or beans production. Yamano and Jayne (2004) use two-year panel of rural Kenyan households and study the effects of prime-age adult mortality on households' size and composition, crop production, asset levels and off-farm income. They find the mortality decreases households' size, area under high-valued crops, gross and net outputs, farm equipment, small animals, and off-farm income. They find that the death of a male household head is associated with a 68% reduction in the net value of the household crop production implying large negative effects of the mortality on households welfare and that channels of the causality are decreases in productive inputs above. Chapoto and Jayne (2008) use nationally representative 3-year panel data in Zambia and find the results similar to Yamano and Jayne (2004).

These production studies show that the negative effects of prime-age adult mortality on household income and channels of the causality. HIV/AIDS also increases an household's expenditure for medical care for the sick and funeral for the deceased. Tibaijuka (1997) finds that this expenditure is almost equivalent to the cash income for the 10 households in her data from Kagera, Tanzania. We can think that decreased income and increased expenditure for health care and funeral due to HIV/AIDS and prime-age adult mortality contribute to the decreased consumption which is found in the consumption studies above. Households hit by HIV/AIDS have to face tighter budget constraints and invest less in productive assets than the other households. Smaller investment in productive assets brings smaller income in the future.

We contribute to the literature with the following three points. First, we provide an answer to the question whether prime-age adult mortality decreases total factor productivity (TFP) in the long run. Previous studies do not ask this question although it is an important question to study the channels from prime-age adult mortality to decreased income and welfare. This question is closely linked to the question how important an adult's knowledge stock of agriculture is for his/her household income generation. Since subsistence agriculture in Kagera, Tanzania depends on weather and is erratic, the knowledge may be important. On the other hand, its agriculture is traditional and does not depend on new technologies and new market opportunities so much, the knowledge may not be important. If the knowledge is important, prime-age adult mortality destroys not only household members but also the quality of household as an agricultural enterprise.

Second, we decompose the agricultural output growth into TFP growth and the contribution of each productive asset. Previous production studies analyze the effects of prime-age adult mortality on each productive asset separately and cannot show how much change in each productive asset due to prime-age adult mortality contributes to change in agricultural income or output. We quantify this channel from change in each productive asset to change in agricultural output by estimating an agricultural production function and decomposing the long-term change in agricultural output growth into TFP growth and change in contribution of each productive asset for households with and without prime-age adult mortality.

Third, we study the effects of prime-age adult mortality on long-term agricultural production and link the previous studies on long-term consumption with the previous studies on short-term change in production mentioned above in this section.

3. Model and hypothesis

Agricultural production is represented by the following function:

$$Y_{jt} = A_{jt} M_{jt}^{\theta_m} K_{jt}^{\theta_k} S_{jt}^{\theta_s} \quad (1)$$

where Y_{jt} is agricultural output, A_{jt} is unobserved productivity (total factor productivity, TFP), M_{jt} is the number of household members, K_{jt} is land in square meter, and S_{jt} is monetary value of livestock for household j at year t. We will discuss the relevancy of this specification of production function in Section 4.4 below after outlining the data. The purpose of this section is to make clear our model, framework of empirical methods and hypothesis without referring application to the particular data. By taking logarithm of both sides of the production function (1), we have

$$y_{jt} = a_{jt} + \theta_m m_{jt} + \theta_k k_{jt} + \theta_s s_{jt} \quad (2)$$

where lower letter is logarithm of upper letter, for example, $y_{jt} = \ln Y_{jt}$. For notational simplicity, we denote productive assets and their coefficients as follows:

$$x_{jt} = (m_{jt}, k_{jt}, s_{jt})' \quad (3)$$

$$\theta = (\theta_m, \theta_k, \theta_s)' \quad (4)$$

By estimating θ, we can recover a_{jt}. Then, we compute change in each variable over time for households with and without prime-age adult mortality. Denote the dummy variable for prime-age adult mortality for household j at period t

by d_{jt}. In order to test whether prime-age adult mortality affects agricultural production, we test the average change in each variable for households without the mortality is larger than one for households with the mortality. For example, we test the null hypothesis $\frac{1}{n_0}\sum_{d_{j,t-1}=0}\Delta y_{jt} = \frac{1}{n_1}\sum_{d_{j,t-1}=1}\Delta y_{jt}$ against the alternative hypothesis $\frac{1}{n_0}\sum_{d_{j,t-1}=0}\Delta y_{jt} > \frac{1}{n_1}\sum_{d_{j,t-1}=1}\Delta y_{jt}$ where n_1 and n_0 are the numbers of households with and without the mortality, respectively. We expect we can reject the null hypothesis for each variable, that is, prime-age adult mortality decreases both TFP growth and the accumulation of each productive asset and thus it decreases agricultural output growth. We are also interested in quantifying difference in contribution of TFP growth and the accumulation of each productive asset to agricultural output growth between households with and without the mortality.

The empirical method in Section 5.2 below tests the hypothesis that prime-age adult mortality decreases TFP within estimation step as well as decomposition step above. We divide a_{jt} into production shock ϵ_{jt} and productivity state variable ω_{jt} and then estimate transition function of ω_{jt}. We test whether the coefficient of prime-age adult mortality $d_{j,t-1}$ is statistically significantly negative in the estimated transition function of ω_{jt}.

Our goals are (1) to test whether prime-age adult mortality decreases TFP as endowment of knowledge or efficiency of a household as an agricultural enterprise, (2) to test whether households without prime-age adult mortality have larger increase in each variable; agricultural output, TFP, each productive asset than households with the mortality, (3) to quantify differences between households with and without the mortality in decomposition of agricultural output growth into TFP growth and the contribution of the accumulation of each asset in order to quantify channels from the mortality to decreased agricultural output growth. When perusing these goals by the framework above, we have to take into accounts (1) the endogeneity of observed productive asset level, (2) the endogeneity of observed prime-age adult mortality, and (3) that TFP A_{jt} may include not only endowment of knowledge but also other productive assets or state variables which affect agricultural production.

Prime-age adult mortality decreases human capital as the number of household members M_{jt} directly. This is the direct negative effect on long-term agricultural production. There are other channels through which prime-age adult mortality affects agricultural production. Tibaijuka (1997) emphasizes households lose family labor not only due to death but also long-period of illness before death and care for the sick adult. Since our production function has the number of household members instead of labor input into agricultural production, this negative effect is captured by smaller productivity A_{jt}. Beegle (2005) studies short-term labor response to prime-age adult mortality as outlined in Section 2. If household members work more in farm when one of its members is sick or died, productivity A_{jt} captures this labor response in our specification. If a deceased adult has more knowledge about agricultural production or management skill than other members, the mortality decreases productivity A_{jt} as the quality of human capital as well as the quantity of human capital M_{jt}.

Prime-age adult mortality affects the accumulation of productive assets in the following two ways: First, some of land and livestock may be inherited from the deceased adult to its children who residing outside of the deceased adult's household. Second, prime-age adult mortality changes the household investment decisions. We can categorize channels through which the mortality changes the investment decision into two: First, the household changes future asset accumulation path as a response to changes in current asset levels due to the mortality and inheritance. For example, the household may sell land and livestock in order to achieve efficient and smaller productive asset level as a response to decreased household members and productivity due to the mortality. Second, the household's budget constraint becomes tighter due to the mortality and the household has to change its allocation of income into consumption and investment over time. The household lost labor for income generation since the member who was sick and deceased did not and will not contribute to the household as labor and other members take care of the sick and thus the household income decreases (Tibaijuka (1997)). Furthermore, the household faces expenditure for medical care and funeral. Tibaijuka (1997) finds that this expenditure is almost equivalent to the cash income for the 10 households in her study.

4. Data

In this section, first we briefly describes that the original KHDS data intends to sample households affected by HIV/AIDS more than other households (Section 4.1). Then, we explain the characteristics of prime-age adult mortality and discuss about the age range of prime-age for our analysis (Section 4.2). In Section 4.3, we explain how we select subset of households for our analysis from the original KHDS data and construct the data for our analysis. Then, we discuss the relevancy of our specification of agricultural production function (1).

4.1. The Original Data

We use Kagera Health and Development Survey (KHDS), Tanzania. KHDS has 5 waves of survey. The first 4 waves were collected between 1991 and 1994 and the last wave (wave 5) was collected in 2003. Wave 1 and wave 5 are annual surveys and asked households about the past 12 months. On the other hand, wave 2, 3, and 4 are half-year surveys and asked households about the past 6 months.

KHDS is not a survey which represents Kagera or Tanzania. The original objective of the survey (in 1991-1994, wave 1-4) was to study effects of HIV/AIDS on households. Kagera is one of the regions in Tanzania where households are severely affected by HIV/AIDS. When investigators chose households in Kagera, first they chose clusters (places) and then chose households in each cluster. Each cluster is categorized into the following four agronomic zones: (1) tree crop zone, (2) riverine zone, (3) annual crop zone, and (4) urban zone and each zone has

10-15 clusters. 26 out of 51 clusters are selected from the wards (administrative area that is smaller than districts) where adult mortality rate is very large (more than 90 % quantile) and the remaining 25 clusters are selected from the remaining wards with smaller adult mortality rates. At each cluster, each household is categorize as "sick" if the household had an adult death due to illness in the past 12 months, adult too sick to work at the time of survey, or both. If not, the household is categorized as "well". At each cluster, 14 households were randomly selected from the "sick" households and 2 households were randomly selected from the "well" households. Total 816 households (= 16 households times 51 clusters) were enumerated in 1991 (wave 1). See World Bank (2004) for the details of wave 1 to 4.

The data is very unique in the sense that investigators in 2003 (wave 5) try to trace out all household members in 1991-1994 (wave 1 to 4). Households split over a decade and the number of households increased from 816 to 2,774 between wave 1 and 5. See Beegle et al. (2006) for the detail of wave 5.

4.2. Prime-age adult mortality

4.2.1. The age range for prime-age adult

Although there is no consensus on what adult age range we should use to study the effects of adult mortality on household welfare[1], we set the age range for prime-age adults is from 15 and 50. In this subsection, we discuss the relevancy of this age range. Our focus is the effects of prime-age adult mortality on agricultural production. We will focus on prime-age adult's death rather than other household members' deaths since prime-age adults contribute to their household as main labor force for agricultural production and they are the age group who are affected by HIV/AIDS directly.

We set the lower bound of prime-age adult to be 15 since 15 year old individuals is physically adult and start to face the risk of HIV/AIDS through heterosexual sex. Although under 15 year old children can contribute to their households with their labor, we do not think that decreasing the lower bound would change the results since most of them do not die due to HIV/AIDS shown below.

On the other hand, we set the upper bound of the age range at 50. Figures 1 and 2 show the distribution of age by gender in the data. Figures 3 and 4 show the distribution of deceased individuals' age by gender. Figures 5, 6, 7, and 8 show the distribution of age of deceased individuals due to HIV/AIDS

[1]Ainsworth et al. (2005) and Ainsworth and Dayton (2003) study the impacts of adult mortality with data from wave 1 to 4 of KHDS and use the same age range, from 15 to 50. On the other hand, Beegle et al. (2008) study the impacts of adult mortality with data from wave 1 to 5 of KHDS and use the different age range, from 20 to 55. Carter et al. (2007) study the impacts of adult mortality in KwaZulu-Natal region in South Africa and use the different age range, from 20 to 50. Yamano and Jayne (2004) study the effects of adult mortality with two-year panel of rural Kenyan households and define age range of working-age adult as 15-54 for men and 15-49 for women. Chapoto and Jayne (2008) study the effects of adult mortality with nationally representative 3-year panel data in Zambia and use age range of 15-59.

