***NOT FOR PUBLICATION**

Online Appendix: "The Consumption, Income and Wealth of the Poorest: An Empirical Analysis of Economic Inequality in Rural and Urban Sub-Saharan Africa for Macroeconomists"

By Leandro de Magalhães and Raül Santaeulàlia-Llopis

Contents

A LSMS-ISA Data

The Integrated Surveys on Agriculture (ISA) are conducted under the umbrella of the Living Standards Measurement Study (LSMS). The LSMS are representative household surveys with a particular focus on recovering the distribution of living standards and inequality (Grosh and Deaton, 2000).

The current ISA builds on extensive experience of the World Bank in data collection and improves previous LSMS data. For example, Grosh and Glewwe (2000) propose a substantial set of improvements to the early LSMS conducted in the 1980s and 1990s that ISA has incorporated. One particular concern of previous LSMS work was the estimation of agricultural income—the main source of household income in poor Sub-Saharan Africa households—and assets. The ISA component is specifically constructed to address this issue (Carletto et al., 2010). ISA incorporates a new and comprehensive agricultural module questionnaire that has been created to keep track of all crops produced and inputs used per plot separately and for each crop season. For example, for the case of Malawi this is separately recorded for the rainy season, the dry season (dimba) and for permanent crops. This information is collected for each and all plots cultivated by each household.[1] The first LSMS-ISA data set made available to the public is the Malawi household survey conducted in 2010-2011 and made public in 2013. Most of the discussion of this paper focuses on this survey.

Other recent improvements of the LSMS beyond ISA are described in the special issue of the Journal of Development Economics for the Symposium on Measurement Survey and Design published in 2012. These include improvements in the three economic dimensions that we study: consumption (Beegle et al., 2012b), income (Deininger et al., 2012; Beegle et al., 2012a), and wealth (Carletto et al., 2013). The high quality of the data can be seen in the response rate of the Malawi 2010-11 survey. The response rate is 99.9% (i.e., a total of 12,271 households). Of those 12,271 households, only 6% of households are replacements.[2] Only a handful of these 12,271 households did not complete the full income, consumption, and wealth sections of the questionnaire that we require to be included in our analysis. This high response rate largely mitigates nonresponse biases (e.g., the concern that richer individuals may be reluctant to take part in the survey). Finally, it is also important to note that the ISA data are not top coded. This helps us capture the very rich in these poor countries, subject to the caveat of potential underreporting, in particular of income, which we discuss in the main text at length.

In this appendix we provide a detailed discussion of how our data set is built from the LSMS-ISA. We focus on Malawi LSMS-ISA 2010/2011 available through the World Bank website. The construction of the variables for the other years and for Tanzania and Uganda are similar, except where small changes are needed to accommodate differences in the questionnaires. These differences can be seen in the do-files related to this paper.

Survey details The sample size in the Malawi ISA 2010-2011 is 12,271 households (56,397 individuals). The survey includes four main questionnaires: household (H), agricultural (AG), fishery (F), and community (C).[3] The sample is representative and consists of 9,024 cross-sectional households (in 768 enumeration areas, EAs) that were visited once, and 3,247 panel households (in 204 EAs) that were visited twice. The households that were visited once received the household questionnaire in full, as well as agriculture and fishery questionnaires when applicable. As part of the agriculture questionnaire, these cross-sectional households reported information on the last completed rainy and dimba (dry) seasons.[4] Depending on the harvesting being completed or not at the time of the interview, the reference rainy season for these households corresponds to the 2009/2010 (79% of all cases

[1]Most rural households represent small-scale farms that sell none or only a small fraction of what they produce in the open market.

[2]The main reason for replacement was that no household member was found when the survey team arrived to the household location. Only 14 households actually refused to answer the survey.

[3]See MNSO (2012) for further detailed information on the design and implementation of the IHS3.

[4]Typically, the rainy season (months in which it rains) comprehends two consecutive years in Malawi from November to March. This way, the start date for planting is generally some time after the rains start in November and ends around January. Farmers generally complete their planting before January 15—usual cut-off date for rainfall-season planting. Harvesting usually starts in March/April. There is some variation, though not significant, of start and end dates across regions. Harvests are possible earlier/later depending on the type of crop, year-specific rainfall and clime, other location-specific agronomic conditions. Earlier harvesting before

