7-31-2013

Cohort Profile: The Malawi Longitudinal Study of Families and Health (MLSFH)

HANS-PETER KOHLER
University of Pennsylvania, Dept. of Soc, HPKOHLER@POP.UPENN.EDU

Susan C. Watkins
University of Pennsylvania, swatkins@pccpr.ucla.edu

Jere R. Behrman
University of Pennsylvania, jbehrman@econ.upenn.edu

Philip Anglewicz
Tulane University of Louisiana

Iliana V. Kohler
University of Pennsylvania, iliana@pop.upenn.edu

See next page for additional authors

Author(s)
HANS-PETER KOHLER, Susan C. Watkins, Jere R. Behrman, Philip Anglewicz, Iliana V. Kohler, Peter Fleming, Rebecca L. Thornton, James Mkandawire, Linda Kalilani-Phiri, Hastings Honde, Augustine Hawara, Ben Chilima, Chiwoza Bandawe, Victor Mwapasa, and MLSFH Study Team

Cohort Profile: The Malawi Longitudinal Study of Families and Health (MLSFH)

Hans-Peter Kohler* Susan C. Watkins Jere R. Behrman
Philip Anglewicz Iliana V. Kohler Peter Fleming
Rebecca L. Thornton James Mkandawire Linda Kalilani-Phiri
Hastings Honde Augustine Hawara Ben Chilima
Chiwoza Bandawe Victor Mwapasa MLSFH Study Team

July 31, 2013

Summary
The Malawi Longitudinal Study of Families and Health (MLSFH) is one of very few long-standing publicly-available longitudinal cohort studies in a sub-Saharan African (SSA) context. It provides a rare record of more than a decade of demographic, socioeconomic and health conditions in one of the world's poorest countries. With data collection rounds in 1998, 2001, 2004, 2006, 2008, 2010 and 2012 for up to 4,000 individuals, the MLSFH permits researchers to investigate the multiple influences that contribute to HIV risks in sexual partnerships, the variety of ways that people manage risk within and outside of marriage, the possible effects of HIV prevention policies and programs, and the mechanisms through which poor rural individuals, families, households, and communities cope with the impacts of high morbidity and mortality that are often—but not always—related to HIV/AIDS. The MLSFH been used to document (*i*) the influence of social networks on HIV-related behaviors and perceptions, (*ii*) the HIV prevention strategies employed by individuals in rural high-HIV prevalence contexts, (*iii*) the relationship between life-course transitions and HIV infection risks, (*iv*) the acceptability of HIV testing and counseling (HTC) and the consequences of HTC on subsequent behaviors, and (*v*) the health and well-being across the life-course of individuals facing multiple challenges resulting from high disease burdens and widespread poverty.

WHY WAS THE COHORT SET UP?

The ***Malawi Longitudinal Study of Families and Health (MLSFH)*** is one of a very few long-standing longitudinal cohort studies in a poor sub-Saharan African (SSA) context. The MLSFH cohorts were selected to represent the rural population of Malawi, where the vast majority of Malawians live in conditions that are similar to those in the rural areas of other countries with high HIV prevalence: health conditions are poor, health facilities and schools are over-burdened and under-staffed, standards of living are low and nutritional needs of adults, children and the elderly are often not met. With major data collection rounds in 1998, 2001, 2004, 2006, 2008, 2010, and 2012 for up to 4,000 individuals, as well as ancillary surveys and qualitative studies, the MLSFH has been a premier dataset for research on health, family

Figure 1: MLSFH study locations in Malawi

dynamics, social networks, and HIV infection risks in a rural SSA context. Providing public-use data on the socioeconomic context, demographics and health of individuals and their families in Malawi over more than a decade (Figure 1), the MLSFH has been the basis of more than 150 publications and working papers submitted for publication. Importantly, the MLSFH has also informed health policy discussions in Malawi and elsewhere in SSA.

The MLSFH (1998–2012) is a collaboration of the University of Pennsylvania with the College of Medicine and the Demography Unit, Chancellor College, both at the University of Malawi. It subsumes earlier research under the Malawi Diffusion and Ideational Change Project (MDICP, 1998–2004), which focused on the influence of social networks on the adoption of family planning and on AIDS-related attitudes and behaviors.[1,2] Subsequently (2006–2012) the goals of the MLSFH expanded. Overall, the MLSFH has (1) provided a rare record of more than a decade of demographic, socioeconomic and health conditions in one of the world's poorest countries through the collection of longitudinal cohort data, and (2) analyzed these data to investigate the multiple influences that contribute to HIV risks in sexual partnerships, the variety of ways in which people manage risk within and outside of marriage and other sexual relationships, the possible effects of HIV prevention policies and programs, and the mechanisms through which poor rural individuals, families, households, and communities cope with the impacts of high morbidity and mortality that are often—but not always—related to HIV/AIDS. The data collection and research conducted by MLSFH was approved by the IRB at the University of Pennsylvania and, in Malawi, by the College of Medicine Research Ethics Committee (COMREC) or the National Health Sciences Research Committee (NHSRC).

WHO IS IN THE COHORT?

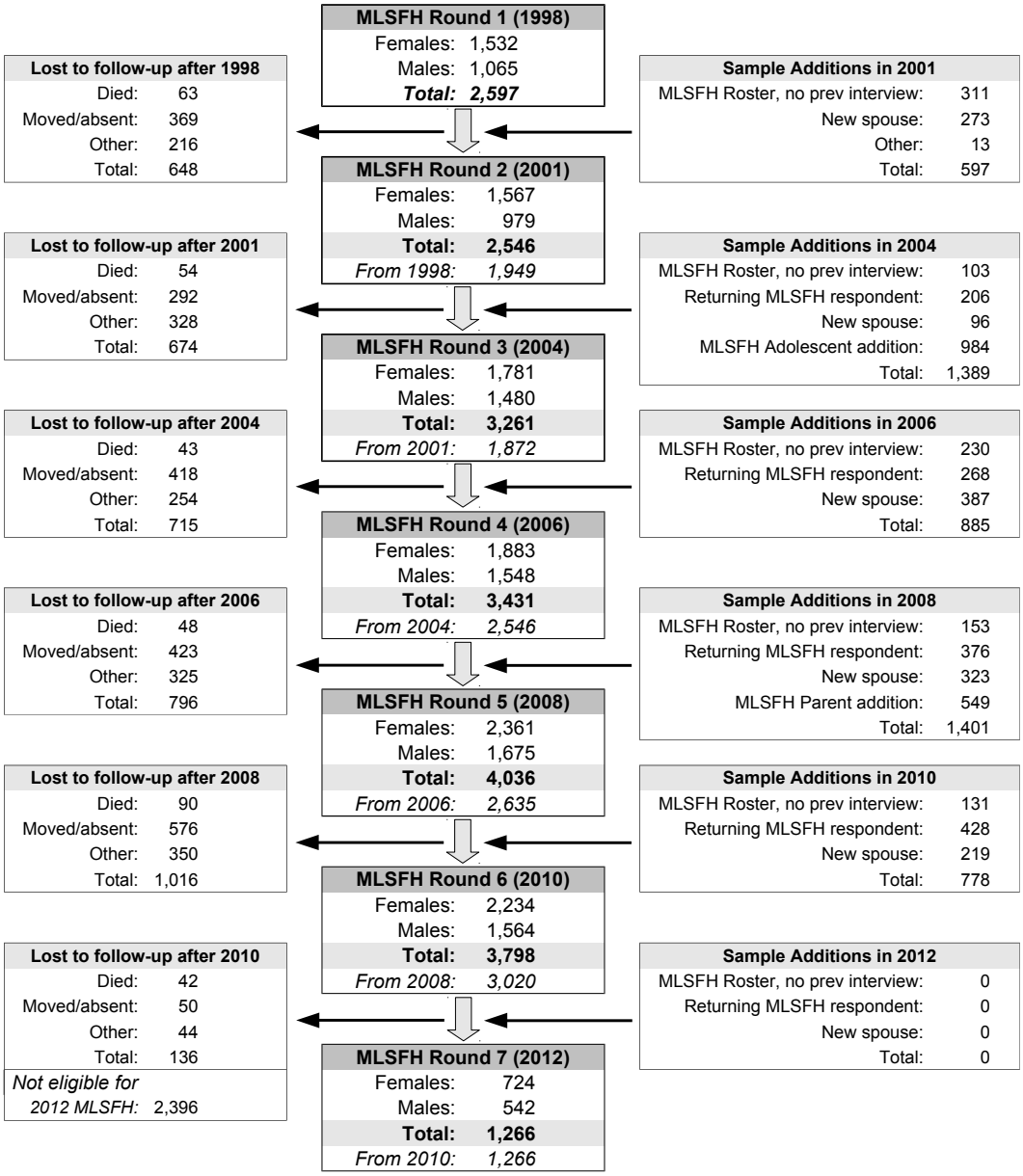
Study Context: Malawi is ranked 153 of 169 countries in terms of the Human Development Index (HDI).[3] The large majority of the population (84.7%) is rural. Pop-

ulation growth continues to be relatively rapid. The Malawi population increased at an average rate of 2.8% from 10.7 to 15.9 million during 1998–2012, the period covered by the MLSFH, and a tripling to 48 million is projected within the next 50 years (UN medium variant).[4,5] Life expectancy at birth was 51 for men and 55 for women in 2010, and healthy life expectancy at birth was 44 years for males and 46 years for females.[6] About 15% of the population is considered "ultra-poor", i.e., with an estimated food consumption below the minimum level of dietary energy requirement.[7] While per capita income is below the SSA average, Malawi is similar to other SSA countries and countries in the World Bank low-income country (LIC) group in terms of life expectancy, infant mortality, children's malnutrition, access to clean water, literacy and schooling enrollment.[8,9] In rural areas, where the MLSFH study population is based, the majority of individuals engage in home production of crops, primarily maize, which is the dietary staple and is highly influenced by the vagaries of the weather and the availability of fertilizer: during the period of the MLSFH, there were several years with "hunger months", when maize production was insufficient. Subsistence agriculture is complemented by some smallholder cash crops (primarily tobacco and cotton), casual agricultural labor and small-scale market activities, such as selling second-hand clothing and vegetables. Malawi has the globally 9th highest prevalence of HIV in the adult population with an estimated 2010 HIV prevalence among 14–49 year olds of 8.9% (women: 10.5%; men: 7.1%) in rural and 17.4% (women: 22.7%; men: 12.0%) in urban areas.[10,11] HIV incidence is estimated to have peaked in the mid-1990s, and by 2012 had fallen to .44, well below replacement level.[10] Nevertheless, the HIV epidemic had, and continues to have, major effects on virtually all aspects of life, many of which were documented by the MLSFH (see below). With aid from international donors, access to antiretroviral treatment (ART) in Malawi expanded during the past decade, attaining a 67% ART coverage (with eligibility for treatment based on WHO 2010 guidelines) in 2011, resulting in significant reductions in adult mortality.[10,12] Tuberculosis, malaria, and endemic parasites (e.g., soil-transmitted helminths (STH) and schistosomia mansoni) also have a relatively high prevalences,[13,14] as do some chronic diseases such as hypertension.[15]

Study Locations: The MLSFH is based in three districts in Malawi—Rumphi in the North, Mchinji in the Center, and Balaka in the South (Figure 1)—and major MLSFH data collection has been conducted in 1998, 2001, 2004, 2006, 2008, 2010 and 2012 (Figure 2). While the three MLSFH study regions are generally similar in terms of their overall epidemiological, socioeconomic and subsistence-agriculture characteristics,[13,16] the regions are heterogeneous in terms of marriage patterns,[17] religious affiliations,[18] schooling,[19] patrilineal vs. matrilineal inheritance and landownership,[20] and HIV prevalence.[11,21] MLSFH Respondents ($N_{2010} \approx 3,800$) are evenly split among the three study locations and clustered in 121 villages.

MLSFH Study Population: The core of the MLSFH is a longitudinal survey, augmented by qualitative data on specific topics (Figure 2). We summarize essential features of the MLSFH study population and data collection below; a de-

Figure 2: MLSFH Sample Flow 1998-2012



Notes: The MLSFH is based in three districts in Malawi—Rumphi in the North, Mchinji in the Center, and Balaka in the South (Figure 1). MLSFH sampling and related relevant data collection procedures are described in Appendix A2. MLSFH Study instruments are described in Table 4. In addition to the major MLSFH waves noted above, the MLSFH also conducted a migration follow-up in 2007 (Appendix A2.2), a 2006–07 MLSFH Incentive Study (Appendix A6.6) that collected repeated sexual diaries, and a 2009 MLSFH Biomarker Study collecting biomarkers for cardiovascular risk, organ/metabolic function and inflammation (Appendix A6.7). The MLSFH survey data are complemented by extensive qualitative and ethnographic data that has been collected during 1998–2012.

Table 1: First available MLSFH Round for MLSFH participants in 2010 and 2012

First available MLSFH Round	MLSFH 6 (2010) Respondents			MLSFH 7 (2012) Respondents		
	Females	*Males*	*Total*	*Females*	*Males*	*Total*
1998	39.79%	35.81%	38.15%	50.14%	54.61%	52.05%
2001	9.13%	7.10%	8.29%	11.60%	10.89%	11.30%
2004	11.82%	23.53%	16.64%	0.97%	6.64%	3.40%
2006	9.18%	11.45%	10.11%	2.49%	3.87%	3.08%
2008	20.41%	13.55%	17.59%	34.81%	23.99%	30.17%
2010	9.67%	8.57%	9.22%	–	–	–
N	2,234	1,564	3,798	724	542	1,266

tailed discussion of the MLSFH data collection and fieldwork procedures across all MLSFH waves is provided in the Appendix. Additional information is available on the MLSFH project website at http://www.malawi.pop.upenn.edu.

The MLSFH began in 1998 with a sample of 1,532 ever-married women aged 15–49 and 1,065 of their spouses (Figure 2). Details of the initial sampling procedure are described in Appendix A2.1. In 2001, respondents were re-interviewed, along with any new spouses since 1998. In 2004, in addition to re-interviewing the 1998 and 2001 study population and new spouses, the MLSFH added a sample of approximately 1,000 adolescents aged 15–24 to compensate for the aging of the initial MLSFH sample and the underrepresentation of unmarried individuals at adolescent and young adult ages. During the 2008 MLSFH round, a sample of parents of the original MLSFH respondents was added to the MLSFH to increase the suitability of the MLSFH for studying intergenerational aspects and the health of older individuals in Malawi. This study population was re-interviewed in 2010. The 2012 MLSFH round was restricted to mature adults, defined as individuals aged 45 and over. Up to the 2001 round of data collection, the MLSFH attempted to re-interview all of the initial MLSFH respondents and their current spouses; after 2001, if an individual was interviewed for the MLSFH once, for instance after being enrolled as a new spouse, he/she was considered a member of the MLSFH cohort and the MLSFH attempted to re-interview him/her at all subsequent waves (exceptions were members of the 2004 adolescent sample, all of whom were attempted to be interviewed during the 2006 MLSFH independently of whether they were successfully interviewed or not in 2004; and the 2008 parents sample, all of whom were included for attempted interview in 2010 regardless of whether they were interviewed in 2008).

To highlight the potential of long-term longitudinal analyses with the MLSFH, Table 1 reports the *first* available MLSFH survey round for participants in the 2010 and 2012 MLSFH Rounds (MLSFH 6 & 7). It shows, that for more than 46% of the 2010 MLSFH participants, and for more than 63% of the 2012 MLSFH participants,

Table 2: Number of MLSFH survey rounds available for MLSFH respondents

Number of available MLSFH Surveys	Among MLSFH Respondents:					
	With at least one MLSFH Interview		MLSFH 6 (2010) Participants		MLSFH 7 (2012) Participants	
	%	Cum %	%	Cum %	%	Cum %
7	8.0%	8.0%	13.4%	13.4%	40.1%	40.1%
6	10.7%	18.7%	18.0%	31.3%	18.6%	58.7%
5	7.9%	26.6%	9.4%	40.7%	6.8%	65.5%
4	12.9%	39.5%	16.0%	56.7%	4.3%	69.8%
3	19.9%	59.4%	23.4%	80.1%	30.2%	100.0%
2	17.2%	76.6%	10.7%	90.8%	–	–
1	23.4%	100.0%	9.2%	100.0%	–	–
N	6,369		3,798		1,266	

initial data are available from either 1998 or 2001. Hence, for close to two-thirds of the mature adults interviewed in 2012, more than a decade of longitudinal MLSFH data are available, and for 63% of the 2010 MLSFH participants, initial MLSFH data are available from at least 2004 onward. For about 40% of respondents who were interviewed by the MLSFH at least once, four or more rounds of MLSFH data are available (Table 2). Among 2010 MLSFH participants, more than 40% have data available from five or more MLSFH Rounds, and more than 80% have data from three or more MLSFH Rounds. Due to the specific selection criteria used for the 2012 MLSFH data collection for mature adults, all 2012 MLSFH participants have at least three rounds of available MLSFH data; for about 2/3rds, data are available from five or more rounds, and for 40% of the 2012 participants, data are available from seven MLSFH rounds covering the period 1998–2012.

Table 3 provides summary statistics for the 2010 MLSFH study population. This population had a mean age of about 42 years. Due to the aging of the MLSFH sample since the inception of the MLSFH in 1998 and the addition of the adolescent sample in 2004, few MLSFH respondents remained below age 20 in 2010 (though information on many respondents' children younger than 20 was collected from respondents). The vast majority of the 2010 MLSFH respondents had been married at least once, and MLSFH respondents had 5.8 children on average (children ever born). Schooling attainment was relatively low. About 21% of MLSFH respondents had no formal schooling, and only about 15% attended secondary or higher schooling levels. Consistent with the general socioeconomic context, the MLSFH study population was relatively poor: most lived in a house made of mud with a thatch roof, and only 21% of MLSFH respondents had a house with a metal roof, a sign of relative wealth. The study population was about evenly split among the three study regions, and about 26% of the 2010 study population was Muslim (concentrated mostly in the southern region). Despite the high burden of disease faced by

Table 3: Summary statistics for the MLSFH Round 6 (2010) study population

	Females mean (sd)	**Males** mean (sd)	**Total** mean (sd0
# of observations	2,234	1,564	3,798
Respondent's age (in 2010)	41.43	43.42	42.25
	(16.83)	(16.63)	(16.78)
Age group (in 2010)			
< 20	0.022	¡ .01	0.013
20–29	0.279	0.281	0.280
30–39	0.221	0.186	0.207
40–49	0.184	0.190	0.186
50–59	0.136	0.149	0.141
60–69	0.077	0.110	0.090
70+	0.081	0.084	0.082
Marital status			
Married/living together	0.770	0.870	0.811
Separated	0.019	0.008	0.015
Divorced	0.079	0.028	0.058
Widowed	0.120	0.018	0.078
Never married	0.012	0.076	0.038
Children ever born	5.66	5.96	5.79
	(3.36)	(4.59)	(3.91)
Schooling attainment			
No formal schooling	0.273	0.125	0.212
Primary schooling	0.631	0.641	0.635
Secondary or higher	0.096	0.233	0.153
Wealth indicator: House has metal roof	0.213	0.207	0.211
Region of residence			
Central (Mchinji)	0.317	0.339	0.326
South (Balaka)	0.351	0.323	0.339
North (Rumphi)	0.332	0.338	0.335
Religion			
Christian	0.687	0.687	0.687
Muslim	0.262	0.248	0.256
Other/none	0.051	0.065	0.057
Subjective health			
Poor	0.022	0.013	0.018
Fair	0.055	0.038	0.048
Good	0.237	0.181	0.214
Very good	0.437	0.404	0.424
Excellent	0.249	0.365	0.297
Subjective likelihood of being infected with HIV			
No likelihood	0.570	0.656	0.605
Low	0.300	0.269	0.287
Medium	0.091	0.055	0.076
High	0.040	0.020	0.032
At least one MLSFH HIV test result	0.796	0.809	0.801
HIV positive*	0.064	0.040	0.054

Note: * HIV+ = at least one MLSFH HTC had a HIV+ test result

MLSFH respondents, about 30% of respondents in 2010 self-rated their health as excellent, and 42% as very good. About 40% of respondents in 2010 also expressed some likelihood (low/medium/high) of being infected with HIV. MLSFH HIV test results were available for about 80% of the 2010 study population, and 5.4% of the 2010 MLSFH study population had previously tested HIV-positive during one of the MLSFH rounds that included HIV testing (2004, 2006 and 2008).

Qualitative data augment the longitudinal MLSFH survey data. From 1999 to the present the MLSFH conducted an ethnographic study, using local participant observers to capture public conversations on AIDS in informal social networks. In addition, qualitative interviews with subsamples of MLSFH respondents were conducted on data quality, intra-familial transfers, the status of women and intra-familial power, women's extramarital partners and responses of religious leaders to the epidemic. Additional qualitative studies were conducted in the MLSFH study sites on voluntary counseling and testing, sex work, condoms, schooling, the role of chiefs, the activities of Community Based Organizations (CBOs) and perceptions of non-governmental organizations (NGOs).

HOW OFTEN HAVE THEY BEEN FOLLOWED UP?

Major MLSFH rounds were collected in 1998 (MLSFH 1), 2001 (MLSFH 2), 2004 (MLSFH 3), 2006 (MLSFH 4), 2008 (MLSFH 5), 2010 (MLSFH 6) and 2012 (MLSFH 7) (Figure 2 and Tables 1–4). The MLSFH 7 in 2012 interviewed only mature adults (aged 45 and older), whereas all the earlier MLSFH rounds included the full MLSFH sample (Figure 2). In addition to the major MLSFH Rounds 1–7, a subset of MLSFH respondents participated in 2006–07 in the collection of "sexual diaries" as part of a conditional cash transfer program that offered financial incentives to men and women to maintain their HIV status for approximately one year (Table 4 and Appendix A6.6).[22] A migration follow-up that tried to trace all ever-interviewed MLSFH respondents not interviewed during MLSFH 4 (2006) due to migration and/or temporary absence was conducted in 2007 (to be updated in 2013) (Table 4 and Appendix A2.2).[23] In addition, in 2009 the MLSFH collected blood-plasma based biomarkers of cardiovascular and related health risks for a subset of MLSFH respondents in Balaka (Table 4 and Appendix A6.7).[24,25]

WHAT HAS BEEN MEASURED?

Across all rounds the MLSFH measured and documented the health, social, economic and demographic context of the MLSFH study population (Table 4). While the specific MLSFH focus topics evolved over time, the seven rounds of MLSFH data provide longitudinal data on aspects such as: ***household structure and family change*** (household/family rosters, marriage and partnership histories), ***human capital*** (health, schooling, nutritional status), ***social capital*** (social networks, intra-familial/intergenerational and community transfers, social participation), ***sexual behaviors*** (sexual relations and networks, HIV/AIDS risk behaviors and prevention strategies), ***subjective expectations and well-being*** (SF12 module, subjective

Table 4: MLSFH survey content and other study components, by year

MLSFH Round	Measurements
All MLSFH Rounds	***Survey data:*** demographic and socioeconomic characteristics of respondents, social and economic context, linkages with spouse data, vital status and migration/absence at time of MLSFH survey. Geocoded respondent residences since 2004; regional market and rainfall data.
MLSFH 1 (1998)	***Survey data:*** Childbearing and fertility desires; attitudes about and use of family planning methods; conversational networks about family planning and HIV/AIDS (see Appendix A6.1); gender attitudes and female autonomy; HIV/AIDS-related knowledge, and risk-perceptions; sexual behaviors and HIV risk reduction strategies.
MLSFH 2 (2001)	***Survey data:*** Mostly identical to MLSFH 1 (1998), plus social participation, marriage and sexual partnership histories (see Appendix A6.4).
MLSFH 3 (2004)	***Survey data:*** Similar to MLSFH 2 (2001), plus: household rosters with data on household membership and health/schooling/morbidity and marital status of household members; measures of religious activities/affiliations; social capital and basic data on transfer and exchange networks; time use; household expenditures on health and schooling; AIDS-related stigma. No longer included: fertility histories and childbearing desires; detailed data about use of and attitudes about family planning methods. ***Biomarkers:*** Testing and counseling for HIV, gonorrhea, chlamydia and trichomonas (see Appendix A3.1). ***Other:*** Randomized experimental design offering financial incentives for learning HIV status (see Appendix A3.1).
MLSFH 4 (2006)	***Survey data:*** Similar to MLSFH 3, plus: SF-12 self-reported health questionnaire; probabilistic expectations about health and HIV-risks (see Appendix A6.3); subjective discount rate; HIV testing history; audio-CASI interview for sensitive behaviors; intergenerational transfers and transfers with community members (see Appendix A6.2); mortality of household/family members. No longer included were conversational networks about family planning and religion. ***Biomarkers:*** Testing and counseling for HIV (see Appendix A3.2). ***Other:*** Initiation of MLSFH Incentive Study, a randomized experiment offering financial incentives for maintaining HIV-negative status (see Appendix A6.6).
MLSFH Incentive Study 2006–07	***Survey data:*** Four rounds of sexual diaries providing detailed day-to-day data on sexual behaviors during a 9-day period prior to the interview (see Appendix A6.6). ***Biomarkers:*** Testing and counseling for HIV in 2007 (after 3rd round of sexual diary collection).
MLSFH 4 migration follow-up (2007)	Migration follow-up with all ever-interviewed MLSFH respondents not interviewed during MLSFH 4 (2006) due to migration and/or temporary absence (see Appendix A2.2). ***Survey data:*** Similar to MLSFH 4, plus detailed questions about migration history and migration reasons. ***Biomarkers:*** Testing and counseling for HIV (see Appendix A3.3).

Continued on next page

well-being and mental health, HIV risk perceptions, mortality and HIV infection risks), ***household production and consumption*** (standard of living, household assets and income, expenditures on health and schooling, time use, migration), and ***mortality and migration*** of MLSFH participants and family members (verbal au-

Table 4: MLSFH survey content and other study component, by year

Continued from previous page

MLSFH Round	Measurements
MLSFH 5 (2008)	***Survey data:*** *Human capital*, including schooling, self-reported SF-12 questionnaire module on physical and mental health, subjective well-being; *household production and consumption*, including standard of living, household assets, time use, health expenditure; *social capital*, including intergenerational/intrafamilial transfers and help relationships, transfer relationships with community members, participation in community associations; *expectations, risk perceptions and attitudes*, including probabilistic expectations about HIV risks and survival (see Appendix A6.3), and AIDS-related attitudes and knowledge; *biographic information*, including marriage and partnership histories (see Appendix A6.4), extra- and pre-marital relationships, partner characteristics. No longer included: conversational networks about AIDS. ***Biomarkers:*** Testing and counseling for HIV (see Appendix A3.3).
MLSFH Biomarker Study (2009)	Focused on selected subset of MLSFH respondents residing in Balaka ($N = 982$) and collected biomakers of inflammation, cardiovascular risks, metabolic processes and organ function (see Appendix A6.7). ***Survey data:*** Health; illnesses and pain experienced by respondent, and household members; illnesses experienced by family members; water source; diet/nutrition. ***Biomarkers:*** Biomarkers for wide-range CRP, total cholesterol, LDL, HDL, total protein, urea, albumin, blood urea nitrogen, creatinine, random blood glucose and HbA1c, collected using the LabAnywhere (previously Demecal) System (LabAnywhere, Haarlem, The Netherlands)[26] ***Other:*** Height, weight and BMI
MLSFH 6 (2010)	***Survey data:*** Same as MLSFH 5 (2008).
MLSFH 7 (2012)	Follow-up survey focused on MLSFH *mature adults* aged 45 and older focused on mental heath, cognition and well-being (see Appendix A6.8) ***Survey data:*** Similar to MLSFH 6 (2010), plus measures of depression and anxiety, measures of spatial/temporal orientation and language, measures of memory/recall and executive functioning, alcohol consumption. ***Biomarkers:*** Testing and counseling for HIV (see Appendix A3.3). ***Other*** Grip strength (both hands), height, weight and BMI.
MLSFH qualitative and related contextual data:	Qualitative interviews on sexual attitudes and behaviors; qualitative data on VCT, HIV-risks and investments in children; ethnographic studies of public conversations about AIDS in informal settings; detailed village characteristics and local infrastructure data; GPS data for all respondents; condom prices and local market prices of key crops and commodities.

topsies, migrant tracking). Data quality in the MLSFH is generally high.[27–31] The MLSFH data have been geocoded since 2004, and since 2006, the MLSFH has included a module about probabilistic expectations—i.e., subjective expectations that can be interpreted as probabilities—for multiple HIV and health-related outcomes (see Appendix A6.6). MLSFH survey data also include linkages between spouses (updated at each round), parent-children linkages, and longitudinal linkages of children listed on the family and household roster (see Appendix A6.5).

Particularly noteworthy aspects of the MLSFH survey data collection—such as

the MLSFH data on social networks, the family/household rosters, the probabilistic expectations, marital and partnership histories, and the recent MLSFH data on mental health and cognition—are described in more detail in Appendix A6.1–A6.5.

In addition to collecting extensive survey data, the MLSFH has also conducted repeated ***HIV testing and counseling (HTC)*** at respondents' homes (see Appendix A3 for details). In 2004, HIV testing was conducted using oral swab specimens that were then sent for analysis to the University of North Carolina laboratory in Lilongwe using ELISA and confirmatory Western blot tests. MLSFH HIV testing was conducted using finger-prick rapid tests from 2006 onward (see Appendix A3 for additional details). In 2004, the MLSFH tested respondents also for chlamydia, gonorrhea and trichomoniasis in addition to HIV, but tests of these sexually transmitted diseases were not repeated given their low prevalence in the 2004 MLSFH study population. The MLSFH has also collected *anthropometric data* (height, weight and BMI) in 2008 and 2012 (Appendix A6.7 and A6.8), and selected biomarker-based indicators of health (CRP, HDL, LDL and others) in 2009 for a subset of MLSFH respondents residing in Balaka (Appendix A6.7).

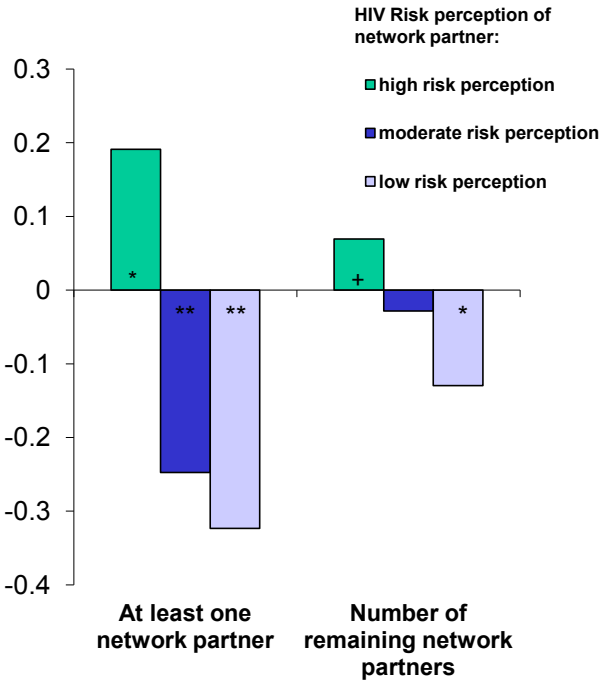
The MLSFH has also implemented ***randomized experimental designs*** related to HIV prevention. In 2004, the MLSFH HTC was combined with an experimental design that offered randomized financial incentives for individuals to learn their HIV status (see Appendix A3.1 for additional detail), and during 2006–07, the MLSFH offered randomized conditional cash transfer to a subset of MLSFH respondents to encourage maintaining their HIV status for approximately one year (see Appendix A6.6).[22,32]

WHAT HAS THE MLSFH FOUND? KEY FINDINGS AND PUBLICATIONS

The MLSFH has been used to make important contributions in several areas related to social networks, HIV/AIDS, fertility and reproductive health, marriage and family dynamics, intergenerational relations, religion, and survey methodologies. Key findings across important topical areas include:

Social interaction and social networks: The MLSFH is one of very few sources of longitudinal data in SSA, both survey and qualitative, on informal social interactions. MLSFH analyses, for example, have shown that social interactions about HIV/AIDS and prevention strategies were a frequent topic of conversation among Malawians, and an important contribution of the MLSFH was to demonstrate that strategies of AIDS prevention were formed collectively and "worked out" in social networks.[33–35] Women reported worrying most about their husbands as possible sources of infection, discussing with them the importance of avoiding infection; increasingly, they used divorce to reduce their risk. Men reported worrying most about risk from their extramarital partners and said they had adopted preventive strategies such as fewer partners and more careful partner selection. Male MLSFH respondents who believed that their best friends had extra-marital sexual partnerships (EMSPs) were significantly more likely to report having had EMSPs themselves.[36] These social interactions related to HIV/AIDS were found

Figure 3: Effect of social network partners' HIV/AIDS risk perceptions of MLSFH respondent's own HIV/AIDS risk perceptions



Notes: The MLSFH survey measured perceived HIV/AIDS risk using the question "*How worried are you that you might catch AIDS?*", with three response categories ranging from "*not worried at all*" (coded as 1) to "*worried a lot*" (coded as 3)." The respondent was asked a corresponding question about his/her social network partners' HIV/AIDS risk perceptions (for up to four social network partners). The graph shows the effect of the network partners' risk perception (by number of network partners in each subjective risk category) on the respondent's own risk perception, estimated based on longitudinal MLSFH 1–2 data using an instrumental-variable fixed-effect regression technique that controls for unobserved respondent characteristics and the potential selective reporting of network partners by respondents. The graph shows that social interactions with network partners who have high HIV risk perceptions increase the respondent's own risk perceptions about HIV/AIDS, and this effect is particularly pronounced for the first member in a respondent's network with high risk perceptions. Network partners with moderate or low HIV risk perceptions tend to reduce respondent's own worries about HIV/AIDS.
Source: Based on estimation results in Kohler *et al.* [41]

to have important—and often *causal*—effects on AIDS-related risk perceptions and behaviors (Figure 3),[37–41] confirming and extending related earlier findings on the influence of social network partners on fertility, HIV testing and related behaviors.[37,42–45] These effects of social networks on HIV-risk perceptions extended to spousal communication about AIDS risk; interactions with network partners—independent of network partners' risk assessments—tended to increase the probability of husband-wife communication about the disease.[38,41]

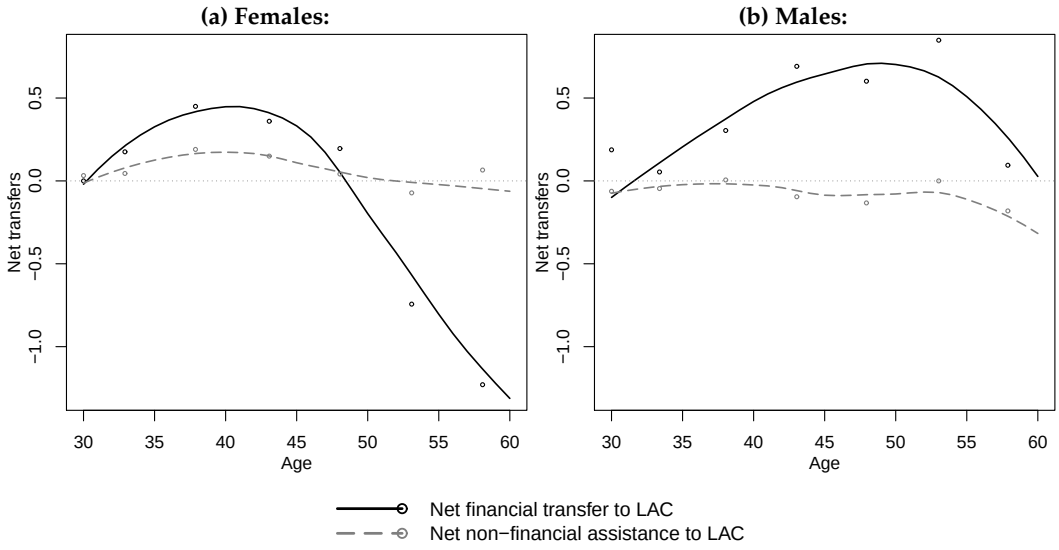
The MLSFH conversational networks that exerted strong influences on respondents risk perceptions and behaviors were found to exhibit a relatively dense network structure (most members knew each other as well as the respondent), and

they were characterized by a relatively stable network *structure* despite high turnover in the specific persons to whom respondents talked; networks were highly gendered (men talked with men, women with women).[33,40,46–48] In such dense networks, married men's expectations about the prevalence of extramarital sexual relationships in the network were shown to have a substantial influence on extramarital sexual behaviors.[49] In addition to providing key information about HIV risks and prevention strategies, MLSFH respondents also reported on sensitive topics such as the extramarital partnerships of their network partners and their best friend.[50]

These findings from the survey and ethnographic data about the role of social interactions on HIV risk perceptions and HIV-related behaviors are of central importance for understanding the dynamics of the HIV epidemic and its behavioral determinants. In general, these MLSFH studies have documented that social interactions, both in local social networks and through participation in group activities (e.g., attendance at funerals and going to bars), constitute important determinants of the strategies that individuals and couples developed for coping with the disease.[35,51] In particular, MLSFH studies have shown social networks exerted systematic and strong influences on risk perceptions and the probability of spousal communication about HIV/AIDS risks, and that these influences were in addition to other factors such as program interventions that disseminated knowledge about the disease, provided access to condoms, and advocated changes in sexual behaviors within and outside marriage. Social networks were also likely to amplify program efforts aimed at increasing individuals' information about HIV/AIDS and their assessments of their own risks as well as the risk they face from their spouses. Thus, social interactions were likely to have had a substantial impact on the course of the epidemic and the magnitude of its consequences, and these should be taken into consideration in understanding and predicting behaviors in such high-prevalence contexts and in devising program interventions with respect to the HIV/AIDS epidemic. A failure to do so may lead to misunderstanding the dynamics of behavioral change in response to the epidemic and to interventions related to the epidemic.[35,41,52]

Besides their role in the diffusion of information and the consideration of acceptable and unacceptable strategies for HIV prevention, social interactions were also important in terms of providing mutual insurance and resources. This is particularly important in contexts such as Malawi where formal insurance programs and financial markets are often absent. To understand these patterns of mutual insurance and transfers, the MLSFH collected detailed data on financial and non-financial transfers among family members and within broader community networks (Appendix A6.2). For example, financial and non-financial transfers occurring in familial social networks have been shown in the MLSFH to have been an important resource for individuals and families to ameliorate the implications of the HIV/AIDS epidemic.[53–56] While these transfers were widespread and a key characteristic of family relationships (Figure 4), contrary to expectations, in-

Figure 4: Net financial/non-financial transfers to living adult children (LAC)



LAC = living adult children. Net financial and non-financial transfers are calculated based on transfers given/received during 2-years prior to the 2008 MLSFH. Positive values indicate a transfer from respondents to their children, and negative values indicate transfers from children to the respondents. For non-financial transfers, the Figure shows that, despite the fact that there is considerable mutual non-financial exchange between parents and their children, the *net* resource flows as a result of these non-financial transfers seem to be relatively small and there is *no* marked age pattern for either male or female respondents. In contrast, net resource flows as a result of financial transfers between respondents and their living adult children follow a marked age-pattern that indicates important differences in the flow of resources between respondents and their children across the life course. Around age 30, the net transfers to living adult children are very small because respondents tend to have a very small number of living adult children. At somewhat older ages, for both male and female respondents, net financial transfers towards children rise. In contrast to female respondents, adult children remain recipients of net financial transfers from male respondents until about respondent's age of 60.
Source: Kohler *et al.* [53]

tergenerational wealth flows did not always differ by kinship systems (matriliny or patriliny), nor were they generally related to health status. This is particularly surprising since the HIV/AIDS epidemic increased uncertainty among individuals about their current and future health status and their survival, and as a consequence, one would expect that the high disease-risk environment prevailing in rural Malawi and other SSA contexts would have affected transfer motivations and behavior among family members.[53] The transfers were, however, importantly constrained by the availability of transfer partners (parents or adult children), which were strongly age-patterned and often affected by (AIDS-related) mortality.