(by gender and by whether diagnosed by a health professional or thought due to HIV/AIDS by the deceased's family, respectively). These figures show that most of deceased due to HIV/AIDS are 22-45 years old males and 15-50 years old females. This observation and the fact that age 15-50 are main labor for household production are the two main reasons why we set the age range to be from 15 to 50.

Another reason why we set the upper bound of the age range at 50 is that KHDS did not ask mortality or illness for below 15 or above 50 when KHDS chose sample households. As we discuss in the following subsection, 33% of prime-age adult mortality in the data is enumerated when KHDS chooses sample households. We need to set the upper bound at 50 or less to include these data into our analysis consistently.

Figures 1-8 will be inserted around here.

Figures 1-8 will be inserted around here.

4.2.2. Descriptive statistics of prime-age adult mortality

Here, we show the characteristics of prime-age adult mortality in the data. There are 6,681 individuals are surveyed in wave 1, 2, 3, or 4 (1991, 92, or 93). Out of these 6,681 individuals, 988 died between 1991 and 2004 and their deaths are recorded in the KHDS. Note that since wave 5 in 2003 asks mortality only for individuals who were household members in wave 1-4 (1991-1994), there can be other deaths which are not recorded in the KHDS. While these 6,681 individuals have individual ID for KHDS, KHDS records other 377 individuals who do not have individual ID since some of them (347) died in the 12 months just before wave 1 and others (30) joined a survey household and died between waves. Thus, KHDS records the details of total 1,365 (=988+347) deaths.

Among 1,365 deaths, 844 deaths are deaths of individuals whose ages are between 15 and 50 when they died. Out of these 844 prime-age adult deaths, 743 deaths are as the result of illness. Out of these 743 illnesses, 398 illnesses are diagnosed by a health professional and 188 are reported as HIV/AIDS. Thus, 47.2% (= 188/398) of diagnosed illnesses are reported as HIV/AIDS. KHDS also asks a respondent in a household what illness the respondent think the died person was suffering from. Out of 743 illnesses, 36.7% (273) illnesses are thought as HIV/AIDS. Out of 844 prime-age adult deaths, 32% (273) deaths are due to HIV/AIDS although respondents may not have enough knowledge about health to understand the cause of death correctly.

As mentioned above, KHDS intended to sample households hit by adult mortality more than other households. KHDS calls the sampling stage before main survey as "enumeration". The enumeration before wave 1 asks whether any adult with age of 15-50 has died in the past 12 months. Then, if so, it asks the ages of each adult and the cause of the death. The cause of the death has only 4 categories: illness, accident, child birth, and other. It does not ask gender of each adult nor any further individual characteristics.

The enumeration recorded 499 deaths. We checked the duplication of deaths between one in the enumeration and one in wave 1. The enumeration was implemented between March 15 and June 13, 1991 while wave 1 was implemented between September 30, 1991 and May 10, 1992[2]. We found 83 duplications although we could rely on only household ID and the age of died adult to find duplications. Thus, the enumeration before wave 1 provides information on 416 (=499-83) adult deaths. Figure 9 shows the age distribution of these died adults. Out of these 416 died adults, 413 adults died due to illness. Figure 10 shows the age distribution of these adults died due to illness.

Figures 9-10 will be inserted around here.

Figures 9-10 will be inserted around here.

We think we should include these mortality in analysis since our focus is effects of adult mortality and there are huge numbers of adult mortality in the enumeration and before wave 1[3]. As we mentioned in the previous subsection, one of the reasons why we set upper limit of prime-age adult at 50 is that the enumeration does not record mortality of individuals whose ages are more than 50.

The reasons why we do not distinguish adult mortality due to HIV/AIDS and one due to other causes are (1) the sample size is not so large, (2) whether the cause is HIV/AIDS is not clear, and (3) the enumeration does not ask whether the cause is HIV/AIDS. Previous studies mentioned that HIV/AIDS is more harmful than other mortality or illness since a household suffers from the longer period of sick before death and other members' care for the sick. Since we do not think we have proper data to study the difference in the effects of HIV/AIDS and those of other illness and mortality, we focus on the effects of prime-age adult mortality on long-term agricultural production.

Table 1 shows the number of prime-age of adult deaths by cause and by year. Most of deaths recoded in the data are in 1990 and 1991. This characteristic is due to KHDS's unique sampling strategies. First, KHDS intentionally sample households which suffered from prime-age adult mortality, more precisely, 14 out of 16 households have prime-age adult mortality in the last 12 months, prime-age adult who is too sick to work or both in the enumeration. Second,

[2]Rigorously speaking this time period is for passage 1 rather than wave 1. KHDS surveyed household with wave 1 questionnaire in passage 2 and passage 3 too. Passage 2 is implemented between April 23, 1992 and November 30, 1992 and Passage 3 is implemented between November 14, 1992 and May 25, 1993. However, most of households are surveyed with wave 1 questionnaire in passage 1.

[3]Beegle et al. (2008) do not include both mortality in the enumeration and mortality of individuals without ID in their analysis, more precisely they do not include 416 deaths in the enumeration and 377 deaths of individuals without individual ID in their analysis.

in wave 5 (in 2003), KHDS does not ask death of individuals who were not household members in previous waves (in 1991, 92, or 93) even if an individual was a household member when he or she deceased. We should take into account that even we call prime-age adult mortality between 1990 and 2003, most of death occurred in 1990 and 1991.

Table 1 will be inserted around here.

Table 1 will be inserted around here.

Table 2 shows the number of households by year and by number of prime-age adult death. As we explain in Section 4.3, we use 401 households out of all households in the original data. There are households which suffer multiple deaths. The number of households which has 0, 1, 2, 3, 4, 5, and 6 deaths are 152, 117, 82, 38, 10, 1, and 1, respectively as shown in Table 2. 56% (249) households have prime-age adult mortality between 1990 and 2003. This table also show that most of prime-age adult death in the data occurred in 1990 and 1991, which is due to KHDS's sample selection scheme as mentioned above.

Table 2 will be inserted around here.

Table 2 will be inserted around here.

Wave 5 of KHDS (in 2003) asked households (1) whether each of the past ten years was a very bad year or not, (2) if so, why it was, and (3) if so, how did they cope with it. As the answer to (2) for year 2003, 25% of 376 individual singled out death of family member, 22% did poor harvest due to weather and 20% did serious illness. As the answer to (3), each individual could answer at most two and there are 525 answers for 2003 from 376 individuals. The content and percentage of each answer is as follows: rely on support from family and friends (30%), reduce consumption (19%), take casual employment (14%), introduce other crops (7%), sell livestock (6%), sell other assets (6%), start other business (5%), start selling processed food (3%), and sell land (2%). These results imply that mortality and illness are the most serious negative economic shock for the households and households respond to it in various ways. We do not study short-term responses although Beegle (2005) studies short-term labor responses to prime-age adult mortality as mentioned in Section 2. Instead, we study the long-term consequences in agricultural production after being hit by prime-age mortality and responding to it.

4.3. Household subsample

We need homogeneity in households in the sense that households solve the same or at least a similar economic problem. In this subsection, we discuss what subsample of households we choose from the original data. In summary, we choose households which engage in agriculture mainly and we exclude households which emigrate from the original location and new households which split from the original households over a decade from our analysis.

Wave 5 of KHDS (in 2003) tracks households and their members who emigrated between 94 and 03. However, investigators do not ask those emigrated households about their agriculture less than non-emigrated households in order to reduce work load for tracking phase and thus the data on agriculture are much less complete compared to non-emigrated households. Since the data on agricultural outputs and productive assets for emigrated households are not collected, we simply drop emigrated households from our analysis. Unfortunately, the number of emigrated household are large: there are 1,413 emigrated households out of all 2,774 households in 2003. However, we should not say 51% (1,413 out of 2,774) households emigrated. First, these 2,774 households in 2003 includes split households from the original 919 households in 1991 and 1992. Second, 540 out of 1,413 emigrated households emigrated to nearby villages. If we take household unit in 1992, total 830 households are resurveyed in 2003[4]. Out of them, 733 households have at least one new household unit (household unit in 2003) which remained in the same village. 46 households do not have any new household units which remained in the same village but have at least one new household unit which emigrated to a nearby village. The remaining 51 households emigrated in the most restricted definition, that is, do not have any new household units which remained in the same village or emigrated to a nearby village.

We exclude households in the most urbanized four clusters since the model does not have occupational choice and poverty dynamics in urban area is very different from the one in rural area we study. The ratio of employment income compared to agricultural output increased a lot in these four most urbanized clusters from 1994 to 2003. Although one fourth of households in wave 1 live in urban zone as mentioned above, we include households in urban zone except households in the most urbanized four clusters since urban zone except the most urbanized four clusters seems to be as agriculture-oriented as other zones in 1991-1994[5]. We drop 55, 51, and 41 households in these four clusters in 1991, 1992, and 2003, respectively.

In order to focus on agricultural households, we drop households whose non-agricultural income or transfer income is larger than agricultural output. We also drop households which have outliers in variables used in our analysis.

We exclude households which split from the original household between 1992 and 2003 and which do not seem to be continuing households from 1992. More particularly, we exclude the following households: If there is a main household

[4]Figure 1 of Beegle et al. (2006) says that there are 912 original households although we find 919 households. Out of them, 17 households deceases, 63 households untraced, and 832 households are re-interviewed but we find 830 reinterviewed households.

[5]de Weerdt (2006) also excludes four urban clusters from his analysis. Since he did not mention what clusters they are, we could not confirm his four clusters are the same as our four clusters.

where household head is the same over 1992 and 2003 and there is another household which was split from the main household between 1992 and 2003, for example, a son's new household, we exclude the split household and focus on the main household. If a household head passed away between 1992 and 2003 and there are two households in 2003, for example, older brother's new household and younger brother's new household, we choose only one household as the continuing household and exclude the other household from our analysis. Table 3 shows the results of this selection of households. See Appendix Appendix A.1 for the detail on how to choose a continuing household.

Table 3 will be inserted around here.

Table 3 will be inserted around here.

4.4. The relevancy of the specification of agricultural production function

In this subsection, we discuss the relevancy of our specification of agricultural production function (1). We exclude a household from analysis if its agricultural output is smaller than non-agricultural income in order to focus on household income generation with subsistence agriculture. We think household members, land and livestock are the three main productive factors/assets for the agricultural production in Kagera region. We use the number of household members instead of labor hour input into agricultural production. Although main labor input is household member's labor, some household use hired labor. For example, in the original KHDS data, 26% of (231 out of 888) and 33.3% of (553 out of 1,663) households used hired labor on their shamba (land for crop) in the past 12 month in wave 1 (1991) and wave 5 (2003), respectively. Also, 10.9% of (121 out of 1,106) households used paid labor for herding in the past 12 month in wave 5 (2003). Although we do not take into account 1) that household members use some labor hours in non-agricultural activity and 2) the differences in gender and age among household members, we do not think it is a shortcoming for our purpose. Our objective is to understand the effects of prime-age adult mortality on long-term income generating power of subsistence agricultural households and production function (1) is a reduced form of household income generation. If households are homogeneous in the sense that they engage in a similar income generation and face a similar economic problem, the estimated production function provides quantitative insights on the effects of the mortality on income generation and the decomposition of the effects into productivity and each productive asset. Since productivity includes all heterogeneities among households except those in the number of household members, land and livestock, if we included more heterogeneities in input, productivity would become less ambiguous in what it includes. However, we do not think it is our primal objective.