(AG:c0a)) or the 2008/2009 season. Analogously, the reference dry season corresponds to the year 2009/2010 (55% of all cases, AG:j0a) or year 2008/2009. The households that received two visits (about 1/3 of the total sample) were visited during the post-planting period of the 2009/2010 rainy season (i.e., in March/April which ensures that planting is finished) and revisited about 3 months later (i.e., June/July) during the post-harvest period. That is, by construction panel households always refer to the 2009/10 rainy and dry season. Further, during the first visit panel households reported information on the 2009/10 rainy season pre-harvest related issues, including cultivated area, input use, associated household and hired labor, costs, etc. During the second visit farming households reported information on 2009/10 rainy season production and post-harvest related matters, as well as complete information on the 2009/2010 dry season.[5,6]

The unit of analsyis: Households In Malawi, as in many other developing countries, household size is large with extended families in which several generations live together in a single household. Household members potentially include family (e.g. children, spouses, siblings, and parents) and also non-relatives (e.g. lodgers and servants). We define household members as individuals that have lived in the household at least 9 months in the last 12 months. While we focus on households as the baseline unit of analysis, we also study individual characteristics of household members *per se* in several parts of our study, for example, to analyze total household labor supply. Some important economic variables are available only at the individual level (e.g. demographic variables, labor income, and fertility) and this implies that we need to carefully aggregate individual data across all household members to obtain household-level variables.

Household characteristics We attribute to the household the household head demographic characteristics. The household head is the person who makes economic decisions in the household. Individual demographic characteristics include sex (H:b03), age (H:b05), birth year (H:b06), and the set of variables that we describe next including educational attainment, health behavior and status, marital status, household structure, risk and insurance mechanisms, food security, geographical variables, and migration characteristics. The household roster collects this information for all household members.[7]

optimal crop maturity (as early as February) might also be generated due to household needs to satisfy minimum subsistence consumption and avoid hunger (see the collection of USAID Malawi Food Security Outlooks and Updates http://www.fews.net/docs/Publications/Malawi_FSU_February_2010_final.pdf).

[5]This subsample of IHS3 households that receive two visits is designed to be representative at national-, regional- and urban/rural-level—the subsample belongs to a set of enumeration areas (EAs) that were randomly selected prior to the start of the field work. Further, these selected EAs/households will be tracked and resurveyed in 2013 as part of a new IHS3-panel component. In our current exercise we focus exclusively on the cross-sectional dimension of the IHS3 that is available. That is, we use all currently available information incorporating all households independently of whether they were visited once or twice in IHS3. Note that the IHS3 is not linked to the previous sample waves, IHS1 or IHS2, in any panel dimension. It is the IHS3 that for the first time serves as a baseline for a panel set to be followed-up in 2013. Further, for the IHS3 panel subcomponent, retrospective information on the 2008/2009 rainy season is also provided, see Module AG:b.

[6]In order to collect consumption data in an evenly spread manner across the 12-month period, the work was organized in a way that approximately 64 EAs were subject to consumption data collection each month. To accomplish this in the context of paying two visits to the panel households in the first six months of the field work, it was decided that when the panel households were visited for the first time during the first quarter of the fieldwork, only half of them (Panel Group A) received the household questionnaire in full, and if applicable, the visit 1 components of the agriculture questionnaire and the fishery questionnaire. The rest of the panel subsample (Panel Group B) were administered only the household roster, the filter module for the agriculture questionnaire, and the visit 1 components of the agriculture questionnaire, if applicable, when they were visited for the first time. In the second quarter of the field work, Panel Group B were administered the remaining parts of the household questionnaire, and the visit 2 components of the agriculture questionnaire and the fishery questionnaire, if applicable, while Panel Group A only received a household roster update and the visit 2 components of the agriculture questionnaire, if applicable.

[7]The questionnaire allows for the data collection of up to 12 household members. If the household has more than 12 members, a second questionnarie is used to complete the information on each and all members. Individual information for each and all household members is available on, among others, demographic characteristics, labor

Geographic variables. Information on the district where the household lives (H:a01), town (H:a02), and place/village name (H:a05) are provided. Further, region can be identified from the first enumeration code of the district variable: North ($r = 1$), Center ($r = 2$), and South ($r = 3$).

A.1 Consumption

Household-z consumption includes nondurables (e.g. food, clothing, services, utility bills, school, and medical expenditures) and durables (e.g. housing services and furniture).[8] All consumption quantities are annualized after controlling for seasonality (see section A).