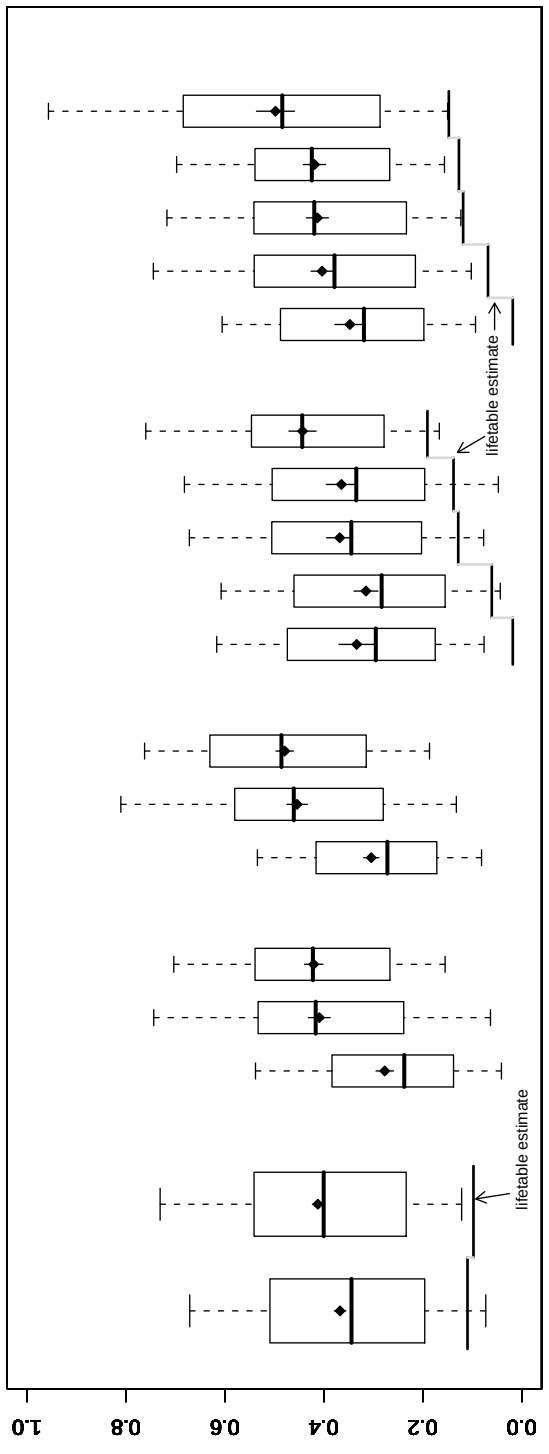
Subjective expectations about HIV infection and related health risks: For the period before 2006, when testing and treatment were not widely available in rural Malawi, the MLSFH has shown that the heuristics used by rural Malawians to assess their HIV risks often resulted in overestimates of their own likelihood of current HIV

infection, as well as that of their spouses.[57–59] Independent of HIV status, these high perceived risks of being infected with HIV had a strong negative effect on mental health and subjective well-being in rural Malawi.[60] Higher subjective HIV infection risks were also significantly associated with the behaviors that were perceived as being most risky in terms of HIV infection,[33,34] and the uncertainty about survival and future health that stems from high perceived HIV infection risks was negatively associated with children's school enrollment.[19] Respondents' HIV risk perceptions were importantly informed by events that individuals perceived as being related to HIV/AIDS, such as seeing a relative or a friend die from AIDS, or being advised to have a blood test before taking the serious step of ending a marriage.[61,62] This is likely related to the perception that testing will only confirm one's expectation of being found positive. An analysis of the ethnographic data shows that the primary reason for not testing is the expectation that one will be found positive and, as a result, will experience profound psychological distress, a result that might lead to suicide.[63] Even among the minority who supported HIV testing, the explanations were based on an expectation of an HIV-positive diagnosis: testing was said to be good because one would get counseling on how to live positively with AIDS.

Probabilistic expectations—that is, expectations that are measured on a well-defined numerical scale, are comparable across domains, and can be consistently interpreted as probabilities—have been collected in the MLSFH since 2006 based on interactive elicitation techniques (Appendix A6.3). These probabilistic expectations, which are preferable to subjective expectations based on Likert scales, indicate that MLSFH respondents were generally aware of differential HIV and other health risks.[64] For example, individuals with lower incomes and less land felt at greater risk of HIV infection than people with higher socioeconomic status (SES), and those who were divorced or widowed rightly perceived a greater risk of being infected with HIV than currently married individuals. Many expectations—including the probability of a newborn child dying within its first year of life and an individual's own probability of being currently infected with HIV—were well-calibrated compared to actual probabilities, but mortality expectations that measured the respondents' own risk of death over a 1, 5 or 10-year horizon were substantially overestimated compared to life table estimates (Figure 5).[64–66] This overestimation of mortality risks may lead individuals to underestimate the benefits of adopting HIV risk-reduction strategies. However, the arrival of ART services in the MLSFH study regions seems to have reduced subjective mortality risks, including among HIV-negative individuals who have not directly benefited from ART.[67]

HIV testing and counseling (HTC): The MLSFH first implemented home-based testing in 2004, and was the first large-scale survey to do so in Malawi. An earlier survey of the acceptability of HTC in Malawi during 2002–03 had concluded that Malawians were not ready to be tested,[68] and in the 2004 MLSFH survey, only 18.1% of respondents reported having had an HIV test. However, when the MLSFH started to offer HTC in 2004, over 90% of respondents in 2004 accepted the HIV test. Re-

Figure 5: Subjective probability of dying within a 5-year time period based on MLSFH probabilistic expectations



Notes: This boxplot-like graph displays the mean (dot) and median (center line) of the reported expectations, as well as the 10th (lower whisker), 25th (bottom of box), 75th (top of box), and 90th (upper whisker) percentiles of the distribution. Region is coded as: N = North (Rumphi), C = Center (Mchinji), S = South (Balaka). The key finding of the analyses is that both men and women in rural Malawi substantially *overestimated* their mortality risk, and they were much more pessimistic regarding their own survival as is warranted given current estimates of actual mortality rates. Across all ages in the sample, the median perceived 5-year mortality risk exceeded the corresponding lifetable estimate by a factor of 3.2 for males and 4.1 for females. The "bias" was the most severe for younger respondents, where the median five-year mortality probability reported by 15–19 year olds which was 16–17 times greater than the lifetable estimate. The discrepancy between life table estimates and subjective mortality expectations, however, decreased with age: above age 30, for men the median subjective probability of dying exceeded the corresponding lifetable estimate by a factor of 2.5–2.7 (5-year period) and 1.7–2.1 (10-year period); for women, the overestimation of mortality risk above age 30 was by a factor of 3.2–3.5. Despite this general overestimation, however, respondents mortality perceptions correctly reflected regional differences and age-gradients in mortality.

Source: Delavande & Kohler[64]

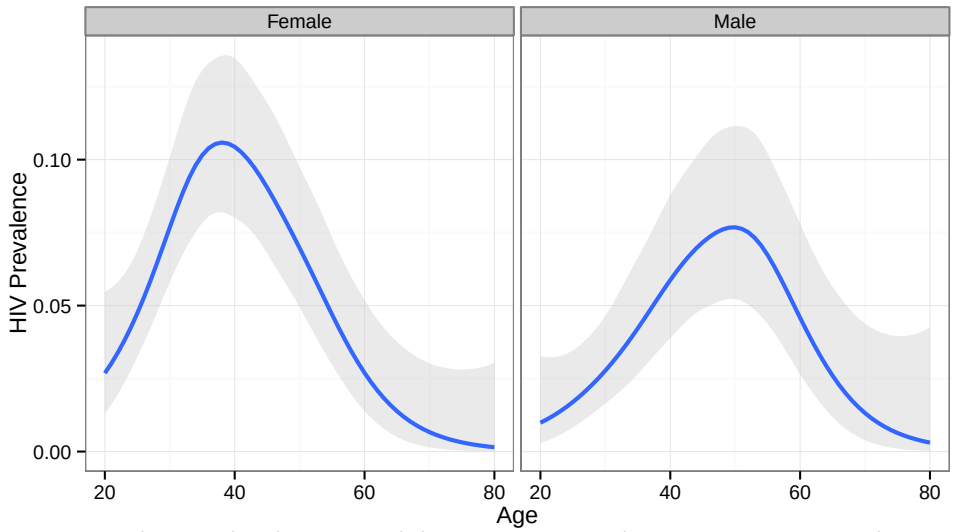
spondents were also offered tests for three other STIs—chlamydia, gonorrhea and trichomoniasis—response rates were similarly high. There was little variation in the fraction of respondents accepting the HIV and other STI tests by gender or by age (adults vs adolescents). Similarly high acceptance of HTC continued in 2006, despite variations in HIV testing and counseling protocols.[21,69,70]

In 2004, due to the use of lab-based HIV testing (rather than rapid HIV tests), the test results were made available to respondents 2–4 months after the sample collection in local HTC centers established by the MLSFH. The percentage of individuals who obtained their test results was 67% in 2004. It varied from about 34% among MLSFH respondents who were not offered a monetary incentive for learning their HIV test result to close to 80% among those who were offered an incentive.[32] 98% of MLSFH respondents in 2006 wanted to learn their HIV status when the HTC test results were available immediately given the use of rapid HIV testing kits. HTC participation in 2006 was 96% among those who learned their HIV status as part of the MLSFH in 2004, and it was 83% and 91% respectively among those tested for the first time by the MLSFH and those who did not learn their HIV status in 2004. The high acceptance rate of HTC during the MLSFH was importantly related the fact that home-based HTC offered credible information on HIV status through a transparent process. Home-based HTC was also perceived as convenient and confidential, which was not necessarily the case for HTC offered at clinics.[63,69] Because these concerns contributed to the low uptake of HTC services offered at government clinics, substituting home-based HTC for clinic-based HIV testing can be one important approach for eliminating socioeconomic inequalities in access to and utilization of HTC.[71]

Adult HIV prevalence in the MLSFH in 2004–06 was stable at around 7%.[21,72] Among 2008 MLSFH respondents aged 15–49, 5.8% (women: 7.0%; men: 3.7%) were HIV-positive (based on at least one positive MLSFH HIV test result during 2004–08 among 2008 respondents with at least one valid MLSFH HIV test). HIV prevalence varies strongly with age (Figure 6), peaking for women around age 35 and for men around 50, which is similar to the pattern of HIV prevalence observed in the 2010 Malawi DHS.[11] HIV prevalence in the MLSFH is also substantially higher among female MLSFH respondents who have experienced a marital separation/divorce or entered widowhood as compared to MLSFH respondents who have not experienced a termination of their marriage (Figure 7). In terms of regional differences among the MLSFH study sites, consistent with DHS findings about regional variation in prevalence, HIV prevalence was found to be highest in Balaka where men are commonly circumcised (which has been shown to reduce HIV infection risks).[73]

The MLSFH documents 50 HIV incident cases during 2004–08, 45 of which occurred among MLSFH respondents aged 25–49 in 2006. The HIV incidence rate observed among MLSFH respondents during 2004–08 was 0.63 per 100 person years (95% CI: 0.47–0.84), higher among women (incidence rate = 0.74 per 100 person years, 95% CI: 0.52–0.011) than among men (incidence rate = 0.47 per 100 person

Figure 6: HIV prevalence among 2008 MLSFH respondents by age



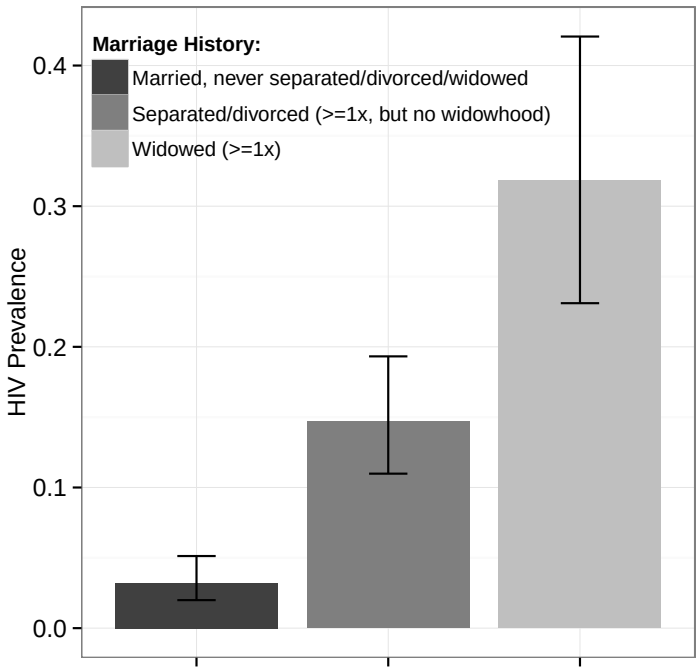
Notes: For respondents with at least one valid MLSFH HIV test during 2006–08. Respondents with at least one HIV-positive MLSFH HIV test during 2006–2008 are considered as being HIV-positive, all others are considered being HIV-negative at the 2008 MLSFH Round (MLSFH 5).

years, 95% CI: 0.28–0.79), although this difference in incidence rates is not statistically significant ($p = .15$).

In addition to establishing HIV prevalence and incidence, HTC has received considerable recent attention that as a method for the effective prevention of the spread of HIV. For example, an Op-Ed piece in the New York Times declared that HTC is the "missing weapon" in the battle against AIDS, based on the argument that those who learn they are *not* infected will be more strongly motivated to avoid infection in the future, and those who learn that they *are* infected will be motivated to avoid infecting others.[75] The evidence supporting this claim, however, remains somewhat mixed.[76,77] Studies based on the MLSFH have shown that sexually-active HIV-positive individuals who learned their results during 2004 MLSFH HTC were three times more likely to purchase condoms two months later than sexually-active HIV-positive individuals who did not learn their results.[32] There was no significant effect of learning HIV-negative status on the purchase of condoms. The interest of individuals in learning their HIV status, however, was importantly influenced by peer influences and social networks, with MLSFH respondents whose neighbors attended the MLSFH HTC clinic being significantly more likely to learn their HIV status than respondents whose neighbors did not learn their HIV status.[45] In addition, disclosure of HIV status by respondents to their spouses and other community members was found to be relatively common among rural Malawians (and while common among both, HIV-negative individuals disclosed their HIV status more frequently than HIV-positive individuals).[78]

Studying the medium-term consequences of learning HIV status during the

Figure 7: HIV prevalence for women aged 35 with different marital histories



Notes: Marital histories are measured as: (*a*) women who are married and have never experienced a separation, divorce or widowhood (pred. HIV prevalence = 3.2%, 95% CI: 2.0%–5.1%); (*b*) women who have experienced at least one marital separation or divorce, but not widowhood (pred. HIV prevalence = 14.7%, 95% CI: 11.0%–19.3%); and (*c*) women who have entered widowhood at least once (pred. HIV prevalence = 31.8%, 95% CI: 23.1%–42.1%). Predicted HIV status at age 35 is obtained from a logistic regression of HIV status on age and marital history using 947 ever-married women aged 25–45 who were interviewed in the 2008 (MLSFH 5) (57.4% were married and have never experienced a separation, divorce or widowhood; 31.9% had experienced at least one marital separation or divorce, but not widowhood; and 10.7% had entered widowhood at least once). Marital histories up to 2008 were constructed using data from the 2006, 2008 and 2010 MLSFH and were cleaned for consistency (Appendix A6.4). Only respondents with recorded marital histories and at least one valid MLSFH HIV test during 2006–08 are included. Respondents with at least one HIV-positive MLSFH HIV test during 2006–2008 are considered as being HIV-positive, all others are considered being HIV-negative at the 2008 MLSFH Round (MLSFH 5).
Source: own calculations based on reconstructed marriage histories provided by Chae[74]

2004 MLSFH HTC on subsequent HIV/AIDS-related expectations and sexual behaviors also revealed that MLSFH respondents who received an HIV-negative test result in 2004 reported—somewhat paradoxically—higher and less accurate subjective expectations about being HIV-positive after two years.[65] HIV-positive individuals who learned their status in 2004 reported having fewer partners in 2006 and having used condoms more often during 2004–06 than those who did not learn their status. Also, the desire to have more children decreased after receiving a positive HIV-test result. Among married MLSFH respondents in HIV-negative

couples, learning only one's own status increased risky behavior, while learning both statuses decreased risky behavior. Learning the HIV status in 2004 did not seem to affect chances of divorce for either HIV-negative or HIV-positive MLSFH respondents after 2004, while it reduced the number of sexual partners among HIV-positive respondents, reduced fertility, and increased condom use with spouses for both HIV-negative and HIV-positive respondents. There were also relatively few differences after two years in terms of savings, income, expenditures, and employment between MLSFH respondents who learned and did not learn their status as part of the 2004 MLSFH HTC.[65,79–84] However, HIV/AIDS-related morbidity and mortality at the household-level resulted in a diversification of income sources, with women (but not men) reallocating their time, generally from work-intensive (typically farming and heavy chores) to cash-generating tasks (such as casual labor).[85] In addition, the arrival of ART services in the MLSFH study regions has contributed to both better mental health and higher agricultural productivity, including among HIV-negative individuals, because access to ART reduced uncertainty about survival and resulted in more forward-looking decision-making.[67,86]

Sexual behaviors, HIV risks and HIV prevention strategies: Only 14.8% of female MLSFH respondents had ever used a condom with their husband, but 62% of men report having ever used a condom with a non-marital partner. Married couples with more children were more likely to use condoms, and having been informed by experts about AIDS prevention at home induced men and women to overreport condom use within marriage.[87] There was no association between religion, frequency of church attendance and HIV infection on condom use within the current or most recent marriage.[88] However, for young people transitioning to marriage, condom use was found to be increasingly acceptable.[89,90] Moreover, in longitudinal observations among adolescents and young adults, hoping to marry later (rather than earlier) was correlated with a later initiation of sexual activity, less recent sexual activity, and a lower total number of lifetime sexual partners.[89] The transition into marriage was also associated with pronounced attitudinal shift regarding the acceptability of condom use within marriage, suggesting that attitudes about and use of condoms are susceptible to change in the context of important life-course transitions.[91] Contrary to a common expectation, qualitative MLSFH data also revealed that monetary exchanges and gifts in premarital sexual partnerships were as much about the expression of love and commitment as they were about meeting the financial needs of girls or the acquisition of sex for boys.[90,92] Polygyny and HIV infection were found to be positively associated on the individual-level; on the ecological level, however, the northern region—where polygyny is more common—had a lower HIV prevalence compared to the southern and central region where polygyny is less widespread. The positive individual-level association was due to men in polygynous marriages having more extramarital sex than men in monogamous unions and an adverse selection of HIV-positive women into polygynous unions. The negative ecological correlation was related to the distinctive structure of sexual networks produced by polygyny, the disproportionate recruitment

of HIV-positive women into marriages with a polygynous husband, and the lower coital frequency in conjugal dyads of polygynous marriages.[93,94]

Although sexual behavior generally declined with age, the MLSFH also revealed that there continued to be considerable levels of sexual activity and HIV infection risks among older Malawians. Hence, while older individuals are often excluded in surveys in SSA, these populations can be highly relevant to studies of sexual behavior and HIV risk, especially as ART improves the health and expands the life expectancy of HIV-positive individuals.[95,96]

Migrants originating from rural areas were also shown to be more likely than non-migrants to be HIV-positive and to have engaged in HIV risk behavior. However, while this association between migration and HIV risk or prevalence has been shown in other contexts as well, the pre- and post-migration data available in the MLSFH also provided evidence that HIV-positive individuals were more likely migrate than those who are HIV-negative.[23] This finding, which contradicts the conventional perception about the relationship between HIV risks and migration, resulted from migration related to marital instability that was more common for HIV-positive individuals and individuals with high HIV infection risks.

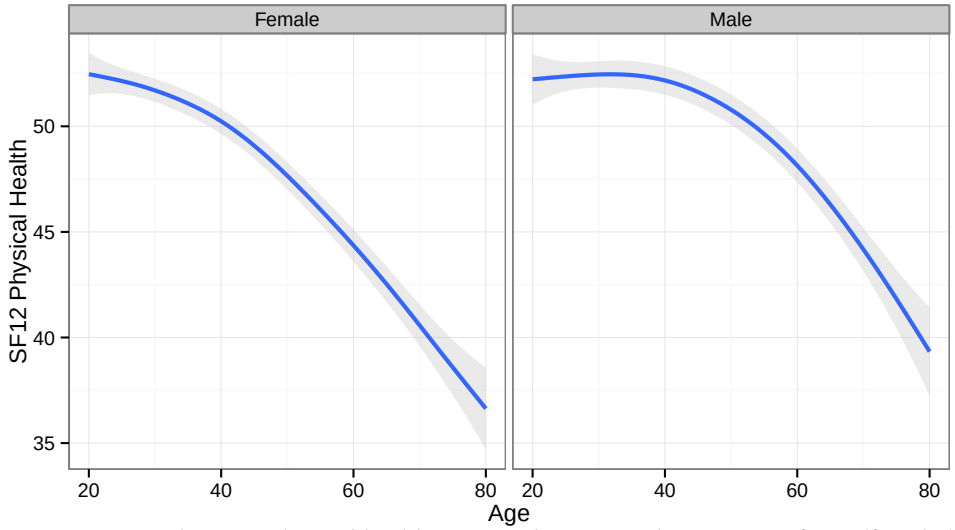
The MLSFH has also provided essential insights into the strategies of prevention that women and men in rural Malawi were using to reduce their HIV infection risks.[33,50,97–99] For many within marriage, the extremes of complete fidelity or consistent condom use were not considered possible or acceptable, and there was considerable struggle to find strategies that were personally, and socially acceptable. The emerging compromise strategy that has been documented using qualitative and quantitative MLSFH data was to avoid extramarital sex "as much as possible"; when it was not possible, select a partner who was likely to be "safe"; when that was not possible, use a condom.[33,89,100] For married individuals, a primary strategy was to try to persuade the spouse to be faithful.[33,46] For both women and men, divorce was also increasingly seen as an appropriate response to the threat that a spouse "will bring AIDS into the family".[17,33,101] Life table probabilities of divorce ranged from 40 to 65%, being among the highest on the continent, and divorce rates in rural Malawi were also often higher for HIV-positive individuals and women who were divorced or widowed have increasingly become less desirable marriage partners.[17,102] In addition, women who both delayed sexual debut and did not marry their first partner were, once married, more likely to have experienced marital disruption and to be HIV-positive.[103] Marital partner choice and divorce were therefore found to be two important behavioral strategies—distinct from the ABC strategies that emphasize abstinence, condom use and faithfulness—that women in rural Malawi deployed to manage their exposure to HIV.

Conditional cash transfers (CCTs), which have received considerable attention as a potentially innovative and effective approach to the prevention of HIV,[104–106] have been shown in the MLSFH to significantly increase the demand for learning one's HIV status and the respondents participation in counseling subsequent to HIV tests.[32] However, a MLSFH conditional cash transfer program that offered

financial incentives to men and women to maintain their HIV status for approximately one year (with rewards ranging from zero to approximately 4 months wages) found no effects of the offered incentives on HIV status or on reported sexual behavior,[22] and these results somewhat question the "unconditional effectiveness" of CCT programs for HIV prevention that has been suggested by related studies and some optimistic media reports.[104–106] In interpreting the above results that are based on self-reported data on sexual behaviors, which is often seen as a sensitive topic, it is important to highlight that comparing MLSFH survey responses about sexual behaviors across different interview modes (face-to-face vs. audio computer-assisted self-interview vs. qualitative interviews) have not found clear patterns of differential reporting of sensitive behaviors by interviewer mode.[107,108] There is therefore no clear indication that the self-reported sexual behavior in surveys misrepresents actual behaviors in systematic ways. This conclusion is consistent with related studies based on sociocentric sexual network studies that have shown that misreporting of sexual behaviors, even where common, biases prevalence estimates of sexual partnership patterns—such as concurrent sexual partners—not necessarily towards zero, but rather in unknown directions.[109]

Health and mortality: Mortality levels among MLSFH respondents have been shown to be similar to that of the Malawi population, including mortality differences by gender, region and HIV status.[110–113] In addition to investigating mortality, a relatively recent research focus of the MLSFH has been analyses of the epidemiological transition in SSA, that is, the transition of disease patterns where the primary causes of morbidity and mortality increasingly shift from communicable to non-communicable diseases. The age group of mature adults (defined here as adults aged 45+) in this context deserves particular attention as physical health declines rapidly with age (Figure 8), and so does mental health.[114] The relatively poor health of mature adults, however, is important as in Malawi and similar SSA low-income countries, 80% of the additional persons-years lived among adults aged 25+ as a result of increasing life expectancies during the next 50 years will occur among individuals aged 45+: 4.1 additional years, or 38% of the overall adult life expectancy gain, will occur among individuals aged 45–64, and 5.1 years, or 47% of the adult life expectancy gain, will occur among individuals aged 65+ (based on UN projections).[115] The limited existing evidence on this topic, however, suggests that chronic and disabling conditions among the mature adult population, resulting from the cumulative effects of poor nutrition and frequent exposure to infectious disease, led to significant levels of functional limitations in day-to-day activities and a substantial gap between potential and actual economic productivity.[15,116–119] The MLSFH shed light on these issues in two distinct domains. First, analyses of the MLSFH biomarkers (total cholesterol, LDL, HDL, ratio of total cholesterol to HDL, albumin, creatinine and wr-CRP) found that only small proportions of MLSFH respondents had biomarker values in the critical range as defined by developed country standards. While the correlational patterns among the biomarkers were consistent with observations from developed countries, the comparison

Figure 8: SF12 physical health score among 2010 MLSFH respondents by age



Notes: At age 20–40, the SF12 physical health score in the MLSFH has a mean of 51.6 (females) and 52.4 (males), with a standard deviation of 7.2 and 6.1 respectively. The average SF12 physical health score for a 60 year old women is therefore more than 1 SD below the mean of 20–40 year old women, with health rapidly declining further with age; the average SF12 physical health score of a 60 year old male is .6 SDs below below the mean of 20–40 year old men, again, with further substantial declines at older ages.

with other low-income populations showed remarkably similar age-specific patterns of the biomarkers despite differences in the mode of blood sampling. The biomarkers exhibited also only very modest associations with measures of socioeconomic status (SES), indicating that commonly-found associations between SES and biomarker-based risk factors for age-related diseases among prime-aged and elderly individuals may not necessarily hold in contexts such as rural Malawi where individuals have been exposed to frequent infectious diseases and undernutrition.[24,25] These MLSFH findings point to a potentially important "puzzle" in understanding non-communicable diseases in Malawi: despite strong hypotheses for the existence of SES differentials in health (and the MLSFH biomarkers in particular), our results provided only weak evidence for variation in the MLSFH biomarkers for cardiovascular risk, non-specific inflammation, and renal or liver functioning by socioeconomic status. It is also possible that specific contexts of individuals in poor high-morbidity SSA environments importantly affected the distribution of these biomarkers and their association with SES and other behavioral/contextual covariates, an aspect that is currently not well understood.