Manure from livestock is important for agriculture in Kagera. Smith (2001) documents that (1) farmers use manure sparingly and efficiently, they mix ash, mulch and composted manure into the holes in which coffee and banana trees are planted and (2) farmers who optimize their use of manure can produce yield up to five times higher than their neighbors who cannot afford cows (p.51). In his data, three-fifths of male farmers use manure (p.167) and all the farmers interviewed wanted to buy a cow to increase their herd in order to improve farm productivity (p.126). The importance of manure is due to the fact that most of Kagera farmer do not use fertilizer. For example, in original KHDS data, only 5.3% (47 out of 888) and 3.2% (53 out of 1,666) households use fertilizer in wave 1 (1991) and wave 5 (2003), respectively. Complementarity between crop production and livestock is mainly due to manure since households do not use cattle for plowing. Complementarity between land and livestock is weak since a household uses communal land for grazing instead of its own household's private land. Complementarity between land and the number of household members is also weak since households use a cattle owner association called *omukondo* which has twenty or so member households, pasture area, and a herd manager and each household does not have to use its own household member for herding.

In our constructed data with total 401 households, there are 160, 119, and 138 households who have zero monetary value of livestock in 1991, 1992 and 2003, respectively. In order to accommodate these household into our analysis, we define livestock S_{jt} is the real monetary value of livestock plus one.

Agricultural output Y_{jt} does not include sale, purchase, and own consumption of livestock. Because of this specification, livestock contributes to agricultural output by providing manure, egg and milk. We may underestimate income generating power of livestock. On the other hand, if we included sale, purchase, and own consumption of livestock, we might overestimate household permanent income generating power when the household is hit by negative economic shock and sells livestock in order to smooth income and consumption. Also, we might underestimate household permanent income generating power when the household purchase livestock as investment for future income generation. We keep this topic as one of our future research topics.

5. Empirical Methods

5.1. Difference-in-difference estimates

We use difference-in-difference estimates in order to study the differences between households with and without prime-age mortality in agricultural output growth and the accumulation of each productive asset. Denote the dummy variable for prime-age adult mortality for household j at period t by d_{jt}. In order to test whether prime-age adult mortality affects agricultural production, we test the average change in each variable for households without the mortality is larger than one for households with the mortality. For example, we test the null hypothesis $\frac{1}{n_0}\sum_{d_{j,t-1}=0}\Delta Y_{jt} = \frac{1}{n_1}\sum_{d_{j,t-1}=1}\Delta Y_{jt}$ against the alternative hypothesis $\frac{1}{n_0}\sum_{d_{j,t-1}=0}\Delta Y_{jt} > \frac{1}{n_1}\sum_{d_{j,t-1}=1}\Delta Y_{jt}$ where n_1 and n_0 are the numbers of households with and without the mortality, respectively. We expect we can reject the null hypothesis for each variable, that is, prime-age adult mortality decreases the accumulation of each productive asset and thus it decreases agricultural output growth. Note that we need justification for taking d_{jt} as exogenous shock rather than endogenous shock. We will discuss on this point with the data in Section 6.1 below.

5.2. Estimating the agricultural production function and TFP growth

In this section, we show two methods for estimating the production function and recovering total factor productivity. The first objective is to estimate the agricultural production function, to compute TFP based on the estimates, and to apply difference-in-difference method to estimated TFP in order to test whether prime-age adult mortality decreases TFP growth. The second objective is to decompose the growth of (the logarithm of) agricultural output Δy_{jt} into TFP growth Δa_{jt} and the contribution of the accumulation of each productive asset, for example, the one of household members $\Delta\theta_m m_{jt}$ and study the difference between households with and without the mortality in each of the decompositions. This comparison allows us to quantify the magnitude of each channel from the mortality to decreased agricultural output growth. The third objective is to estimate a transition function of productivity state variable ω_{jt} and test whether the mortality affects the transition. The second method allows this test.

5.3. Unobserved time-invariant heterogeneity in productivity

We introduce new variables and notations for prime-age adult mortality. $\tilde{d}_{j,-\tau,t}$ takes 1 if household j experiences one or multiple prime-age adult deaths τ years ago counting from year t and takes 0 otherwise. $d_{j,-\tau,t} = \tilde{d}_{j,-\tau,t}1\{t = \tilde{t}\}$ where $\tilde{t} = 1991, 1992$ or 2003 and $1\{\cdot\}$ is integer function. For example, $d_{j,-1,1991}$ takes 1 if the year when we estimate production is 1991 and household j experience one or multiple prime-age adult deaths in 1990, one year ago counting from 1991. We include this $d_{j,-\tau,t}$ into productivity term for year t as an explanatory variable and test whether the coefficient of $d_{j,-\tau,t}$ is negative and statistically significantly different from zero and quantify the negative effects of prime-age adult mortality on productivity growth.

We divide productivity a_{jt} into a constant ω, residual ϵ_{jt}, negative effects of mortality on productivity growth $\gamma_{-\tau,t}d_{j,-\tau,t}$, and year dummies: In order to take into account the endogeneity of x_{jt} due to heterogeneity in productivity, we allow heterogeneity in time-invariant term of productivity among households:

$$y_{jt} = \omega_j + \theta' x_{jt} + \gamma_{-\tau,t}d_{j,-\tau,t} + \theta_{1992}1\{t = 1992\} + \theta_{2003}1\{t = 2003\} + \epsilon_{jt}. \tag{5}$$

where $1\{\cdot\}$ is integer function, $1\{t = \tilde{t}\}$ is the year dummy for year $\tilde{t}$, $\theta_{\tilde{t}}$ is its coefficient, and $d_{j,-\tau,t}$ is a vector which has factors with different $-\tau$ and t and $\gamma_{-\tau,t}$ is its coefficient vector. ω_j captures not only heterogeneity in time-invariant term of productivity but also heterogeneity in household characteristics which affect how likely a household has prime-age adult mortality.

We specify $\gamma_{-\tau,t}d_{j,-\tau,t}$ as follows:

$$\begin{aligned}\gamma_{-\tau,t}d_{j,-\tau,t} &= \gamma_{0,1991}d_{j0,1991} + \gamma_{-1,1991}d_{j,-1,1991} \\ &+\gamma_{0,1992}d_{j0,1992} + \gamma_{-1,1992}d_{j,-1,1992} + \gamma_{-2,1992}d_{j,-2,1992} \\ &+\gamma_{0-2,2003}d_{j,0-2,2003} + \gamma_{-3-6,2003}d_{j,-3-6,2003} \\ &+\gamma_{-7-10,2003}d_{j,-7-10,2003} + \gamma_{-11-13,2003}d_{j,-11-13,2003}\end{aligned}$$

where d_{j0t} takes 1 if a household experience one or multiple prime-age adult deaths in year t and production equation is one for year t. Similarly, $d_{j,-1,t}$ takes 1 if a household experiences death(s) in year $t-1$ and production equation is one for year t. $d_{j,-3-6,t}$ takes 1 if a household experiences death(s) between year $t-3$ and year $t-6$ and production equation is one for year t.

Note that Table 2 shows that we do not have so many observations on death that we can estimate production with dummies for death in each year except 1990 and 1991. Note also that our observation on death is restricted only to death between 1990 and 2003 and death of individuals who are household members between 1990 and 1993 since KHDS in 2003 ask mortality for those original household members between 1990 and 1993. These are reasons why we specify $\gamma_{-\tau,t}d_{j,-\tau,t}$ as above by aggregating death over 3 or 4 years and by using different coefficient for different year t.

Note that it is better to add $\gamma_{-\tau,2003}d_{j,-\tau,t}$ since we found households with prime-age adult mortality accumulate less productive assets and it implies that they have lower productive growth between 1991 and 2003. If we did not add $\gamma_{-\tau,2003}d_{j,-\tau,2003}$, $\mathrm{E}\left[x_{j2003}\epsilon_{j2003}\right] = 0$ would not hold. Since we control negative effects of prime-age adult mortality on productivity growth, the estimation equation satisfy $\mathrm{E}\left[x_{j2003}\epsilon_{j2003}\right] = 0$ more likely than a regression without $\gamma_{-\tau,2003}d_{j,-\tau,2003}$.

5.4. Relax time-invariant ω_j

The main limitation on empirical methods above is that change in productivity over time is modeled with random term ϵ_{jt} and we do not allow the systematic productivity growth over time. In order to relax this limitation, we introduce a more structural method based on Olley and Pakes (1996). The method utilizes the relationship between another variable such as investment or intermediate input and unobserved productivity which is based on household's dynamic optimization (income maximization) in order to recover unobserved productivity. The main advantage of this method compared to dynamic panel methods, another method which allows the systematic productivity growth over time, is that we do not have to specify the functional form of transition function of state variable component of productivity a_{jt}.

We divide productivity a_{jt} into ω_{jt} and ϵ_{jt}. $\omega_{j,t}$ is the term of productivity which is a state variable and it is related with next period value $\omega_{j,t+1}$. On the other hand, ϵ_{jt} is the one-time shock component of productivity and it is assume to be statistically independent from other variables.

The specification of production function is as follows:

$$y_{jt} = \omega_{jt} + \theta' x_{jt} + \epsilon_{jt}. \tag{6}$$

Productivity evolves exogenously as follows:

$$\omega_{j1992} = \mathrm{E}\left[\omega_{j1992} | \omega_{j,1991}\right] + \xi_{j1992} \tag{7}$$

$$\omega_{j2003} = \mathrm{E}\left[\omega_{j2003} | \omega_{j,1992}, d_{j,0-13,2003}\right] + \xi_{j2003} \tag{8}$$

where $d_{j,0-13,2003}$ takes 1 if household j experiences one or multiple prime-age adult deaths $0-13$ years ago counting backwards from year 2003 and takes 0 otherwise. We test whether the coefficient of $d_{j,-\tau,t}$ is negative and statistically significantly different from zero and quantify the negative effects of prime-age adult mortality on productivity growth. Note that Table 2 shows that most of deaths occurred in 1990 and 1991, our observation on death is restricted only to death between 1990 and 2003, and death of individuals who are household members between 1990 and 1993 since KHDS in 2003 asked mortality only for those original household members between 1990 and 1993.

On the other hand, asset evolves endogenously as follows:

$$M_{j,t+1} = M_{jt} + I_{Mjt}$$
$$K_{j,t+1} = K_{jt} + I_{Kjt}$$
$$S_{j,t+1} = S_{jt} + I_{Sjt}$$

For notational simplicity, denote column vectors as follows:

$$X_{jt} = (M_{jt}, K_{jt}, S_{jt})'$$
$$I_{jt} = (I_{Mjt}, I_{Kjt}, I_{Sjt})'$$

A household decides investment based on asset stock and productivity:

$$I_{jt} = I(x_{jt}, \omega_{jt})$$

Note that we are assuming that productivity ω_{jt} and x_{jt} are state variables which explain investment I_{jt} so well that we do not have to add mortality $d_{j,0-13,2003}$ as determinant of investment as well as ω_{jt} and x_{jt}. If monotonicity of $I(\cdot,\cdot)$ with respect to x_{jt} is satisfied, we can invert $I(\cdot,\cdot)$ with respect to x_{jt}:

$$\omega_{jt} = \omega(x_{jt}, I_{jt}) \tag{9}$$

where $\omega(\cdot,\cdot)$ is inverted function of $I(\cdot,\cdot)$ with respect to x_{jt}[6].

[6]Pakes (1994) shows that provided $I_{jt} > 0$, $I(x_{jt}, \omega_{jt})$ is strictly increasing in ω_{jt}. Rigorously speaking, we need to show that $I(x_{jt}, \omega_{jt})$ is increasing in ω_{jt} even if I_{jt} is negative or zero. We keep it as a future research topic to study under what conditions this monotonicity condition and invertibility condition are satisfied.

We will utilize the following condition:

$$\mathrm{E}[\epsilon_{jt} + \xi_{jt}] = 0. \tag{10}$$

In order to do so, we compute $\epsilon_{jt} + \xi_{jt}$ as follows: By (6), (7), and (8), we have

$$\epsilon_{j1992} + \xi_{j1992} = y_{j1992} - \theta' x_{j1992} - \mathrm{E}\left[\omega_{j1992}|\omega_{j,1991}\right] \tag{11}$$

$$\epsilon_{j2003} + \xi_{j2003} = y_{j2003} - \theta' x_{j2003} - \mathrm{E}\left[\omega_{j2003}|\omega_{j,1992}, d_{j,0-13,2003}\right]. \tag{12}$$

Thus we need $\mathrm{E}\left[\omega_{j1992}|\omega_{j1991}\right]$ and $\mathrm{E}\left[\omega_{j2003}|\omega_{j1992}, d_{j,0-13,2003}\right]$ and we obtain them as follows: By substituting (9) into (6), we have

$$y_{jt} = \phi(x_{jt}, I_{jt}) + \epsilon_{jt} \tag{13}$$

$$\text{where } \phi(x_{jt}, I_{jt}) = \theta' x_{jt} + \omega(x_{jt}, I_{jt})$$

We estimate $\phi(\cdot,\cdot)$ as the second order polynomial function and denote the estimated function by $\hat{\phi}(\cdot,\cdot)$. By (6) and (13), we have

$$\omega_{jt} = \phi(x_{jt}, I_{jt}) - \theta' x_{jt} \tag{14}$$

Given unknown parameter value of θ, $\hat{\phi}(\cdot,\cdot)$, and (14), we can calculate the estimate of productivity $\hat{\omega}_{jt}(\theta)$ as follows:

$$\hat{\omega}_{jt}(\theta) = \hat{\phi}(x_{jt}, I_{jt}) - \theta' x_{jt}$$

Denote estimates of $\mathrm{E}\left[\omega_{j1992}|\omega_{j1991}\right]$ and $\mathrm{E}\left[\omega_{j2003}|\omega_{j1992}, d_{j,0-13,2003}\right]$ by $\mathrm{E}\left[\omega_{j1992}\widehat{|\omega_{j1991}}\right]$ and
$\mathrm{E}\left[\omega_{j2003}|\omega_{j1992}\widehat{, d_{j,0-13,2003}}\right]$, respectively. In order to obtain them, we regress the following simple transition functions:

$$\hat{\omega}_{j1992}(\theta) = \theta_{\omega1992,0} + \theta_{\omega1992,1}\hat{\omega}_{j1991} + \epsilon_\omega \tag{15}$$

$$\hat{\omega}_{j2003}(\theta) = \theta_{\omega2003,0} + \theta_{\omega2003,1}\hat{\omega}_{j1992} + \theta_{\omega,0-13,2003} d_{j,0-13,2003} + \epsilon_\omega. \tag{16}$$

and denote estimated coefficients by $\hat{\theta}_{\omega1992,\cdot}$ and $\hat{\theta}_{\omega2003,\cdot}$. Then, $\mathrm{E}\left[\omega_{j1992}\widehat{|\omega_{j1991}}\right]$ and
$\mathrm{E}\left[\omega_{j2003}|\omega_{j1992}\widehat{, d_{j,0-13,2003}}\right]$ are obtained as

$$\mathrm{E}\left[\omega_{j1992}\widehat{|\omega_{j1991}}\right] = \hat{\theta}_{\omega1992,0} + \hat{\theta}_{\omega1992,1}\hat{\omega}_{j1991} \tag{17}$$

$$\mathrm{E}\left[\omega_{j2003}|\omega_{j1992}\widehat{, d_{j,0-13,2003}}\right] = \hat{\theta}_{\omega2003,0} + \hat{\theta}_{\omega2003,1}\hat{\omega}_{j1992} + \hat{\theta}_{\omega,0-13,2003} d_{j,0-13,2003} \tag{18}$$

By substituting (11), (12), (17) and (18) into the condition (10), we can construct the following sample analogue of the condition:

$$J(\theta) = \sum_{j=1}^{n} \sum_{t=1992,2003} (\epsilon_{jt} + \xi_{jt})^2$$

where n is the number of households (= 411). Our estimates based on Olley and Pakes (1996) method are θ which minimize the objective function $J(\theta)$. We obtain standard errors of estimates by bootstrap (See Appendix Appendix B.2 for the detail).

6. Results

6.1. Difference-in-difference estimates

In this subsection, we will show the descriptive statistics in each productive asset and agricultural output in 1991 for households with and without prime-age adult mortality in order to check how these two groups of households are different in 1991 and whether the data support us in taking prime-age adult mortality between 1990 and 2003 as an exogenous shock. Then, we will show the difference-in-difference estimates of the change in each variables from 1991 to 2003.

Table 4 show the mean of each variable in 1991 for households with and without prime-age adult mortality between 1990 and 2003. We divide households simply into households with mortality and those without mortality. The table shows that there is not clear difference in productive assets and agricultural output in 1991 between households with and without prime-age adult mortality. We test the null hypothesis that the mean of each variable for households without death is the same as one for households with death and we cannot reject any of that hypotheses even with 10% significance level. These results support us in taking prime-age adult mortality as an exogenous shock.

Table 4 will be inserted around here.

Table 4 will be inserted around here.

Table 5 shows average productive asset accumulation and output growth between 1991 and 2003 for households with and without adult mortality. Note that output growth measured in the period between 1991 and 2003 but prime-age adult mortality is recorded from 1990 and 2003 and only for the household members before 1994. Average agricultural output for both households with and without mortality decreased from 1991 to 2003 and this result is consistent with that per capita food production in Tanzania decreased from 1991 to 1997 by 6% (FAO 1999; UNDP Human Development Report 1999; JICA 2001 p.2-46). Note that the households without death decreased agricultural output by 38% ($= -(-71342)/187207$) on average while the households with death decreased it by 49% ($= -(-101852)/208725$) on average. On average, mortality decreases agricultural output by 15% ($= -(-71342-(-101852))/200229$ where the denominator 200229 is average agricultural output for all households in 1991).

Table 5 will be inserted around here.

Table 5 will be inserted around here.

Apparently, households without mortality accumulate total assets and increase total agricultural output more than households with mortality. We test the null hypothesis that average change in each variable for household without death is the same as one for households with death against the alternative hypothesis that the former is larger than the latter. The test for each variable rejects the null hypothesis at 5% significance level (1% for the number of household members and livestock).

However, there are not clear differences in change of per capita land, livestock, and agricultural output between households with and without mortality. We cannot reject the null hypothesis that change in per capita land and agricultural output and rent for households without death is the same as one for households with death. In order to check whether our observation in Table 5 is robust, we make figures of distribution of change in each variable for households with and without mortality and the figures confirm our observation. These results shows the possibility that households hit by mortality endogenously respond to the negative shock and adjust productive asset level in order to improve efficiency. Thus, it is interesting to ask whether there is the difference in productivity growth between households with and without mortality. Note that we can reject the null hypothesis that change in per capita livestock for households without death is the same as one for households with death in favor of the alternative hypothesis that the former is larger than the latter. These results imply that households hit by adult death kept per capita land but per capita livestock to improve or keep per capita agricultural output.

6.2. Estimates of the production function and transition function of productivity state variable

In this subsection, we will show the estimates of the agricultural production function based on the empirical method outlined in Section 5.2. We will also show the estimate of the transition function of productivity state variable ω_{jt} and explore the question whether prime-age adult mortality decreases TFP growth.

The coefficients for $d_{-1,1992}$ and $d_{-7-10,2003}$ are negative and statistically significantly different from zero. However, the other coefficients for the effects of prime-age adult death on productivity are positive and statistically significantly different from zero except one for $d_{-1,1991}$.

Table 6 shows the results based on the fixed effect method. The coefficients for the effects of prime-age adult death on productivity are negative as we expected except those for the immediate effects ($d_{0,1991}$ and $d_{0,1991}$). Only the coeffcients for $d_{0-2,2003}$ and $d_{-7-10,2003}$ are statistically significantly different from zero, though.

Table 6 will be inserted around here.

Table 6 will be inserted around here.

Tables 7, 8, and 9 show the results based on Olley and Pakes (1996). We use net accumulation of livestock within a period (the last 12 months) as a proxy for the investment vector $I_{jt} = (I_{Mjt}, I_{Kjt}, I_{Sjt})'$. This variable has 148, 79, and 167 zero values in 1991, 1992 and 2003, respectively and it is far from an ideal variable for our purpose since the monotonicity condition of $I(\cdot,\cdot)$ with respect to x_{jt} and invertibility of $I(\cdot,\cdot)$ with respect to x_{jt} may not be satisfied. However, estimates are reasonable compared to the results based on the fixed effect method above. The coefficient for the effects of prime-age adult mortality on productivity growth ($\theta_{\omega,0-13,2003}$ in Table 8) is negative and statistically significantly different from zero. This result supports the hypothesis that prime-age adult mortality negatively affects agricultural productivity growth.

Tables 7-9 will be inserted around here.

Tables 7-9 will be inserted around here.

6.3. Decomposing the agricultural output growth

Our next question is how large the negative effect of prime-age adult mortality on productivity is compared to the negative effects of it on productive asset accumulation. We decompose the factors of increase in logarithm of agricultural output between 1991 and 2003 into the contribution of productivity growth and the contribution of accumulation of each productive asset. The logarithm of productivity is computed as $a_{jt} = y_{jt} - \theta x_{jt}$ based on estimates of coefficient θ. Table 10 shows the results. The table has 4 rows for each of the two results based on fixed effects and Olley and Pakes methods: The first row with 0 shows average change in each variable for households without prime-age adult mortality while the second row with 1 shows one for households hit by prime-age adult mortality. The third row has values of the first row minus the second row, that is, the difference of average change in each variable between these two types of households. The forth row has p-value for the test on the null hypothesis that the difference is zero against the alternative hypothesis that the difference is positive. See Appendix Appendix B.1 for the detail of this test.