While these biomarker-based analyses did not indicate widespread chronic conditions in the MLSFH study population, MLSFH data on self-reported functional limitations (Table 5) showed that disabilities among mature adults were common, and that the physical health may have been an important limiting factor for individuals' social and economic activities. For example, in 2010, close to one third

Table 5: Disability, work efforts, pain interfering with work, and subjective well-being for mature adults (aged 45+)

	Disability Status				% (2010) **worked** for income last 2 weeks		% (2010) **pain interfered** with work last 4 weeks		% (2010) somewhat/ very **unsatisfied** with life	
	2008		2010							
	N	%	N	%	N	%	N	%	N	%
Age 45–64										
Functional Limitation Classification										
Healthy	537	67.4	428	56.8	165	38.5	64	14.9	16	3.7
Moderately limited	200	25.1	239	31.7	82	34.3	120	50.2	23	9.6
Severely limited	41	5.1	64	8.5	16	26.2	40	65.6	16	26.2
Deceased	19	2.4	23	3.1	–	–	–			
Proportion Male	362	45.4	340	45.1	–	–	–			
Age 65+										
Functional Limitation Classification										
Healthy	120	41.0	87	24.6	23	26.4	18	20.6	5	5.7
Moderately limited	108	36.9	142	40.1	34	23.9	75	52.8	14	9.9
Severely limited	62	21.2	93	26.3	8	8.99	69	77.5	26	29.2
Deceased	3	1.0	32	9.0	–	–	–			
Proportion Male	142	48.5	170	48.0	–	–	–			

Notes: Functional limitation classification is based on MLSFH questions (*i*) "Do you have any health problems that limit you in carrying out moderate activities?" and (*ii*) "Do you have any health problems that limit you in carrying out strenuous activities?", with each question providing a list of moderately/strenuous activities and response categories being "not limited", "limited a little" and "limited a lot". Individuals who indicate that they had no limitations in either set of activities are classified as healthy, those who respond "somewhat limited" on either question are classified as *moderately limited*, and individuals who respond "limited a lot" on either question are classified as *severely limited*. Deceased refers to mortality between survey waves among respondents who were interviewed in the MLSFH 2006 and/or 2008. Pain interfering with work is based on the question "During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?" Individuals responding "moderately", "quite a bit", or "extremely" are classified as being limited by pain. Individuals responding "somewhat unsatisfied" or "very unsatisfied" to the question "How satisfied are you with your life, all things considered?" are classified as having low life satisfaction.
Source: Payne *et al.* [110]

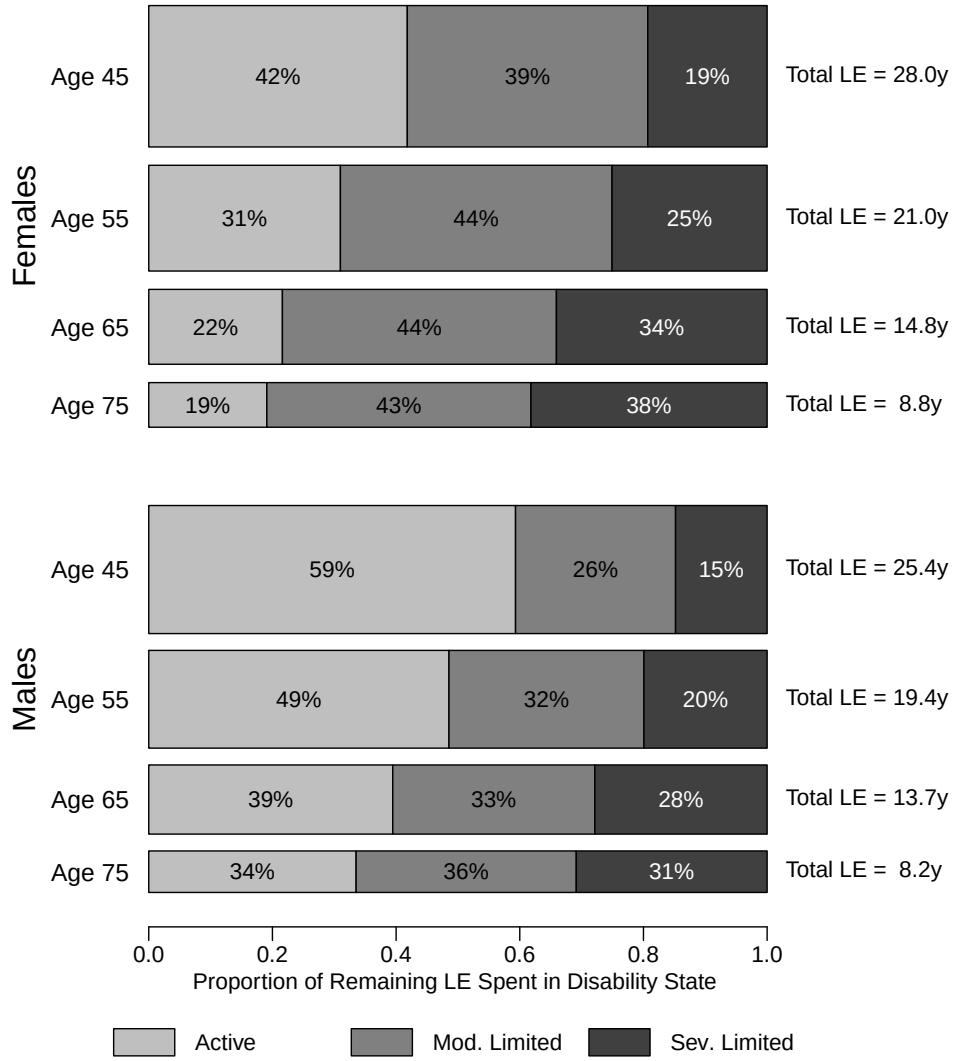
of respondents aged 45–64 indicated that they had moderate functional limitations, and 8.5% reported being severely limited in their physical activities, and both physical limitations were substantially more common among individuals aged 65+. Functional limitations were also strongly associated with reduced social and/or economic activities (Table 5). For example, the percentage of individuals working for income within the past week decreased steadily with increasing disability, and individuals who reported limitations on physical activity also reported that their work efforts (both within and outside the household) were substantially limited by pain. Functional limitation was also associated with substantially lower subjective well-being, with more than a quarter of severely limited individuals responding that they were "somewhat unsatisfied" or "very unsatisfied" with their lives, as compared to less than 4% of healthy individuals 45–64 and less than 6% of healthy individuals 65+. Moreover, age patterns of the onset of functional limitations and the transitions-over-time between different disability states in the MLSFH revealed that the age-specific risks of experiencing an onset of functional limitations due to poor physical health were high in this population compared to more developed contexts, and onset of persistent disabilities happened considerably earlier in life. For example, the MLSFH suggested that 45-year old women in Malawi can expect to spend 58% of their remaining 28 years of life with functional limitations, while 45-year old men can expect to live 41% of their remaining 25.4 years subject to such limitations (Figure 9). Disabilities related to functional limitations had a substantial negative effect on individuals' labor activities, a major concern in this agrarian context, and were negatively related to subjective well-being.[110] However, health status varied strongly with income in this agrarian context, and a doubling of income (causally) increased general health status by 10.2% and well-being by 12.5%.[120]

What are the main strengths and weaknesses?

The MLSFH cohort was selected to represent the rural population of Malawi, where the majority of Malawians live in conditions that are similar to those in the rural areas of other countries in high HIV prevalence: health conditions are poor, health facilities and schools are over-burdened and under-staffed, standards of living are low and nutritional needs of adults, children and the elderly are often not met. In addition to this focus on poor rural individuals that constitute the majority of the Malawi and SSA population, main strengths of the MLSFH data include the relatively large sample size, generally high data quality, the longitudinal design covering more than a decade of health conditions and socioeconomic changes in a rural SSA context with high HIV prevalence, and the broad focus of the MLSFH that provides information about health (including biomarkers for HIV), social networks, social and economic contexts, sexual behaviors, marriage and marital transitions and household structures and dynamics.

Several weaknesses of the MLSFH are noteworthy. First, the MLSFH does not have a nationally representative sample design, in part related to the considerable

Figure 9: Distribution of remaining life expectancy (LE) by disability state



The figure shows the proportion of remaining life an average individual will spend in healthy, moderately limited, and severely limited life at age 45, 55, 65, and 75, for females (top panel) and males (bottom panel). The height and area of each bar is proportional to the overall remaining life expectancy of the synthetic cohorts with initial ages of 45, 55, 65 and 75 years, and the differently shaded areas represent the distribution of the remaining life expectancy across the three disability states: healthy, moderately limited and severely limited. The bars do not necessarily reflect the ordering of these life-years by disability states as individuals in our analysis can recover and relapse between disability states, so not all years of limitation are spent at the end of life. Analyses are based on MLSFH respondents from 2006, 2008 and 2010, using longitudinal data to estimate age-patterns of functional limitations and the transitions-over-time between different disability states using a discrete-time hazard model. Based on these transition rates, multi-state life tables (MSLTs) are estimated using microsimulation approaches to estimate the above LEs by disability state.
Source: Payne *et al.* [110]

costs such a study design would entail. As a result, urban contexts—where about 15% of the Malawi population live—are not reflected in the MLSFH, and the rural MLSFH study population is from only three study sites (Balaka, Mchinji and Rumphi). Nevertheless, despite this limitation, the MLSFH reflects the considerable heterogeneity of social and demographic contexts across rural Malawi, and comparison of the MLSFH study with the rural samples of nationally representative studies have shown few substantively-relevant differences in the composition of MLSFH and national representative study populations (Appendix A3.4). A second important concern in the MLSFH pertains to attrition. As is expected, the MLSFH study population is subject to considerable attrition as a result of migration, temporary absences, and mortality (Figure 2). Attrition was sometimes reversed as attriters at one waves were reinterviewed again at subsequent MLSFH waves. While the MLSFH made some efforts to follow migrants who left the MLSFH study villages,[23] this migration follow-up was not comprehensive and did not cover the most recent waves (a new project to update the migration follow-up is scheduled for 2013). Our analyses of attrition indicated that even though respondent characteristics often differ significantly between those who were lost to follow-up and those who were re-interviewed and attrition was often predicted by key respondent characteristics, the coefficient estimates for standard family background variables in regressions and probit equations for the majority of the outcome variables were not affected significantly by attrition (see Appendix A5 and related attrition analyses[2,80,121]). Thus, the attrition levels observed in the MLSFH may not necessarily represent a general problem for obtaining consistent estimates of the coefficients of interest for most of these outcomes. These results, which are very similar to those documented in other contexts,[121–123] suggest that multivariate estimates of behavioral relations may not be biased due to attrition and thus support the collection of longitudinal data.

CAN I GET HOLD OF THE DATA? WHERE CAN I FIND OUT MORE?

Public-use version of the MLSFH data without identifying individual or village information are made publicly available with some delay after data collection. MLSFH data up to 2010 (MLSFH 6) can currently be requested on the project website at http://www.malawi.upenn.edu, and these data are also processed for inclusion at the ICPSR at the University of Michigan. Researchers interested in using MLSFH data that have not (yet) been made available as part of the MLSFH public use data files can submit a two-page proposal (including an analysis plan and IRB plan) to the MLSFH principal investigator (mailto:hpkohler@pop.upenn.edu). If deemed scientifically sound and not overlapping with ongoing MLSFH research projects, researchers will then be asked to sign a Data Use Agreement to be able to access and utilize the MLSFH data that are not part of the public-use data sets. All analyses of the restricted MLSFH data are conducted in collaboration with members of the MLSFH study team.

FUNDING

The first wave of the MLSFH was funded by the Rockefeller Foundation. Subsequent funding has been provided by the NICHD (grants R01 HD053781, R01 HD/MH041713, R01 HD37276, R01 HD044228, R21 HD050652, R03 HD058976, R21 HD 071471), and has been supported by pilot grants from the Population Studies Center (PSC), Population Aging Research Center (PARC), the Boettner Center for Pensions and Retirement Security, the Institute on Aging, and the Center for AIDS Research (CFAR), all at the University of Pennsylvania and supported among other sources by NIH grants NICHD R24 HD044964, NIA P30 AG12836, NIAID AI 045008. The project also received funding from the University of Pennsylvania Research Foundation.

KEY MESSAGES

- The Malawi Longitudinal Study of Families and Health (MLSFH) provides a rare record of more than a decade of demographic, socioeconomic and health conditions in one of the world's poorest countries.
- MLSFH Data collection rounds in 1998, 2001, 2004, 2006, 2008, 2010 and 2012 for up to 4,000 individuals in three rural regions of Malawi.
- With more than 150 publications/working papers, the MLSFH is one of the premier datasets for research on health, family dynamics, social/sexual networks and human capital in SSA.

Author affiliations:
H.-P. Kohler is Frederick J. Warren Professor of Demography, Department of Sociology and Population Studies Center, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6299, USA; *Email:* hpkohler@pop.upenn.edu.

Watkins is Professor of Sociology, Department of Sociology and Population Studies Center, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6299, USA; *Email:* swatkins@pccpr.ucla.edu.

Behrman is William R. Kenan Jr. Professor of Economics and Sociology, Department of Economics and Population Studies Center, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104, USA; *Email:* jbehrman@econ.upenn.edu.

Anglewicz is Assistant Professor, School of Public Health & Tropical Medicine and Department of Global Health Systems and Development, Tulane University, 1440 Canal St., Suite 2200, New Orleans, LA 70112-2737; *Email:* panglewi@tulane.edu.

Fleming is Market Researcher at Facebook, Menlo Park, CA, and was Co-Founder of Invest in Knowledge (IKI), Zomba. *Email:* pcfleming@gmail.com.

I. Kohler is Associate Director, Population Aging Research Center (PARC) and PSC Research Associate, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104, USA; *Email:* iliana@pop.upenn.edu.

Thornton is Assistant Professor, Department of of Economics, University of Michigan, Ann Arbor, MI 48109; *Email:* rebeccal@umich.edu.

Mkandawire is Research Director, Invest in Knowledge (IKI), Zomba, Malawi;
Email: james.mkandawire@investinknowledge.org.

Kalilani-Phiri is Epidemiologist, Department of Community Health and Research Support Centre, College of Medicine, University of Malawi, Blantyre, Malawi; *Email:* lkalilani@medcol.mw.

Honde is Researcher and Fieldwork Manager, Invest in Knowledge (IKI), Zomba, Malawi; *Email:* hastings.honde@investinknowledge.org.

Haware is Researcher and Fieldwork Manager, Invest in Knowledge (IKI), Zomba, Malawi; *Email:* augustineharawa@yahoo.com.

Chilima is Director, Community Health Sciences Unit, Ministry of Health and Population, Lilonge, Malawi; *Email:* bchilima2@yahoo.com.

Bandawe is Associate Professor, Department of Mental Health, College of Medicine, University of Malawi, Blantyre, Malawi. *Email:* cbandawe@medcol.mw.

Mwapasa is Associate Professor of Epidemiology and Community Health, College of Medicine, University of Malawi, Blantyre, Malawi. *Email:* vmwapasa@medcol.mw.

MLSFH Study Team: In addition to the authors of this paper, members of the MLSFH Study Team who have made important contribution to the design and/or implementation of the MLSFH during 1998–2012 include: Jimi Adams, Thomas Anderson, Nicole Angotti, Afua Appiah-Yeboah, Simona Bignami-Van Assche, Crystal Biruk, Jennifer Browning, Kevin Burke, Ben Capistrant, Ruben Castro, Sophia Chae, Davie Chagunda, Li-Wei Chao, Kondwani Chavula, Edgar Chihana, Raphael Chikadza, Abdullah Chilungu, Twaina Chilungu, Agnes Chimbiri, Rachael Chimbwete, Wyson Chimesya, Angela Chimwaza, Brian Chin, Jesman Chintsanya, Davie Chitenje, Shelley Clark, Elisabetta De Cao, Adeline Delavande, Kim Deslandes, Kim Dionne, Henry Doctor, Theresa Fedor, David Frankenfield, Emily Freeman, Margaret Frye, Patrick Gerland, Monica Grant, Sylvester Guuru, Mike Harhay, Stéphane Helleringer, Ning Hsieh, Heide Jackson, Humphreys Kabota, Liney Kachama, Amy Kaler, Jonathan Kandodo, Doreen Kanyika, Nancy Luke, Sydney Lungu, Denview Malagasi, Christopher Manyamba, Aaron Mapondera, Winford Masanjala, Chiyembekezo Mbewe, Tara McKay, Andy Mguntha, Esnat Mkandawire, Mathero Mkhalamba, Placid Mpeketula, Grace Msefula, Patrick Msukwa, Macgivens Msumba, James Mwera, Tyler Myroniuk, Blessings M'bwerazino, Flora Nankhuni, Ken Ndala, Oleosi Ntshebe, Julius Nyambo, Gift Nyasulu, Francis Obare Onyango, Collin

Payne, Joel Phiri, Michelle Poulin, Mark Regnerus, George Reniers, Julio Romero Prieto, Anat Rosenthal, Gifton Saizi, Raül Santaeulàlia-Llopis, Enid Schatz, Chikondi Singano, Kristen Smith, Beth Soldo, Erica Soler-Hampejsek, Sarah Spell, Frank Taulo, Linda Tawfik, Evance Tchale Rebbeca Tesfai, Jenny Trinitapoli, Ina Warriner, Alexander Weinreb, Anika Wilson, Sara Yeatman, Eliya Zulu.

*Corresponding author: Hans-Peter Kohler, hpkohler@pop.upenn.edu.

Acknowledgements: We gratefully acknowledge the generous support for the Malawi Longitudinal Study of Families and Health (MLSFH) by the Rockefeller Foundation, the National Institute of Child Health and Development (grant numbers R03 HD05 8976, R21 HD050652, R01 HD044228, R01 HD053781), the National Institute on Aging (grant number P30 AG12836), the Boettner Center for Pensions and Retirement Security at the University of Pennsylvania, and the National Institute of Child Health and Development Population Research Infrastructure Program (grant number R24 HD-044964), all at the University of Pennsylvania. We are also grateful for pilot funding received through the Penn Center for AIDS Research (CFAR), supported by NIAID AI 045008, and the Penn Institute on Aging. We also are grateful for research assistance provided by Sophia Chae, Theresa Fedor and Collin Payne for the analyses in this paper.

REFERENCES

[1] Watkins S, Behrman JR, Kohler HP, Zulu EM. Introduction to "Research on demographic aspects of HIV/AIDS in rural Africa". *Demographic Research* 2003; **Special Collection 1(1)**:1–30.

[2] Anglewicz P, Adams J, Obare F, Kohler HP, Watkins S. The Malawi Diffusion and Ideational Change Project 2004–06: Data collection, data quality and analyses of attrition. *Demographic Research* 2009; **20(21)**:503–540.

[3] UNDP. *Human Development Report 2010: The Real Wealth of Nations—Pathways to Human Development*. New York, NY: United Nations Development Programme (UNDP), 2010. http://hdr.undp.org/en/reports/global/hdr2010.

[4] UN Population Division. World Population Prospects, the 2012 revision: Standard (median) forecasts, 2012. United Nations, Department of Economic and Social Affairs, Population Division. http://esa.un.org/unpd/wpp/.

[5] Kohler HP. Population growth. In *Global Problems, Smart Solutions*, editor Lomborg B, Cambridge, MA: Cambridge University Press. 2013. Earlier version is available as University of Pennsylvania PSC Working Paper 12-03 (http://repository.upenn.edu/psc_working_papers/34).

[6] Salomon JA, Wang H, Freeman MK, Vos T, Flaxman AD, Lopez AD, *et al.* Healthy life expectancy for 187 countries, 1990–2010: A systematic analysis for the Global Burden Disease Study 2010. *Lancet* 2012; **380(9859)**:2144–2162.

[7] UNDP. *2010 Malawi Millennium Development Goals Report*. New York, NY, and Lilongwe, Malawi: United Nations Development Programme (UNDP) and Malawi Ministry of Development Planning and Cooperation, 2010.

[8] World Bank. Malawi at a glance, 2013. World Bank Country Profile. http://devdata.worldbank.org/AAG/mwi_aag.pdf.

[9] WHO. Malawi: Health profile, 2013. Geneva: World Health Organization. http://www.who.int/gho/countries/mwi.pdf.

[10] UNAIDS. *Global Report: UNAIDS Report on the Global AIDS Epidemic*. New York: World Health Organization and UNAIDS, 2012. http://www.unaids.org/globalreport/.

[11] Malawi DHS. *Malawi Demographic and Health Survey 2010 (Final Report)*. Zomba, Malawi, and Calverton, Maryland, USA: National Statistical Office (NSO) and ICF Macro, 2011. http://www.measuredhs.com.

[12] Jahn A, Floyd S, Crampin AC, Mwaungulu F, Mvula H, Munthali F, *et al.* Population-level effect of HIV on adult mortality and early evidence of reversal after introduction of antiretroviral therapy in Malawi. *Lancet* 2008; **371(9624)**:1603–1611.

[13] Guebbels E, Bowie C (editors). *The Epidemiology of Malawi*. College of Medicine, University of Malawi: Special Issue of the *Malawi Medical Journal*, 2006. http://www.medcol.mw/commhealth/publications/epi%20book/epidemiology%20book.htm.

[14] Chitsulo L, Engels D, Montresor A, Savioli L. The global status of schistosomiasis and its control. *Acta Tropica* 2000; **77(1)**:41–51.

[15] Msyamboza KP, Ngwira B, Dzowela T, Mvula C, Kathyola D, Harries AD, *et al.* The burden of selected chronic non-communicable diseases and their risk factors in Malawi: Nationwide STEPS survey. *PLoS ONE* 2011; **6(5)**:e20316–.

[16] Malawi NSO. *Malawi: An Atlas of Social Statistics*. Lilongwe, Malawi: National Statistics Office of Malawi, 2002. http://www.ifpri.org/sites/default/files/publications/malawiatlas.pdf.

[17] Reniers G. Divorce and remarriage in rural Malawi. *Demographic Research* 2003; **Special Collection 1(6)**:175–206.

[18] Trinitapoli J, Regnerus MD. Religious involvement and HIV/AIDS risk in rural Malawi. *Journal for the Scientific Study of Religion* 2006; **45**:505–528.

[19] Grant MJ. Children's school participation and HIV/AIDS in rural Malawi: The role of parental knowledge and perceptions. *Demographic Research* 2008; **19(45)**:1603–1634.

[20] Kerr R. Food security in northern Malawi: Gender, kinship relations and entitlements in historical context. *Journal of Southern African Studies* 2005; **31(1)**:53–74.

[21] Obare F, Fleming P, Anglewicz P, Thornton R, Martinson F, Kapatuka A, *et al.* Acceptance of repeat population-based voluntary counseling and testing for HIV in rural Malawi. *Sexually Transmitted Infections* 2009; **85(2)**:139–144.

[22] Kohler HP, Thornton R. Conditional cash transfers and HIV/AIDS prevention: Unconditionally promising? *World Bank Economic Review* 2012; **26(2)**:165–190.

[23] Anglewicz P. Migration, marital dissolution, and HIV infection in Malawi. *Demography* 2012; **49(1)**:239–265.

[24] Kohler IV, Soldo BJ, Anglewicz P, Chilima B, Kohler HP. Association of blood lipids, creatinine, albumin and CRP with socioeconomic status in Malawi. *Populaton Health Metrics* 2013; **11(4)**.

[25] Kohler IV, Anglewicz P, Kohler HP, Mcabe J, Chilima B, Soldo BJ. Evaluating health and disease in sub-Saharan Africa: Minimally invasive collection of plasma in the Malawi Longitudinal Study of Families and Health. *Genus* 2012; **68(2)**:1–27.

[26] Gootjes J, Tel RM, Bergkamp FJM, Gorgels JPMC. Laboratory evaluation of a novel capillary blood sampling device for measuring eight clinical chemistry parameters and HbA1c. *Clinica Chimica Acta* 2009; **401(1–2)**:152–157.

[27] Bignami-Van Assche S. Are we measuring what we want to measure? An analysis of individual consistency in survey response in rural Malawi. *Demographic Research* 2003; **Special Collection 1(3)**:77–108.

[28] Bignami-Van Assche S, Reniers G, Weinreb AA. An assessment of the KDICP and MDICP data quality: Interviewer effects, question reliability and sample attrition. *Demographic Research* 2003; **Special Collection 1(2)**:31–76.

[29] Castro R. Inconsistent respondents and sensitive questions. *Field Methods* 2013; E-publication before print, Dec 11, 2012.

[30] Biruk C. Seeing like a research project: Producing "high-quality data" in AIDS research in Malawi. *Medical Anthropology* 2012; **31(4)**:347–366.

[31] Chae S. Forgotten marriages? measuring the reliability of retrospective marriage histories, 2012. Presented at a workshop on Quality and Comparability of Demographic Data in Sub-Saharan Africa, INED, March 2012.

[32] Thornton RL. The demand for learning HIV status and the impact on sexual behavior: Evidence from a field experiment. *American Economic Review* 2008; **98(5)**:1829–1863.

[33] Watkins SC. Navigating the AIDS epidemic in rural Malawi. *Population and Development Review* 2004; **30(4)**:673–705.

[34] Smith KP, Watkins SC. Perceptions of risk and strategies for prevention: Responses to HIV/AIDS in rural Malawi. *Social Science and Medicine* 2005; **60(3)**:649–660.

[35] Kohler HP, Helleringer S, Behrman JR, Watkins SC. The social and the sexual: Networks in contemporary demographic research. In *Population in the Human Sciences: Concepts, Models, Evidence*, editors Kreager P, Ulijaszek S, Winney B, Capelli C, Oxford, UK: Oxford University Press. forthcoming. Available as PSC Working Paper 13-01 at http://repository.upenn.edu/psc_working_papers/41/.

[36] Clark S. Extra-marital sexual partnerships and male friendships in rural Malawi. *Demographic Research* 2010; **22(1)**:1–28.

[37] Behrman JR, Kohler HP, Watkins SC. Social networks and changes in contraceptive use over time: Evidence from a longitudinal study in rural Kenya. *Demography* 2002; **39(4)**:713–736.

[38] Helleringer S, Kohler HP. Social networks, risk perceptions and changing attitudes towards HIV/AIDS: New evidence from a longitudinal study using fixed-effect estimation. *Population Studies* 2005; **59(3)**:265–282.

[39] Bühler C, Kohler HP. Talking about AIDS: The influence of communication networks on individual risk perceptions of HIV/AIDS infection and favored protective behaviors in South Nyanza District, Kenya. *Demographic Research* 2003; **Special Collection 1(13)**:398–438.

[40] Watkins SC, Warriner I. How do we know we need to control for selectivity? *Demographic Research* 2003; **Special Collection 1(4)**:109–142.

[41] Kohler HP, Behrman JR, Watkins SC. Social networks and HIV/AIDS risk perceptions. *Demography* 2007; **44(1)**:1–33.

[42] Bongaarts J, Watkins SC. Social interactions and contemporary fertility transitions. *Population and Development Review* 1996; **22(4)**:639–682.

[43] Kohler HP, Behrman JR, Watkins SC. The density of social networks and fertility decisions: Evidence from South Nyanza District, Kenya. *Demography* 2001; **38(1)**:43–58.

[44] Kohler HP. *Fertility and Social Interactions: An Economic Perspective*. Oxford: Oxford University Press, 2001. http://books.google.com/books?id=03WVNSjf1KkC.

[45] Godlonton S, Thornton R. Peer effects in learning HIV results. *Journal of Development Economics* 2012; **97(1)**:118–129.

[46] Zulu EM, Chepngeno G. Spousal communication about the risk of contracting HIV/AIDS in rural Malawi. *Demographic Research* 2003; **Special Collection 1(8)**:247–277.

[47] Bignami-Van Assche S. Network stability in longitudinal data: A case-study from rural Malawi. *Social Networks* 2005; **27(3)**:231–247.

[48] Weinreb AA. Change and instability: A multilevel analysis of AIDS-related conversational networks among Malawian women. *Demographic Research* 2003; **Special Collection 1(12)**:373–396.

[49] Cordero Coma J. When the group encourages extramarital sex: Difficulties in HIV/AIDS prevention in rural Malawi. *Demographic Research* 2013; **28(30)**:849–880. http://www.demographic-research.org/volumes/vol28/30/.

[50] Tawfik L, Watkins SC. Sex in Geneva, sex in Malawi. *Social Science and Medicine* 2007; **64(5)**:1090–1101.

[51] Gerland P. Effects of social interactions on individual AIDS-prevention attitudes and behaviors in rural Malawi, 2005. Unpublished Doctoral Dissertation, Office of Population Research, Princeton University, Princeton, NJ.

[52] Behrman JR, Kohler HP, Watkins SC. Lessons from empirical network analyses on matters of life and death in East Africa. In *The Network Challenge: Strategy, Profit, and Risk in an Interlinked World*, editors Kleindorfer PR, Wind Y, Upper Saddle River, NJ: Wharton School Publishing, chap. 28, pp. 495–512. 2009. http://books.google.com/books?id=gjq4mC3hW1QC.

[53] Kohler IV, Kohler HP, Anglewicz P, Behrman JR. Intergenerational transfers in the era of HIV/AIDS: Evidence from rural Malawi. *Demographic Research* 2012; **27(27)**:775–834.

[54] Mtika MM, Doctor HV. Matriliny, patriliny, and wealth flow variations in Malawi. *African Sociological Review* 2004; **6(2)**:71–97.

[55] Weinreb AA. Lateral and vertical intergenerational exchange in rural Malawi. *Journal of Cross-Cultural Gerontology* 2002; **12(2)**:1–38.

[56] Potter GE, Handcock MS. A description of within-family resource exchange networks in a Malawian village. *Demographic Research* 2010; **23(6)**:117–152.

[57] Anglewicz P, Kohler HP. Overestimating HIV infection: The construction and accuracy of subjective probabilities of HIV infection in rural Malawi. *Demographic Research* 2009; **20(6)**:65–96.

[58] Anglewicz P, Bignami-Van Assche S, Clark S, Mkandawire J. HIV risk among currently married couples in rural Malawi: What do spouses know about each other? *AIDS and Behavior* 2010; **14(1)**:103–112.

[59] Bignami-Van Assche S, Chao LW, Anglewicz P, Chilongozi D, Bula A. The validity of self-reported likelihood of HIV infection among the general population in rural Malawi. *Sexually Transmitted Infections* 2007; **83(1)**:35–40. PMCID: PMC1800875.

[60] Hsieh N. Perceived risk of HIV infection and mental well-being in rural Malawi. *Demographic Research* 2013; **28(13)**:373–408.

[61] Watkins SC, Swidler A. Hearsay ethnography: Conversational journals as a method for studying culture in action. *Poetics* 2009; **37(2)**:162–184.

[62] Watkins SC, Swidler A. Hearsay ethnography: A method for learning about responses to health interventions. In *Handbook of the Sociology of Health, Illness & Healing*, editors Pescosolido B, McLeod J, Martin J, Rogers A, New York: Springer, chap. 22, pp. 431–445. 2011.

[63] Kaler A, Watkins S. Asking god about the date you will die: HIV testing as a zone of uncertainty in rural Malawi. *Demographic Research* 2010; **23(32)**:905–932.

[64] Delavande A, Kohler HP. Subjective expectations in the context of HIV/AIDS in Malawi. *Demographic Research* 2009; **20(31)**:817–874.

[65] Delavande A, Kohler HP. The impact of HIV testing on subjective expectations and risky behavior in Malawi. *Demography* 2012; **49(3)**:1011–1036.

[66] Delavande A, Giné X, McKenzie D. Eliciting probabilistic expectations with visual aids in developing countries: how sensitive are answers to variations in elicitation design? *Journal of Applied Econometrics* 2011; **26(3)**:479–497.

[67] Baranov V, Bennett D, Kohler HP. The indirect impact of antiretroviral therapy, 2012. Population Studies Center, University of Pennsylvania, Working Paper #12-08. http://repository.upenn.edu/psc_working_papers/39.

[68] Yoder SP, Matinga P. Voluntary counselling and testing (VCT) for HIV in Malawi: Public perspectives and recent VCT experiences, 2004. Calverton, Maryland, USA: ORC Macro.

[69] Angotti N, Bula A, Gaydosh L, Zeev Kimchi E, Thornton RL, Yeatman SE. Increasing the acceptability of HIV counseling and testing with three C's: Convenience, confidentiality and credibility. *Social Science & Medicine* 2009; **68(12)**:2263–2270.

[70] Angotti N. Working outside of the box: How HIV counselors in Sub-Saharan Africa adapt western hiv testing norms. *Social Science & Medicine* 2010; **71(5)**:986–993.

[71] Weinreb AA, Stecklov G. Social inequality and HIV-testing: Comparing home- and clinic-based testing in rural Malawi. *Demographic Research* 2009; **21(21)**:627–646.

[72] Obare F. Response to population-based voluntary counseling and testing for HIV in rural Malawi, 2007. Ph.D. Dissertation, Graduate Group in Demography, University of Pennsylvania, Philadelphia, PA.

[73] Poulin M, Muula AS. An inquiry into the uneven distribution of women's HIV infection in rural Malawi. *Demographic Research* 2011; **25(28)**:869–902.

[74] Chae S. Essays on family structure and marriage in sub-Saharan Africa, 2013. PhD Dissertation, Graduate Group in Demography, University of Pennsylvania.

[75] Holbrooke R, Furman R. A global battle's missing weapon. *New York Times*

2004; Op-Ed, February 10, 2004. Op-Ed,February10,2004.

[76] Denison J, O'Reilly K, Schmid G, Kennedy C, Sweat M. HIV voluntary counseling and testing and behavioral risk reduction in developing countries: A meta-analysis, 1990–2005. *AIDS and Behavior* 2008; **12(3)**:363–373.

[77] Fonner VA, Denison J, Kennedy CE, O'Reilly K, Sweat M. Voluntary counseling and testing (VCT) for changing HIV-related risk behavior in developing countries. *Coachrane Database of Systematic Reviews* 2012; **9**.

[78] Anglewicz P, Chitsanya J. Disclosure of HIV status between spouses in Malawi. *AIDS Care* 2011; **23(8)**:998–1005.

[79] Thornton RL. HIV testing, subjective beliefs and economic behavior. *Journal of Development Economics* 2012; **99(2)**:300–313.

[80] Fedor TM, Kohler HP, Behrman JR. The impact of learning HIV status on marital stability and sexual behavior within marriage in Malawi, 2013. Population Studies Center, University of Pennsylvania, Working Paper #13-02. http://repository.upenn.edu/psc_working_papers/42/.

[81] Yeatman S. HIV infection and fertility preferences in rural Malawi. *Studies in Family Planning* 2009; **40(4)**:261–276.

[82] Yeatman SE. The impact of HIV status and perceived status on fertility desires in rural Malawi. *AIDS and Behavior* 2009; **13(Issue 1 Supplement)**:S12–S19.

[83] De Paula A, Shapira G, Todd P. How beliefs about HIV status affect risky behaviors: Evidence from Malawi. *Journal of Applied Econometrics* forthcoming; Earlier version of this paper is available as PIER Working Paper No. 11-005. URL: http://ssrn.com/paper=1766000.

[84] Shapira G. How subjective beliefs about HIV infection affect life-cycle fertility: Evidence from rural Malawi, 2013. World Bank Policy Research Working Paper #6343.

[85] Bignami-Van Assche S, Fleming P, Van Assche A, van de Ruit C, Anglewicz P. HIV/AIDS and time allocation in rural Malawi. *Demographic Research* 2011; **24(27)**:671–708.

[86] Baranov V, Kohler HP. The impact of AIDS treatment on savings and human capital investment in Malawi, 2013. Paper presented at the Annual Meeting of the Population Association of America, New Orleans, April 11–13, 2013,. http://paa2013.princeton.edu/abstracts/131999.

[87] Cordero-Coma J, Breen R. HIV prevention and social desirability: Husband-wife discrepancies in reports of condom use. *Journal of Marriage and Family* 2012; **74(3)**:601–613.

[88] Muula AS. Marriage, not religion, is associated with hiv infection among women in rural Malawi. *AIDS and Behavior* 2010; **14(1)**:125–131.

[89] Clark S, Poulin M, Kohler HP. Marital aspirations, sexual behaviors, and HIV/AIDS in rural Malawi. *Journal of Marriage and the Family* 2009; **71(2)**:396–416.

[90] Poulin MJ. Sex, money, and premarital relationships in southern Malawi.

Social Science and Medicine 2007; **65(11)**:2383–2393.

[91] Anglewicz P, Clark S. The effect of marriage and HIV risks on condom use acceptability in rural Malawi. *Social Science & Medicine* 2013; p. forthcoming.

[92] Swidler A, Watkins SC. Ties of dependence: AIDS and transactional sex in rural Malawi. *Studies in Family Planning* 2007; **38(3)**:147–162.

[93] Reniers G, Tfaily R. Polygyny and HIV in Malawi. *Demographic Research* 2008; **19(53)**:1811–1830.

[94] Reniers G, Watkins SC. Polygyny and the spread of HIV in sub-Saharan Africa: a case of benign concurrency. *AIDS* 2010; **24(2)**:299–308. Ahead of print publication available online.

[95] Freeman E, Anglewicz P. HIV prevalence and sexual behavior at older ages in rural Malawi. *International Journal of STD & AIDS* 2012; **23(7)**:490–496.

[96] Freeman E, Coast E. Sex in older age in rural Malawi. *Ageing and Society* 2013; Preprint Publication March 2013.

[97] Poulin MJ. The sexual and social relations of youth in rural Malawi: Strategies for AIDS prevention, 2006. Ph.D. Dissertation, Department of Sociology, Boston University, Boston, MA.

[98] Tavory I, Swidler A. Condom semiotics: Meaning and condom use in rural Malawi. *American Sociological Review* 2009; **74(2)**:171–189.

[99] Behrman JR, Kohler HP. Sexual transmission of HIV. In *Rethink HIV: Smarter Ways to Invest in Ending HIV in Sub-Saharan Africa*, editor Lomborg B, Cambridge, MA: Cambridge University Press, chap. 1, pp. 11–60. 2012. An earlier working paper version of this paper is available online at http://www.rethinkHIV.com.

[100] Chimbiri A. The condom is an intruder in marriage: Evidence from rural Malawi. *Social Science and Medicine* 2007; **64(5)**:1102–1115.

[101] Kaler A. AIDS-talk in everyday life: HIV/AIDS in men's informal conversations in southern Malawi. *Social Science and Medicine* 2004; **59(2)**:285–297.

[102] Reniers G. Marital strategies for regulating exposure to HIV. *Demography* 2008; **45(2)**:417–438.

[103] Boileau C, Clark S, Poulin M, Bignami-Van Assche S, Reniers G, Watkins SC, *et al.* Sexual and marital trajectories and HIV infection among ever-married women in rural Malawi. *Sexually Transmitted Infections* 2009; **85(Supplement 1)**:i27–i33.

[104] Baird SJ, Garfein RS, McIntosh CT, Özler B. Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: A cluster randomised trial. *Lancet* 2012; **379(9823)**:1320–1329.

[105] de Walque D, Dow WH, Nathan R, Abdul R, Abilahi F, Gong E, *et al.* Incentivising safe sex: a randomised trial of conditional cash transfers for HIV and sexually transmitted infection prevention in rural tanzania. *BMJ Open* 2012; **2(1)**.

[106] Dugger CW. African studies give women hope in H.I.V. fight, 2010. *New York Times*, July 19, 2010. http://www.nytimes.com/2010/07/20/world/africa/

20safrica.html.

[107] Mensch BS, Hewett PC, Gregory R, Helleringer S. Sexual behavior and STI/HIV status among adolescents in rural Malawi: An evaluation of the effect of interview mode on reporting. *Studies in Family Planning* 2008; **39(4)**:321–334.

[108] Poulin M. Reporting on first sexual experience: The importance of interviewer-respondent interaction. *Demographic Research* 2010; **22(11)**:237–288.

[109] Helleringer S, Kohler HP, Kalilani-Phiri L, Mkandawire J, Armbruster B. The reliability of sexual partnership histories: implications for the measurement of partnership concurrency during surveys. *AIDS* 2011; **25(254)**:503–511.

[110] Payne C, Mkandawire J, Kohler HP. Disability transitions and health expectancies among adults 45 years and older mature in Malawi: A cohort modeling approach. *PLOS Medicine* 2013; **10(5)**:e1001435.

[111] Doctor HV. Adult mortality in rural Malawi. *Southern African Journal of Demography* 2004; **9(1)**:49–66.

[112] Chin B. Essays on health, mortality, and intergenerational transfers in rural malawi, 2011. Ph.D. Dissertation, Graduate Group in Demography, University of Pennsylvania, Philadelphia, PA.

[113] Gerland P, Kaphuka J, Mandere G, Misiri H, Fleming P. Adult mortality estimations from cohort and census/survey data: A comparison of direct and indirect methods in rural Malawi, 2007. http://uaps2007.princeton.edu.

[114] Kohler IV, Kalilani-Phiri L, Blank M, Hennessy M, Kohler HP. Mental health among the elderly in Sub-Saharan Africa—a neglected health dimension, 2012. Paper presented at the 2012 annual meetings of the Population Assocation of America, San Francisco, CA, May 3–5, 2012. http://paa2012.princeton.edu.

[115] UN Population Division. World Population Prospects, the 2010 revision: Standard (median) forecasts, 2010. United Nations, Department of Economic and Social Affairs, Population Division. http://esa.un.org/unpd/wpp/.

[116] Kandzandira J. Rural ageing and livelihood challenges: A case study of the elderly in rural areas of Zomba district, Malawi, 2007. University of Malawi, Center fo Social Research: Report prepared for the Food and Agriculture Organization (FAO), Gender and Population Division.

[117] FAO and BSF. Population ageing in Malawi: Understanding challenges, responding to opportunities, 2008. FAO, Rome: Proceedings of the meeting organized jointly by the Bingu Silvergrey Foundation for the Elderly (BSF) and the Food and Agriculture Organization (FAO) of the UN in Lilongwe, Malawi, 28-29 November 2007. ftp://ftp.fao.org/docrep/fao/010/ai581e/ai581e01.pdf.

[118] Scholten F, Mugisha J, Seeley J, Kinyanda E, Nakubukwa S, Kowal P, *et al.* Health and functional status among older people with HIV/AIDS in Uganda. *BMC Public Health* 2011; **11(1)**:886.

[119] Abegunde DO, Mathers CD, Adam T, Ortegon M, Strong K. The burden and costs of chronic diseases in low-income and middle-income countries. *Lancet* 2007; **370(9603)**:1929–1938.

[120] Chin B. Income, health, and well-being in rural Malawi. *Demographic Research* 2010; **23(35)**:997–1030.

[121] Alderman H, Behrman JR, Kohler HP, Maluccio J, Watkins SC. Attrition in longitudinal household survey data: Some tests for three developing country samples. *Demographic Research* 2001; **5(4)**:79–123.

[122] Fitzgerald J, Gottschalk P, Moffitt R. An analysis of sample attrition in panel data – the Michigan Panel Study of Income Dynamics. *Journal of Human Resources* 1998; **33(2)**:251–299.

[123] Falaris EM. The effect of survey attrition in longitudinal surveys: Evidence from Peru, Côte d'Ivoire and Vietnam. *Journal of Development Economics* 2003; **70(1)**:133–157.

APPENDIX

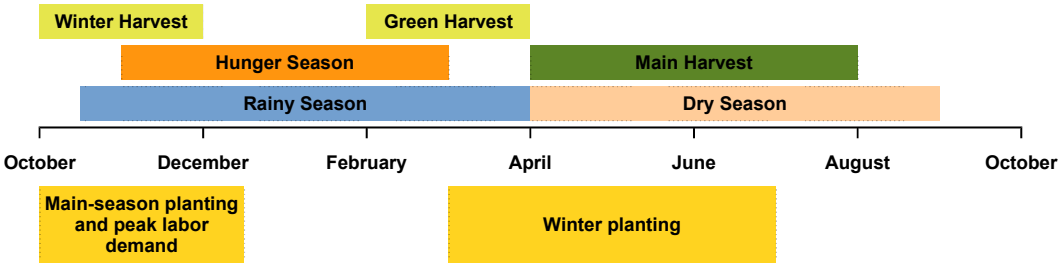
This Appendix provides additional information about the MLSFH study areas and study contexts, the sampling for the MLSFH and the refreshment of the MLSFH sample over time, and the procedures for HIV testing and counseling that were implemented as part of the MLSFH. This Appendix also provides comparisons of the MLSFH study populations with nationally representative datasets, analyses of attrition in the MLSFH sample, and discussions of some specific features of the MLSFH data that have been widely used across many MLSFH-based papers. Some of the information provided in this Appendix was previously published, but often scattered across multiple publications. It is integrated and combined here for the first time.

Content:

A1. MLSFH study areas and context

The MLSFH is based in three districts in rural Malawi that have been the study sites since 1998: Rumphi in the north, Mchinji in the center, and Balaka in the south (Figure 1). In all of these three regions, the primary source of livelihood for MLSFH respondents is subsistence agriculture, augmented with smallholder cash crops, small-scale trade of agricultural products and other goods, and casual labor.

Figure A1: Seasonality of harvest and labor demand in Malawi



Source: Adapted from USAID & FEWS NET [124]

Transportation networks are relatively rudimentary with paved primary roads and generally unpaved secondary roads, which may be impassable during the rainy season. Communication infrastructure has importantly changed during the period observed by the MLSFH. Cell phones were absent when the MLSFH was initiated in 1998, but have spread rapidly since, and 37% of MLSFH respondents owned a cell in 2010.

Marriage is nearly universal in rural Malawi, with more than 96% of women having ever married by age 25–29, and more than 95% of men having ever married by age 30–34.[11] While the broad demographic, socioeconomic and epidemiological conditions are fairly similar across the three MLSFH study regions, and also across other parts of rural Malawi, some noteworthy differences across the MLSFH regions include the following. Rumphi District, located in the northern region of the country, follows the patrilineal system of kinship and lineage where residence is primarily patrilocal, inheritance is traced through sons, and parents of a groom pay bridewealth. The northern district, inhabited primarily by Tumbukas, is predominantly Protestant. Mchinji District, located in the central region, follows a less rigid matrilineal system whereby residence may be matrilocal or patrilocal or neither (among MLSFH participants in Mchinji, about 75% follow a patriolocal tradition). The Center is primarily inhabited by Chewas, with almost equal proportions of Catholics and Protestants. Balaka District, which is located in the southern region, is primarily inhabited by Lomwes and Yaos and has the highest proportion of Muslims. The region follows a matrilineal system of kinship and lineage system where residence is ideally matrilocal, although it is not uncommon for wives to live at least some period of time in their husband's village. The Balaka region also exhibits a lower age of sexual debut and larger numbers of lifetime sexual partners than the other MLSFH study regions, and residents tend to be less educated and poorer than those living in the north, leading to higher levels of migration. HIV/AIDS prevalence in the southern region is significantly higher than in the northern and central region.

Work effort in Malawi is highly seasonal (Figure A1).[124] The peak labor demand season occurs during the rainy season, which coincides with the hunger season, a

Table A1: Size, age range (25th and 75th percentile) and gender distribution of the MLSFH study population 1998–2012

MLSFH	*N*	Age (percentile) 25th	75th	Prop female	*N* from previous round	% re-interviewed at next round
1998	2,597	25	42	0.59	–	75.0
2001	2,546	28	44	0.62	1,949	73.5
2004	3,261	22	43	0.55	1,872	78.1
2006	3,431	24	44	0.55	2,546	76.8
2008	4,036	27	53	0.58	2,635	74.8
2010	3,798	28	54	0.59	3,020	90.3
2012	1,266	50	67	0.57	1,266	–

time when the poorest households many households may be reduced to one meal of watery porridge a day. An ethnographer working in village in southern Malawi wrote that towards the end of the hunger season, "*farmers' eyes grow increasingly hollow, their faces shrunken, and their bodies frail. [...] Activities are reduced to a minimum; villagers lie listlessly in the shade of their huts, waiting for the hours to pass and the maize to mature*."[125] Because the hunger season is also the rainy season, it is the height of the malaria season, when people are more likely to be ill.

A2. MLSFH sampling methods and related relevant data collection procedures

A2.1. MLSFH-sampling

The initial MLSFH sample was established in 1998. The MLSFH study sample was augmented in 2004 by adding the MLSFH Adolescent Sample, and in 2008 by adding the MLSFH Parent Sample. In addition, ongoing additions occurred as a result of enrolling new spouses of respondents. Table A1 reports the size, age range (25th and 75th percentile) and gender distribution of the MLSFH study population during 1998–2012 (see also Figure 2). The details of the MLSFH sampling procedures are described below.

A2.1.a. Initial MLSFH Sample: The original 1998 MLSFH target sample was 500 ever-married women age 15–49 in each district, plus their husbands (for additional information, see http://malawi.pop.upenn.edu/malawi-documentation-sampling). The sampling strategy adopted for the three districts differed in order to permit comparison with earlier surveys. In Mchinji and Rumphi districts the sample was designed to cover Census Enumeration Areas (CEAs) included in the 1988 Traditional Methods of Child Spacing in Malawi (TMCSM) survey. However, since the TMCSM sampled women regardless of their marital status, the CEAs included in the TMCSM survey had fewer ever-married women than the MLSFH target sample of 500 women in each district. Three neighboring CEAs covered by the 1988 survey were thus added to the MLSFH Round 1 sample. In each district a cluster sampling strategy was used in all villages in the selected CEAs. Household lists of those nor-

mally resident in those villages were compiled during the week prior to fieldwork, and a sample of eligible women was then randomly selected. Since villages varied in size, sampling fractions were inversely proportional to village populations, such that a higher proportion of eligible women in the smaller villages was sampled. In Balaka district, a somewhat different procedure was followed to allow the evaluation of a Community Based Distribution (CBD) initiative that was conducted in this area at the time, following an earlier baseline survey conducted by the German aid agency GTZ (now GIZ) with 1098 women and men in 1993. A random subset of 4/7 of the CBD villages and 5/11 of the non-CBD villages from this study were selected as MLSFH study villages. A random 1 in 4 sample of women of reproductive age (15- 49) and their husbands was then drawn from these villages to yield a target sample of 500 women and their husbands. To further increase the number of MLSFH respondents who participated in the 1993 GTZ survey, an additional 260 women and 125 men were randomly drawn from the GTZ sampling lists (divided equally between the CBD and non-CBD areas) and enrolled in the MLSFH. In total, across all three regions, the MLSFH Round 1 in 1998 enrolled a sample of slightly more than 1,500 ever-married women aged 15–49 and close to 1,100 of their spouses residing in about 120 study villages (Table A2 and Figure 2).

The sampling strategy was not designed to be representative of the national population of rural Malawi. As Table A3 shows, however, our sample characteristics closely match the characteristics of the rural population of the 1996 Malawi Demographic and Healtersh Survey (MDHS). We do not expect perfect alignment with the rural MDHS sample, since the MDHS clusters are not identical to a MDICP village; moreover, the MDHS includes small trading centers, which the MLSFH does not, thus making the MLSFH less urban.

A2.1.b. MLSFH Respondent follow-up, migration and vital status: The MLSFH returned to the study areas in 2001, 2004, 2006, 2008 and 2010 to reinterview the MLSFH study population. For this purpose, the MLSFH maintained a respondent database that contained previously collected identifying information for each respondent (respondents name, compound name, village name and GPS coordinates, etc.). Using this existing identifying information, MLSFH interviewers attempted to contact and reinterview MLSFH participants in each of the follow-up years. If MLSFH participants were absent at the first interviewer visit, up to two additional follow-up visits were made. Except for a migration follow-up study in 2007, and one that is currently (2013) in the field, MLSFH respondents were not followed if they had migrated outside of the MLSFH study villages. However, they remained in the MLSFH sampling frame, and were interviewed at subsequent MLSFH waves if they returned to a MLSFH study village (as was common since a significant amount of migration was temporary). On average, the MLSFH succeeded during 2001–1998 in re-interviewing between about 73–78% of the respondents interviewed at the previous MLSFH wave (Figure 2). When a MLSFH participant could not be found and contacted for a MLSFH follow-up interview, the MLSFH conducted a short interview with family members and/or neighbors to obtain essential infor-

Table A2: Summary statistics for the MLSFH Round 1 (1998) study population

	Females mean (sd)	**Males** mean (sd)	**Total** mean (sd)
# of observations	1,532	1,065	2,597
Respondent's age (in 1998)	32.45	38.78	35.04
	(13.16)	(12.21)	(13.15)
Age group (in 1998)			
< 20	0.123	0.008	0.076
20–29	0.381	0.271	0.336
30–39	0.270	0.293	0.279
40–49	0.136	0.248	0.182
50–59	0.039	0.108	0.067
60–69	0.022	0.060	0.038
70+	0.029	0.012	0.022
Marital status			
Married	0.870	0.990	0.919
Separate	0.026	0.005	0.017
Divorced	0.076	0.006	0.047
Widowed	0.029	0.000	0.017
Children ever born	4.21	5.17	4.60
	(3.00)	(4.10)	(3.52)
Schooling attainment			
No formal schooling	0.358	0.218	0.301
Primary schooling	0.589	0.634	0.607
Secondary or higher	0.053	0.148	0.092
Religion			
Christian	0.764	0.745	0.756
Muslim	0.217	0.225	0.220
Other/none	0.020	0.030	0.024
Wealth indicator: House has metal roof	0.078	0.077	0.078
Region of residence			
Central	0.353	0.358	0.355
South	0.332	0.335	0.333
North	0.315	0.307	0.312
Worried about getting AIDS			
Not worried at all	0.167	0.258	0.204
Worried a little	0.208	0.190	0.200
Worried a lot	0.625	0.553	0.596

mation about the vital status and migration of the MLSFH respondent. Based on this information, the respondent's status in the MLSFH was recorded as classified as *dead*, *migrated*, *refused*, *hospitalized*, *temporarily absent*, *other*, and *unknown*. Conditional on successfully contacting a MLSFH respondents, refusals to participation in the MLSFH have been very low across all MLSFH waves ($<$ 3% up to 2008, and $<$ 5% in 2010).

Table A3: Comparison between 1998 MLSFH and 1996 Malawi Demographic and Health Survey (MDHS)

	% ever-married women aged 15–49	
	MLSFH (1998)	MDHS (1996)
Age: *15–19*	8.4	10.3
20–24	22.3	21.2
25–29	21.8	16.5
30–34	15.1	16.0
35–39	15.4	12.2
40–44	9.7	14.5
45–49	7.2	9.2
Schooling: *None*	33.5	48.1
Primary	60.8	50.4
Secondary or higher	5.7	1.5
Number of surviving children: *0*	3.8	12.5
1	21.6	18.5
2	18.2	17.4
3	15.4	14.5
4	13.0	12.6
5	11.3	7.8
6	7.5	7.7
7+	9.2	9.0
Owns a radio	57.4	43.1
Ever used contraception	32.0	38.2
Currently using contraception	22.5	16.4
Observations (*N*)	1478	1123

Notes: The Malawi Demographic and Health Survey (MDHS) 1996 [126] was a nationally representative sample survey conducted in 1996 and designed to provide estimates of family planning and health indicators for the three administrative regions of the country, urban and rural areas, and Malawi as a whole. For the comparison with the 1998 MLSFH, the MDHS is restricted to rural subsample.
Source: Watkins *et al.* [1]

A2.1.c. MLSFH Sample Additions: Additions to the MLSFH have occurred primarily through three mechanisms: new spouses, the 2004 adolescent sample, and the 2008 parent sample. We discuss these three mechanisms in turn. (*i*) ***New spouses:*** The initial MLSFH sample in 1998 included 1,532 ever-married women aged 15–49 and their spouses. In the 2001 round of data collection, the MLSFH attempted to re-interview all of these initial MLSFH respondents and their current spouses; that is, if a MLSFH respondent divorced and remarried, or in the case of polygamous men, added an additional wife, the MLSFH added the current wife (all current wives) of the initial MLSFH participants. However, spouses who were not part of the initial MLSFH sample were not followed and retained in the 2001 MLSFH

if they divorced or their spouses died. Starting with the 2004 MLSFH, the study retained all MLSFH study participants; that is, from 2004 onward, once an individual was interviewed for the MLSFH once, for instance after being enrolled as a new spouse, the MLSFH made an attempt to re-interview the respondent at all subsequent waves. (*ii*) ***2004 Adolescent Sample:*** In 2004, to compensate for the aging of the initial MLSFH sample and the underrepresentation of unmarried individuals at adolescent and young adult ages, the MLSFH added an adolescent sample in 2004 ($N = 998$). For this purpose, two household rosters were collected in each sampled community as part of the 2004 MLSFH data collection. The first was collected from all households in the sampled villages—that is, MLSFH and non-MLSFH households—during a household listing interview in which all members of all households in the MLSFH were enumerated along with basic demographic characteristics. The second household roster was incorporated into the primary MLSFH survey instrument administered to all female MLSFH participants to enumerate all eligible adolescents who were part of existing MLSFH households. To allow for intergenerational analyses, all adolescents aged 15–25 listed as members of the existing MLSFH households and residing in the MLSFH study villages were enrolled into the MLSFH adolescent sample, constituting about 1/3 of the adolescent sample. The remaining members of the MLSFH adolescent sample were selected from the household listing conducted for non-MLSFH households using an age-stratified sampling strategy that adjusted for the differential ages at marriage between gender and MLSFH study regions (for additional information, see http://malawi.pop.upenn.edu/malawi-documentation-sampling). (*iii*) ***2008 MLSFH Parent Sample:*** To increase the suitability of the MLSFH to study intergenerational aspects and the health of older individuals in Malawi, a parent sample was added to the MLSFH in 2008. This new sample of parents of MLSFH respondents was drawn from family listings from MLSFH respondents in 2006 (because of the respondents' young age, parents of MLSFH respondents in the 2004 adolescent sample were not included). All living biological parents who resided in the same village as the respondent were included in the 2008 MLSFH new sample of parents. Based on this approach, approximately 800 parents of MLSFH respondents living in the MLSFH study villages were added to the 2008 MLSFH sample ($N = 549$). As a result of adding the MLSFH parent sample, the age range covered by the MLSFH was substantially extended in 2008 (Table 3). Moreover, since a parent enrolled through this process could be the parent of multiple MLSFH respondents (some of the MLSFH respondents are siblings), a manual data cleaning was used to identify all duplicate parent nominations and correct parent-child linkages were established, which as a side effect, also enables us to identify sibling MLSFH respondents in the data.

Among approximately 3,800 respondents interviewed in the 2010 MLSFH, 44.1% were from the original MLSFH sample drawn in 1998, 19.5% were from the 2004 adolescent sample, 12.5% from the 2008 parent sample, and the remainder (23.9%) was new spouses that have been added during 2001–2010.

A2.2. MLSFH 2007 Migration Follow-up

The MLSFH 2007 migration follow-up aimed to collect data on respondents who were interviewed by the MLSFH prior to the 2006 waves, but could to be located at the 2006 round of the MLSFH.[23] Specifically, the 2006 MLSFH interviewed approximately 70% of the target sample members. Absence due to migration (as reported by family members or neighbors) was the most frequent reason why individuals were not interviewed: approximately 18% of the 2006 MLSFH sample moved sometime between the first wave in 1998 and the fourth wave in 2006. Of these migrants, 11% moved outside of Malawi and no attempts were made to reach them.

The target sample for the migration study consisted of 718 men and women who had been interviewed at least once by the MLSFH prior to 2006 and who had subsequently relocated permanently within Malawi (to an urban or rural area). Of the 718 migrants in the target sample, the 2007 migration study team traced approximately 60% and interviewed 56% (N = 400) (the remaining 4% were dead, were hospitalized, or refused to be interviewed). Of respondents who were not traced by the migration team, approximately 28% were not found at the location described in their migration autopsy. Often, the family members or neighbors could provide only a general location, which is not surprising because street names and house numbers are rare even in urban areas of Malawi. When information was specific, it was occasionally incorrect. The default was to search by name, which was problematic because migrants sometimes changed their name after migration and were therefore not known at their place of destination. Background information for the 718 migrants that compose the MLSFH migration study target sample and the 400 migrants found by the migration study team in 2007 are shown in Table A4. Differences in migration patterns reflect differences in migration by region, sex, and age. In the target sample, more men from the matrilocal South migrated (46%) than men from the other two regions, and more women from the patrilocal North migrated (40%) than women from the Center or South. Although either the husband or the wife may move at marriage, women typically marry at younger ages than men.[17] The age and sex distribution of the migrants who were located is roughly similar to the age distribution of the migration target sample. Of the 718 MLSFH respondents who moved within Malawi, 20% (146 migrants) moved to an urban area. The most common urban destination was Lilongwe, the centrally located nation's capital, where approximately 31% of all rural-urban MLSFH migrants were living. For MLSFH respondents, a slightly larger percentage of male migrants moved to an urban area than female migrants (23% for men and 19% for women). Statistical tests show no significant differences in urban residence between migrants who were located by the migration team and those who were not.

Approximately 31% of migrants moved for marriage-related reasons (divorce, widowhood, or new marriage), compared with 39% who moved for work. Reasons for migration in Table A4 were asked directly of migrants interviewed by the migration study team in 2007. Women were more likely to move for marriage than

Table A4: Background characteristics for migration study respondents: Target sample and respondents interviewed by migration team

	Target Sample			Sample Interviewed		
Characteristics	Female	Male	Total	Female	Male	Total
Age Distribution (%)						
10–19	12.4	8.1	10.4	12.6	10.0	11.5
20–29	31.4	22.2	27.1	34.6	22.4	29.4
30–39	31.1	29.2	30.2	29.9	30.0	29.9
40–49	18.8	21.9	20.2	17.3	19.4	18.2
50–59	5.0	14.1	9.2	4.3	13.5	8.2
60–69	1.4	4.6	2.9	1.3	4.7	2.7
Region of Origin (%)						
Central	30.1	25.9	28.3	28.6	27.7	28.2
South	29.9	46.2	37.0	27.7	41.8	33.7
North	40.0	27.9	34.7	43.7	30.5	38.1
Reason for Migration (%)						
Marriage-related	41.0	17.1	30.6	49.3	16.5	35.4
Work-related	29.4	51.3	39.0	20.4	45.3	27.9
Other	29.6	31.6	30.4	30.3	38.2	36.7
Rural-Urban Migration (%)	18.5	22.8	20.4	19.1	21.2	20.0
N	402	316	718	231	171	402

Notes: Reasons for migration for migrants not interviewed are from the migration autopsies, which were administered to relatives or friends of the migrant.
Source: Anglewicz [23]

for work, and men were more likely to move for work than for marriage. The "other" category groups all reasons for migration that did not fit into the above categories—for example, to attend school, to visit a relative, to follow parents or relatives to a new location, and because of imprisonment.

An update of the 2007 MLSFH Migration Follow-up is being conducted during 2013, focusing on MLSFH respondents who were interviewed at least once during 2004–2008, but not in 2010 (the most recent MLSFH round covering all MLSFH respondents). According to the migrant tracking information collected in 2013, migration patterns among these ever-interviewed MLSFH respondents are highly clustered ($N \approx 1,150$): 64% had moved within the same district (often related to marriage/divorce); 12% had moved to one of Malawi's four largest cities (Blantyre, Lilongwe, Mzuzu, Zomba); and 24% had moved to another rural/peri-urban area. MLSFH Survey data, using a study instrument similar to the 2010 MLSFH Questionnaire and augmented with additional questions related to migration, and updated HIV status information will be collected for these MLSFH migrants in 2013 (with data collection ongoing at the time of this writing).

A2.3. Longitudinal identification and linkage of MLSFH respondents

Ensuring a correct longitudinal identification of MLSFH respondents was challenging in rural Malawi due to the absence of well-defined addresses, frequent mobility

of individuals, and relatively common marriage/divorce that often results in migration.[127] The MLSFH also encountered community members to who claimed to be MLSFH study participants ("imposters"), even though they were not (often a family member was).

To maintain a high quality of the MLSFH longitudinal linkages and overcome these challenges, the MLSFH employed several steps in its fieldwork and data collection, including: (1) relying on fieldwork personnel who have been working with MLSFH consistently for several years to identify and address problems in the field during data collection, (2) employing our knowledge of the local setting, including identifying villages where challenges are greatest, and becoming aware of these challenges in advance, and (3) using our longitudinal data during data collection, in which background characteristics (such as spouse's name, level of education, birthplace, father's name) from current MLSFH data collection is compared with the same information from previous waves to ensure that the correct respondent has been interviewed, and correcting immediately if not. To provide these identifying data during fieldwork, the MLSFH maintains a MLSFH Respondent Database that contains previously collected identifying information for each respondent (respondent's name and ID, previously taken pictures of respondents (if available), GPS coordinates of previous residence of respondents (since 2004), name of respondent's parents and current husband, selected respondent characteristics (age, sex, education), and name of village headsman). During MLSFH fieldwork and data collection, daily interviewer lists were created for the interviewers containing contact information of respondents to be interviewed on a particular day. Interviewers used this information to locate respondents, and verify the identity of the respondent using the identifying information provided from the respondent database (including the printed picture of the respondent). Interviewers recorded the interview outcomes (interview completed, refused, respondent not present, respondent moved) on MLSFH Survey Log Sheets that were provided from the respondent database for each day. At the end of each day, the respondent database was updated with a log of the interview outcomes, and if applicable, the respondent database was updated in case that there have been any changes in a respondent's identifying or contact information (e.g., respondent has moved). The picture of each respondent that was taken as part of the 2006, 2008 and 2012 MLSFH was uploaded to the respondent database to replace any previously taken picture, and the questionnaire cover sheet (containing respondents name and other contact information) was removed from the remaining questionnaire that contains merely the respondent ID number (and no other identifying information). Using the above process, the MLSFH has been able to maintain a relatively high retention rate of respondents across waves (Figure 2), and incorrect identification of MLSFH study participants over time are rare.

A2.4. Common and distinctive features of the MLSFH fieldwork during 1998–2012

Most of the MLSFH data collections during MLSFH rounds in 1998, 2001, 2004, 2006, 2008, 2010 and 2012 were conducted during May–August of the respective

years. During this period, which coincides with the harvest seasons, individuals eat three meals a day and work effort is relatively low. MLSFH time use data show, for example, that 48.9% of MLSFH respondents performed agricultural labor during the period of the MLSFH data collection, 22.1% performed non-agricultural labor, and 69.9% performed domestic labor.[85]

Since 2006, the MLSFH has collaborated with Invest in Knowledge (IKI, www.http://www.investinknowledge.org), a Malawi NGO that has founded by members of the MLSFH research team and specializes in research capacity building and data collection. For the survey data collections, interviewers were recruited for each MLSFH data collection in the each of the MLSFH study regions, often recruiting interviewers who have previously worked for the MLSFH or related IKI projects. During each round, interviewers received extensive training in survey data collection and the specific MLSFH survey instruments prior to the fieldwork. The MLSFH obtained approvals from the District Commissioner, District Health Officer, local police, Traditional Authorities and village headmen in each of the MLSFH study villages prior to any data collection in a village. Once a MLSFH study participant was located using the identifying information in the MLSFH Respondent Database, informed consent was obtained and the survey was conducted in the respondents' home using paper-and-pencil techniques. To ensure the confidentiality of the data, interviewers were instructed to select a location for the interviews that guarantees the privacy of the information provided by the respondent. At the end of each fieldwork day, each survey was reviewed and checked for inconsistencies and/or omissions, and interviewers returned to MLSFH participants when necessary to obtain missing information.

For HIV testing and counseling (and the biomarker collection in 2009), the MLSFH recruited and trained Ministry of Health-certified HTC counselors (and to ensure the confidentiality of HTC, only counselors from outside the MLSFH study villages were recruited). HTC was conducted, usually after the MLSFH survey, at the respondent's home (see Appendix A3 for additional detail).

In addition to the evolution of the topics covered by the MLSFH survey over time (Table 4), some distinctive features of some specific MLSFH rounds are noteworthy:

A2.4.a. MLSFH 3 (2004): The 2004 MLSFH refreshed the MLSFH study population at younger ages by enrolling the MLSFH Adolescent Sample (Appendix A2.1.c). For this purpose, prior to the main MLSFH survey, a household listing was conducted in all MLSFH study villages to obtain information about the resident population and the members of all village households in each of the MLSFH study villages. The sampling of the MLSFH Adolescent Sample is described above (Appendix A2.1.c). The 2004 MLSFH is also noteworthy because it was the first MLSFH round that collected biomarkers for HIV and other sexually transmitted diseases (Appendix A3.1). The 2004 MLSFH also implemented a randomized experiment that offered financial incentives to respondents who decided to learn their HIV test results (Appendix A3.1).

A2.4.b. MLSFH 4 (2006): As in 2004, the 2006 MLSFH included both survey data collection and testing for HIV (Appendix A3.2). To accommodate a substantially expanded MLSFH questionnaire (see Table 4 for a summary of MLSFH 4 survey measurements), the survey team of the MLSFH 4 (2006) was split up into a household-listing team that located and identified respondents and then collected the newly introduced extensive household/family rosters that asked respondents about their resident and non-resident household and family members along with information on their health and transfer/exchange relations (Appendix A6.2). A MLSFH survey team then followed up within a few days of the household listing team to collect additional survey data from each MLSFH respondent that was successfully located and identified by the household listing team. Finally, a team of HTC counselors visited all MLSFH respondents with a completed household listing to conduct HIV testing and counseling (Appendix A3.2). In contrast to the 2004 HTC, the MLSFH offered both individual and couple HTC as part of the 2006 MLSFH HTC. Subsequent to the 2006 MLSFH data collection and HTC, the MLSFH implemented an experimental design that offered financial incentives to respondents who maintained their HIV status during 2006–07 (Appendix A6.6). As part of this study, the MLSFH also collected "sexual diaries" that provide detailed day-to-day data on sexual behaviors for four 10-day periods during 2006–07. In addition, a 2007 MLSFH Migration Follow-up study was conducted to trace, survey and HIV test all ever-interviewed MLSFH respondents not interviewed during MLSFH 4 (2006) due to migration and/or temporary absence (Appendix A2.2).

A2.4.c. MLSFH 5 (2008): The 2008 MLSFH expanded the MLSFH study sample at older ages by adding the MLSFH Parent Sample (Appendix A2.1.c). The complete MLSFH study population was then contacted by a single survey team (see Table 4 for a summary of MLSFH 5 survey measurements), which was followed by the HTC team for HIV testing.

A2.4.d. MLSFH 6 (2010): The 2010 MLSFH was largely identical to the previous 2008 MLSFH round (see Table 4 for a summary of MLSFH 6 survey measurements), except that no HIV testing was conducted given the observed low HIV incidence in the MLSFH study population and the already comprehensive MLSFH data on the HIV status of study population.

A2.4.e. MLSFH 7 (2012): To develop a stronger aging-related MLSFH research agenda, the research team conducted in 2012 a MLSFH mature adults survey on mental health and well-being. This survey focused on mature adults, that is, MLSFH respondents aged 45 and older, who had previously been interviewed in the 2008 and 2010 MLSFH. A total of 1,266 MLSFH mature adults were interviewed (Figure 2) using a questionnaire that continued key elements of the 2008 and 2010 data collections (Table 4) and newly added detailed measures of mental health, cognitive function, and physical performance (see Appendix A6.8) for additional information).

A3. MLSFH HIV testing and counseling (HTC)

HIV testing was conducted as part of the MLSFH in 2004, 2006, 2008 and 2012 using HTC counselors certified by the Malawi Ministry of Health. The HTC procedures followed guidelines given by the Malawi Ministry of Health and the WHO,[128,129] and written consent was obtained from all HTC participants prior to HTC. In 2004, HIV testing was conducted through the collection of oral swab specimens that were analyzed in a central lab in Lilongwe using ELISA and confirmatory Western blot tests. MLSFH HIV testing was conducted using finger-prick rapid tests from 2006 onward. The different HIV testing and counseling protocols, and the 2004 experimental design that offered randomized financial incentives for individuals to learn their HIV status, are described below. To ensure the confidentiality of HTC and the HIV test results, the MLSFH implemented several privacy and data protection measures, including the use of separate IDs and data file for survey data and HTC-related data, non-local HTC counselors who had never lived nor had close relatives or friends in the MLSFH study villages, a secure storage of consent forms, the separation of identifying information from all study materials containing HIV test results and related information, and adequate protections to ensure the privacy of the in-home HTC sessions.

As part of the MLSFH HTC procedures, all HIV tests were preceded and followed by a counseling session. The pre-test counseling emphasized privacy and informed consent. The respondent chose the venue for the counseling that he/she considered most private; in order to provide a foundation for informed consent, counselors explained the procedures to be followed during testing, as well as the implications of learning one's own HIV status. Post-test counseling emphasized the results of the test, the window period and importance of retesting, and appropriate behavior for the future. Starting in 2008, when antiretroviral treatment (ART) had become available in the MLSFH study regions, HIV-positive received referrals to district hospitals for confirmatory testing and determining of eligibility for ART.