Table 10 will be inserted around here.

Table 10 will be inserted around here.

As we have already seen in Table 5, the table also shows that average agricultural output growth from 1991 to 2003 is negative for both types of households and households hit by prime-age households mortality experienced more severe decrease in agricultural output than households without the mortality. The table shows how much productivity growth and accumulation of each productive asset contribute to this negative agricultural output growth. Since the results based on TOW differenct estimation methods are almost the same, we will discuss the results based Olley and Pakes methods.

The decomposition of average agricultural output growth for the households without prime-age adult mortality (the first row with 0) shows that the decrease is mostly due to the decrease of productivity rather than the decrease in productive assets. The percentage of contribution of the decrease in productivity is 91% $(= (-0.3470)/(-0.3810))$. The decomposition of output growth for the households hit by prime-age adult mortality (the second row with 1) shows that the households with mortality increase less every component of the decomposition than the households without the mortality. The percentage of the contribution of the decrease in productivity for the households with mortality is 81% $(= (-0.5210)/(-0.6414))$, which is smaller than one for the households without mortality (91%). These results imply that on average, households without mortality could kept their productive assets but the households hit by mortality could not.

The third row shows that a half of the difference in agricultural output growth between households with and without prime-age adult mortality is due to the difference in productivity growth. The percentage of how much the difference in productivity growth explains the difference in agricultural output growth is 67% $(= 0.1740/0.2604)$.

The third row also shows that how much the difference in the accumulation of each productive asset consists of the difference in agricultural output growth between households with and without prime-age adult mortality. The difference in the accumulation of household members consists the most and those of land and livestock follows. The difference in the accumulation of household members explain more than a half of the difference in accumulation of all three productive assets; the number is 58% $(= 0.0499/0.0864)$. This is reasonable since prime-age adult mortality decreases the accumulation of household members directly and may decrease accumulation of land and livestock indirectly. We could interpret that the difference in the accumulation of household members is direct negative effects of prime-age adult mortality on agricultural production. Note that although we call it as direct negative effects, we do not mean it excludes households' endogenous response to prime-age adult mortality. "Direct" means just that adult death directly decreases the number of household members. On the other hand, differences in productivity growth and the accumulation of land and livestock are indirect negative effects. Surprisingly, the results show that direct effects do not count for the largest part in the difference in agricultural output growth between households with and without prime-age adult mortality. Instead, the difference in productivity growth plays the largest role to explain the difference in agricultural output growth. The percentage for the difference in productivity growth is 67% $(= 0.1740/0.2604)$ as we mentioned above while the percentage for the difference in accumulation of household members is 19% $(= 0.0499/0.2604)$.

These results imply that households hit by prime-age adult mortality could not cope with it and not accumulate not only household members but also land and livestock as much as households without death could. Furthermore, households with the mortality could not increase productivity as much as the other households could. Surprisingly, the negative effects on productivity growth are larger than negative effects on productive asset accumulation.

The fourth row shows whether each variable for households without death is statistically significantly larger than one for households with death. Productivity growth for households without death is statistically significantly larger than one for households with death. The increase in income generating power due to accumulation of all productive assets as a whole and household members only for households without death is statistically significantly larger than one for households with death.

We can say that households hit by prime-age adult mortality could not increase income generating power in every factor among productivity growth and the accumulation of each productive asset as much as households without mortality could. A surprising result is that the difference in the accumulation of household members between households with and without mortality is not the largest factor in explaining the difference in agricultural output growth. This result implies the following two things: First, households hit by mortality could not increase or keep productivity and productive assets, land and livestock as much as households without mortality could. Thus, mortality destroys not only household human capital but also land, livestock and productivity indirectly. Second, a household hit by mortality responds to and mitigates the decrease in household members due to mortality somehow. We may think that the household tries to increase its household members or at least try to keep them by accommodating a new member through marriage or keeping current members who would move out of the household if there was no mortality. A households hit by mortality adjusts its amount of each productive asset after mortality in order to improve productivity. However, the results show that the magnitude of negative effects of prime-age adult mortality is so large that we can observe the differences in productivity growth and accumulation of each productive asset between households with and without mortality even in a long term of 13 years.

7. Conclusion

We study the effects of prime-age adult mortality on agricultural production based on asset-based poverty dynamics. Agricultural production is modeled as income generation utilizing three productive assets; household members, land, and livestock and unobserved productivity. We found that prime-age adult mortality decreases long-term agricultural output growth by decreasing both the long-term accumulation of productive assets and long-term productivity growth. We also found that the decrease in productivity growth consists of a larger share of the decrease in agricultural output growth than the decrease in the accumulation of household members although prime-age adult mortality decreases household members directly.

The direction of future research is categorized in 3 ways. First, we may study the generality of our results by applying the method to data from other regions where households engage in traditional, labor-intensive, subsistence agriculture and face prime-age adult mortality due to HIV/AIDS epidemic. Second, we may study what factors consist of unobserved agricultural productivity and understand the channels through which prime-age adult mortality affects agricultural production. Third, we utilize only a subset of the data: accumulated level of productive assets, agricultural outputs, and net accumulation of livestock. Except net accumulation of livestock, these data are outcomes of household behaviors rather than behaviors themselves. We may make a further step to understand poverty dynamics more structurally by utilizing observed accumulation of productive assets, consumption and expenditure for HIV/AIDS as endogenous behaviors as well as the outcomes of the behaviors.

References

Ainsworth, M., Beegle, K., Koda, G., April 2005. The impact of adult mortality and parental deaths on primary schooling in northwestern tanzania. Journal of Development Studies 41 (3), 412–439.

Ainsworth, M., Dayton, J., 2003. The impact of the aids epidemic on the health of the elderly in tanzania. World Development 31 (1), 131–148.

Beegle, K., 2005. Labor effects of adult mortality in tanzanian households. Economic Development and Cultural Change 53 (3).

Beegle, K., de Weerdt, J., Dercon, S., March 2006. Kagera health and development survey 2004 basic information document, mimeo. The World Bank.

Beegle, K., de Weerdt, J., Dercon, S., 2008. Adult mortality and consumption growth in the age of hiv/aids. Economic Development and Cultural Change 56 (2), 299–326.

Carter, M. R., May, J., Aguero, J., Ravindranath, S., 2007. The economic impacts of premature adult mortality: panel data evidence from kwazulu-natal, south africa. AIDS 21.

Chapoto, A., Jayne, T. S., January 2008. Impact of aids-related mortality on farm household welfare in zambia. Economic Development and Cultural Change 56, 327–374.

de Weerdt, J., October 2006. Moving out of poverty in tanzania's kagera region, mimeo.

JBIC, August 2006. Poverty profile, united republic of tanzania, japan Bank for International Cooperation (JBIC). In Japanese.

Olley, G. S., Pakes, A., November 1996. The dynamics of productivity in the telecommunications equipment industry. Econometrica 64 (6), 1263–1297.

Pakes, A., 1994. Dynamics structrual models, problems and prospects part ii: Mixed continuous-dscrete control problems, and market interactions. In: Sims, C. (Ed.), Advances in Econometrics. Cambridge University Press, Cambridge.

Smith, C. D., 2001. Ecology, civil society and the informal economy in north west Tanzania. Ashgate.

TACAIDS, 2005. Tanzania hiv/aids indicator survey 2003-04, lverton, Maryland, USA: Tanzania Commission for AIDS (TACAIDS), National Bureau of Statistics (NBS), and ORC Macro.

Tanzania NBS, 2002. Tanzanian Household Budget Survey (HBS) 2000/01. Tanzania NBS, tanzania National Bureau of Statistics (NBS). `http://www.tanzania.go.tz/hbs/HomePage_HBS.html`.

Tibaijuka, A. K., June 1997. Aids and economic welfare in peasant agriculture: Case studies from kagabiro village, kagera region, tanzania. World Development 25 (6), 963–975.

World Bank, 1992. Tanzania aids assessment and planning study, washington, DC: World Bank.

World Bank, December 2004. User's guide to the kagera health and development survey datasets, mimeo.

Yamano, T., Jayne, T. S., January 2004. Measuring the impacts of working-age adult mortality on small-scale farm households in kenya. World Development 32 (1), 91–119.

Table 1: Prime-age adult mortality by cause and by year

year	all	ill	HIV/AIDS(pro)	HIV/AIDS(family)
1990	416	413	-	-
1991	345	329	54	83
1992	83	70	22	32
1993	54	47	12	21
1994	22	19	9	14
1995	41	33	6	10
1996	24	18	6	12
1997	37	32	16	20
1998	34	28	8	11
1999	31	28	10	14
2000	41	34	8	12
2001	45	35	9	11
2002	36	31	7	10
2003	36	28	18	17
2004	15	11	3	6
total	1260	1156	188	273

Each cell has number of prime-age adult mortality death. Age range for prime-age adult is between 15 and 50. The second column "all" includes all prime-age adult's deaths. The third column "ill" includes prime-age adult's deaths due to illness. The forth column "HIV/AIDS(pro)" includes prime-age adult's deaths due to HIV/AIDS which diagnoses by a health professional. The forth column "HIV/AIDS(family)" includes prime-age adult's deaths due to HIV/AIDS which deceased's family think the cause of death is HIV/AIDS (including HIV/AIDS diagnoses by a health professional).

Table 2: Prime-age adult mortality by year, number of households

year	# of death						
	0	1	2	3	4	5	6
1990	236	100	61	4	0	0	0
1991	337	60	3	0	1	0	0
1992	381	19	1	0	0	0	0
1993	379	20	2	0	0	0	0
1994	395	6	0	0	0	0	0
1995	391	10	0	0	0	0	0
1996	391	10	0	0	0	0	0
1997	391	10	0	0	0	0	0
1998	391	10	0	0	0	0	0
1999	394	7	0	0	0	0	0
2000	393	8	0	0	0	0	0
2001	394	7	0	0	0	0	0
2002	389	12	0	0	0	0	0
2003	391	10	0	0	0	0	0
2004	397	4	0	0	0	0	0
1990-2004	152	117	82	38	10	1	1

Each cell has number of households. Total number of households is 401. Age range for prime-age adult is between 15 and 50. The (2,2) cell has 236, which means that 236 households have 0 prime-age adult death in 1990. Similarly, the (2,3) cell has 100, which means that 100 households have 1 prime-age adult death in 1990.

Table 3: Household split and attrition

1991	from 1991 889												total 899
1992	same 774			new 55			d 60			from 1992 30			919
2003 based on 1992	same	new	d	same	new	d	same	new	d	same	new	d	
	465	254	55	27	25	3	21	11	28	16	12	2	919
identified	465	224		27	17		21	10		16	11		772
not identified	0	30		0	8		0	1		0	1		59
based on 2003	1647	764		92	71		44	21		58	43		2740
the data	275	109	22	15	2		6	2		8	5		444
use?	Y	Y	N	Y	Y		N	N		N	N		401

The number of each cell is the number of households in each category.
"based on 1992" means based on household unit (ID) in 1992, that is, we aggregate households in 2003 over the same original household in 1992.
"same" and "new" represent same head and new head, respectively.
"d" represents disappear.
"identified" and "not identified" shows whether one household is identified as the continuing household among households from the same original household or not.
"based on 2003" means based on household unit (ID) in 2003 and it show how many household splits happen between 1992 and 2003. For example, the second column shows that there are 465 households with the same head over 1991 and 2003 based on household unit in 1992 and those households split and the total number of household increased from 466 in 1992 to 1,647 in 2003.
"the data" is the constructed data for our analysis.
In "use?" row, each cell has "Y" if we use households in that category in our analysis and has "N" otherwise. We do not use households which do not have all 3 time period observations and the total number of households we analyze is 401.
Note that the total number of household in "based on 2003" row is 2,740 although in data there are 2,774 households in 2003. Thus, 34 households are missing in the table, mainly due to missing data on which member is household head. Although total number of households in 1992 is 919 in this table, Beegle et al. (2006) p. 27 say that it is 912. See also Table VI.12 on page 61 of World Bank (2004) for the detail.

Table 4: Descriptive statistics of asset and agricultural output in 1991 with and without prime-age adult mortality

death	**N**	**mean**	**p25**	**p50**	**p75**
number of household members					
0	152	6.0	4	6	8
1	249	6.3	4	6	8
p-value		0.29			
land					
0	152	20456	10217	16347	28608
1	249	20810	8991	15121	26973
p-value		0.83			
livestock					
0	152	17922	0	4105	13069
1	249	28206	0	747	8458
p-value		0.18			
agricultural output					
0	159	187207	73749	127797	210658
1	242	208725	86781	150202	249426
p-value		0.33			
per capita land					
0	152	3878	1761	2835	4700
1	249	3705	1635	2725	4700
p-value		0.64			
per capita livestock					
0	152	2759	0	705	2550
1	249	4162	0	128	1452
p-value		0.23			
per capita agricultural output					
0	159	34333	14546	21993	40076
1	242	35399	16801	24913	40746
p-value		0.79			

"death" takes 1 if a household has prime-age adult mortality between 1990 and 2003 and takes 0 otherwise. "N" is number of households. "mean", "p25", "p50", and "p75" are mean, 25%, 50%, and 75% quantiles of each variable in 1991. "p-value" is the one for testing H_0 : means are the same between households with and without death against H_1 : means are different. Age range of prime-age adult is between 15 and 50. The measure of land is square meter (m^2). The measure of livestock and agricultural output and rent is Tanzania shilling (real price based on 1991). 1 United States dollar (USD) is equivalent to 219.2 Tanzania shilling (TZS) in 1991.

Table 5: Change in assets and agricultural output between 1991 and 2003 with and without prime-age adult mortality

death	**N**	**mean**	**p25**	**p50**	**p75**
the number of household members					
0	152	-0.4	-2	0	1
1	249	-1.3	-3	-1	1
p-value		0.000			
land					
0	152	-1369	-10217	-409	6743
1	249	-3849	-10217	-1635	5313
p-value		0.034			
livestock					
0	152	10560	-6373	0	13229
1	249	-8255	-2392	0	4754
p-value		0.001			
agricultural output					
0	159	-71342	-111306	-33586	22280
1	242	-101852	-152372	-63401	1548
p-value		0.029			
per capita land					
0	152	325	-1691	234	2043
1	249	317	-1499	204	2043
p-value		0.490			
per capita livestock					
0	152	2053	-912	208	2675
1	249	-376	-337	0	1102
p-value		0.008			
per capita agricultural output					
0	159	-10152	-21305	-5103	9386
1	242	-10739	-22157	-4936	4793
p-value		0.423			

"death" takes 1 if a household has prime-age adult mortality between 1990 and 2003 and takes 0 otherwise. "N" is number of households. "mean", "p25", "p50", and "p75" are mean, 25%, 50%, and 75% quantiles of change in each variable from 1991 and 2003. "p-value" is the one for testing H_0 : means are the same between households with and without death against H_1 : the mean for households without death is larger than the mean for households without death. The measure of land is square meter (m^2). The measure of livestock and agricultural output and rent is Tanzania shilling (real price based on 1991). 1 United States dollar (USD) is equivalent to 219.2 Tanzania shilling (TZS) in 1991.

Table 6: Production function with prime-age adult mortality by $-\tau$ and t and year dummies, fixed effects, bootstrap

Variable	Estimate	S.E.	P-value
Member	0.3437	0.0421	0.000
Land	0.1611	0.0291	0.001
Livestock	0.0228	0.0048	0.000
$d_{0,1991}$	0.1523	0.0924	0.217
$d_{-1,1991}$	-0.0343	0.0530	0.852
$d_{0,1992}$	0.0750	0.0843	0.576
$d_{-1,1992}$	-0.0631	0.0563	0.662
$d_{-2,1992}$	-0.0799	0.0418	0.433
$d_{0-2,2003}$	-0.1462	0.0689	0.051
$d_{-3-6,2003}$	-0.0394	0.0570	0.927
$d_{-7-10,2003}$	-0.1297	0.0541	0.047
$d_{-11-13,2003}$	-0.0627	0.0402	0.780
$1\{t = 1992\}$	-0.1702	0.0618	0.005
$1\{t = 2003\}$	-0.3782	0.0622	0.000

The estimated equation is

$$\begin{aligned} y_{jt} = \omega_j + \theta' x_{jt} + \gamma_{0,1991} d_{j0,1991} + \gamma_{-1,1991} d_{j,-1,1991} \\ + \gamma_{0,1992} d_{j0,1992} + \gamma_{-1,1992} d_{j,-1,1992} + \gamma_{-2,1992} d_{j,-2,1992} \\ + \gamma_{0-2,2003} d_{j,0-2,2003} + \gamma_{-3-6,2003} d_{j,-3-6,2003} \\ + \gamma_{-7-10,2003} d_{j,-7-10,2003} + \gamma_{-11-13,2003} d_{j,-11-13,2003} \\ + \theta_{1992} 1\{t = 1992\} + \theta_{2003} 1\{t = 2003\} + \epsilon_{jt} \end{aligned}$$

where d_{j0t} takes 1 if a household experience one or multiple prime-age adult deaths in year t and production equation is one for year t. Similarly, $d_{j,-1,t}$ takes 1 if a household experiences death(s) in year $t-1$ and production equation is one for year t. $d_{j,-3-6,t}$ takes 1 if a household experiences death(s) between year $t-3$ and year $t-6$ and production equation is one for year t. The number of households is 401. The number of observed years is 3. The measure of land is logarithm of square meter (m^2). The measure of livestock and agricultural output and rent is logarithm of Tanzania shilling (real price based on 1991). 1 United States dollar (USD) is equivalent to 219.2 Tanzania shilling (TZS) in 1991.

Table 7: Production function with prime-age adult mortality, Olley and Pakes (1996), least square

Variable	Estimate	S.E.	P-value
Member	0.3208	0.0108	0.0000
Land	0.2102	0.0066	0.0000
Livestock	0.0094	0.0047	0.7510

The estimated equation is

$$y_{jt} = \omega_{jt} + \theta' x_{jt} + \epsilon_{jt}$$

The number of households is 401. The number of observed years is 3. The number of observations for each estimation is 1,203 (= 401 households × 3 years). The measure of land is logarithm of square meter (m^2). The measure of livestock and agricultural output and rent is logarithm of Tanzania shilling (real price based on 1991). 1 United States dollar (USD) is equivalent to 219.2 Tanzania shilling (TZS) in 1991. Households includes continuing households but exclude split households. Standard errors and p-values are obtained by bootstrap. The number of iterations for bootstrap is 1,000. In each iteration, we draw households with replacement. Note that we draw 3 time observations of the same household at the same time instead of drawing a particular time observation. The sample size for each iteration is the same as original sample, that is, we have a bootstrapped sample of 401 households with duplications. It takes 1,545 seconds to compute estimates and their standard errors by fortran 90 (gfortran on Dell Latitude 2100). Simulated Annealing method for minimizing $\sum(\epsilon_{jt} + \xi_{jt})^2$. 72 iterations for obtaining the minimum value $\sum(\epsilon_{jt} + \xi_{jt})^2 = 287.18$.

Table 8: Productivity transition function with prime-age adult mortality, Olley and Pakes (1996), least square

Coefficient	Estimate	Std. Err.	P-value
from 1991 to 1992			
$\theta_{\omega 1992,0}$	1.818	0.424	0.000
$\theta_{\omega 1992,1}$	0.796	0.052	0.000
from 1992 to 2003			
$\theta_{\omega 2003,0}$	5.160	0.303	0.000
$\theta_{\omega 2003,1}$	0.426	0.039	0.000
$\theta_{\omega 2003,d}$	-0.041	0.009	0.000

The estimated equations are (15) and (16). Standard errors obtained by bootstrap. The number of iterations for bootstrap is 1,000. In each iteration, we draw households with replacement. Note that we draw 3 time observations of the same household at the same time instead of drawing a particular time observation. The sample size for each iteration is the same as original sample, that is, we have a bootstrapped sample of 401 households with duplications. P-values are asymptotic p-value, that is, $2(1 - \Phi(|\hat{\theta} - 0|/s(\hat{\theta})))$ where $\Phi(\cdot)$ is cumulative density function of standard normal distribution and $s(\hat{\theta})$ is the standard error of $\hat{\theta}$. It takes 1,545 seconds to compute estimates and their standard errors by fortran 90 (gfortran on Dell Latitude 2100). Simulated Annealing method for minimizing $\sum(\epsilon_{jt} + \xi_{jt})^2$. 72 iterations for obtaining the minimum value $\sum(\epsilon_{jt} + \xi_{jt})^2 = 287.18$.

Table 9: Estimates of $\phi(x_{jt}, I_{jt})$, Olley and Pakes (1996)

Variable	Coefficient	Std. Err.	P-value
1	11.0936	1.4540	0.000
m	0.9847	0.3201	0.001
k	-0.3776	0.3230	0.258
s	-0.1021	0.0380	0.007
i	15.2011	6.6567	0.020
m^2	0.0736	0.0461	0.119
k^2	0.0348	0.0183	0.045
s^2	0.0090	0.0013	0.000
i^2	7.1447	3.6783	0.083
$m \times k$	-0.0631	0.0393	0.112
$m \times s$	-0.0159	0.0069	0.014
$m \times i$	0.2454	0.6331	0.701
$k \times s$	0.0057	0.0044	0.204
$k \times i$	-0.6724	0.6002	0.245
$s \times i$	-0.5460	0.4198	0.183

The estimated equation is (13). The number of observations for each estimation is 1,203 (= 401 households $\times$ 3 years). Standard errors and p-values are obtained by bootstrap. The number of iterations for bootstrap is 1,000. In each iteration, we draw households with replacement. Note that we draw 3 time observations of the same household at the same time instead of drawing a particular time observation. The sample size for each iteration is the same as original sample, that is, we have a bootstrapped sample of 401 households with duplications.