A3.1. 2004 MLSFH HTC and MLSFH experimental design offering financial incentives for learning HIV status in 2004

The 2004 MLSFH asked all respondents—the vast majority (82%) of whom had not previously participated in HCT—to provide a biomarker sample for a lab-based HIV test as part of a randomized experiment to study the determinants of HCT uptake.[21,32,130,131] Rapid HIV testing and counseling had not yet been approved for use in Malawi at this time, and therefore an approach combining home-based collection of biomarkers, centralized lab-based testing of the specimen, and subsequent dissemination of HIV test results in MLSFH-established local HTC clinics was chosen. Specifically, between May and August of 2004, nurses from outside each area offered respondents free tests in their homes for HIV and three other sexually transmitted infections (STIs) (gonorrhea, chlamydia, and trichomoniasis). At the time that the HIV tests were offered, respondents were given pre-test counseling about HIV prevention strategies. Samples were taken through OraSure™

oral swabs (OraSure Technologies, USA) to test for HIV and through urine samples (for men) or self-administered vaginal swabs (for women) to test for other STIs. The oral swabs were tested for HIV in a central lab in Lilongwe using ELISA and confirmatory Western blot tests. Across the three districts, 2,894 of the 3,185 respondents who were offered accepted an HIV test (91%). The prevalence of STIs in 2004 was very low (3.0% positive and 3.1% inhibitory/inconclusive for gonorrhea, .25% positive and 3.3% inhibitory/inconclusive for chlamydia, and 2.4% positive for trichomoniasis).[21] HIV prevalence was 6.5% (plus .5% inconclusive results that may indicate a recent HIV infection), with significant regional variation (8.2% in Balaka, 6.6% in Mchinji and 4.7% in Rumphi) and gender difference (7.1% for women, 5.7% for men). After taking the HIV test samples, nurses gave each respondent vouchers redeemable upon obtaining either HIV or STI results. Voucher amounts were randomized by letting each respondent draw a token out of a bag indicating a monetary amount. In Mchinji and Balaka each respondent received two vouchers, one for obtaining HIV results, and one for obtaining STI results. In Rumphi, respondents received only one voucher redeemable by returning for either HIV or STI results. Analyses of these data generally used the combined HIV and STI incentive (the sum of the HIV and STI incentives).[32,45,65,79] The combined vouchers ranged between zero to 300 Kwacha (between zero to 3 Dollars at the exchange rates at the time), with an average total voucher amount (including zeros) of 101 Kwacha (1.01 Dollars), worth approximately a day's wage. The distribution of vouchers was carefully monitored to ensure that each nurse followed the rules of randomization. Each voucher included the amount, a respondent ID, and the nurse's signature; a carbon copy was made to prevent forgeries. Respondents who drew a zero token received no voucher; 22% received no incentive to return for either HIV or STI results. Drawing a "zero" may have had a demotivating effect on individuals wanting to attend the HTC, center which may have had an impact on attendance. Because all of the respondents participated in the "lottery" draw, it was impossible to estimate the potential effect of disappointment. However, this is likely to have been minimal.

Descriptive statistics for the MLSFH population selected for the 2004 MLSFH experimental design offering financial incentives for learning HIV status are reported in Table A5, along with information about the average incentive offered and distance to HTC center.

Two to four months after sample collection, test results became available and temporary test results (HTC) centers, consisting of small portable tents, were placed randomly throughout the districts. Based on their geospatial (GPS) coordinates, respondents' households in villages were grouped into zones, and a location within each zone was randomly selected to place a tent. The average distance to a center was two kilometers and over 95% of those tested lived within five kilometers. Distance to the HTC center was calculated as a straight line and does not account for roads or paths. In most cases, tents were placed in the exact randomly selected location and paths were created for easy accessibility. Baseline characteristics were

Table A5: Descriptive statistics for participants in the 2004 MLSFH experimental design offering financial incentives for learning HIV status in 2004

Panel A: Respondent characteristics		
Male	0.46	(0.50)
Age	33.4	(13.66)
Married	0.71	(0.45)
Years of education	3.6	(3.70)
Owns land	0.73	(0.44)
Panel B: Health		
HIV positive	0.063	(0.24)
Gonorrhea positive	0.032	(0.18)
Chlamydia positive	0.003	(0.06)
Trichomoniasis positive	0.024	(0.15)
Ever had an HIV test (before 2004)	0.181	(0.385)
Thinks treatment will be available in five years	0.341	(0.474)
Reported having sex during 2004	0.761	(0.43)
Reported using condoms during 2004	0.210	(0.41)
Panel C: Incentives, distance, and attendance at results centers		
Monetary incentive (dollars)	1.01	(0.90)
Proportion receiving incentive > 0	0.78	(0.41)
Monetary incentive (dollars, if incentive > 0)	1.29	(0.92)
Distance to HTC center (km)	2.02	(1.27)
Attended HTC center	0.69	(0.46)
Attended HTC center (if incentive = 0)	0.34	(0.47)
N	2,812	–

Notes: The analyses included respondents who accepted a test for HIV in 2004 and had basic demographic data available in the MLSFH. The monetary incentive is a sum of an incentive for learning HIV results and an incentive for learning other STI results (in Mchinji and Balaka). Distance from assigned testing centers to respondents' homes is a straight-line spherical distance measured in kilometers.
Source: Thornton [32]

similar across groups receiving any incentive amount (including zero) and living within various HTC zones. Although there were some statistically significant differences among these groups, they were small in magnitude.[32]

Respondents were personally informed of the hours of operation and location of their assigned center and centers were operational for approximately one week. Respondents were allowed to attend any of the HTC centers but were informed only of the location and hours of operation of their assigned center (fewer than 6% of respondents went to a center other than the one to which they were assigned). When they obtained their test results, respondents also received counseling. On average, nurses spent 30 minutes counseling each respondent about safe sexual practices, including abstinence and condom use, regardless of respondent's HIV

test results. Couples were given their test results verbally and were informed of their results separately. Respondents could redeem their vouchers only after hearing their results. Those who were HIV- positive were referred to the nearest permanent clinic for further counseling. Those who were positive for other sexually transmitted diseases were also given free treatment at that time, which may have provided additional incentive to attend VCT centers, over and above the monetary incentive.

Approximately two months after results were available, respondents who tested for HIV in two districts, Balaka and Rumphi, were reinterviewed in their homes by interviewers who had no part in the testing and did not know the respondents' HIV status. Both those who had obtained their results and those who had not were approached for this follow-up interview. During this interview, respondents were asked about their sexual behavior in the prior two months and their attitudes toward condom use. At the end of the interview, respondents were given approximately 30 cents as appreciation for participation and were offered the opportunity to purchase condoms at half the subsidized retail price: five cents for a package of three condoms or two cents for a single condom. Respondents were allowed to purchase condoms only from the 30 cents they had just been given in order to prevent condom purchases from being correlated with any monetary incentive received two months prior at the results center.

A3.2. 2006 MLSFH HIV testing and counseling (HTC)

Starting in 2006, the MLSFH HTC used home-based rapid HIV testing procedures using parallel Determine HIV/1-2™ (Abbott Laboratories, USA) and UniGold™ HIV (Trinity Biotech, Ireland) test kits. MLSFH respondents were approached by the MLSFH HTC team after the completion of the 2006 MLSFH survey. The HTC team only approached MLSFH respondents who completed the 2006 MLSFH survey (and specifically, both the household/family rosters and the survey; but due to coordination problems between teams, also about 100 MLSFH respondents were tested for HIV for whom no household/family rosters and/or survey was collected). A total of three attempts were made to locate each MLSFH respondent for HTC.

HIV testing using rapid HIV testing kits was offered at the respondents' homes. After the informed consent process, a brief survey on prior HIV testing was conducted, and blood was collected by a finger prick and immediately tested for HIV using the parallel Determine™ and UniGold™ HIV test kits. For MLSFH respondents who were still minors (age 17 and younger), informed consent was obtained from the parents and assent was obtained from the respondent. HTC participants were given a choice of receiving post-test counseling and results at their home immediately after the HIV test (results were available after about 20 minutes) or during a few days subsequent to the HIV test at a mobile clinic set up by the MLSFH in the study area. Virtually all of those who chose to receive their results did so at their homes. When the results of both tests were either concordant positive (reactive) or concordant negative (non-reactive), the HIV test results were given to the

respondent. Although no discordant tests occurred in 2006, such test results would have been declared *inconclusive* and the respondent would have been referred to a nearby laboratory for subsequent confirmatory HTC.

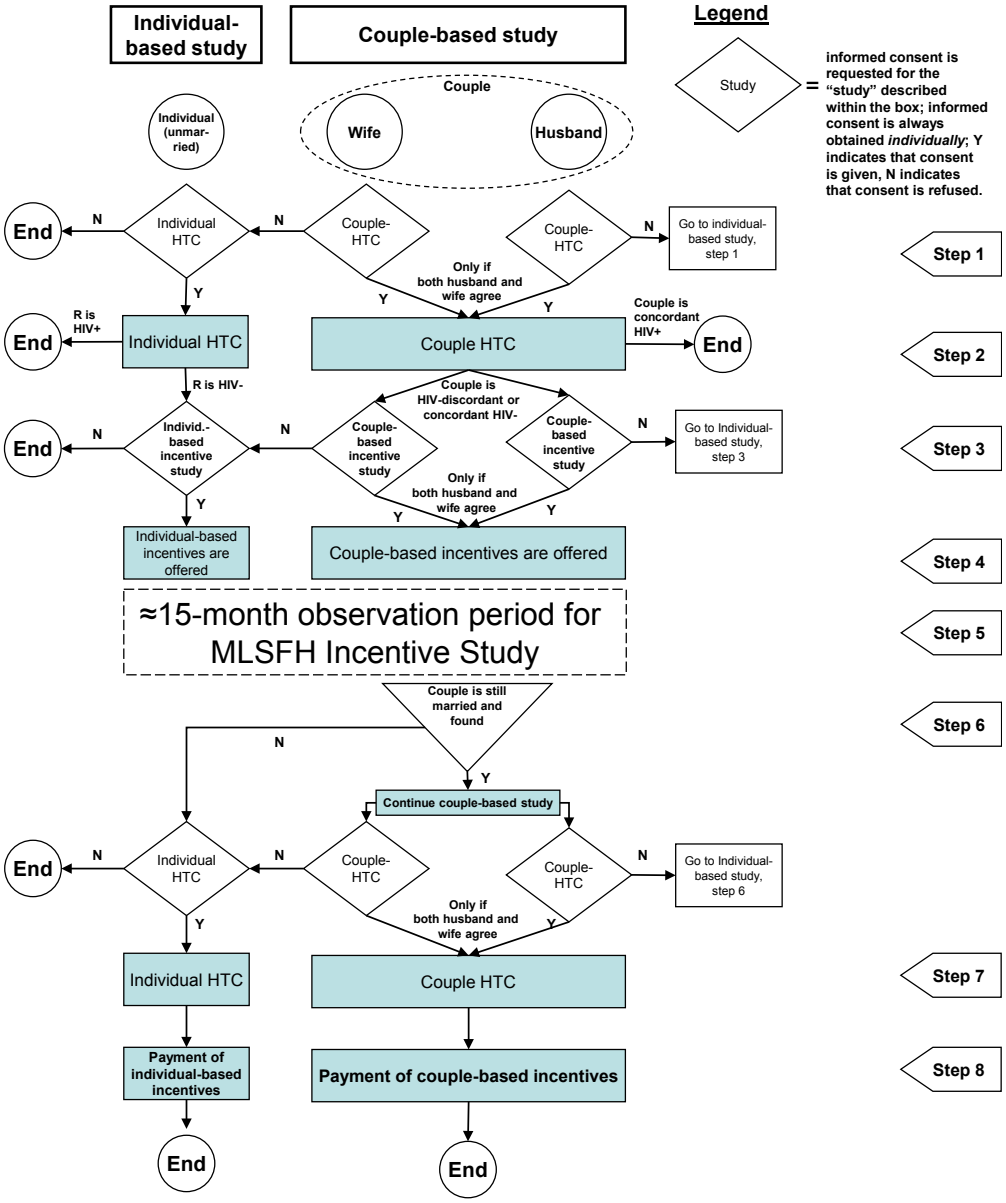
A total of 2,987 respondents were successfully contacted and offered a HIV test; 2,758 (92%) were tested. Of those MLSFH sample members who were successfully contacted in 2006 for HTC, 26% had not been tested in 2004 because they refused (5%), were away at the time of the survey (4%) or were included in 2006 as new sample members—that is, new spouses to those already in the sample (17%). In addition, about a third (32%) of those who accepted a HIV test in 2004 were not tested in 2006 primarily due to mobility (12%), refusal (4%), death (1%) and inability to trace the respondent (15%). Loss to follow-up was somewhat higher in the South compared with the other two sites due to higher mobility and frequent name changes among respondents.

The specific procedures and logistics of the 2006 MLSFH HTC differed slightly across MLSFH respondents who were eligible for couple HTC as part of the MLSFH Incentive Study (Appendix A6.6), and MLSFH respondents who were offered only offered individual HTC. The difference are discussed below. Participants of the MLSFH Incentive Study were followed until March–August 2007, when they participated in a follow-up HTC that following the same HTC procedures as in 2006.

A3.2.a. Couple HTC for the 2006 MLSFH Incentive Study sample: Subsequent to the 2006 MLSFH HTC, the MLSFH implemented an Incentive Study that offered financial incentives to a subset of MLSFH respondents for maintaining their 2006 HIV status during an approximately 15-month period during 2006–07.[22] The MLSFH Incentive Study sample included all MLSFH individuals in discordant couples and a random subset of MLSFH respondents. A total of 1,407 adult individuals were offered to participate in the MLSFH Incentive Study during the 2006 MLSFH HTC (mean age = 35.8, SD = 13), and 1,307 accepted. The details of the study population and the experimental design are given below (Appendix A6.6). Because of the specific structure of this MLSFH Incentive Study, it was determined during the IRB approval process that participants in this study need to be offered the opportunity to participate in couple HTC. That is, married couples had to be given the opportunity to participate in HTC jointly, and through this couple HTC, learn both their own HIV status and that of their spouse. As a result, married individuals who were living with their spouse were offered couple-HTC. Individual HTC was offered to individuals in married couples if the spouse was absent or one of the two spouses declined to participate in couple HTC (either because he/she preferred individual HTC or didn't want to participate in HTC at all). A flow-diagram outlining the assignment to individual or couple HTC is shown in Figure A2. In polygamous marriages selected for the MLSFH Incentive Study, the husband and a randomly selected wife were selected for participation. In addition, couples were only assigned to the MLSFH Incentive Study and couple HTC if both partners were 18 or older.

For the individual HTC, the HTC followed essentially the same procedures as

Figure A2: MLSFH Incentive Study 2006–07: Individual-based and couple-based study design and informed consent procedure



the HTC for MLSFH respondents who were not part of the MLSFH Incentive Study (see below). Individuals who did not give consent to HTC were no longer eligible for participation in MLSFH incentive study. The participants of the MLSFH Incentive study were also informed that if they choose not to participate in the individual-HTC at the follow-up visit, the change in their HIV status during 2006–07 could not be determined and they would not be eligible to receive any incentive payments during this study.

For couple HTC, eligible MLSFH respondents were first asked *individually* about their informed consent to participate in couple HTC. The informed consent process for couple HTC explicitly mentioned that participation in couple HTC would involve revealing the HIV status to the their spouse, and respondents could opt out of couple HTC and choose either individual HTC or no HTC at all. As soon as one member of a married couple opted out of couple HTC, individual HTC was continued for both husband and wife.

For married couples where both agreed to couple HTC during the individual consent process, both were tested for HIV using parallel Determine™ and UniGold™ HIV test kits. The couple was then brought together for a post-test counseling that involved the discussion of the HIV status of both husband and wife.

After couple-HTC, all individuals irrespective of their and their spouse's HIV status received a description of the couple-based incentive study (see Appendix A6.6 for additional details). This description included that: (*a*) the study team will return in 12 months to determine the couple's HIV status; (*b*) the couple will receive a reward, determined by a lottery drawing, if the couple was willing to have their HIV status determined during a couple-HTC session at our second visit and if the couple maintains its current HIV status until this second visit in about 12 months. The couple was also informed that if they choose not to participate in the couple-HTC in 12 months, the change in their HIV status since the previous visit cannot be determined and they were not eligible to receive any reward payments, and they were informed about the available options if the couple separated, the spouse was not available at the follow-up HTC or refused to participate in the follow-up HTC.

The MLSFH study team returned to participants in the MLSFH Incentive Study after about 15 months with a follow-up HTC, following the same procedures for couple and individual HTC as described above (Figure A2).

A3.2.b. Individual HTC during the 2006 MLSFH: Individual HTC was offered to MLSFH respondents who (*a*) completed 2006 MLSFH survey both the household/family rosters and the survey) and were *not* selected for the MLSFH Incentive Study, or (*b*) individuals who were initially selected for the MLSFH Incentive Study but then assigned to individual HTC during the study (see Figure A2). In individual HTC, the consent process, pre-test counseling, HIV testing using parallel Determine™ and UniGold™ HIV test kits, and post-test counseling were all conducted individually at the respondents home in an area that ensured adequate privacy. Post-test counseling included a discussion of the meaning of the result and HIV prevention strategies, a discussion of disclose of HIV test results to spouses, and for HTC participants who tested positive for HIV, a referral respondent to the nearest district hospital or HTC clinic for a confirmatory test and an assessment of the possibilities of treatment with antiretroviral treatment.

MLSFH respondents who participated in individual HTC and were selected for the MLSFH Incentive Study were informed about this study subsequent to HTC, and offered to participate (see Appendix A6.6 for additional detail). The partici-

pants in the MLSFH Incentive study were also informed that if they choose not to participate in the individual-HTC at the follow-up visit, the change in their HIV status during 2006–07 could not be determined and they would not be eligible to receive any incentive payments during this study.

A3.3. 2008 and 2012 MLSFH HTC

The 2008 MLSFH HTC followed the same procedure of the 2006 individual HTC outlined above. All MLSFH respondents who completed the 2008 MLSFH survey were approached by the HTC team. In 2012, essentially the same HTC procedures were repeated, except that only MLSFH respondents age 45 who were eligible for the 2012 MLSFH survey were approached.

A3.4. Comparison of MLSFH HIV prevalence to other population-based estimates

Both the 2004 and the 2006 MLSFH estimates of HIV prevalence are considerably lower than the estimates for rural Malawi based on data collected in 2003 from all the rural antenatal clinics (ANCs) in the national HIV surveillance system (15%). They are also lower than the estimates based on the 2004 Malawi Demographic and Health Survey (MDHS).[132] Age standardization, using the MDHS 2004 age distribution as the standard, did not significantly change the MLSFH estimates.[21] The 2008 MLSFH HIV prevalence is also lower as both the rural DHS estimates for 2004 and 2010. A potential explanation for the variations in the HIV prevalence estimates between the MLSFH and the MDHS is sampling variability coupled with the geographic variation in HIV prevalence. HIV prevalence has, for instance, been found to be higher near the market centers than in the rural villages.[133] The MLSFH sample probably consists of a larger proportion of individuals from the rural villages than the MDHS or ANCs; hence, the lower prevalence.

A4. Comparisons of the MLSFH with national representative samples:

While the initial sampling strategy of the MLSFH was not designed to be representative of the national population of rural Malawi (Appendix A2.1), the initial sample characteristics closely matched the characteristics of the rural population of the 1996 Malawi Demographic and Health Survey (MDHS) (Table A3).[1] After three rounds of longitudinal data collection during 1998–2004, despite attrition and the enrollment of new subjects, the 2004 MLSFH sample remained in close agreement in observable characteristics with the nationally-representative 2004 MDHS (rural sub-population).[2] We update these earlier comparisons based the 2010 MLSFH and MDHS. Since the MDHS is restricted to ages 15–49 (for women), we focus on the respective age range in the MLSFH as well. Given the rural nature of the MLSFH, we also restrict the MDHS to the rural sub-sample.

Table A6 compares the age distribution of the 2010 MDHS and MLSFH sample populations. As is expected, even in the 15–49 age range, the MLSFH study population is significantly older than the MDHS study population. The MLSFH contains a significantly smaller fraction of respondents at ages 15–19, while older ages are overrepresented in the MLSFH (plus, the MLSFH contains a substantial number of respondents older than age 49, see Table A1, that are not included in the compar-

Table A6: Age distribution of the 2010 MDHS and 2010 MLSFH sample populations

	North		*Central*		*South*		*All Sites*	
Age	MDHS	MLSFH	MDHS	MLSFH	MDHS	MLSFH	MDHS	MLSFH
15–19	23.5	2.5	22.7	1.3	22.5	2.0	22.7	1.9
20–24	18.7	20.3^{ns}	19.1	15.1	17.9	21.9	18.5	19.1^{ns}
25–29	17.5	21.1	17.1	23.2	18.4	20.9^{ns}	17.8	21.7
30–34	13.3	15.0^{ns}	13.6	16.5	14.8	14.2^{ns}	14.1	15.2^{ns}
35–39	11.3	14.4	11.6	16.6	11.7	13.7^{ns}	11.6	14.9
40–44	8.1	14.1	8.2	16.1	7.8	14.1	8.0	14.8
45–49	7.7	12.7	7.8	11.4	6.9	13.2	7.4	12.4
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
N	4,746	896	9,109	863	11,929	844	25,784	2,603

Notes: (*a*) T-tests for statistically significant differences between MDHS and MLSFH are significant at $p < 0.05$ or higher for all categories except those labeled *ns* = not significant. (*b*) *MDHS:* The 2010 Malawi Demographic and Health Survey (2010 MDHS)[11] was implemented by the National Statistical Office (NSO) from June through November 2010, with a nationally representative sample of more than 27,000 households. All eligible women age 15–49 in these households and all eligible men age 15–54 in a subsample of one-third of the households were individually interviewed. The primary objectives of the 2010 MDHS project were to provide up-to-date information on fertility levels; nuptiality; sexual activity; fertility preferences; awareness and use of family planning methods; breastfeeding practices; nutritional status of mothers and young children; early childhood mortality; maternal mortality; maternal and child health; malaria; awareness and behavior regarding HIV/AIDS and other sexually transmitted infections; and HIV prevalence. (*c*) Table is restricted to ages 15–49 for both MDHS and MLSFH, and MDHS includes only rural subsample.

isons with the MDHS). This difference in the age distribution of the MLSFH and the representative MDHS is expected given the aging of the MLSFH study population over time, and the fact that the MLSFH was last refreshed at younger ages in 2004 through the MLSFH adolescent sample—the members of which are now at least 23 years old. In terms of characteristics other than age, Table A7 shows that, the 2010 MLSFH sample population is more likely married and has a larger number of children than the MDHS study population. The differential age distribution is an important factor contributing to these differences between the MLSFH and MDHS study populations. Despite these age differences, however, there are no marked differences in schooling levels between the MLSFH and MDHS study populations.

Because of the differential age distributions of the MDHS and MLSFH, Table A8 compares the characteristics of the MDHS and MLSFH study populations by age groups, and Table A9 provides a comparison between the MDHS and MLSFH study populations with the latter being reweighted to match the 2010 MDHS age distribution (separate by gender, restricted to ages 15–49). Controlling for the differential age distribution significantly reduces the differences between the MLSFH and MDHS study populations, with the MLSFH study population being somewhat more likely to be currently married and more likely to have been married more than once; fertility is also slightly higher in the MLSFH study population as compared

Table A7: 2010 MDHS and 2010 MLSFH sample characteristics

	North		*Central*		*South*		*All Sites*	
	MDHS	MLSFH	MDHS	MLSFH	MDHS	MLSFH	MDHS	MLSFH
Panel A: Females								
Primary+ schooling	96.2	98.5	82.4	81.2^{ns}	80.3	60.8	84.0	80.1
Currently married	61.6	87.4	60.3	83.5	58.9	86.3	59.9	85.8
Married > once	17.5	18.9^{ns}	22.9	27.8	28.8	39.1	24.7	28.6
# children ever born	3.2	4.2	3.3	5.0	3.2	4.7	3.3	4.6
# living children	2.8	3.6	2.8	3.8	2.7	3.8	2.7	3.7
Panel B: Males								
Primary+ schooling	97.8	99.2	91.6	90.4^{ns}	92.1	82.5	93.0	91.1
Currently married	47.8	78.5	53.7	86.0	52.5	87.7	52.1	83.9
Married > once	24.1	35.4	27.6	37.4	33.1	40.0	29.5	37.6
# children ever born	2.8	3.5	3.0	4.4	2.9	4.2	2.9	4.0
# living children	2.4	3.1	2.5	3.7	2.4	3.5	2.5	3.4

Notes: Restricted to ages 15–49 for both MDHS and MLSFH. MDHS includes only rural subsample. Primary+ schooling = completed at least some primary schooling. T-tests for statistically significant differences between MDHS and MLSFH are significant at $p < 0.05$ or higher for all categories except those labeled *ns* = not significant.

to the MDHS. Once the differential age structures in the MDHS and MLSFH are controlled for, the remaining differences between the MDHS and MLSFH study populations tend to be relatively small, with the potential exception of marital status, where MLSFH respondents remain more likely to be married than MDHS respondents even after differences in the age distribution of these two surveys are accounted for. The higher likelihood of MLSFH respondents to be married, as compared to MDHS respondents, is likely due to the initial 1998 MLSFH sample that focused on ever-married women and their spouses and the fact that peri-urban regions are missing in the MLSFH (Appendix A2.1). Nevertheless, overall the remaining differences in sample characteristics are mostly not substantively significant and/or indicative of important distortions in the MLSFH study populations as compared to the nationally-representative MDHS rural sample.

In order to provide a comparison of the MLSFH with a representative sample that extends to older ages, Table A10 compares the 2010 MLSFH study population with the rural subsample of the 2010 Malawi Integrated Household Survey (IHS3), which is a nationally representative survey conducted by the Malawi National Statistical Office to monitor progress towards the Millennium Development Goals (MDGs). As already documented in the comparison with the MDHS, the MLSFH study population is more likely to be married than the nationally representative IHS3 study population (rural subsample), and the MDHS respondents have slightly more schooling. The differences in religion result from the fact that the MLSFH is based in only three regions, one of which (Balaka) is predominantly Muslim. MLSFH respondents also are less likely to reside in a house with

Table A8: 2010 MDHS and 2010 MLSFH sample characteristics, by age and gender

	North		*Central*		*South*		*All Sites*	
	MDHS	MLSFH	MDHS	MLSFH	MDHS	MLSFH	MDHS	MLSFH
Females								
Panel A: Age 20–29								
Primary+ schooling	97.9	99.5*	89.8	89.8^{ns}	87.7	83.2^{ns}	90.2	90.8^{ns}
Currently married	73.7	87.8***	72.1	85.6***	69.7	86.5***	71.3	86.6***
Married > once	13.3	14.1^{ns}	14.0	18.8^{ns}	22.3	29.2*	17.8	20.8^{ns}
# children ever born	2.4	2.2^{ns}	2.3	2.9***	2.6	2.9*	2.5	2.7**
# living children	2.2	2.1^{ns}	2.1	2.4*	2.2	2.5^{ns}	2.2	2.3*
Panel B: Age 30–39								
Primary+ schooling	95.0	98.1*	72.7	82.0**	70.6	51.9***	75.6	77.5^{ns}
Currently married	74.0	86.9***	75.1	87.4***	69.7	90.5***	72.3	88.3***
Married > once	22.5	19.0^{ns}	30.5	27.7^{ns}	34.7	42.6^{ns}	31.1	29.7^{ns}
# children ever born	5.0	4.9^{ns}	5.2	5.6**	4.9	5.7***	5.0	5.4***
# living children	4.4	4.4^{ns}	4.3	4.6^{ns}	4.0	4.6***	4.2	4.5***
Panel C: Age 40–49								
Primary+ schooling	88.8	97.2***	60.3	65.8^{ns}	55.0	35.0***	63.3	65.8^{ns}
Currently married	68.8	85.8***	70.0	76.7^{ns}	64.7	79.6***	67.4	80.9***
Married > once	25.9	26.2^{ns}	36.5	43.7^{ns}	42.9	50.7^{ns}	37.4	40.0^{ns}
# children ever born	6.5	6.7^{ns}	7.3	7.5^{ns}	6.3	6.7^{ns}	6.7	6.9^{ns}
# living children	5.4	5.4^{ns}	5.6	5.4^{ns}	5.0	5.0^{ns}	5.3	5.2^{ns}
Males								
Panel A: Age 20–29								
Primary+ schooling	98.4	100.0*	94.7	96.8^{ns}	94.2	87.8*	95.3	94.9^{ns}
Currently married	45.1	63.1***	50.7	67.5***	52.7	78.9***	50.4	69.8***
Married > once	8.9	16.4^{ns}	13.2	22.4^{ns}	16.4	30.7**	13.9	23.7***
# children ever born	1.1	1.1^{ns}	1.1	1.5**	1.4	2.3***	1.3	1.7***
# living children	1.0	1.0^{ns}	1.0	1.5***	1.3	1.9***	1.1	1.4***
Panel B: Age 30–39								
Primary+ schooling	96.8	98.98^{ns}	88.4	87.5^{ns}	89.0	84.2^{ns}	90.1	90.6^{ns}
Currently married	72.1	89.8***	80.6	96.4***	81.2	98.7***	79.5	94.8***
Married > once	28.6	46.88**	29.7	40.2*	36.8	41.9^{ns}	32.9	42.9**
# children ever born	4.2	4.1^{ns}	4.3	4.8*	4.4	4.7^{ns}	4.3	4.5^{ns}
# living children	3.8	3.9^{ns}	3.7	3.9^{ns}	3.7	4.2^{ns}	3.7	4.0^{ns}
Panel C: Age 40–49								
Primary+ schooling	98.3	97.87^{ns}	83.2	86.2^{ns}	84.1	72.1*	86.5	85.8^{ns}
Currently married	82.8	93.6**	84.1	96.3***	80.1	92.9***	82.1	94.43***
Married > once	34.7	44.7^{ns}	38.6	46.3^{ns}	47.0	52.4^{ns}	41.5	47.6^{ns}
# children ever born	6.6	7.0^{ns}	7.4	7.1^{ns}	6.7	6.9^{ns}	6.9	7.0^{ns}
# living children	5.7	5.9^{ns}	6.1	5.9^{ns}	5.5	5.5^{ns}	5.8	5.8^{ns}

Notes: Primary+ schooling = completed at least some primary schooling. T-tests for statistically significant differences between MDHS and MLSFH ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; *ns* = p not significant. MDHS includes only rural subsample. MDHS sample sizes are 7,421 (females) and 1,916 (males) at age 20–29, 5,131 (females) and 1,494 (males) at age 30–39, and 3,042 (females) and 917 (males) at age 40–49. MLSFH sample sizes are 598 (females) and 430 (males) at age 20–29, 485 (females) and 286 (males) at age 30–39, and 404 (females) and 289 (males) at age 40–49.

Table A9: 2010 MDHS and 2010 MLSFH sample characteristics, with MLSFH weighted to match MDHS age distribution

	Females		Males	
	MDHS	MLSFH	MDHS	MLSFH
Primary+ schooling	84.0	84.2^{ns}	93.0	94.0^{ns}
Currently married	59.9	87.3	52.1	60.9
Married > once	24.7	24.7^{ns}	29.5	36.3
# children ever born	3.3	3.7	2.9	2.7^{ns}
# living children	2.7	3.0	2.5	2.3^{ns}

Notes: Restricted to ages 15–49 for both MDHS and MLSFH, with MLSFH weighted to match MDHS age distribution (separately by gender). MDHS includes only rural subsample. Primary+ schooling = completed at least some primary schooling. T-tests for statistically significant differences between MDHS and MLSFH are significant at $p < 0.05$ or higher for all categories except those labeled *ns* = not significant.

a metal/tile roof, possibly related to the fact that peri-urban areas are included in the rural IHS3 sample, but not the MLSFH. The comparison with the IHS3 also provide information on health-related outcomes, which are of particular interest at somewhat older ages (age 45+), and both datasets reveal relatively high levels of disabilities at ages 45 and over. Overall, similar to our comparison with the DHS above, there are selected differences in several of the variables reported in Table A10 between the 2010 MLSFH and IHS3 study population, but overall these differences do not seem to be substantially significant and indicative of important biases/distortions in the MLSFH study population—and where they exist, they can be related to the specific study design of the MLSFH and controlled for analyses that aim at estimating population-level characteristics based on the MLSFH.

In summary, therefore, our comparisons of the 2010 MLSFH study population with the rural samples of the MDHS and IHS3 surveys reveal that the MLSFH study population continues to closely match the characteristics of nationally-representative cross-sectional surveys, despite the fact that the initial MLSFH sample was not selected to be nationally representative and the MLSFH has been subject to attrition over time (see below). Neither the initial sample selection that restricted the MLSFH to three rural region, nor the MLSFH attrition and enrollment of new MLSFH respondents over time, seem to have importantly affected the MLSFH in terms of its ability to represent the rural population of Malawi. The MLSFH is different from nationally-representative rural samples in terms of its age distribution, and where appropriate, the MLSFH can be weighted to match the age distribution of rural Malawi. The MLSFH also contains a larger fraction of respondents who are currently married, which is likely due to the initial 1998 MLSFH sample that focused on ever-married women and their spouses and the fact that peri-urban regions are missing in the MLSFH. Where appropriate, analyses can adjust for this over-representation of married individuals in the MLSFH.

Table A10: 2010 IHS3 and 2010 MLSFH sample characteristics

Age range	**20–29**				**30–44**			
	MLSFH 2010		**IHS3 2010–11**		**MLSFH 2010**		**IHS3 2010–11**	
	N	%	N	%	N	%	N	%
Demographic and Socioeconomic Characteristics								
Male	435	41.3%	3,228	45.7%	454	39.0%	3,289	50.0%
Any schooling	941	89.4%	6,118	86.6%	918	78.9%	4,947	75.2%
Married	809	76.9%	4,767	67.5%	1027	88.3%	5,625	85.5%
Religion:*Christian*	612	58.2%	6,047	85.6%	728	62.6%	5,518	83.9%
Muslim	218	20.7%	778	11.0%	200	17.2%	732	11.1%
Other	222	21.1%	238	3.4%	235	20.2%	330	5.0%
Metal/tile roof	158	15.0%	1,806	25.6%	248	21.3%	1,913	29.1%
Health Indicators								
Functional limitations and disability state								
Moderate Limitation	124	11.8%	–	–	202	17.4%	–	–
Severe Limitation	24	2.3%	–	–	43	3.7%	–	–
ADL disabled	–	–	350	5.0%	–	–	506	7.7%
Average Age	24.7		24.3		36.8		35.8	
Total	1052		7,063		1163		6,580	

Age range	**45–64**				**65+**			
	MLSFH 2010		**IHS3 2010–11**		**MLSFH 2010**		**IHS3 2010–11**	
	N	%	N	%	N	%	N	%
Demographic and Socioeconomic Characteristics								
Male	467	43.6%	1,924	48.1%	201	45.0%	785	43.2%
Any schooling	729	68.0%	2,483	62.1%	263	58.8%	745	41.0%
Married	850	79.3%	3,071	76.8%	255	57.0%	914	50.3%
Religion:*Christian*	628	58.6%	3,319	83.0%	233	52.1%	1,418	78.0%
Muslim	200	18.7%	441	11.0%	101	22.6%	241	13.3%
Other	244	22.8%	241	6.0%	113	25.3%	158	8.7%
Metal/tile roof	254	23.7%	1,253	31.3%	100	22.4%	537	29.6%
Health Indicators								
Functional limitations and disability state								
Moderate Limitation	314	29.3%	–	–	176	39.4%	–	–
Severe Limitation	86	8.0%	–	–	113	25.3%	–	–
ADL disabled	–	–	783	19.6%	–	–	895	49.3%
Average Age	53.5		53.1		74.1		74.5	
Total	1072		4,001		447		1,817	

Notes: (*a*) IHS3 data description: The Integrated Household Survey is one of the primary instruments implemented by the Government of Malawi through the National Statistical Office (NSO) roughly every 5 years to monitor and evaluate the changing conditions of Malawian households. The IHS data have, among other insights, provided benchmark poverty and vulnerability indicators to foster evidence-based policy formulation and monitor the progress of meeting the Millennium Development Goals (MDGs) as well as the goals listed as part of the Malawi Growth and Development Strategy (MGDS). The Third Integrated Household Survey (IHS3) was conducted by National Statistical Office (NSO) in March 2010-March 2011.[134] A stratified two-stage sample design was used for the IHS3. The IHS3 sampling frame is based on the listing information and cartography from the 2008 Malawi Population and Housing Census (PHC); includes the three major regions of Malawi,

(Continued on next page)

Table A10: 2010 IHS3 and 2010 MLSFH sample characteristics

(Note to Table A10, continued from previous page)
namely North, Center and South; and is stratified into rural and urban strata. The rural subsample of the IHS3, which is used for the above analyses, includes residents from each of the 27 districts of Malawi, except those living in the urban centers of Lilongwe City, Blantyre City, Mzuzu City, and the Municipality of Zomba, and except for residents of the island of Likoma on Lake Malawi. The sampling frame excludes the population living in institutions, such as hospitals, prisons and military barracks. (*b*) Health indicators: There are no directly comparable disability/health indicators in the MLSFH and IHS3. Functional limitations and disability states for the MLSFH are defined as follows: respondents who answered "somewhat limited" on either of the two MLSFH SF-12 question about physical limitations are classified as *moderately limited*, and respondents who answered "limited a lot" on either question are classified as *severely limited*. *ADL disabled* in the IHS3 is defined as having difficulty in any one of the following five activities of daily living (ADLs): Seeing, hearing, walking, remembering/concentrating, self-care (bathing/dressing). (*c*) Comparisons between the IHS3 and the MLSFH are based on IHS3 and the MLSFH unweighted samples. IHS3 includes only rural subsample. All differences between the MLSFH and IHS3, except for the proportion with any schooling among 20–30 year olds, proportion with any schooling and proportion married for 30–45 year-olds, proportion married for 45–64 year olds, and proportion male 65+, are significant ($p < .05$) according to chi-square tests.
Source: Modified from Payne *et al.* [110]

A5. Analyses of attrition in the MLSFH

All longitudinal data collection projects face the inherent problem of sample attrition: the failure to find or reinterview individuals who were surveyed in an earlier wave of the study.[121,122,135–138] Attrition leads to decrease in sample sizes, which can reduce power in statistical analysis. More importantly, however, attrition may bias subsequent analyses if those who leave the sample are substantially and systematically different from those who do not—particularly on unobserved characteristics.[121,122,135–138] Numerous events can lead to sample attrition, including short- or long-term mobility, mortality, failures to recontact respondents in the absence of reliable addresses, or refusal of respondents to participate in follow-up waves of the study. In rural sub-Saharan Africa, rates of attrition are often found to be relatively high due to high levels of mobility which is often work-related or related to marriage and/or divorce.[23,121,139] The MLSFH is no exception to this pattern (Figure 2 and Table A1), with the fraction of MLSFH respondents who were successfully reinterviewed at a subsequent MLSFH round ranging from 73.5% (in 2001) to 78.1% (in 2004); a outlier is the 2012 MLSFH, where 90.3% of eligible respondents from the 2010 MLSFH (age 45+ and interviewed in both the 2008 and 2010 MLSFH) were reinterviewed.

Earlier analyses of attrition in the MLSFH focused on attrition up to the 2006 MLSFH.[2,80] These analyses concluded that, even though respondent characteristics often differ significantly between those who were lost to follow-up and those who were re-interviewed and attrition was often predicted by key respondent characteristics, the coefficient estimates for standard family background variables in regressions and probit equations for the majority of the outcome variables were not

Table A11: Reasons for attrition in 2010 MLSFH

	Reason (%) for attrition in 2010 among MLSFH respondents interviewed in	
	2006	**2008**
Refused 2010 MLSFH survey	10.0	9.1
Hospitalized at time of MLSFH survey	0.1	0.4
Deceased since prior MLSFH survey	6.5	8.9
Person unknown	4.9	9.1
Temporarily absent at time of MLSFH survey	12.9	14.1
Moved outside of MLSFH study village/region	45.9	42.6
Other reasons	19.8	15.9
Number of respondents lost to follow-up (and thus *N* for above tabulations)	970	1,016
Number successfully interviewed (2006/2008)	3,431	4,036
% lost to follow-up by 2010	28.3	25.2

affected significantly by attrition.

We update these earlier analyses of attrition in the MLSFH by focusing on attrition during 2006–10 and 2008–10, i.e., attrition among the most recent complete MLSFH surveys. Following our earlier analyses of attrition, after describing the primary reasons leading to a loss-to-follow-up in the MLSFH, we conduct three sets of analyses to assess concerns about attrition-related biases in the MLSFH. First, we compare observable characteristics of respondents who were interviewed by the MLSFH in an earlier but not in a subsequent wave, with the characteristics of respondents who were observed in both MLSFH waves. Second, to identify possible predictors of attrition, we report logistic regressions of the probability to attrit after a respondent was interviewed in the 2006 or 2008 MLSFH round. Finally, we perform a series of OLS and logistic regressions predicting several outcomes of interest from the 2006 and 2008 data, which are chosen based on their ability to reflect a broad range of topics investigated with the MLSFH. To assess if attrition potentially results in these estimated relationships, these regressions include interactions of explanatory variables with an indicator that a respondent attrited subsequent to the 2006 or 2008 MLSFH (this approach for investigating the potentially distorting effect of attrition is sometimes referred to as a BGLW test).[136]

We focus in our analyses on attrition of respondents who were interviewed as part of the 2006 or 2008 MLSFH, but were not successfully reinterviewed in the 2010 MLSFH. Table A11 reports the recorded reasons of why MLSFH respondents who were interviewed in 2006 or 2008 were not reinterviewed during the 2010 MLSFH (see also Figure 2, where the same information is reported with less detail). In the majority of cases, attrition is due to migration, which in most cases is related to work or marital transitions (marriage/divorce). Refusal rates in the MLSFH study remain remarkably low, and refusals are not a major source of attrition. And while

mortality rates are relatively high in this rural SSA context, and MLSFH mortality levels are comparable to that of the general population,[110,112,140] mortality is also not a primary reason of why MLSFH respondents are lost to follow-up.

Tables A12 and A13 compare individual characteristics and selected key outcome variables for MLSFH respondents who have been interviewed in 2006 or 2008, but not in 2010 as a result of attrition. These descriptive comparisons are based on observed characteristics from the initial MLSFH wave (2006 or 2008 in our analyses) and, as expected, indicate some differences between respondents who were retained in the 2010 MLSFH and those who were lost to follow-up. Those who were not reinterviewed were more likely to be male, were somewhat younger and had fewer children, and were more likely from Balaka (where mobility is higher). Attriters don't differ markedly in education levels from those who were re interviewed. And while attriters were less likely to agree to HTC, more likely to be HIV+ (conditional on being tested), and had more sexual partners (2006 only). There are no substantively-relevant differences between attriters and non-attriters in terms of SF12 physical or mental health scores, the ever-use of condoms, the worries about HIV or the subjective expectations of being HIV+ or becoming infected in the future. These patterns of differences between attriters and non-attriters are essentially confirmed in Tables A14 and A15, which report logistic regressions of the probability to attrit on 2006/08 individual characteristics and outcome variable. The first column in this table reports bivariate relationships between the indicated variables (measured at the initial MLSFH wave in 2006 or 2008) and attrition status by 2010. For attrition during 2006–10 and 2008–10, gender, age, region of residence, number of children, agreeing to HTC and being HIV+ are significantly associated with a attrition; marital status and the subjective likelihood of being HIV+ are significantly associated with attrition during 2006–10 but not for the shorter time horizon 2008–10.

Finally, Tables A16 and A17 report regressions of selected outcome variables—SF12 physical and mental health score, number of sexual partners, being HIV-positive, condom use, worries about HIV and subjective HIV infection probability—on individual characteristics (measured in 2006 or 2008), including an interaction of all included characteristics with an indicator for attrition in the 2010 MLSFH. If the estimated relationships for these outcome variables differ between MLSFH respondents who are retained in the sample and those who are lost to follow-up, the interaction effects with attrition would be individually or jointly significant (this is referred to as the BGLW test for selective attrition[136]). In Tables A16 and A17 we therefore report the interaction effects with attrition for all individual characteristics included in the regressions (main effects are omitted), and tests for the individual and joint significance of the interaction effects.

Very few of the individual interactions are statistically significant. For the majority of outcomes considered in Tables A16 and A17, including SF12 physical health score, being HIV-positive, condom use and worries about HIV, the null-hypotheses that the interactions effects with attrition are jointly zero for all in-

Table A12: MLSFH attrition 2006–2010: 2006 descriptive statistics by attrition status in 2010

MLSFH 2010 survey outcome	*Interviewed*		*Attrited*		*Difference*		
Panel A: Key respondent characteristics (in 2006)	Mean	SD	Mean	SD	Means	*t*-value	*p*-value
Male	0.43	0.50	0.50	0.50	-0.07	-3.44	0.00
Age	35.73	13.50	33.47	13.76	2.26	4.36	0.00
Currently married	0.84	0.37	0.74	0.44	0.10	6.46	0.00
Number of living children	3.93	3.61	3.73	3.22	0.72	5.39	0.00
Region: *Mchinji (center)*	0.32	0.47	0.34	0.47	-0.02	-1.10	0.27
Balaka (south)	0.33	0.47	0.38	0.48	-0.05	-2.64	0.01
Rumphi (north)	0.35	0.48	0.28	0.45	0.07	3.76	0.00
Schooling: *No schooling*	0.24	0.43	0.23	0.42	0.01	0.62	0.54
Primary	0.63	0.48	0.61	0.49	0.02	1.16	0.25
Secondary or more	0.13	0.34	0.16	0.37	-0.03	-2.38	0.02
Missing data on key respondent characteristics	0.00	0.05	0.02	0.13	-0.01	-4.42	0.00
Observations (*N*)	2,461	71.7%	970	28.3%			
Panel B: Key outcomes of interest (in 2006)	Mean	SD	Mean	SD	Means	*t*-value	*p*-value
SF12: physical health score (2,205; 816)	52.55	7.22	52.32	7.71	0.23	0.77	0.44
SF12: mental health score (2,205; 816)	55.64	7.91	55.44	8.07	0.20	0.62	0.53
HIV test result available (2,454; 954)	0.90	0.30	0.80	0.40	0.10	7.84	0.00
HIV positive in 2006 (2212; 766)	0.05	0.21	0.10	0.30	-0.06	-5.60	0.00
Ever used condom, with any of last 3 partners (1,987; 668)	0.29	0.45	0.32	0.47	-0.03	-1.49	0.14
Lifetime number of sexual partners (1,991; 671)	2.92	3.94	3.43	4.78	-0.51	-2.75	0.01
HIV/AIDS worry (1,984; 672); *Not at all*	0.58	0.49	0.55	0.50	0.03	1.43	0.15
A little	0.26	0.44	0.25	0.43	0.01	0.44	0.66
A lot	0.16	0.37	0.20	0.40	-0.04	-2.40	0.02
Subj. likelihood of being HIV+ (2,059; 724): *None*	0.72	0.45	0.69	0.46	0.03	1.66	0.10
Low	0.19	0.40	0.20	0.40	0.00	-0.11	0.91
Medium	0.05	0.22	0.06	0.24	-0.01	-1.14	0.25
High	0.03	0.18	0.05	0.22	-0.02	-2.39	0.02

Notes: Missing data on key respondent characteristics = 1 if one of the following variables has missing value in the data: age, currently married, number of living children, and schooling (there are no missing values for gender and region). In Panel B, the number of observations varies by outcome; *N*s are reported in parentheses as (*N* reinterviewed; *N* attrited). HIV test result available = at least one HIV test result is available from 2004 and/or 2006 MLSFH; HIV+ = at least one of the available test results is positive.

Table A13: MLSFH attrition 2008–2010: 2008 descriptive statistics by attrition status in 2010

MLSFH 2010 survey outcome	*Interviewed*		*Attrited*		*Difference*		
Panel A: Key respondent characteristics (in 2008)	Mean	SD	Mean	SD	Means	*t*-value	*p*-value
Male	0.40	0.49	0.45	0.50	-0.05	-2.88	0.00
Age	41.57	16.66	39.65	17.98	1.91	3.03	0.00
Currently married	0.83	0.38	0.81	0.39	0.02	1.24	0.22
Number of living children	4.45	4.02	3.77	4.17	0.68	4.50	0.00
Region: *Mchinji (center)*	0.32	0.47	0.37	0.48	-0.05	-2.96	0.00
Balaka (south)	0.34	0.47	0.41	0.49	-0.07	-3.99	0.00
Rumphi (north)	0.34	0.48	0.22	0.42	0.12	7.15	0.00
Schooling: *No schooling*	0.25	0.43	0.24	0.43	0.01	0.60	0.55
Primary	0.62	0.48	0.62	0.49	0.01	0.32	0.75
Secondary or more	0.12	0.33	0.14	0.35	-0.02	-1.24	0.22
Missing data on key respondent characteristics	0.03	0.18	0.05	0.22	-0.02	-2.65	0.01
Observations (*N*)	3,020	74.8%	1,016	25.2%			
Panel B: Key outcomes of interest (in 2008)	Mean	SD	Mean	SD	Means	*t*-value	*p*-value
SF12: physical health score (2,427; 558)	51.84	8.45	51.39	8.02	0.46	1.16	0.24
SF12: mental health score (2,427; 558)	54.10	9.00	54.26	8.81	-0.16	-0.37	0.71
HIV test result available (2,923; 965)	0.88	0.32	0.71	0.46	0.18	13.32	0.00
HIV positive in 2008 (2,586; 683)	0.05	0.22	0.08	0.27	-0.03	-3.09	0.00
Lifetime number of sexual partners (2,573; 837)	3.24	4.75	3.23	3.06	0.01	0.05	0.96
HIV/AIDS worry (2,907; 955): *Not at all*	0.51	0.50	0.53	0.50	-0.03	-1.37	0.17
A little	0.24	0.43	0.25	0.43	0.00	-0.18	0.86
A lot	0.25	0.43	0.22	0.42	0.03	1.78	0.08
Subj. likelihood of being HIV+ (2,862; 944): *None*	0.53	0.50	0.51	0.50	0.02	0.87	0.38
Low	0.30	0.46	0.30	0.46	0.00	-0.10	0.92
Medium	0.11	0.31	0.13	0.34	-0.02	-1.68	0.09
High	0.06	0.24	0.06	0.23	0.01	0.62	0.53

Notes: Missing data on key respondent characteristics = 1 if one of the following variables has missing value in the data: age, currently married, number of living children, and schooling (there are no missing values for gender and region). In Panel B, the number of observations varies by outcome; *N*s are reported in parentheses as (*N* reinterviewed; *N* attrited). HIV test result available = at least one HIV test result is available from 2004, 2006 and/or 2008 MLSFH; HIV+ = at least one of the available test results is positive.

Table A14: MLSFH attrition 2006–2010: Logistic regressions predicting attrition in 2010 among respondents interviewed in 2006 (Odds ratios)

	uni-variate	Multivariate analyses					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Missing data	5.880***						
	(2.674)						
Male	1.285***	1.300**	1.265**	1.519***	1.500***	1.607***	1.573***
	(0.098)	(0.104)	(0.102)	(0.160)	(0.161)	(0.172)	(0.171)
Age	0.988***	0.988***	0.997	0.989**	0.992	0.993	0.996
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Region (*Mchinji*)							
Balaka	1.066	1.105	1.078	1.132	1.015	1.111	0.998
	(0.097)	(0.104)	(0.102)	(0.127)	(0.117)	(0.126)	(0.115)
Rumphi	0.749**	0.709***	0.687***	0.667**	0.670**	0.672**	0.672**
	(0.071)	(0.072)	(0.071)	(0.085)	(0.086)	(0.086)	(0.087)
Schooling (*No schooling*)							
Primary	1.008	1.019	1.047	0.982	0.964	0.993	0.974
	(0.093)	(0.105)	(0.108)	(0.116)	(0.116)	(0.118)	(0.118)
Secondary+	1.292*	1.361*	1.285	1.163	1.112	1.138	1.094
	(0.163)	(0.206)	(0.196)	(0.225)	(0.218)	(0.221)	(0.215)
# of living children	0.930***		0.964*			0.966*	0.972
	(0.012)		(0.015)			(0.017)	(0.017)
Currently married	0.558***		0.645***			0.619**	0.646**
	(0.051)		(0.065)			(0.099)	(0.105)
SF12: physical health	0.996			0.993	0.993	0.993	0.993
	(0.005)			(0.006)	(0.007)	(0.006)	(0.007)
SF12: mental health	0.997			0.996	0.996	0.997	0.997
	(0.005)			(0.006)	(0.006)	(0.006)	(0.006)
Ever used condom, with	1.146			1.057	1.075	1.065	1.086
any of last 3 partners	(0.110)			(0.112)	(0.117)	(0.114)	(0.118)
Lifetime number of	1.026*			1.008	1.007	1.008	1.006
sexual partners	(0.011)			(0.011)	(0.011)	(0.011)	(0.011)
HIV/AIDS worry	1.129*			0.978	0.974	0.978	0.976
	(0.065)			(0.073)	(0.074)	(0.074)	(0.075)
Subj. likelihood of	1.154*			1.161*	1.109	1.159*	1.110
being HIV+	(0.065)			(0.086)	(0.084)	(0.086)	(0.085)
HIV+	2.355***				2.564***		2.402***
	(0.365)				(0.449)		(0.424)
HIV test result	0.426***				0.388***		0.389***
available	(0.044)				(0.052)		(0.053)
Observations		3,413	3,408	2,604	2,604	2,600	2,600

Notes: Odds ratios are reported. Constant is omitted. (1) = univariate logistic regressions of attrition in 2010 in variables listed in the respective row. (2-7) = multivariate logistic regressions of attrition in 2010, with various specifications. Reference groups are stated in parentheses. Std errors in parentheses. Respondents with at least one HIV-positive MLSFH HIV test during 2004–2006 are considered as being HIV positive, all others are considered being HIV negative at the 2006 MLSFH Round (MLSFH 4). HIV/AIDS worry is codes as: 0 = not at all, 1 = a little, and 2 = a lot. Subjective likelihood of being HIV+ (at the time of the MLSFH survey) is coded as: 0 = none, 1 = low, 2 = medium and 3 = high. HIV+ status is as compared to HIV– in column 1 and as compared to HIV- or not tested in columns 5 and 7, with HIV test result available reflecting the effect of participation in HTC. p-values: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

cluded coefficients is not rejected. For attrition during 2006–10 (Table A16), this is also the case for subjective HIV infection probability. The estimated relationships (coefficients) seem to differ between attriters and non-attriters for none of the outcome variables when attrition during 2006–10 is considered (Table A16), and it differs only for the subjective HIV infection probability when attrition during

Table A15: MLSFH attrition 2008–2010: Logistic regressions predicting attrition in 2010 among respondents interviewed in 2008 (Odds ratios)

	uni-variate	Multivariate analyses					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Missing data	1.593**						
	(0.282)						
Male	1.236**	1.202*	1.239**	1.337*	1.367**	1.298*	1.327*
	(0.090)	(0.095)	(0.099)	(0.156)	(0.161)	(0.152)	(0.157)
Age	0.994**	0.995*	0.999	0.981***	0.980***	0.991	0.990
	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)
Region (*Mchinji*)							
Balaka	1.057	1.098	1.066	0.921	0.866	0.890	0.842
	(0.088)	(0.097)	(0.095)	(0.114)	(0.108)	(0.111)	(0.106)
Rumphi	0.543***	0.508***	0.508***	0.454***	0.452***	0.452***	0.450***
	(0.052)	(0.052)	(0.052)	(0.066)	(0.066)	(0.066)	(0.066)
Schooling (*No schooling*)							
Primary	1.029	1.151	1.162	1.351*	1.392*	1.336	1.379*
	(0.091)	(0.116)	(0.118)	(0.202)	(0.210)	(0.201)	(0.209)
Secondary+	1.164	1.520**	1.429*	1.593*	1.605*	1.437	1.458
	(0.146)	(0.233)	(0.221)	(0.347)	(0.351)	(0.318)	(0.324)
# of living children	0.935***		0.945***			0.926**	0.926**
	(0.013)		(0.015)			(0.025)	(0.026)
Currently married	0.894		0.906			0.802	0.824
	(0.084)		(0.092)			(0.116)	(0.120)
SF12: physical health	0.994			0.991	0.992	0.992	0.993
	(0.006)			(0.008)	(0.008)	(0.008)	(0.008)
SF12: mental health	1.001			0.996	0.996	0.997	0.997
	(0.005)			(0.006)	(0.006)	(0.006)	(0.007)
Lifetime number of sexual partners	1.001			0.988	0.986	0.991	0.989
	(0.009)			(0.014)	(0.015)	(0.014)	(0.014)
HIV/AIDS worry	0.928			0.973	0.978	0.974	0.979
	(0.042)			(0.073)	(0.073)	(0.073)	(0.074)
Subj. likelihood of being HIV+	1.033			1.025	0.983	1.023	0.983
	(0.043)			(0.066)	(0.064)	(0.066)	(0.065)
HIV+	1.857***				2.010***		1.879**
	(0.294)				(0.398)		(0.377)
HIV test result available	0.306***				0.416***		0.420***
	(0.027)				(0.101)		(0.102)
Observations		3,900	3,888	2,591	2,591	2,584	2,584

Notes: Odds ratios are reported. Constant is omitted. (1) = univariate logistic regressions of attrition in 2010 in variables listed in the respective row. (2-7) = multivariate logistic regressions of attrition in 2010, with various specifications. Reference groups are stated in parentheses. Std errors in parentheses. Respondents with at least one HIV-positive MLSFH HIV test during 2004–2008 are considered as being HIV positive, all others are considered being HIV negative at the 2008 MLSFH Round (MLSFH 5). HIV/AIDS worry is codes as: 0 = not at all, 1 = a little, and 2 = a lot. Subjective likelihood of being HIV+ (at the time of the MLSFH survey) is coded as: 0 = none, 1 = low, 2 = medium and 3 = high. HIV+ status is as compared to HIV– in column 1 and as compared to HIV- or not tested in columns 5 and 7, with HIV test result available reflecting the effect of participation in HTC. p-values: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

2008–10 is considered (Table A17).

In summary, therefore, our analyses of MLSFH attrition during 2006–10 (Tables A12, A14 and A16) and during 2008–10 (Tables A13, A15 and A17) confirm our earlier findings. MLSFH respondents who are lost to follow-up differ significantly in important observed characteristics—including gender, age, region of residence, number of children and HIV status—from those who are retained in the MLSFH;

Table A16: MLSFH attrition 2006–2010: OLS, ordered logit, and logit models for selected key outcome variables in 2006, with interaction for respondents who subsequently attrited during 2006–10

	SF12 mental health score	SF12 physical health score	Lifetime number of sexual partners	HIV positive	Used condom with recent partners	HIV/ AIDS worry	Subj. likelihood of being HIV+
	OLS	OLS	OLS	Logit	Logit	Ordered logit	Ordered logit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Interactions with attrition:							
Male × attrition	0.195	-0.096	0.020	-0.323	0.092	0.144	-0.237
	(0.669)	(0.612)	(0.380)	(0.352)	(0.234)	(0.192)	(0.219)
Age × attrition	0.013	-0.048	0.041**	0.012	-0.014	-0.002	0.011
	(0.031)	(0.028)	(0.016)	(0.013)	(0.012)	(0.008)	(0.009)
Region (*Mchinji*)							
Rumphi × attrition	1.206	0.161	-0.085	-0.024	0.152	0.454	-0.088
	(0.861)	(0.788)	(0.476)	(0.439)	(0.283)	(0.243)	(0.266)
Balaka × attrition	-0.392	0.898	0.566	0.353	0.097	-0.200	-0.200
	(0.794)	(0.726)	(0.428)	(0.387)	(0.264)	(0.214)	(0.238)
Schooling (*No schooling*)							
Primary × attrition	-2.276**	-0.271	0.784	-0.273	0.339	-0.536*	0.174
	(0.865)	(0.791)	(0.455)	(0.412)	(0.299)	(0.223)	(0.254)
Secondary × attrition	-3.102*	-0.917	0.337	-0.018	0.053	-1.149**	0.056
	(1.252)	(1.145)	(0.734)	(0.673)	(0.435)	(0.388)	(0.425)
Children × attrition	0.015	0.160	-0.159*	-0.015	0.006	-0.011	0.016
	(0.124)	(0.113)	(0.063)	(0.063)	(0.040)	(0.032)	(0.035)
Married × attrition	-1.844*	-0.395	0.759	0.596	-0.160	0.004	0.198
	(0.848)	(0.776)	(0.599)	(0.402)	(0.384)	(0.297)	(0.307)
Attrition (effect on constant)	2.045	0.902	-2.106*	0.165	0.310	0.579	-0.352
	(1.392)	(1.273)	(0.904)	(0.691)	(0.579)	(0.449)	(0.491)
Observations (*N*)	3,021	3,021	2,662	2,978	2,655	2,656	2,642
χ^2***-tests (F-tests for OLS) for joint effects of attrition on:***							
Constants only	2.16	0.50	5.43*	0.06	0.29	1.66	0.51
	[0.142]	[0.479]	[0.020]	[0.811]	[0.593]	[0.198]	[0.474]
Coefficients only	1.84	0.85	1.84	7.50	4.04	11.84	5.10
	[0.066]	[0.55]	[0.066]	[0.483]	[0.854]	[0.158]	[0.747]
Constants and coefficients	1.82	1.04	1.77	38.66***	4.94	14.27	10.33
	[0.060]	[0.40]	[0.069]	[0.000]	[0.839]	[0.113]	[0.324]

Notes: Results of OLS/logit regressions of different outcomes on key individual characteristics, all measured in 2008, with all coefficients interacted with an indicator for subsequent attrition during 2008–10. Only interaction effects are shown, first-order effects are not reported. Standard errors are in round parentheses. Constants and cut-points (for ordered logit models) are not reported. Children = # of living children. Married = currently married. Reference categories: HIV/AIDS worries: *not at all*; Subjective likelihood of being HIV+: *none*. Numbers in brackets [ ] represent p-values for χ^2-tests or F-tests. p-values: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

however, and perhaps contrary to expectations, several key outcome measures—including SF12 physical and mental health, lifetime number of sexual partners, condom use, HIV status, HIV worries and risk perceptions—do not seem to be substantially different between attriters and non-attriters. While attrition is therefore predicted by several individual characteristics and outcome variables, the coefficient estimates in relationships between key outcome variables and individual characteristics are not necessarily affected by attrition. Specifically, for the majority of outcomes in Tables A16 and A17, the null-hypothesis that the estimated

Table A17: MLSFH attrition 2008–2010: OLS, ordered logit, and logit models for selected key outcome variables in 2008, with interaction for respondents who subsequently attrited during 2008–10

	SF12 mental health score	SF12 physical health score	Lifetime number of sexual partners	HIV positive	HIV/ AIDS worry	Subj. likelihood of being HIV+
	OLS	OLS	OLS	Logit	Ordered logit	Ordered logit
	(1)	(2)	(3)	(4)	(5)	(6)
Interactions with attrition:						
Male × attrition	0.943	1.470	-0.696*	-0.814*	0.177	0.420**
	(0.860)	(0.777)	(0.348)	(0.398)	(0.154)	(0.152)
Age × attrition	0.010	-0.038	0.000	-0.003	-0.004	-0.004
	(0.030)	(0.027)	(0.016)	(0.014)	(0.006)	(0.006)
Region (*Mchinji*)						
Rumphi × attrition	1.129	-0.961	0.478	0.602	0.141	0.243
	(1.086)	(0.982)	(0.456)	(0.479)	(0.201)	(0.201)
Balaka × attrition	0.318	-0.584	0.163	0.394	0.065	0.165
	(0.984)	(0.889)	(0.397)	(0.430)	(0.176)	(0.172)
Schooling (*No schooling*)						
Primary × attrition	-2.226	-0.844	-0.410	-0.355	0.140	0.060
	(1.136)	(1.027)	(0.464)	(0.457)	(0.198)	(0.197)
Secondary × attrition	-3.818*	-1.674	-0.501	0.301	-0.025	0.400
	(1.706)	(1.542)	(0.663)	(0.674)	(0.293)	(0.291)
Children × attrition	0.013	0.220	0.071	0.071	0.043	-0.012
	(0.187)	(0.169)	(0.081)	(0.078)	(0.032)	(0.032)
Married × attrition	1.042	-0.819	0.329	0.944*	-0.020	0.088
	(1.080)	(0.976)	(0.475)	(0.425)	(0.205)	(0.206)
Attrition (effect on constant)	-0.211	1.381	-0.226	-0.284	-0.222	-0.207
	(1.874)	(1.694)	(0.807)	(0.803)	(0.347)	(0.349)
Observations (*N*)	2,982	2,982	3,408	3,265	3,857	3,800
χ^2-tests (F-tests for OLS) for joint effects of attrition on:						
Constants only	0.01	0.66	0.08	0.12	0.41	0.35
	[0.910]	[0.415]	[0.780]	[0.724]	[0.522]	[0.552]
Coefficients only	1.28	1.00	1.11	13.29	5.37	17.33*
	[0.251]	[0.431]	[0.350]	[0.102]	[0.718]	[0.027]
Constants and coefficients	1.13	1.16	1.09	24.43**	5.45	18.09*
	[0.334]	[0.320]	[0.364]	[0.004]	[0.793]	[0.034]

Notes: Results of OLS/logit regressions of different outcomes on key individual characteristics, all measured in 2008, with all coefficients interacted with an indicator for subsequent attrition during 2008–10. Only interaction effects are shown, first-order effects are not reported. Standard errors are in round parentheses. Constants and cut-points (for ordered logit models) are not reported. Children = # of living children. Married = currently married. Reference categories: HIV/AIDS worries: *not at all*; Subjective likelihood of being HIV+: *none*. Numbers in brackets [ ] represent *p*-values for χ^2-tests or *F*-tests. *p*-values: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

coefficients in these relationships are identical for attriters and non-attriters is not rejected. For SF12 physical health score, SF12 mental health score, HIV-positive status, number of sexual partners, condom use and worries about HIV this null hypothesis is not rejected when either longer-term attrition during 2006–10 or shorter-term attrition during 2008–10 are considered. In none of the outcomes in Tables A16–A17 is the null-hypothesis rejected for both shorter- and longer-term attrition.

The analyses of attrition reported here therefore confirm our previous findings that, while attrition in the MLSFH is substantial and predicted by several observ-

able characteristics, attrition does not necessarily bias the coefficients of estimated relationships. Thus, the attrition levels observed in the MLSFH may not necessarily represent a general problem for obtaining consistent estimates of the coefficients of interest for most of these outcomes. These results, which are very similar to those documented in other contexts,[121–123] suggest that multivariate estimates of behavioral relations may not be biased due to attrition and thus support the collection of longitudinal data. And while the attrition analyses reported here do not substitute for analyses of the potential biases caused by attrition in the context of specific MLSFH research projects, the results reported here substantially alleviate concerns about attrition-related biases in the MLSFH. Despite this conclusion, however, the MLSFH has made efforts to re-contact and re-interview respondents who were lost to follow-up. The 2007 MLSFH Migration Follow-up (Appendix A2.2) has previously traced and interviewed MLSFH respondents who attrited in 2006, and an ongoing migration follow-up project in 2013 traces and reinterviews MLSFH respondents who attrited in 2010. This follow-up is likely to find a large number of MLSFH respondents who have been lost to follow-up as a substantial proportion of the migration leading to attrition is relatively local or to a small set of destinations: based on the migrant tracking information collected in 2013 ($N \approx 1,150$), 64% of ever-interviewed MLSFH respondents who were lost-to-follow-up in 2010 remained within the same district, while 12% had moved to one of Malawi's four largest cities, and 24% had moved to another rural/peri-urban area. The forthcoming availability of this 2013 migration follow-up will further reduce the concerns about attrition in the MLSFH, as well as enable more detailed analyses of the processes leading to migration and attrition.

A6. Specific features of the MLSFH data and study design

In the subsequent sections, we provide detailed information about some specific features of the MLSFH data and the MLSFH study design that have been relevant to a broad set of MLSFH analyses.

A6.1. Social network data in the MLSFH

A unique aspect of the MLSFH is the inclusion of longitudinal data on social networks that measure women's and men's social interactions about family planning or the HIV/AIDS epidemic.[33–41] In particular, the data include information on egocentric networks, that is, networks that contain the respondent and network partners with whom the respondent had chatted about family planning (MLSFH Rounds 1998 and 2001) or HIV/AIDS (MLSFH Rounds 1998, 2001, 2004 and 2006), with detailed information on up to four network partners. The name generator for these ego-centric conversational networks is shown in Figure A3. This MLSFH questionnaire module on social networks began by first asking the respondents about how many people they had chatted with about these respective topics, where the term "*chat*" was used to indicate informal conversations rather than lectures at clinics. Names (or nick-names of up to four network partners were recorded. Subsequent questions then asked for each of these named conversational network

Figure A3: Name generator for ego-centric conversational networks

Next, I'd like to ask you some questions about people you've chatted with about AIDS

A23a *How many people have you chatted with about AIDS? I mean people other than your husband or partner.* **IF LESS THAN FOUR ARE NAMED, PROBE**: "Can you think of anyone else? How about sitting in on a conversation, even if you yourself didn't say anything?"	Total number named..... [_____] **If none are named <u>after probing</u>, skip to A39a**
A23b *Could you please give me the names of four of these? As I said earlier, this information will be completely confidential. You can also make up names, if you feel more comfortable.* **WRITE THE FOUR NAMES, AND START ASKING THE QUESTIONS BELOW FOR EACH PARTNER NAMED ON THE RIGHT**	NAME: #1. _____ #2. _____ #3. _____ #4. _____

Fill in the names for A24 as follows. Check the [__] box after each task.

a. [__] Copy the first name listed on LINE #1 in A23b to column "NWP #1" in Question A24
b. [__] Copy the names on LINES #2, #3 and #4 in A23b to columns "NWP #2", "NWP #3", "NWP #4" in Question A24
c. [__] Copy the names in A24 to the first row of the continuation pages of the below table with questions A24a to A38d Keep the same sequence of names.