Table 10: Decomposition of agricultural output growth, average among households with or without prime-age adult mortality, prime-age adult mortality by t and $-\tau$ and year dummies as an explanatory variable,

dpaam	log of agri-cultural output Δy_{jt}	produc-tivity Δa_{jt}	assets $\Delta(\theta' x_{jt})$	member $\Delta(\theta_m m_{jt})$	land $\Delta(\theta_k k_{jt})$	livestock $\Delta(\theta_s s_{jt})$
Fixed effects						
0	-0.3810	-0.3601	-0.0209	-0.0350	-0.0080	0.0220
1	-0.6414	-0.5291	-0.1123	-0.0885	-0.0313	0.0075
0-1	0.2604	0.1690	0.0914	0.0535	0.0233	0.0145
p-value	0.004	0.016	0.085	0.100	0.284	0.148
Olley and Pakes						
0	-0.3810	-0.3470	-0.0340	-0.0327	-0.0104	0.0090
1	-0.6414	-0.5210	-0.1203	-0.0826	-0.0408	0.0031
0-1	0.2604	0.1740	0.0864	0.0499	0.0305	0.0060
p-value	0.004	0.023	0.089	0.076	0.218	0.468

"dpaam" takes 1 if a household has prime-age adult (age 15-50) mortality between 1990 and 2003 and takes 0 otherwise. The first and second row show the average change in each variable between 1991 and 2003. For example, the (1,1) cell has -0.389, which is average change of logarithm of agricultural output from 1991 to 2003 for households without prime-age adult mortality. The last row show the difference between the first row and the second row. "p-value" is the one for testing H_0 : means are the same between households with and without death against H_1 : the mean for households without death is larger than the mean for households without death. The basic form of estimated equation is
$y_{ajt} = a_{jt} + \theta_m m_{jt} + \theta_k k_{jt} + \theta_s s_{jt}$. The number of households is 401. The number of observed years is 3. The measure of land is logarithm of square meter (m^2). The measure of livestock and agricultural output and rent is logarithm of Tanzania shilling (real price based on 1991). 1 United States dollar (USD) is equivalent to 219.2 Tanzania shilling (TZS) in 1991.

Table 11: score of relevancy as a continuing household

score	current head	same place	original head	spouse	son	daughter	N
23	original head	1		1			215
22	original head	0		1			32
21	original head	1		0			215
20	original head	0		0			83
19	spouse	1	1				1
18	spouse	0	1				0
17	spouse	1	0				76
16	spouse	0	0				34
15	son	1	1				1
14	son	0	1				1
13	son	1	0	1			4
12	son	0	0	1			1
11	son	1	0	0			156
10	son	0	0	0			249
9		1	1				5
8		0	1				16
7		1	0	1			6
6		0	0	1			30
5		1	0	0	1		22
4		0	0	0	1		64
3		1	0	0	0	1	39
2		0	0	0	0	1	480
1		1					130
0		0					914

The second column "current head" shows who is the current head. The third column "same place" takes 1 if a household answer that it lives the same house or plot in 2003 as the one in 1991 or 1992. The fourth to the last columns show the existence of each kind of original members. For example, a cell in the third column "original head" has 1 if the original head exists in a household and has 0 otherwise. The last column "N" represents the number of households.

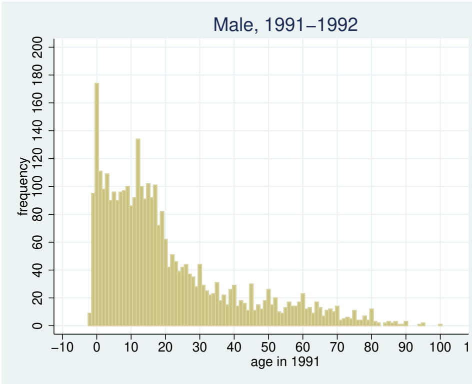

Figure 1: Distribution of male age

Horizontal line represents age in 1991. Vertical line represents frequency. We include all individuals who are surveyed at least once in wave 1-4. We include all individuals whose death are recorded and who do not have individual ID for main survey. Age -1 and -2 in 1991 means an individual is bone in 1992 and 1993, respectively.

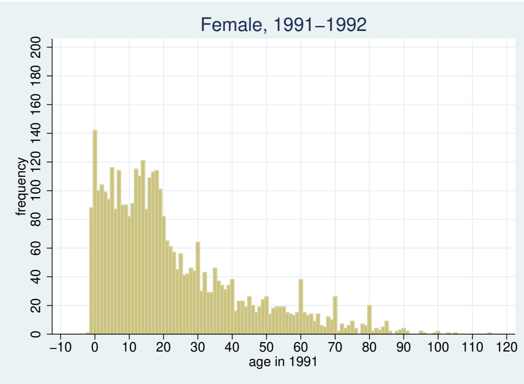

Figure 2: Distribution of female age

Horizontal line represents age in 1991. Vertical line represents frequency. We include all individuals who are surveyed at least once in wave 1-4. We include all individuals whose death are recorded and who do not have individual ID for main survey. Age -1 and -2 in 1991 means an individual is bone in 1992 and 1993, respectively.

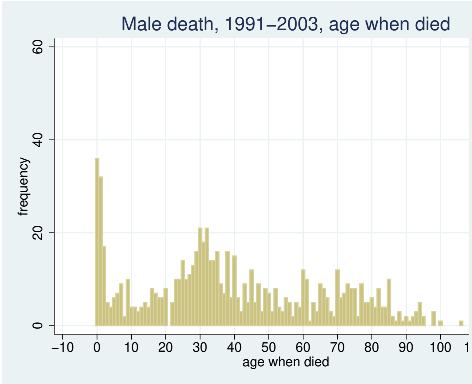

Figure 3: Distribution of male age when died, 1991-2003

Horizontal line represents age when died. Vertical line represents frequency. We include all individuals who are surveyed at least once in wave 1-4. We include all individuals whose death are recorded and who do not have individual ID for main survey.

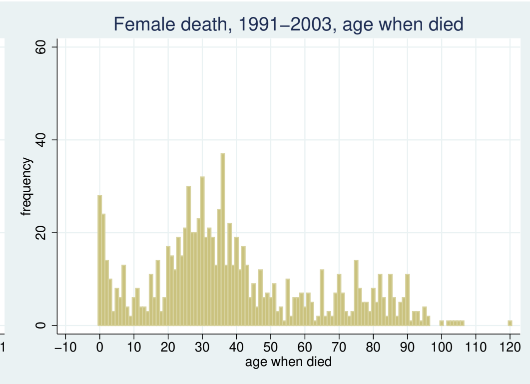

Figure 4: Distribution of female age when died, 1991-2003

Horizontal line represents age when died. Vertical line represents frequency. We include all individuals who are surveyed at least once in wave 1-4. We include all individuals whose death are recorded and who do not have individual ID for main survey.

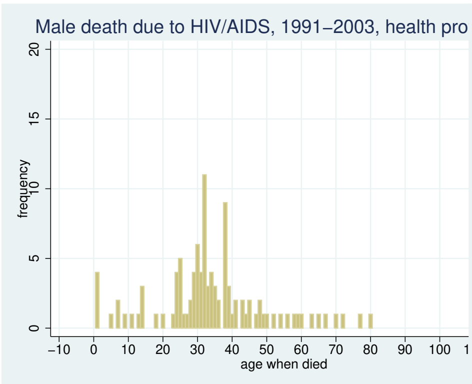

Figure 5: Distribution of age when male died due to HIV/AIDS, professional diagnosis, 1991-2003

Horizontal line represents age when died. Vertical line represents frequency. We include all individuals who are surveyed at least once in wave 1-4. We include all individuals whose death are recorded and who do not have individual ID for main survey.

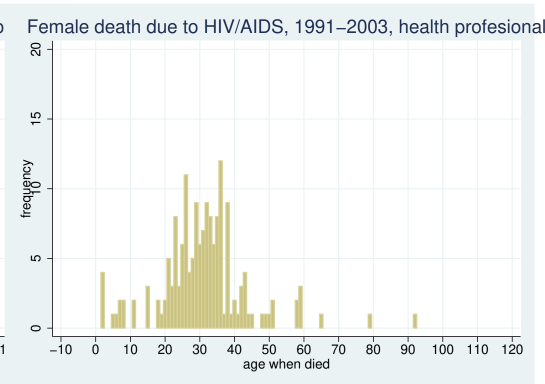

Figure 6: Distribution of age when female died due to HIV/AIDS, professional diagnosis, 1991-2003

Horizontal line represents age when died. Vertical line represents frequency. We include all individuals who are surveyed at least once in wave 1-4. We include all individuals whose death are recorded and who do not have individual ID for main survey.

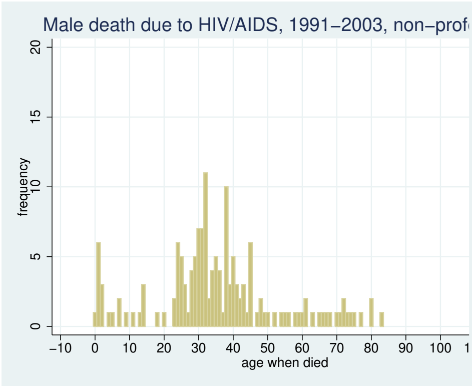

Figure 7: Distribution of age when male died due to HIV/AIDS, non-professional diagnosis, 1991-2003

Horizontal line represents age when died. Vertical line represents frequency. We include all individuals who are surveyed at least once in wave 1-4. We include all individuals whose death are recorded and who do not have individual ID for main survey.

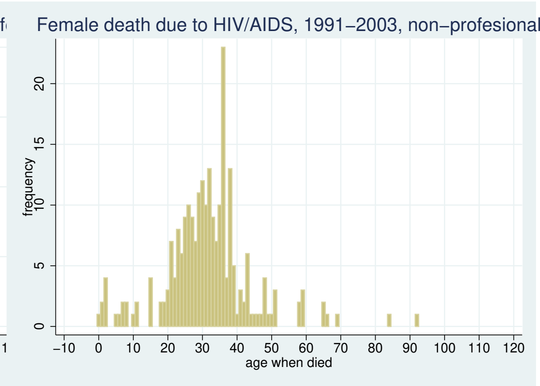

Figure 8: Distribution of age when female died due to HIV/AIDS, non-professional diagnosis, 1991-2003

Horizontal line represents age when died. Vertical line represents frequency. We include all individuals who are surveyed at least once in wave 1-4. We include all individuals whose death are recorded and who do not have individual ID for main survey.

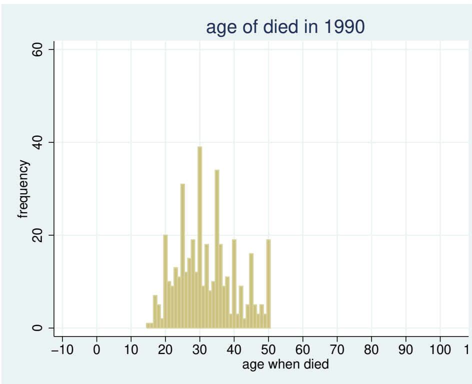

Figure 9: Distribution of age of adults with age 15-50 who died in 1990

Horizontal line represents age when died in 1990. Vertical line represents frequency. We include individuals whose death is recorded in the enumeration before wave 1. We exclude duplication of death between on in the enumeration and one in wave 1.