Starting with NWP#1, <u>column by column</u>, ask questions A24a to A38d for the persons listed under NWP#1, NWP#2, NWP#3 and NPW#4

Question	Code	NWP #1	NWP #2	NWP #3	NWP #4
A24 NAME (copy name from A23b)					
A24a Is [NAME] male or female?	Male	1	1	1	1
	Female	2	2	2	2

Questionnaire social network module continues

Notes: For ego-centric conversational networks about HIV/AIDS from 2006 MLSFH questionnaire; the identical name generator was used during the MLSFH 1998–2006. A corresponding data were also collected for conversational networks about family planning (MLSFH 1998–2001) and for religion (MLSFH 2004)

partners a set of questions, including about (*i*) characteristics of the network partner (gender, age, location of residence, schooling, religion, marital status, wealth), (*ii*) characteristics of the relationship between the respondent and the network partner (familial relationship to respondent, closeness of the relationship, frequency of contact, etc.), and the content of the respondent's conversations with the network partner (for instance, in the context of HIV/AIDS-related conversational networks: the network partner's stated level of worries about HIV/AIDS and his/her subjective infection probability, the network partner's perceptions about infidelity about his/her spouse, etc.; in the context of family-planning-related networks: the network partners reported use of family planning methods and the spousal approval of this use). The specific question regarding the risk perception of the network partners, for example, was phrased as "How worried is *name of network partner* about getting AIDS?" with the same response categories as for the respondent (no risk, moderate risk, great risk). In addition, the MLSFH also asked respondents about the relationship among each of the nominated network partners to facilitate the cal-

Table A18: Summary of MLSFH HIV/AIDS conversational networks 1998–2001

	Females		Males	
	1998	2001	1998	2001
Characteristics of respondents with HIV/AIDS conversational networks				
N	1,179	1,159	806	799
Age	31.1	34.3	37.0	40.4
	(9.26)	(9.39)	(10.43)	(10.96)
Not Married	0.11	0.11	0.01	0.03
Children ever born	4.38	5.11	5.28	6.17
	(3.05)	(2.89)	(4.20)	(3.98)
Perceived AIDS risk, respondent				
Proportion perceiving no risk	0.17	0.29	0.27	0.42
Proportion perceiving moderate risk	0.21	0.23	0.19	0.21
Proportion perceiving great risk	0.61	0.47	0.53	0.37
Descriptive statistics for HIV/AIDS conversational network				
Prop. with at least one nwp in AIDS network	0.83	0.95	0.92	0.97
Uncensored size of AIDS network	4.33	5.84	6.24	7.04
	(5.14)	(5.57)	(6.46)	(6.92)
Censored size of AIDS network	2.53	3.42	3.08	3.56
	(1.50)	(1.09)	(1.26)	(0.95)
Proportion with more than 4 network partners	0.28	0.42	0.43	0.49
Prop. with at least one nwp who perceives great AIDS risk	0.61	0.52	0.67	0.47
Number of nwp who perceive great risk	1.46	1.06	1.77	1.05
	(1.49)	(1.28)	(1.59)	(1.35)
Prop. with at least one nwp who perceives moderate AIDS risk	0.31	0.45	0.32	0.43
Number of nwp who perceive moderate AIDS risk	0.50	0.71	0.54	0.71
	(0.87)	(0.95)	(0.94)	(1.03)
Proportion with at least one nwp who perceives no AIDS risk	0.26	0.57	0.30	0.58
Number of nwp who perceive no AIDS risk	0.48	1.12	0.68	1.24
	(0.94)	(1.23)	(1.20)	(1.32)

Notes: 'nwp(s)' = network partner(s). Uncensored size of the network is the mean response to the question about the number of network partners (Question A23a in Figure A3), and the uncensored network size is the mean number of network partners that were listed with name (censored at four) (Question A23b in Figure A3).
Source: Kohler et al.[41]

culation of network densities and related measures to describe the structure of the conversational networks, which can be important for identifying the mechanisms through which social interactions affect individual behaviors.[43] For illustration of these data on conversational networks in the MLSFH, selected descriptive statistics for the network about HIV/AIDS conversation in the 1998 and 2001 MLSFH are reported in Table A18.

Table A19: Categories of individuals included in the MLSFH Household/family roster from 2008 onward

Individuals listed in MLSFH Household/family roster
1. List the respondent
2. List name of spouse(s) of respondent. If respondent is not currently married, list name of most recently deceased or divorced spouse. For polygamous men: list all wives. If never married proceed to instruction 3, below.
3. List name of respondents parents (list names even if parents are deceased)
4. [if R is married or widowed] List name of spouses parents (list names even if parents are deceased; for polygamous men: list parents of all wives)
5. List the names of all children of the respondent (children ever born; include children who are no longer alive or do not live in respondents household)
6. List the names of any other children who usually live in this household (including non-biological children, grandchildren, nieces & nephews).
7. List the names of all other persons who slept in this household last night
8. List the names of all other persons who usually sleep in this household, but did not last night
9. List the names of all non-related children who are under your care but not living in the household (for example, anyone you have helped with school fees in the last 5 years).

A6.2. Household/family rosters in the MLSFH

An innovation of the 2006 data collection, which was continued in the 2008, 2010 and 2012 MLSFH waves, was the expansion of data on family structure and financial/non-financial transfers that is collected as part of the MLSFH. Specifically, starting in 2006, the MLSFH household and family roster included not only all individuals who currently live in the household as frequently done in other studies, but it also asked information about all parents and children independent of their survival and resident status (Table A19), including selected demographic, socioeconomic characteristics and information about the household/family members health as known to/perceived by the respondent (Table A20).

For all persons listed on the MLSFH household/family roster who were above age 15 and alive at the time of the survey (or had died within less than two years prior to the survey), the MLSFH asked a set of questions about transfers given to and received from the respondent. Since the quantitative measurement of transfers in contexts such as Malawi is inherently difficult, the MLSFH did not attempt to monetize the financial and non-financial transfers between respondents and their children or parents. Instead, for all alive parents and children above age 15, MLSFH respondents were asked a set of questions about financial and non-financial assistance during the last two years, including: (*i*) "*In the past two years, have you given [name] any money or financial assistance?*", with responses ranging from: 0 = no; 1 = yes, a little; 2 = yes, some; and 3 = yes, a lot; (*ii*) "*In the past two years, have you given [name] any non-financial help? This could include help that takes time like collect-*

Table A20: Socioeconomic and health information reported by MLSFH respondent for each indivdual included in the MLSFH Household/family roster (from 2008 onward)

Information about each person listed on the MLSFH household/family roster	
Q2	What is [name's] relationship to you?
Q3	Is [name] male or female?
Q4	Is [name] alive? If [name] is dead, when did he/she die? (*Note:* Questions Q5–16 were not asked for persons who had died)
Q5	How old is [name]? Or, in what year was [name] born?
Q6	Where does [name] usually live?
Q7	Did [name] sleep here last night?
Q8	If a person does not regularly live here: when did [name] move to this place?
Q9	Has [name] been ill in the past 12 months? If yes, for how long?
Q10	How would you rate [name's] health in general?
Q11	How would you compare [name's] health to other people in your village who are the same age and sex?
Q12	What is [name's] current marital status?
Q13–14	What is the highest level of schooling name completed? How many grades (in years) did [name] complete at that level?
Q15	If age $>$ 10: What is [name's] main way of earning money?

ing firewood, cooking, taking care of people, or helping with farming.", with responses ranging from 0 = no; 1 = yes, once; 2 = yes, several times a year; 3 = yes, at least once a month; 4 = yes, at least once a week; and 5 = Yes, daily; (*iii*) "*In the past two years, has [name] given you any money or financial assistance?*", with responses ranging from: 0 = no; 1 = yes, a little; 2 = yes, some; and 3 = yes, a lot; and (*iv*) "*In the past two years, has [name] given you any non-financial help? This could include help that takes time like collecting firewood, cooking, taking care of people, or helping with farming.*", with responses ranging from 0 = no; 1 = yes, once; 2 = yes, several times per year; 3 = yes, at least once a month; 4 = yes, at least once per week; and 5 = yes, daily.

For persons who were reported as having died during the previous two years on the MLSFH household/family roster, the MLSFH also asked more detailed information about when the death occurred, how old the person was when he/she died, the level of schooling and the marital status of the diseased person, the health prior to the dying, and the likelihood (as perceived by the respondent) that the death was due to AIDS. There were also questions about health care and funeral costs incurred by the respondent in connection with the death of the person.

The 2008 and 2010 MLSFH also asked respondents to list up to 10 persons whom they would ask for assistance during a crisis (e.g., famine, health problems or other events that may lead to economic shortages in the household). The respondent was then asked a set of questions about basic demographic/socioeconomic characteristics of each listed person, followed by a set of questions about financial and non-financial transfers that the respondent had given to and/or received from each of the listed persons.

A6.3. Probabilistic expectation data in the MLSFH

Since 2006, the MLSFH has included a module eliciting *probabilistic expectations*, that is, expectations that are measured on a well-defined numerical scale, are comparable across domains, and can be consistently interpreted as probabilities. These expectations data cover domains such as own and village-level mortality risks, risk of becoming infected with HIV during a single intercourse, and respondent's perception about being infected with HIV, and have proven useful for uncovering how respondents in rural Malawi perceive HIV and mortality risk, how they respond to new information about their HIV status, and how perceptions about HIV affect risk-taking behaviors.[64–66,83,84,141,142] In Figure 5, for example, we illustrate respondents' subjective probability of dying within a 5-year time period (by gender, region and age group), and related analyses have shown that these survival expectations are importantly related to HIV risk taking and sexual behaviors.

In order to elicit probabilistic expectations in the relatively low literacy and numeracy context of rural Malawi, the MLSFH developed an interactive elicitation technique that relied on asking respondents to allocate up to ten beans on a plate to express the likelihood that an event will be realized (Figure A4).[64] This bean format has the advantage of being visual, relatively intuitive and fairly engaging for respondents, and can be designed to improve the consistency of answers. Following an introductory text and example (Figure A4), respondents were first asked a training question about the probability of winning in a local board game (Bawo), followed by a question about the likelihood of a newborn baby dying before his first birthday. To evaluate whether respondents understand the concept of probability, respondents were then asked about two *nested* events: going to the market within (*a*) *two days*, and (*b*) *two weeks*. If respondents understand the concept of probability, they should provide an answer for the two-week period that is greater than or equal to the one of the two-day period. Interviewers were instructed to leave the number of beans on the plate after the respondents had responded to the likelihood of going to the market within two days, thereby ensuring that s/he remembered the answer when answering about the two-week period in the next question. If the respondent violated the monotonicity property, the interviewer was instructed to explain the incoherency of the answers by stating that: "*as time goes by, you may find more time to go to the market. Therefore, you should have added beans to the plate.*" And the respondent was invited to reformulate the answer. For this first set of training questions, the interviewers were also instructed to prompt the respondent if s/he allocated 0 or 10 beans in the plate.

Respondents were then asked a set of questions related to economic outcomes, health outcomes, and risk-prevention strategies (Figure A4), including: (*a*) going to market within the next *2 days*; (*b*) going to the market within the next *2 weeks*; (*c*) experiencing a food shortage within the next 12 months; (*d*) having to rely on family members for financial assistance in the next 12 months; (*e*) being infected with HIV now; (*f*) using condom at the next sexual encounter with a spouse; (*g*) using condom at the next sexual encounter with someone other than spouse (not asked if

Figure A4: 2006 MLSFH questionnaire module about subjective expectations

INTERVIEWER: Put the plate and the cup side by side. Recount the number of beans and check that you have 10 beans in the cup [__]. As you provide the explanation below, add the beans into the plate to illustrate what you say.

"I will ask you several questions about the chance or likelihood that certain events are going to happen. There are 10 beans in the cup. I would like you to choose some beans out of these 10 beans and put them in the plate to express what you think the likelihood or chance is of a specific event happening. One bean represents one chance out of 10. If you do not put any beans in the plate, it means you are sure that the event will NOT happen. As you add beans, it means that you think the likelihood that the event happens increases. For example, if you put one or two beans, it means you think the event is not likely to happen but it is still possible. If you pick 5 beans, it means that it is just as likely it happens as it does not happen (fifty-fifty). If you pick 6 beans, it means the event is slightly more likely to happen than not to happen. If you put 10 beans in the plate, it means you are sure the event will happen. There is no right or wrong answer, I just want to know what you think.

Let me give you an example. Imagine that we are playing Bawo. Say, when asked about the chance that you will win, you put 7 beans in the plate. This means that you believe you would win 7 out of 10 games on average if we play for a long time.

INTERVIEWER: Report for each question the NUMBER OF BEANS put in the PLATE. After each question, replace the beans in the cup (unless otherwise noted).

For questions X1a to X1f: If respondent puts 10 (or 0) beans, prompt "Are you sure that this event will almost surely (not) happen?" CIRCLE 1 in column P if you prompted the respondent, and report the final answer only.

X1 Pick the number of beans that reflects how likely you think it is that…	# of beans in plate	Prompt for 0 or 10?
a) *you will win if we play a game of Bawo after this interview*	[____]	1
b) *a baby born in your community this month will die within one year*	[____]	1
c) *you will go to the market at least once within the next 2 days* (LEAVE BEANS IN PLATE)	[____]	1
d) *you will go to the market at least once within the next 2 weeks?*	[____]	1
INTERVIEWER: Did Respondent add any beans between X1c and X1d?	**If yes → X1f**	
e) *Remember, as time goes by, you may find more time to go to the market. Therefore, you should have added beans to the plate. Let me ask you again. Now, add beans in the plate so that the number of beans in the plate reflects how likely you think it is that you will go the market at least once within 2 weeks?*	[____]	1
f) *you will experience shortage of food in the next 12 months?*	[____]	1

Continued on next page

respondent reports sex only with spouse); and (*h*) the respondent dying within (*i*) *1 year*; (*i*) *5 years*; and (*ii*) *10 years*. The mortality questions were designed to ensure that respondents provided answers that would allow us to construct well-defined survival curves. In particular, respondents were first asked to pick the number of beans that reflects how likely it is that they will die within a one-year period beginning today. Then, with the beans of the previous question still on the plate, they were asked to *add* more beans to reflect how likely it is that they would die within a five-year period. The same procedure was followed for the ten-year period mortality question. This ensured that respondents provided weakly increasing answers when the time horizon increased.

The MLSFH expectation module was implemented in 2006, 2008, 2010 and 2012

Figure A4: 2006 MLSFH questionnaire module about subjective expectations

Continued from previous page

For the subsequent questions, no longer prompt for "0" and "10" answers

X2	Pick the number of beans that reflects how likely you think it is that…	# of beans in plate
	a) *you will have to rely on family members for financial assistance in the next 12 months*	[____]
	b) *you are infected with HIV/AIDS now*	[____]
FOR MARRIED RESPONDENTS (INTERVIEWER: If respondent is not married → X2f)		
	c) *your spouse is infected with HIV/AIDS now*	[____]
	d) *you will use condom the next time you have sex with your spouse*	[____]
	e) *you will use condom the next time you have sex with someone else other than your spouse* (INTERVIEWER: If sex only with spouse, write 99)	[____] → X3
FOR UNMARRIED RESPONDENTS		
	f) *your romantic partner is infected with HIV/AIDS now* (INTERVIEWER: If no romantic partner, write 99 and → X2h)	[____]
	g) *you will use condom the next time you have sex with your romantic partner* (INTERVIEWER: if no romantic partner, write 99)	[____]
	h) *you will use condom the next time you have sex with someone you just met* (INTERVIEWER: If no sex with someone just met, write 99)	[____]
	i) *you will be married one year from now*	[____]

Finally, I would like to ask you to consider the likelihood that you may not be alive as time goes by. We hope that nothing bad will happen to you, but nevertheless, something unfortunate may occur over the next years despite all precautions that you may take. If you don't want to, you do not need to answer this question.

INTERVIEWER: If respondent refuses to answer, skip to **X8.**

		# OF BEANS in plate
X6	*Pick the number of beans that reflects how likely you think it is that you will die within a one-year period beginning today.* (LEAVE BEANS ON PLATE)	[____] if 10 →X8
X7	Put additional beans so that the number of beans in the plate reflects how likely you think it is that you…	
	a) *will die within a five-year period beginning today* (LEAVE BEANS ON PLATE; IT IS POSSIBLE TO ADD ZERO ADDITIONAL BEANS)	[____] if 10 → X8
	b) *will die within a ten-year period beginning today* (IT IS POSSIBLE TO ADD ZERO ADDITIONAL BEANS. PUT BEANS BACK IN CUP AFTER RECORDING THE ANSWER)	[____]

(with only minor changes in the content over time), providing up to four observations over seven years of detailed subjective expectations data in the MLSFH. A detailed evaluation of the subjective probability module in the MLSFH concluded that the reported expectations were remarkably consistent with a basic property of probability theory (the monotonicity of nested events), and vary in meaningful ways with individual or contextual characteristics. In addition, respondents had relatively well-calibrated beliefs about infant mortality, but were greatly pessimistic about their own survival—perhaps as an overreaction to the substantial

increases in adult mortality that have occurred as a result of HIV/AIDS in the last decade.[64] As an illustration of these data, Table A21 reports from the 2006 MLSFH the responses in terms of number of beans to the questions about going to the market, experiencing a food shortage, having to rely on family members, infant mortality, being infected with HIV, condom use, and mortality.[64] The implied subjective probabilities in this table were calculated by assuming that each number of beans between zero and ten corresponds to a specific probability interval between zero and one. This approach assumes that respondents choose the number of beans that best represents their subjective probability, and it reflects our beliefs that all respondents who place zero (ten) beans on the plate do not believe literally that this event has a probability of zero (one). Specifically, one plausible approach to assigning probabilities to each allocation of bean is as follows, which was also how the implied subjective probability P_i in Table A21 have been calculated: $P_i < 0.05$ for *zero* beans, $0.05 \leq P_i < 0.15$ for *one* bean, $\ldots$, $\frac{X_i}{10} - 0.05 \leq P_i < \frac{X_i}{10} + 0.05$ for X_i beans, $\ldots$, $0.85 \leq P_i < 0.95$ for *nine* beans, and $P_i \geq 0.95$ for *ten* beans, where X_i is the number of beans allocated by respondent i given his/her underlying subjective probability P_i.

To illustrate the correspondence of subjective probabilities in the MLSFH with the commonly asked verbal scales, Table A22 also compares respondents' answers to the question, "In your opinion, what is the likelihood (chance) that you are infected with HIV/AIDS now?" with the number of beans provided when asked how likely do you think it is that you are infected with HIV/AIDS now. It shows that respondents who provided a higher likelihood of being infected using the verbal scale were also more likely to provide a higher number of beans. For example, the modal answer is 0 beans among respondents who said "no likelihood", 1 bean among those who said "low likelihood" and 5 beans among those who said "medium likelihood." However, Table A22 also highlights that there was a great variation in what probability respondents associate with "low likelihood" or "medium likelihood." For example, a bit more than a quarter of the respondents who said "low likelihood" allocated 1 bean and another quarter allocated 2 beans; 18% allocated 4 beans and 12% allocated 0 beans. This suggests that the bean measure may be more informative to compare risk rankings across individuals.

A6.4. Marital histories and sexual behaviors

In addition to collecting the current marital status at the time of each MLSFH survey, the MLSFH has also obtained marital histories that include all marriages of MLSFH respondents along with key information pertaining to each marriage. Specifically, the MLSFH marriage histories asked respondents to list the start and end dates of all of their marriages (up to ten marriages). Subsequently, respondents were asked a series of questions about their current/most recent spouse, previous spouse, and first spouse. Questions included age at marriage, spousal age difference, how long they had known each other, educational attainment before marriage, where they were staying when they first met, spouse's ethnic group, number of children produced, whether husband had another wife when she mar-

Table A21: Subjective probabilities of various common events in the 2006 MLSFH

# beans	implied subj. prob.	Going to the market within: two days	Going to the market within: two wks	Experiencing food shortage next 12 months	Rely on family for financial assistance next 12 months	Baby dying before first b'day	Being infected with HIV now	Using condom at next sexual encounter with: spouse	Using condom at next sexual encounter with: other than spouse	Own Mortality: Probability of dying within: 1-yr	Own Mortality: Probability of dying within: 5-yrs	Own Mortality: Probability of dying within: 10-yrs
0	0 to .05	6.0	1.1	11.2	22.2	18.4	66.7	64.3	10.1	29.1	6.0	2.0
1	.05 to .15	9.6	1.3	7.7	11.7	22.4	9.6	7.5	1.5	24.2	8.4	2.0
2	.15 to .25	17.7	2.5	10.2	11.9	17.7	7.5	6.7	3.8	16.2	16.2	4.9
3	.25 to .35	15.0	6.3	8.9	8.6	10.7	4.8	4.4	3.6	8.9	16.6	8.0
4	.35 to .45	14.6	7.0	9.2	6.8	7.3	2.3	2.9	3.4	5.4	12.2	10.0
5	.45 to .55	14.5	12.4	17.3	13.2	17.6	5.6	6.1	11.2	13.2	20.2	24.9
6	.55 to .65	7.9	14.1	8.2	6.7	2.8	0.5	2.6	8.2	0.8	7.1	10.1
7	.65 to .75	4.9	15.3	7.4	4.9	1.3	0.5	0.9	5.3	0.4	5.9	11.2
8	.75 to .85	3.5	13.3	7.1	4.9	0.9	0.9	1.4	8.8	0.5	3.3	9.8
9	.85 to .95	1.8	6.7	3.0	2.4	0.4	0.7	0.5	7.4	0.3	0.9	4.9
10	.95 to 1	4.5	20.1	9.8	6.8	0.7	1.0	2.7	36.6	0.9	3.3	12.2
Total (Percent)		100	100	100	100	100	100	100	100	100	100	100
Implied subjective probability												
Mean		0.39	0.67	0.46	0.36	0.25	0.12	0.16	0.67	0.20	0.39	0.58
10^{th} percentile		0.09	0.33	0.05	0.02	0.0	0.01	0.01	0.05	0.02	0.10	0.27
25^{th} percentile		0.20	0.51	0.21	0.07	0.0	0.02	0.02	0.47	0.04	0.22	0.43
Median		0.36	0.69	0.47	0.30	0.20	0.04	0.04	0.78	0.14	0.37	0.54
75^{th} percentile		0.53	0.88	0.68	0.56	0.4	0.14	0.20	0.97	0.31	0.53	0.77
90^{th} percentile		0.75	0.98	0.94	0.83	0.5	0.41	0.52	0.99	0.50	0.71	0.96
N		3,170	3,157	3,145	3,163	3,171	3,162	2,583	582	3,125	3,125	3,123

Source: Delavande & Kohler [64]

Table A22: Comparison of probabilistic expectation and likelihood-based verbal scale about the likelihood of being infected with HIV

Probabilistic expectation:		Response on verbal likelihood scale					
# beans	subjective probability	None	Low	Medium	High	Don't know	Total
0	0 to .05	88.84	11.43	6.29	1.89	23.81	66.75
1	.05 to .15	5.75	27.06	1.26	0.00	28.57	9.50
2	.15 to .25	3.03	25.71	5.03	1.89	19.05	7.47
3	.25 to .35	1.32	18.32	5.66	0.94	9.52	4.78
4	.35 to .45	0.40	9.24	5.66	0.00	0.00	2.31
5	.45 to .55	0.35	7.39	71.07	6.60	19.05	5.57
6	.55 to .65	0.04	0.17	3.77	8.49	0.00	0.54
7	.65 to .75	0.04	0.50	0.63	10.38	0.00	0.51
8	.75 to .85	0.04	0.00	0.00	25.47	0.00	0.89
9	.85 to .95	0.04	0.00	0.63	18.87	0.00	0.70
10	.95 to 1	0.13	0.17	0.00	25.47	0.00	0.98
Total (Percent)		100.00	100.00	100.00	100.00	100.00	100.00
Implied subjective probability							
Mean		0.04	0.22	0.44	0.79	0.20	0.12
10^{th} percentile		0.01	0.04	0.20	0.53	0.02	0.01
25^{th} percentile		0.01	0.10	0.45	0.70	0.05	0.02
Median		0.03	0.20	0.49	0.83	0.14	0.04
75^{th} percentile		0.04	0.31	0.52	0.95	0.29	0.14
90^{th} percentile		0.07	0.43	0.54	0.98	0.50	0.42
N		2,277	595	159	106	21	3,158

Source: Delavande & Kohler [64]

ried him (females only), rank order of wife (females only), and suspected spouse having sexual relations with other women. For marriages that had ended, additional questions were asked about the duration of the marriage, how it ended, and whether the respondent remarried. In 2004, the format of marriage histories changed. Respondents listed up to five marriages starting from their first marriage and answered a series of questions about each marriage. Questions asked included marriage start and end dates (if the marriage had ended), age at marriage, spousal age difference, whether husband had other wives, number of children produced, how the marriage ended, and main reason for divorce/separation.

In 2006, the MLSFH increased the number of marriages that could be listed in marriage histories, up to a maximum of 10, and decreased the number of questions asked about each spouse. For the first time, respondents reported the names of all spouses to whom they were ever married. For each spouse, respondents reported the year marriage began, how many children they had with that spouse, whether

or not they were still married to the spouse, if they had ever talked about HIV with their spouse, and if they knew the HIV status of the spouse at the time they got married. If the marriage had ended, they reported the year it ended and the main reason why it ended. With the exception of a few minor changes, the format of marriage histories in 2010 was similar to that of the previous wave. While the MLSFH did not ask about the number of children produced from the marriage, they did ask if the respondent knew his/her HIV status at the time of marriage and whether the respondent's HIV status was the same as that of his or spouse when they got married. In 2012, the MLSFH did not collect full marriage histories. Instead, they asked respondents if they were still married to the spouse(s) they were married to in 2010. If no longer married, they reported how and when the marriage ended. While the MLSFH did not collect information on new spouses, they did ask respondents how many times they had ever been married. This information, in conjunction with data from the previous wave, could be used to determine whether respondents remarried between the 2010 and 2012 waves.

As part of a study investigating the reliability of marriage histories, reconstructed marriage histories were created for respondents who participated in the 2006 and 2010 MLSFH.[74] They were initially created in response to tabulations that a sizable proportion of respondents reported declines in the number of times married over time in the MLSFH. Due to its potential effect on marriage-related analyses, an attempt was made to reconstruct a more complete set of marriage histories using data from the 2006 and 2010 MLSFH. Although reconstructed marriage histories (RMH) may not be entirely complete, they should represent a lower bound in the true number of marriages in the sample. The process of reconstructing marriage histories consisted of two parts. First, marriages were matched across surveys for each respondent. Because names tend to be spelled differently across survey waves, mostly due to the interpretation of the interviewer, marriages were visually matched on a case-by-case basis. Spouse name and dates of marriage were used to confirm that a marriage listed in 2006 is the same as a marriage listed in 2010. As a next step, all of the marriages reported in 2006 and 2010, even marriages reported only once, are listed in the RMH. These histories contain information on marriage start and end dates (in cases of terminated marriages) and status of marriage. If a marriage is reported only once, then information listed in the original marriage history is used to create the RMH. If a marriage is reported in both waves, then reports of status of marriage, marriage start date, and marriage end date are compared between the two waves. If they are consistent, then they are used to create the RMH. If inconsistent reports are given, then information provided in the earlier survey is used, if reported by the respondent. Data from the earlier survey are used because the marriage in question would have happened closer in time to this survey. If a respondent reports "don't know", then data from the later survey is used, if this information was reported. In cases where marriage dates overlap (e.g., there has been a case in the data where the reconstructed dates of her first and second marriage are 1995–2000 and 1999–2006, respectively), corrections were made

to marriage start and end dates listed in the RMH. In these cases, data from the later survey was used.[74]

A6.5. Spouse, children and parent linkages in the MLSFH

A6.5.a. MLSFH spouse linkages: The MLSFH sample has been based on married couples since its inception in 1998. Thus, establishing and maintaining linkages between spouses has consistently been an important component of MLSFH research and data collection. To create such linkages between husbands and wives in the MLSFH sample, we have taken the following steps. In addition to individual IDs, each married respondent in the MLSFH sample also has a *spouse ID* that links to the husband or wife to whom the individual is married. For polygamous men, linkages for all spouses are included; MLSFH polygamous men have up to six spouse IDs. Given the large amount of marital turnover in Malawi, these linkages can change over time, and maintenance is required in order to update the linkages. To ensure that the spouse linkages are accurate in each wave, the MLSFH has taken two steps: (1) during *data collection*: as described above, MLSFH survey supervisors are given lists of all MLSFH respondents to be interviewed in each village, which also contains identifying information for the respondent, such as the names of their spouses. These lists also contain a section that requires the supervisor to update the spouse link after the interviewer completes the survey. Upon returning with a completed survey, the supervisor indicates whether the marriage from the prior MLSFH wave is still active, and if not, and the respondent has remarried, a new ID is created for the new spouse and this information is updated in the list. This list is then entered at the end of each day of fieldwork and the dataset subsequently updated; and (2) during *data cleaning*: after data is collected for all MLSFH respondents, the spouse IDs are cleaned. This involves using information from marriage rosters to identify individuals who experienced marital dissolution in between waves, dropping spouse IDs for currently unmarried respondents, and ensuring that new spouse IDs are present for all who remarried in between waves.

This task is not without challenges, however. Not infrequently one spouse will not be interviewed during fieldwork, in which case we rely on information for only one individual in the marriage. Such information is not always reliable, as discrepancies between reporting of marriage have been found among MLSFH respondents: for example, in the case that a woman reports a marriage but the man reports being married to a different individual. In this case we include the man's ID as spouse of the woman, but not the woman's ID as spouse of the man.

A6.5.b. MLSFH parent-children linkages: As with spouse linkages, MLSFH also collects and maintains linkages of parents to children. MLSFH respondents could have a linkage to their parent or child through two different samples: (1) the 2004 MLSFH household roster, and (2) the 2008 MLSFH parents' sample. To establish parent-child linkages for each, we kept track of the index respondent in each case; so in 2004 we linked IDs for MLSFH women whose household roster was used to draw a 2004 MLSFH adolescent; and in 2008 we linked respondents who had a

parent added to the sample. In doing so, complications arose for several reasons. First, in many cases we were unaware of siblings in the MLSFH data. In this case two individuals could have listed the same people as parents in the 2008 sample, which would result in duplicates. We cleaned these by name matching and by identifying duplicates during 2010 data collection and removing the extra entry from the dataset. Secondly, some of the 2004 adolescents were drawn only from their mother's household roster, which means that they could be reliably linked with the mother but not the father.

A6.5.c. Longitudinal linkage of children listed in household/family rosters: To allow for longitudinal analyses of the information elicited in the above household and transfer rosters, the data on the respondent's children listed in the 2006–2010 MLSFH family and transfer rosters have been linked using names, ages, sex, and birth order. Because not all data were available in every wave, and because the spelling of names is not always exactly identical across waves, the matching was not undertaken with a computerized algorithm, but was done case-by-case instead. Two processes were undertaken. First, names were designated the principal matching variable; so to be consider matched, a minimum similarity in spelling was required. Second, a quality indicator for the quality of the match was assigned to each matched child, with the match being *low quality*, if no other data than the spelling itself was available to establish the match, and the spelling itself was not sufficiently similar across waves, *medium quality*, if any other variable was available (age, sex, birth order) to establish the match or, if no other data was available but the spelling matched very closely, and *high quality*, if two or more variables were available to establish the match.

To illustrate one example of the longitudinal child linkages in the MLSFH, Table A23 compares adolescent children (aged 15–20) that were listed by female MLSFH respondents in 2008 depending on whether they could be linked or not linked to the 2006 MLSFH household roster. Female respondents in 2008 had listed in the household rosters 952 children age 15–20, of whom the MLSFH data provide longitudinal linkages for 788 (83%), and 164 children (17%) could not be linked across the 2006 and 2008 household rosters. The main reasons for the failure to link include a misreporting of children and/or their names or other essential information in either the 2006 or the 2008 household roster; because for linked children the essential information was consistently reported in two rounds we expect the data quality for these children to be fairly high. However, linked and unlinked children were very similar in terms of their sociodemographic characteristics. Adolescent children were on average 17 years old, and 23% of them were married in 2008. A high proportion (65% of the linked and 60% of the unlinked children) was co-residing with their mothers. More than half of the adolescent children were enrolled in school, however linked children have on average one more grade of schooling compared to the unlinked children. According to the mothers' reports, over 40% of the adolescent children were in excellent health, about 40% in very good and 20% were in good/poor/very poor health.

Table A23: Descriptive statistics about linkage of children aged 15–20 in household rosters between 2006 and 2008 (female MLSFH respondents only)

	Youth child linked in HH roster		Youth not linked in HH roster		Total	
	mean	sd	mean	sd	mean	sd
N	788		164		952	
Youth characteristic in 2008						
Age	17.43	(1.68)	17.38	(1.80)	17.43	(1.70)
Female	0.51	(0.50)	0.49	(0.50)	0.51	(0.50)
Currently married	0.23	(0.42)	0.23	(0.42)	0.23	(0.42)
Coresident with parents†	0.65	(0.48)	0.60	(0.49)	0.64	(0.48)
Enrolled in school	0.58	(0.49)	0.52	(0.50)	0.57	(0.49)
Years of completed schooling	5.96	(2.71)	5.04	(2.88)	5.80	(2.76)
Subjective health (assessed by mother)						
excellent	0.41	(0.49)	0.43	(0.50)	0.41	(0.49)
very good	0.39	(0.49)	0.36	(0.48)	0.39	(0.49)
good/poor/very poor	0.19	(0.40)	0.21	(0.41)	0.20	(0.40)
Subjective health score‡	2.22	(0.75)	2.22	(0.77)	2.22	(0.75)

Notes: †: coresidence = residence in same household or compound
‡: subjective health score: 1 = good/poor/very poor; 2 = very good; 3 = excellent

A6.6. MLSFH Incentives Study: an experimental design offering financial incentives for maintaining HIV status during 2006–07

The MLSFH Incentives Study was an experimental design that offered financial incentives for maintaining HIV status during 2006–07 and was implemented subsequent to the 2006 MLSFH data collection.[22] The time line for this project is described in Table A24. In 2006, the MLSFH offered both couple and individual HIV testing and counseling, with the former being offered first to all married couples, and if one of the spouses opted out of couple HIV testing, individual HTC was offered to both members of a couple (Appendix A3.3). Unmarried individuals were only offered individual HTC. As indicated earlier (Appendix A3.3), 92% of the respondents who were offered an HIV test accepted the test during the 2006 MLSFH HTC. Among these, the HIV prevalence rate was 9.2%. For the 2006–07 MLSFH Incentives Study, adult respondents were selected in two steps: first, all adult individuals in HIV-discordant couples were selected; second, we randomly selected adult individuals from the 2006 MLSFH HTC participants. A total of 1,402 individuals who participated in the 2006 MLSFH HTC were offered to participate in this project. These individuals were approached about 1–2 months after the 2006 MLSFH HTC and introduced to the MLSFH Incentive Study. Couples who participated in couple HTC were offered to participate in the MLSFH Incentive Study as couples, with a fall-back option of joining the study as an individual if one of the spouses didn't want to participate. Individuals who participated in individual

Table A24: Time line for 2006–07 MLSFH experimental design offering financial incentives for maintaining HIV status

2006	May–July	HIV Initial Testing + MLSFH Surveys (2006 Round)
	August–December	Incentives Offered
2007	April–May	Round 1 Sexual Diaries
	July–October	Round 2 Sexual Diaries
2008	March–August	Round 3 Sexual Diaries + HIV Testing + Incentives Given
	March–August, 1–2 weeks after HTC	Round 4 Sexual Diaries

Source: Kohler & Thornton [22]

HTC were offered to participate as individuals. A total of 1,307 (or 93%) individuals were successfully enrolled during the informed consent process (either as part of couples or as individuals). Descriptive statistics for the MLSFH Incentive Study sample are provided in Table A25. 45% of the study population was male, the average age was 36 years, and majority of the sample (84%) were married.