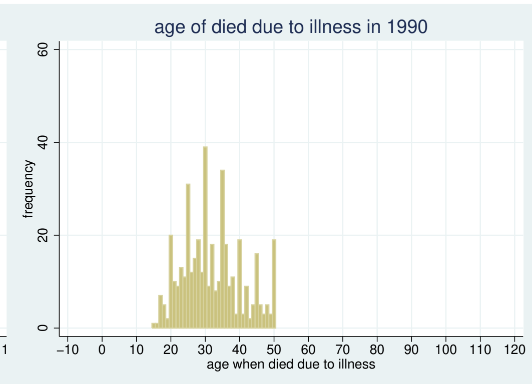

Figure 10: Distribution of age of adults with age 15-50 who died due to ill in 1990

Horizontal line represents age when died due to ill in 1990. Vertical line represents frequency. We include individuals whose death due to ill is recorded in the enumeration before wave 1. We exclude duplication of death between on in the enumeration and one in wave 1.

Appendixes

Appendix A. Construct data set

Wave 1 and wave 5 are annual surveys but wave 2, 3, and 4 are half-year surveys. In order to construct annual data, we combined wave 2 and 3 and threw away data of wave 4. The constructed data have 3 annual observations in 91, 92, and 03.

KHDS 91-94 added new households in the sample set from the middle of survey period (1991-1994) in order to replace drop-out households. If households received wave 1 questionnaires in the first two time intervals, called as "passages" in KHDS, we categorize them in the households from 1991 (889 households in the first row of Table 3). On the other hand, if households received wave 1 questionnaires in passage 3, we categorize them in the households from 1992 (30 households in (2,4) cell in Table 3).

As mentioned in Section 4, we will focus on agricultural households by excluding households in the most urbanized for clusters and emigrated households. We also drop samples if household agricultural output is smaller than non-agricultural income or transfer income.

As mentioned in Section 4.3, we focus on only one continuing household among other households from the same original household. Appendix Appendix A.1 explains how we choose a continuing household. We choose it regardless of whether each household is dropped or not for focusing agricultural households.

After dropping samples in each year and identifying a continuing household, we combine them and construct panel data by excluding households if each of households does not have the identified continuing household or each of households is dropped in any year as outliers or non-agricultural households. Table 3 shows the number of households in the constructed data set ("the data" row).

Appendix A.1. Choosing a continuing household among households from the same original household

There are three main ways for choose main household among split households: (1) choose the household which has the largest number of original household members, (2) choose the household where original head's spouse or son becomes head if household head changes, or (3) choose the household which keeps living in the same dwelling. We checked data and it looks like these three choices give us similar results since there are obvious correlations between these three characteristics of households. We mixed these three choices as follows. The reason why we do not simply rely on only (3) is that there are migrated households and there are multiple split households which answer that they live in the same house or plot in some extended families (households which split from the same household in 1991 or 1992)[7].

[7]175, 489, 125, 33, and 8 extended families have 0, 1, 2, 3, and 4 split households which keep living in the house or plot, respectively.

First, we make score of the relevancy as a candidate for continuing household as Table 11 shows. Second, if there is tie based on the score among households with score 10 or more from the same original household, we choose the household which has the largest number of original household members. Third, if there is no household with score 10 or more from the same original household, we choose the household which has the largest number of original household members. If there is tie based on the household which has the largest number of original household members, we choose the household with the highest score.

Table 3 shows the number of households which are identified and not identified household as a continuing household ("identified" row and "not identified" row)[8].

Table 11 will be inserted around here.

Table 11 will be inserted around here.

Appendix A.2. Construct variables

Appendix A.2.1. Agricultural output

Agricultural output y_{at} includes own consumption of agricultural products. Own consumption of livestock is not included in agricultural output but included in (dis)investment i_{ct}. On the other hand, own consumption of animal products such as milk and egg is included in agricultural output.

Appendix A.2.2. Household members

Household members M_{jt} is just the total number of household members. We have not taken into accounts heterogeneity in gender and age.

Appendix A.2.3. Land

The definition of land K_{jt} is the size (m^2) of land owned and land rented-in.

Land K_{jt} includes all of the four types of land: (1) owned and used (2) owned and fallowed, (3) owned and rented-out, and (4) not-owned and rented-in. 85% of land is type (1) and the most of the remaining 15% is type (4). The reason why we did not divide land into type (1) and type (4) is to decrease the dimension of state variables. We use hectare on land in the end of a period instead of in the beginning of the period since the data does not allow us to recover the hectare of land in the beginning of the period in 1991 and 2003.

Appendix A.2.4. Livestock

Monetary value of livestock in the beginning of period t is denoted by k_{ct} where subscript c is from cattle. KHDS asked the monetary value of each

[8] There are 39 (out of 403) households which do not live the same house or plot in 2003 as the one in 1991 or 1992 but are selected as continuing households.

livestock at the time of survey and we treat the total value of livestock as $k_{ct} + i_{ct} + \epsilon_{kct}$ (the monetary value of livestock in the end of period t) for 1991 and 1992. KHDS asked each household how much it bought, sold and ate and we construct investment i_{ct} based on them. Note that we treat own consumption of livestock as disinvestment or dis-saving instead of consumption or income. KHDS asked how many livestocks were (1) lost, stolen, disinherited, or died and (2) born, received as gift and we construct asset shock ϵ_{kct} based on these. Then we obtain k_{ct} the monetary value of livestock in the beginning of period t by subtracting i_{ct} and ϵ_{kct} from $k_{ct} + i_{ct} + \epsilon_{kct}$.

In 2003 KHDS did not ask households about i_{ct} and ϵ_{kct} in the same way as in 1991 and 1992 and data on i_{ct} and ϵ_{kct} in 2003 are much less complete than those in 1991 and 1992. Thus, we treat the total monetary value of livestock at the time of survey as k_{ct}.

We assume that households use k_{ct} for production at t and i_{ct} and ϵ_{kct} do not affect the production of production at t.

Appendix A.2.5. price index

Economic Development Initiatives (EDI), Kagera, Tanzania construct price index based on KHDS and make it public at `http://www.edi-africa.com/research/khds/introduction.htm`. We use it in order to transform nominal monetary values into real ones.

Appendix B. Details on empirical methods

Appendix B.1. Test mean equivalence by bootstrap

In this subsection we explain how we test that average of estimate for households with death is the same as one for households without death by bootstrap. Denote the difference between average of estimate between households with and without death by θ. For example, θ is average change in contribution of land to agricultural production for households without death minus one for households with death. We would like to test $H_0 : \theta = 0$ against $H_1 : \theta > 0$ at size α, for example, $\alpha = 0.05$. Set $T(\theta) = (\hat{\theta} - \theta)/s(\hat{\theta})$ where $s(\hat{\theta})$ is standard error of $\hat{\theta}$ and reject H_0 in favor of H_1 if $T(\theta) > q(1-\alpha)$ where $q(1-\alpha)$ would be selected so that $P(T(\theta) > q(1-\alpha)) = \alpha$.

We compute $q(1-\alpha)$ and $s(\theta)$ by bootstrap and denote bootstrap estimates for them by $q^*(1-\alpha)$ and $s^*(\hat{\theta})$. The number of iterations for bootstrap is 1,000. In each iteration, we draw households with replacement. Note that we draw 3 time observations of the same household at the same time instead of drawing a particular time observation. The sample size for each iteration is the same as original sample, that is, we have a bootstrapped sample of 401 households with duplications. $s^*(\hat{\theta})$ is computed as follows:

$$s^*(\hat{\theta}) = \frac{1}{B} \sum_{b=1}^{B} \left(\hat{\theta}_b^* - \bar{\hat{\theta}}^* \right)^2$$

where B is the number of iterations for bootstrap ($B = 1{,}000$), $\hat{\theta}_b^*$ is estimate of θ in the bth bootstrap iteration, and $\bar{\hat{\theta}}^* = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}_b^*$. For obtaining $q^*(1-\alpha)$, we compute $T_b^* = (\hat{\theta}_b^* - \hat{\theta})/s(\hat{\theta}_b^*)$, sort them and take $(1-\alpha)$ quantile. Note that in the numerator of T_b^*, we subtract the original estimate $\hat{\theta}$ instead of θ, true value under H_0. Note that we have $s(\hat{\theta}_b^*)$ as the denominator of T_b^* and this is unknown. Although we can compute it by another bootstrap, we substitute it with $s^*(\hat{\theta})$ for decreasing computational time.

Appendix B.2. Standard errors and p-values for coefficient estimates by bootstrap

In this subsection we explain how we compute standard errors and p-values for coefficient estimates by bootstrap. Denote a coefficient by θ. For example, θ is a coefficient for land in production function. We would like to test $H_0 : \theta = 0$ against $H_1 : \theta \neq 0$ at size α, for example, $\alpha = 0.05$. Set $T(\theta) = (\hat{\theta} - \theta)/s(\hat{\theta})$ where $s(\hat{\theta})$ is standard error of $\hat{\theta}$ and reject H_0 in favor of H_1 if $|T(\theta)| > q(\alpha)$ where $q(\alpha)$ would be selected so that $P(|T(\theta)| > q(\alpha)) = \alpha$. In stead of showing whether we can reject H_0 at a particular size of α, we will show the p-value as a measure of the evidence against H_0. The p-value of the statistics $|T(\theta)|$ is $p(|T(\theta)|) = P(x > |T(\theta)|)$ where x is a random draw from the distribution of $|T(\theta)|$. Note that $p(|T(\theta)|) < \alpha$ is equivalent to that we can reject H_0 at size α.

We compute $s(\theta)$ and $p(|T(\theta)|)$ by bootstrap and denote bootstrap estimates for them by $s^*(\hat{\theta})$ and $p^*(|T(\hat{\theta})|)$. The number of iterations for bootstrap is 1,000. In each iteration, we draw households with replacement. Note that we draw 3 time observations of the same household at the same time instead of drawing a particular time observation. The sample size for each iteration is the same as original sample, that is, we have a bootstrapped sample of 401 households with duplications. $s^*(\hat{\theta})$ is computed as follows:

$$s^*(\hat{\theta}) = \frac{1}{B}\sum_{b=1}^{B}\left(\hat{\theta}_b^* - \bar{\hat{\theta}}^*\right)^2$$

where B is the number of iterations for bootstrap ($B = 1{,}000$), $\hat{\theta}_b^*$ is estimate of θ in the bth bootstrap iteration, and $\bar{\hat{\theta}}^* = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}_b^*$. For obtaining $p^*(|T(\hat{\theta})|)$, we compute $|T_b^*| = |\hat{\theta}_b^* - \hat{\theta}|/s(\hat{\theta}_b^*)$ and sort them. Then, we search xth value of the sorted sequence of $|T_b^*|$ which is the closest to $|\hat{\theta} - \theta|/s^*(\hat{\theta})$ and $p^*(|T(\hat{\theta})|)$ is $(B-x)/B$. Note that in the numerator of T_b^*, we subtract the original estimate $\hat{\theta}$ instead of θ, true value under H_0. Note that we have $s(\hat{\theta}_b^*)$ as the denominator of T_b^* and this is unknown. Although we can compute it by another bootstrap, we substitute it with $s^*(\hat{\theta})$ for decreasing computational time.