The hypothesis underlying this project was that financial rewards for maintaining one's HIV status would result in changes in HIV-risk behaviors. And while only HIV-negative individuals could change their HIV status during the period, and thus not receive the offered award, the study was offered to both HIV-positive and HIV-negative individuals in the MLSFH to avoid that the exclusion from study participation could be interpreted as an indication of a persons HIV status by the MLSFH interviewers and/or family or community members. To assign the incentives that would be paid in case HIV status was maintained during the study period, each individual or couple randomly drew a token out of a bag to determine their incentive amount. The incentive amounts included zero, 500 Kwacha (approximately 4 dollars at the time), or 2,000 Kwacha (approximately 16 dollars at the time) for an individual, or zero, 1,000 Kwacha, or 4,000 Kwacha (approximately 32 dollars) for a couple. The incentives were distributed among the three levels, across both couples and individuals, with an equal probability of receiving each incentive amount. In practice, the realized (ex-post) distribution of the incentives resulted in 35% receiving zero, 32% receiving a medium-level incentive, and 33% receiving a high-level incentive. As a point of reference, these financial incentives were significant when compared to the incomes in this subsistence agriculture context where piecework daily wage rates (ganyu) for farm workers were approximately 20 Kwacha for men and 5–10 Kwacha for women.[143]

After drawing the incentive amount, each individual was given a voucher of the financial amount they randomly drew. Couples were told that both members of the couple must maintain their HIV status in order for the couple to receive the money. Couples who divorced, separated, or for whom one member was away, would receive one half of the couple incentives after one year if the individual who tested maintained his/her status. Individuals participating as individuals (rather than as

Table A25: Descriptive statistics for participants in 2006–07 MLSFH experimental design offering financial incentives for maintaining HIV status

	Mean	Std. Dev.
Male	0.450	0.498
Age	35.78	12.96
Married	0.838	0.369
Expenditures	3130	5781
Subjective Health	2.065	0.935
Number of lifetime sexual partners	3.108	3.780
Acceptable to use condom	0.405	0.491
Ever used condom with current partner	0.263	0.440
Fear about HIV	1.593	0.752
Number friends died of HIV	8.197	8.045
Some likelihood of HIV infection (current)	0.287	0.453
Some likelihood of HIV infection (future)	0.566	0.496
HIV positive at baseline	0.087	0.282
Enrolled as a "couple"	0.238	0.426
N	1,307	–

Notes: This table presents baseline summary statistics among 1,307 respondents who participated in the incentives program. Expenditures are measured as household expenditures in the past 3 months (on clothes, schooling, medical expenses, fertilizer, agricultural inputs, and funerals). Subjective health represents self-reported health and was phrased: "In general, would you say your health is: Excellent (1), Very Good (2), Good (3), Fair (4), Poor (5)." Number of lifetime sexual partners includes any partner (long-term or short-term) that the respondent had sex with. Fear about HIV was phrased as: "How worried are you that you might catch HIV/AIDS? Not worried at all (1), Worried a little (2), Worried a lot (3)." Some likelihood of infection was coded one if the respondent answered low, medium, high, or don't know, and zero otherwise. Each variable was measured before incentives were offered.
Source: Kohler & Thornton [22]

member of couples) were told that they must maintain their HIV status in order to receive the money approximately one year later. Due to logistical issues, the second round of HIV testing—based on which the incentives were paid–was conducted several months occurred approximately 15 months—instead of the initially stated 12 months—after enrollment in the MLSFH Incentives Study. Table A26 presents baseline summary statistics among those offered zero, medium, and high amounts of the incentive.

Approximately three to six months after the incentives were offered and vouchers given out, respondents were interviewed in their homes and asked about their recent sexual behavior. In particular, asked about the previous nine days, asking sexual activities and condom use each day. These interviewer administered diaries were collected three times over the period of the study, which we identify as Round 1, Round 2, and Round 3, respectively. These were unannounced visits that occurred approximately every three months; the same questionnaire was administered each time. At the end of the third round, respondents were visited by

Table A26: Descriptive statistics, by level of incentives offered, for participants in 2006–07 MLSFH experimental design

	Zero Incentive ($N = 455$)	***Medium Incentive*** ($N = 420$)	***High Incentive*** ($N = 432$)	p-value of joint test
Male	0.446	0.469	0.435	0.59
Age	34.80	35.52	37.07	0.03
Married	0.844	0.831	0.838	0.87
Expenditures	3013	3131	3250	0.84
Subjective Health	2.031	2.000	2.163	0.03
Number of lifetime sexual partners	2.940	3.349	3.053	0.32
Acceptable to use condom	0.400	0.392	0.424	0.62
Used condom with current partner	0.261	0.257	0.271	0.89
Fear about HIV	1.597	1.579	1.603	0.89
Number friends died of HIV	7.816	8.581	8.222	0.40
Some likelihood of HIV infection (current)	0.294	0.288	0.280	0.92
Some likelihood of HIV infection (future)	0.593	0.557	0.547	0.38
HIV positive at baseline	0.105	0.088	0.067	0.13
Enrolled as a "couple"	0.209	0.240	0.266	0.13

Notes: Standard errors in parenthesis. The table presents baseline demographic statistics by incentives amounts among 1,307 respondents who participated in the incentives program. Expenditures are measured as household expenditures in the past 3 months (on clothes, schooling, medical expenses, fertilizer, agricultural inputs, and funerals). Subjective health represents self-reported health and was phrased: "In general, would you say your health is: Excellent (1), Very Good (2), Good (3), Fair (4), Poor (5)." Number of lifetime sexual partners includes any partner (long-term or short-term) that the respondent had sex with). Fear about HIV was phrased as: "How worried are you that you might catch HIV/AIDS? Not worried at all (1), Worried a little (2), Worried a lot (3)." Some likelihood of infection was coded one if the respondent answered low, medium, high, or don't know, and zero otherwise. Each variable was measured before incentives were offered.
Source: Kohler & Thornton [22]

a project nurse and were offered another HIV test. This HIV test was tied to the financial incentives and thus was required in order to be eligible to receive any of the financial incentives.

At the end of the study, of the 1,076 HIV-negative individuals who took a test at the follow-up, seven were HIV-positive. Approximately 93% of the sample completed round 1 diaries, 89% completed round 2 diaries, and 92% completed round 3 diaries. Individuals who were HIV-positive in 2006 were less likely to complete rounds, and this became more of a factor over time. HIV-positives were 6.6 percentage points less likely to complete round 1 diaries, 9.9 percentage points less likely

to complete round 2 diaries, 10 percentage points less likely to complete round 3 diaries, and 20 percentage points less likely to take the follow-up test. Almost all of the respondents (98%) completed at least one round of diaries, with an average of 2.7 rounds. At the end of the study, 89% of all enrolled respondents obtained a follow-up HIV test after round 3.

A6.7. MLSFH 2009 Biomarker Study

The collection of biomarker-based indicators of adult health is an important addition to socioeconomic surveys since they can provide valuable insights into biological functions, and the complex causal pathways between socioeconomic environments and health outcomes. The MLSFH implemented a 2009 Biomarker Study that included collection of blood-plasma based biomarkers.[24,25]

A6.7.a. MLSFH Biomarkers: The collected MLSFH biomarkers included:[25] a lipids panel consisting of cholesterol, LDL, HDL, and triglycerides, as measures for risk factors for cardiovascular disease; circulating blood glucose and HbA1c (only in cases when the blood glucose was above the normal range) as markers of the metabolic function; markers of organ, specifically renal function and clearance (total protein, uric acid, albumin, urea/blood urea nitrogen (BUN), and creatinine) and wide-range CRP (wrCRP) as a measure of inflammation and the immune function.[144] Few, if any, biomarkers are free-standing reliable diagnostic tools, including those collected as part of the MLSFH. Although the MLSFH biomarkers are generally well-known, and we briefly discuss our reasons for their selection, and the critical levels used for obtaining indicators of health risks.

Total cholesterol (TC), high-density lipoprotein (HDL), low density lipoprotein (LDL), and triglycerides (TG): Lipids are fats that store energy for quick release, and to varying degrees, all lipids are recognized risk factors for cardiovascular disease in the developed world. In the absence of other risk factors, the American Heart Association considers a total cholesterol reading of less than 200mg/dl desirable, 200-230 mg/dl borderline, and in excess of 240mg/dl as conveying a high risk for cardiovascular disease. ***Glucose and HbA1c:*** Random blood glucose, also known as a non-fasting blood sugar, is a biomarker for the efficiency of the metabolic system. Glucose is the main source of energy for the body. Insulin, the hormone that cells use to metabolize the glucose, is produced in the pancreas. It is released into the blood in response to levels of circulating glucose. A random blood glucose (RBG) tests has two advantages: it does not require respondent fasting and it is less expensive. But because fasting is not a prerequisite for the test, the RBG measure is less precise. The normal range for a random blood sugar test is 70–100 mg/dl. HbA1c measures average blood sugar level for the past two to three months rather than measuring blood sugar levels at one point of time. HbA1c below 5 percent is seen as normal level and a target, although it can range from 4.5 to 6 percent. People with diabetes are characterized by elevated HbA1c levels and for them a level of about 7 percent is a target. In our data collection, HbA1c was not measured for the entire sample, but only for respondents who showed elevated

blood glucose levels (i.e., 12 study participants with a mean value of HbA1c of 5.53 and 0.71 std. dev.).

Creatinine: Creatinine is one of the waste products in the blood created by the normal breakdown of muscles and circulating levels of creatinine are fairly reliable indicator of the efficacy of kidneys. Normal levels of creatinine in the blood are approximately 0.6 to 1.2 mg/dl in adult males and 0.5 to 1.1 mg/dl in adult females. Any condition that impairs the function of the kidneys will increase creatinine level in the blood. ***Albumin:*** Like creatinine, serum albumin is used to assess renal and liver function. Albumin is the protein of highest concentration in the blood and maintains oncotic pressure of blood to prevent its leakage into tissue. The normal (U.S.) range for of albumin is 3.5 to 5.5 mg/dl. A low albumin level is correlated with inflammation and malnutrition while high levels signal dehydration. ***Total protein:*** Unlike fats and carbohydrates, proteins are not stored in the body. They are continuously broken down (metabolized) into amino acids that are used as building blocks for other proteins. The LabAnywhere test is a rough measure of all the proteins found in the plasma, principally albumin and globulin. The normal range of the test is 6.0 to 8.3mg/dl. ***Uric Acid:*** Uric acid is produced in the body from purine metabolism and excreted by the kidneys. Elevated uric acid is associated with gout, starvation, metabolic syndrome or kidney stones, and decreased uric acid is associated with multiple sclerosis. Normal values of uric acid range between 3.5 and 7.2 mg/dl. ***Urea/Blood Urea Nitrogen (BUN):*** Blood carries proteins for use by cells throughout the body. After the cells use the protein, the remaining waste products are returned to the blood as urea, a compound containing nitrogen. Healthy kidneys take urea out of the blood and send it to the bladder for excretion. If kidneys are not working well, the urea stays in the blood. Normal blood contains 7 to 20 milligrams of urea per deciliter of blood, and a result of more than 20 mg/dl indicates that kidneys are not functioning normally.

C-reactive protein (CRP): CRP is the most commonly used marker of inflammation and infection. As an acute-phase response protein, CRP can increase as much as 1000-fold in 24 hours. At elevated levels CPR indicates systemic infection or tissue damage, and levels above 3.0 mg/l are generally considered as indicating a high risk for cardiovascular disease. The MLSFH assayed only this biomarker of immune function because of budgetary constraints. The wide-range CRP (wrCRP) assay was used since it detects levels of CRP in the range of 0.012–16.0 mg/l, and thus is sensitive at both very low and very high levels.

A6.7.b. MLSFH biomarker sample and data collection: The MLSFH biomarker sample was restricted to Balaka, and the sample was selected in two stages. First, all respondents who were found HIV-positive in a previous MLSFH round were included in the sample. Next, in addition we drew a random sample of approximately 1,500 respondents (aged $\geq$ 18 years) from the 2500 total respondents in the 2008 MLSFH Balaka sample. Because of weather obstacles and failed attempts to find respondents, we were able to contact 1,031 individuals. Of these, 49 respondents (4.7%) refused to participate, and we collected biomarker specimens for

Table A27: Summary statistics for the 2009 MLSFH biomarker study population

	Females mean (sd)	**Males** mean (sd)	**Total** mean (sd)
# of observations	571	335	906
Age (in 2008)	42.17	43.54	42.68
	(17.75)	(16.87)	(17.43)
Age Group			
< 30	0.296	0.307	0.300
30–39	0.205	0.131	0.178
40–49	0.186	0.152	0.173
50–59	0.144	0.209	0.168
60–69	0.082	0.140	0.104
70+	0.088	0.060	0.077
Married (in 2008)	0.762	0.892	0.809
Muslim (vs Christian/other/none)	0.691	0.706	0.696
Schooling attainment			
No school	0.575	0.320	0.483
Primary level	0.399	0.618	0.478
Secondary level	0.026	0.062	0.039
Body Mass Index (BMI) (2008)			
Underweight (BMI < 18.5*)*	0.143	0.118	0.135
Normal ($18.5 \leq BMI < 25$*)*	0.750	0.837	0.777
Overweight ($25 \leq BMI < 30$*)*	0.092	0.035	0.074
Obese (BMI ≥ 30*)*	0.016	0.010	0.014
BMI unknown	0.215	0.396	0.282
HIV positive	0.084	0.050	0.072
Subjective health			
Fair/Poor	0.158	0.100	0.136
Good	0.307	0.195	0.265
Very good	0.279	0.298	0.286
Excellent	0.256	0.407	0.312
Resp.'s household has			
access to potable water	0.843	0.880	0.857
metal roof on house	0.144	0.159	0.149
pit latrine	0.782	0.838	0.802
mosquito nets	0.816	0.828	0.821
mosquito nets treated with insecticide	0.652	0.662	0.655

Source: Kohler *et al.* [25]

982 respondents, of which approximately 60 cases had previously tested positive for HIV. The characteristics of the MLSFH biomarker sample are reported in Table A27, and Table A28 documents the means, std. deviations and percentiles of the MLSFH biomarkers. The biomarker data collection was approved by the IRB at the University of Pennsylvania (May 9th, 2008) and by the Malawi National Health Sciences Research Council (NHSRC) (December 8th, 2008). The actual field work commenced in mid-January and was completed by early February, 2009.

To avoid the complications associated with dried blood spots (DBS), the MLSFH has tested a new approach for collecting measures of population health and their

Table A28: Summary statistics for the biomarker-based health indicators

			Percentiles				
	Mean	std.	5th	25th	50th	75th	95th
Total cholesterol (TC) (mg/dL)	110.4	29.6	65.6	88.8	108.1	131.3	162.2
High-density cholesterol (HDL) (mg/dL)	32.0	10.8	15.4	23.2	30.9	38.6	50.2
Low-density cholesterol (LDL) (mg/dL)	59.0	22.3	27.0	42.5	57.9	73.4	96.5
Triglecerides (TG) (mg/dL)	59.5	29.6	26.5	35.4	53.1	70.8	115.0
Glucose (RBG) (mg/dL)	75.0	19.5	52.3	61.3	68.5	84.7	113.5
Creatinine (mg/dL)	0.73	0.19	0.45	0.60	0.71	0.83	1.06
Albumin (ALB) (g/DL)	3.64	0.44	2.90	3.36	3.63	3.92	4.34
Total protein (TP) (mg/DL)	6.89	0.83	5.52	6.36	6.86	7.40	8.28
Uric Acid (mg/dL)	4.45	1.18	2.69	3.70	4.37	5.21	6.56
Urea (BUN) (mg/dL)	10.7	3.13	6.16	8.68	10.4	12.3	16.5
C-reactive protein (CRP) (mg/L)	4.50	11.8	0.10	0.20	0.70	2.80	25.0

Notes: Number of observations (N) varies from 845 (CRP) to 905 (Uric Acid)
Source: Kohler *et al.* [25]

adaptability to extreme conditions in tropical zones. Our results indicate the reproducibility of biomarkers obtained from the LabAnywhere (previously Demecal) system (LabAnywhere, Haarlem, The Netherlands),[26] a system for the collection of blood plasma that has been used in other large-scale biomarker collections in developed countries.[145–147] The LabAnywhere system required only a few drops of blood harvested from a lancet puncture of a sanitized fingertip. A sponge device was used for absorbing the drop of blood. After the sponge turns completely red, it was dropped into a container with buffer fluid. A gentle swinging motion for 40 seconds was necessary to release the dilution buffer. A filter was used to separate the red blood cells from the plasma. The distinctive feature of this system was that the blood was pressed through a patented filter that separates out plasma. Unlike a clinic based procedure for obtaining blood plasma, the LabAnywhere system did not require the use of a centrifuge. The reliability, sensitivity, and specificity of the test kits had been demonstrated by LabAnywhere in the Netherlands, and the applications of test specific recovery factors yielded a good correlation with results of venous blood samples.[26] In general, LabAnywhere plasma samples are stable for 4 days at 4°C, 2–3 days at room temperature and 1 day at 37°C. While in the field during the day, the collected specimens were stored in a cooler. Upon returning from the field each day, the biomarker coordinator checked all samples to verify that they were collected and labeled properly; all plasma samples were stored in a -20°C freezer until they were shipped to LabAnywhere. At the end of each week, all biomarker samples were cross-checked with field records, and sent via DHL from Malawi to the LabAnywhere laboratory in the Netherlands for testing. The samples were packed in a special cooler with ice packs provided by LabAnywhere, which were designed specifically for transporting the frozen blood samples, including minimum/maximum thermometers to monitor the cooling conditions. LabAnywhere was able to analyze 92.7%, or 910 of the 982, samples they received.

Table A28 provides summary statistics for the collected MLSFH biomarkers. Upon receiving the test results, MLSFH convened an information session in all participating villages during which potential health concerns identified by the tests were discussed. Individual respondents were given the option to discuss privately their results with a health care counselor. The MLSFH also worked with local health clinics to follow up on any potential health issues that were identified by the biomarker tests. However, except for referrals to local health clinics, no specific treatments were provided as part of the MLSFH biomarker study.

A6.8. 2012 MLSFH mature adults survey on mental health and well-being:

The 2017 MLSFH 7 survey on mature adults (= adults aged 45 and older) focused on on mental health and well-being, including MLSFH respondents aged 45 and older, who had previously been interviewed in the 2008 and 2010 MLSFH. A total of 1,266 MLSFH mature adults were interviewed (Figure 2) using a questionnaire that continued key elements of the 2008 and 2010 data collections (Table 4) and newly added detailed measures of mental health, cognitive function, and physical performance. The inclusion criteria for the 2012 MLSFH restricted the sample to MLSFH respondents who were aged 45 or older in 2012, and who had been successfully interviewed in both the 2008 and 2010 MLSFH (a restriction that ensured that at least three waves of MLSFH data were available for each participant in the 2012 MLSFH).

Specifically, the measures of mental health and well-being collected as part of the 2012 MLSFH included (Table A29): (1) To assess *mental health*, we collected the following data: (*a*) continued measurement of the SF12 mental health score that is available since 2006; (*b*) the depression and anxiety modules of the Patient Health Questionnaire-9 (PHQ-9) that allow to assess both, the presence and the severity of depression and anxiety disorders;[150–152] (*c*) detailed information of alcohol consumption since alcohol is the most commonly used psychoactive substance in rural Malawi (and comparable SSA contexts),[164–168] including among HIV+ individuals.[169–176] (2) To assess *cognitive function and performance*, we collected measurements for: (*a*) spatial/temporal orientation and language based on typical questions used in many different mental status examinations;[153] visual/constructional test to assess space and object perception;[154] (*b*) visual/verbal memory, attention/working memory, memory/immediate and delayed recall and executive functioning that resemble many clinical tests assessing these functions, but with necessary adaptations to low literacy levels. (3) *Grip strength as a measurement of physical performance:* Grip strength was measured in both hands using a mechanic handheld dynamometer.[155] Grip strength is important as an estimate of the isometric strength in the upper extremity, and it correlates highly with other muscle groups and is often seen as a measurement of overall strength and physical performance.[156] It is a strong predictor of functional limitations, limitations in ADL, morbidity and mortality.[177,178] It is preferable to other measures of physical performance such as climbing stairs, walking on a flat surface, etc. that are difficult to collect and/or inappropriate (e.g.,there are no stairs) in rural Malawi.[179] The 2012 MLSFH grip

Table A29: Selected measurements in the 2012 MLSFH mature adult survey on mental health and well-being

Construct Definition	*Measurement/Scales/Items Source*
Mental health and depression	SF12 mental health score;[148,149] PHQ-9 of the Primary Care Evaluation of Mental Disorders (PRIME-MD)[150–152]
Cognitive function	spatial/temporal orientation and language;[153] visual/constructional test;[154] visual/verbal memory, attention/working memory, memory/delayed recall and executive functioning developed by the project team
Physical performance	Hand Grip Strength;[155,156]; Body Mass Index (BMI);[157] Activities of daily living (ADLs)[158–161]
HIV status	Determine HIV/1-2™ or Bioline™ HIV
Alcohol Consumption	Alcohol Use based on the Alcohol Use Disorder Identification Test (AUDIT)[162]
Subjective risk assessments and probabilistic expectations	Interactive probabilistic expectation elicitation method developed for Malawi and low literacy populations[64,65,142]
Social capital and resource networks	MLSFH modules on social capital & family transfer networks[53,163]
Social, demographic and economic background	Modules repeated from MLSFH questionnaire 2008 & 2010 (Table 4)
Work efforts and productivity	Time devoted to different work activities and intensity of work; work efforts and work-related health limitations

strength measurement followed identical field procedures as those used by the Health and Retirement Study (HRS) and SHARE studies, and as a result, the existing 2012 and proposed 2014 MLSFH grip strength measures represent the *first* comparable measurements of physical performance between a SSA mature population and the HRS and SHARE study populations. (4) *BMI and HIV testing:* Body mass index (BMI)—an important indicator of nutritional status—was obtained in 2012 from *measured* height and weight, complementing earlier MLSFH BMI data for 2008. In addition, all mature adults who participated in the 2012 data collection were tested for HIV, updating earlier MLSFH HIV tests from 2004–08. (5) *Additional selected measures of well-being*: we continued in 2012 to collect the MLSFH instruments on subjective risks assessments and probabilistic expectations, social capital and resource networks, social, demographic and economic background, and work efforts, productivity and related income/expenditure measures.

ADDITIONAL REFERENCES CITED IN APPENDIX

[124] USAID, FEWS NET. Malawi food security outlook update (march 2012): High food prices and a delayed harvest as assistance programs end, 2012. Washington, DC: FEWS NET. http://www.fews.net/.

[125] Verheijen JZ. Balancing men, moral, and money: Women's agency between hiv and security in a malawi village, 2013. Ph.D. Dissertation, Amsterdam Institute of Social Science Research.

[126] Malawi DHS. *Knowledge, Attitudes and Practices in Health Survey 1996*. Malawi National Statistical Office and ORC Macro, Laverton, Maryland, 1997. http://www.measuredhs.com.

[127] Adams J, Anglewicz P, Helleringer S, Manyamba C, Mwera J, Reniers G, *et al.* Identifying elusive and eager would-be respondents in longitudinal data collection, 2013. Unpublished working paper, Population Studies Center, University of Pennsylvania, Philadelphia, PA.

[128] Malawi MOH and NAC. Voluntary counseling and testing guidelines for Malawi, 2003. Malawi Ministry of Health and Population and Malawi National AIDS Commission.

[129] World Health Organization. *Rapid HIV Tests: Guidelines for Use in HIV Testing and Counselling Services in Resource-constrained Settings*. Geneva, Switzerland: World Health Organization, 2004. URL:http://www.who.int/hiv/pub/VCT/en/rapidhivtestsen.pdf.

[130] Bignami-Van Assche S, Smith K, Reniers G, Anglewicz P, Thornton R, Chao LW, *et al.* Protocol for biomarker testing in the 2004 Malawi Diffusion and Ideational Change Project, 2004. University of Pennsylvania, Social Networks Project Working Paper No. 6 (June 2004).

[131] Obare F. Nonresponse in repeat population-based voluntary counseling and testing for HIV in rural Malawi. *Demography* 2010; **47(3)**:651–665.

[132] Malawi DHS. *Malawi Demographic and Health Survey 2004 (Final Report)*. Malawi National Statistical Office and ORC Macro, Laverton, Maryland, 2005, 2004. http://www.measuredhs.com.

[133] Boerma JT, Urassa M, Nnko S, Ng'weshemi J, Isingo R, Zaba B, *et al.* Sociodemographic context of the aids epidemic in a rural area in Tanzania with a focus on people's mobility and marriage. *Sexually Transmitted Infections* 2002; **78(Suppl. 1)**:i97–i105.

[134] Malawi IHS. Malawi Integrated Household Survey 2010–11, 2012. Malawi National Statistical Office, Zomba, Malawi, and The World Bank, Washington, DC. http://microdata.worldbank.org/index.php/catalog/1003.

[135] Thomas D, Witoelar F, Frankenberg E, Sikoki B, Strauss J, Sumantri C, *et al.* Cutting the costs of attrition: Results from the Indonesia Family Life Survey. *Journal of Development Economics* 2012; **98(1)**:108–123.

[136] Becketti S, Gould W, Lillard L, Welch F. The Panel Study of Income Dynamics after fourteen years: An evaluation. *Journal of Labor Economics* 1988; **6(4)**:472–492.

[137] Ashenfelter O, Deaton A, Solon G. Collecting panel data in developing countries: Does it make sense?, 1986. Living Standards Measurement Study Working Paper 23, World Bank, Washington D.C. http://www.princeton.edu/~deaton/downloads/Collecting_Panel_Data_in_Developing_Countries.pdf.

[138] Thomas D, Frankenberg E, Smith JP. Lost but not forgotten: Attrition and follow-up in the Indonesia Family Life Survey. *Journal of Human Resources* 2001; **36(3)**:556–592.

[139] Norris SA, Richter LM, Fleetwood SA. Panel studies in developing countries: case analysis of sample attrition over the past 16 years within the birth to twenty cohort in Johannesburg, South Africa. *Journal of International Development* 2007; **19(8)**:1143–1150.

[140] Doctor HV, Weinreb AA. Estimation of AIDS adult mortality by verbal autopsy in rural Malawi. *AIDS* 2003; **17(17)**:2509–2513.

[141] Delavande A, Gine X, McKenzie D. Measuring subjective expectations in developing countries: A critical review and new evidence. *Journal of Development Economics* 2011; **94**:151–163.

[142] Delavande A, Kohler HP. HIV/AIDS-related expectations and risky behavior in Malawi, 2011. Population Studies Center, University of Pennsylvania, Working Paper 11-04. http://repository.upenn.edu/psc_working_papers/28.

[143] Whiteside M. When the whole is more than the sum of the parts: The effect of cross-border interactions on livelihood security in southern Malawi and northern Mozambique, 1998. Report for Oxfam GB. http://www.eldis.org/vfile/upload/1/document/0708/DOC6586.pdf.

[144] Maharshak N, Arbel Y, Gal-Oz A, Rogowski O, Shapira I, Berliner S, *et al.* Comparative analysis of bayer wide-range C-reactive protein (wr-CRP) and the Dade-Behring high sensitivity C-reactive protein (hs-CRP) in patients with inflammatory bowel disease. *Journal of Digestive Diseases* 2008; **9(3)**:140–143.

[145] de Beer M, van Eijsden M, Vrijkotte T, Gemke R. Early growth patterns and cardiometabolic function at the age of 5 in a multiethnic birth cohort: the ABCD study. *BMC Pediatrics* 2009; **9(1)**:23.

[146] van Dijk A, van Eijsden M, Stronks K, Gemke R, Vrijkotte T. Cardiometabolic risk in 5-year-old children prenatally exposed to maternal psychosocial stress: the ABCD study. *BMC Public Health* 2010; **10(1)**:251.

[147] van Stralen M, te Velde S, Singh A, De Bourdeaudhuij I, Martens M, van der Sluis M, *et al.* EuropeaN Energy Balance Research to prevent excessive weight gain among youth (ENERGY) project: Design and methodology of the ENERGY cross-sectional survey. *BMC Public Health* 2011; **11(1)**:65.

[148] Ware JE, Kosinksi M, Keller SD. *SF-12: How to Score the SF12 Physical & Mental Health Summary Scales, Third Edition*. Lincoln, RI: QualityMetric, 2001.

[149] Ware JE, Kosinksi M, Keller SD. A 12-Item Short-Form Health Survey: Construction of scales and preliminary tests of reliability and validity. *Medical*

care 1996; **34(3)**:220.

[150] Spitzer R, Williams J, Kroenke K, Linzer M, deGruy III F, Hahn S, *et al.* Utility of a new procedure for diagnosing mental disorders in primary care: the PRIME-MD 1000 study. *JAMA* 1994; **272(22)**:1749.

[151] Kroenke K, Spitzer R, Williams J. The PHQ-9. *Journal of General Internal Medicine* 2001; **16(9)**:606–613.

[152] PHQ Website. Patient Health Questionnaire (PHQ) Screeners: Screener Overview, 2011. Accessed on March, 5th 2011. http://www.phqscreeners.com/overview.aspx.

[153] DeRenzi E, Vignolo L Brain. The Token Test: A sensitive test to detect receptive disturbances in aphasics. *Brain* 1962; **85**:665–678.

[154] Warrington E, James M. *The Visual Object and Space Perception Battery*. Bury St. Edmunds, Suffolk, Tngland: Thames Valley Test Company., 1991.

[155] Frederiksen H, Hjelmborg J, Mortensen J, Mcgue M, Vaupel J, Christensen K. Age trajectories of grip strength: cross-sectional and longitudinal data among 8,342 Danes aged 46 to 102. *Annals of Epidemiology* 2006; **16(7)**:554–562.

[156] Rantanen T, Era P, Kauppinen M, Heikkinen E. Maximal isometric muscle strength and socioeconomic status, health, and physical activity in 75-year-old persons. *J Aging Phys Activity* 1994; **2**:206–220.

[157] Molini V, Nubé M, van den Boom B. Adult BMI as a health and nutritional inequality measure: Applications at macro and micro levels. *World Development* 2010; **38(7)**:1012–1023.

[158] Gama EV, Damián JE, Pérez de Molino J, López MR, López Pérez M, Gavira Iglesias FJ. Association of individual activities of daily living with self-rated health in older people. *Age and Ageing* 2000; **29(3)**:267–270.

[159] Bedirhan U, Chatterji S, Kostanjsek S, *et al.* WHO assessment schedule for measuring disability: Developing the World Health Organization Disability Assessment Schedule 2.0. *Bulletin ofthe World Health Organization* 2010; .

[160] WHO. Disability Assessment Scale version 11, 2000. Phase 2 Field Trials–Health Services Research: 12-ltem Interviewer Administered Version. Geneva: WHO.

[161] Stineman MG, Xie D, Pan Q, Kurichi JE, Saliba D, Streim J. Activity of daily living staging, chronic health conditions, and perceived lack of home accessibility features for elderly people living in the community. *Journal of the American Geriatrics Society* 2011; **59(3)**:454–462.

[162] Babor T, Higgins-Biddle J, Saunders J, Monteiro M. The Alcohol Use Disorders Identification Test (AUDIT): Guidelines for use in primary care, 2001. 2nd Edition, World Health Organization Department of Mental Health and Substance Abuse. (Accessed: June 4, 2011). http://www.whqlibdoc.who.int/hq/2001/who_msd_msb_01.6a.pdf.

[163] Kohler IV, Kohler HP, Behrman JR, Kalilani-Phiri L. Shocks, family transfers and youth-transitions in rural Malawi, 2011. Unpublished manuscript, Pop-

ulation Studies Center, University of Pennsylvania. Presented at the World Bank and Jacobs Foundation conference on "Children and Youth in Crisis" in Marbach, Germany, May 5–6, 2011.

[164] Parry C. South Africa: alcohol today. *Addiction* 2005; **100(4)**:426–429.

[165] Parry C, Myers B, Morojele N, Flisher A, Bhana A, Donson H, *et al.* Trends in adolescent alcohol and other drug use: findings from three sentinel sites in South Africa (1997-2001). *Journal of Adolescence* 2004; **27(4)**:429–440.

[166] Shisana O, Rehle T, Simbayi L, Parker W, Zuma K, Bhana A, *et al. South African national HIV prevalence, HIV incidence, behaviour and communication survey, 2005*. Human Sciences Research Council, 2005.

[167] Fisher J, Bang H, Kapiga S. The association between HIV infection and alcohol use: A systematic review and meta-analysis of African studies. *Sexually Transmitted Diseases* 2007; **34(11)**:856.

[168] Kalichman S, Simbayi L, Kaufman M, Cain D, Jooste S. Alcohol use and sexual risks for HIV/AIDS in sub-Saharan Africa: Systematic review of empirical findings. *Prevention Science* 2007; **8(2)**:141–151.

[169] Weinhardt L, Carey M. Does alcohol lead to sexual risk behavior? Findings from event-level research. *Annual Review of Sex Research* 2000; **11**:125. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2426779/.

[170] Burnam MA, Bing EG, Morton SC, Sherbourne C, Fleishman JA, London AS, *et al.* Use of mental health and substance abuse treatment services among adults with HIV in the United States. *Arch Gen Psychiatry* 2001; **58(8)**:729–736. http://archpsyc.ama-assn.org/cgi/content/abstract/58/8/729.

[171] Dausey D, Desai R. Psychiatric comorbidity and the prevalence of HIV infection in a sample of patients in treatment for substance abuse. *Journal of nervous and mental disease* 2003; **191(1)**:10.

[172] Whetten K, Reif SS, Napravnik Sea. Substance abuse and symptoms of mental illness among HIV-positive persons in the Southeast. *South Med J* 2005; **98**:9–14.

[173] Klinkenberg W, Sacks S. Mental disorders and drug abuse in persons living with HIV/AIDS. *AIDS Care* 2004; **16(Suppl 1)**:S22–S42.

[174] Stoff DM, Mitnick L, Kalich. Research issues in the multiple diagnosis of HIV/AIDS, mental illness and substance abuse. *AIDS Care* 2004; **16(Suppl. 1)**:S1–S5.

[175] Perkins D, Stern R, Golden R, Murphy C, Naftolowitz D, Evans D. Mood disorders in HIV infection: prevalence and risk factors in a nonepicenter of the AIDS epidemic. *Am J Psychiatry* 1994; **151(2)**:233–236. http://ajp.psychiatryonline.org/cgi/content/abstract/151/2/233.

[176] Atkinson JH, Grant I. Natural history of neuropsychiatric manifestations of HIV disease. *Psychiatr Clin North Am* 1994; **17**:17–33.

[177] Nybo H, Gaist D, Jeune B, McGue M, Vaupel J, Christensen K. Functional status and self-rated health in 2,262 nonagenarians: The Danish 1905 cohort survey. *Journal of the American Geriatrics Society* 2001; **49(5)**:601–609.

[178] Rantanen T, Harris T, Leveille S, Visser M, Foley D, Masaki K, *et al*. Muscle strength and body mass index as long-term predictors of mortality in initially healthy men. *Journal of Gerontology Series A: Biological Sciences and Medical Sciences* 2000; **55(3)**:M168.

[179] Robertson K, Liner J, Heaton R. Neuropsychological assessment of HIV-infected populations in international settings. *Neuropsychology review* 2009; **19(2)**:232–249.