A.1.1 Nondurable consumption

First, food consumption, $c_{f,z}$, includes 135 items distributed into: (a) cereals, grains, and cereal products, (b) roots, tubers, and plantains, (c) nuts and pulses, (d) vegetables, (e) meat, fish, and animal products, (f) fruits, (g) cooked food from vendors, (h) milk and milk products, (i) beverages, and (j) spices and miscellaneous. Information about each food item includes quantity consumed last week (H:g03), quantity purchased (H:g04), quantity consumed from own production (H:g06), quantity received from gifts and other sources (H:g07).[9] We use expenditures on food items (H:g05) and the purchased quantity (H:g04) to infer food prices per kg. and per food item κ,

$$p_{f_\kappa,r} = q^{50}_{z\in r}\left(\frac{exp_{f_\kappa,z}}{c_{f_\kappa,z}}\right),$$

where $q^{50}_{z\in r}$ is the median function and $p_{f_\kappa,r}$ the median price computed from households in region r with positive consumption and expenditures.[10] These imputed prices, p_{f_κ}, for each food item κ are used to compute the monetary value of nonpurchased food consumption from own production and gifts. Second, other nondurable expenditures, $exp_{nf,z}$, include fuel, paper and paper products, personal products, rubber, plastics, textiles, clothing, footwear, mortgage payments, funeral costs, bridewealth costs, and mosquito nets (H:i01-k03). Third, utilities expenditures, $exp_{u,z}$, include the value of purchased firewood (last week) (H:f18), electricity (H:f25-26), MTL telephone (H:f32-33), cell phone (last month) (H:f35), and water (last month) (H:f37). Fourth, school expenditures in the last 12 months, $exp_{s,z}$, include tuition, after school programs and tutoring, books, uniform, boarding fees, transport, etc. (H:c22A-c22I). Fifth, medical expenditures), $exp_{m,z}$, include treatment and prevention activities. Regarding treatment, medical expenditures include hospitalizations costs (H:d14) and stays over traditional healers (H:d19), transport costs (H:d15 and H:d20), and overnight food (H:d16 and H:d21), the amount spent in the last 4 weeks on medicine, tests, and inpatient fees (H:d10), and nonpresciption medicines (e.g. Panadol, Fansidar, cough syrup) (H:d12). Regarding prevention, medical expenditures include care not related to an illness (e.g. preventive care, prenatal visits, check-ups, etc.) (H:d11). This implies that the monetary value of

income, and fertility history.

[8]Our construction of household consumption is similar what is proposed in Deaton and Zaidi (2002), except for how we value the part of production not sold, which we discuss in Section A.2.

[9]These consumption quantities are provided in different units (e.g. kg, 50kg bag, 90kg bag, Pail, n.10 plate, n.12 plate, bunch, piece, heap, bale, basket, ox-cart, litre, cup, tin, gram, millilitre, teaspoon, basin, satchet/tube, other (spec.)). We convert these quantities into kg. We use reported prices to generate the conversion rates. The median unit price (of items reported in the modal unit) for a given item in a given region and season is used to generate household specific conversion rates. We pick the median conversion rate (if there are at least 7) for each item-unit (conversion rates are item specific). With the resulting conversion rates, items are first converted into the modal unit, and then into kg.

[10]The median of rural-urban deflated prices per region are used to estimate monetary values (historical inflation series are available at the Malawi National Statistical Office website). We are unable to estimate the value for 3.8% of household items for own consumption, which is negligible in monetary value as we are able to value the main crops such as maize, tobacco, potatoes, milk, eggs, and so on.

household-z nondurable consumption is

$$p_{nd}c_{nd,z} = \sum_{\kappa} \left(exp_{f_\kappa,z} + p_{f_\kappa,r}(c^{own}_{f_\kappa,z} + c^{rec}_{f_\kappa,z}) \right) + exp_{nf,z} + exp_{u,z} + exp_{s,z} + exp_{m,z},$$

and note that $p_{f_\kappa,r}(c^{own}_{f_\kappa,z} + c^{rec}_{f_\kappa,z})$ are not household expenditures, hence, do not enter the household budget constraint.[11]

A.1.2 Durable consumption

Durable consumption consists of housing services and purchases of other durable goods such as furniture, AC, TV, refrigerator, bicycle, etc. in the last 12 months (H:l06). For households who do not own their dwellings, the rental expenditure is reported (H:f04). Housing services are inferred from the self reported monthly renting value of dwellings owned by the household, $r_h k_{h,z}$, (H:f03). Denote expenditures on other durables as $exp_{od,z}$ (H:l07). This way, the monetary value of household-z durable consumption is,

$$p_d c_{d,z} = r_h k_{h,z} + exp_{od,z},$$

and note that $r_h k_{h,z}$ are not household expenditures, hence, do not enter the household budget constraint.

A.2 Income

Household-z annual income includes labor market income wh_z (A.2.2), agricultural net production $p_a y_{a,z}$ (A.2.1), fishery net production $p_f y_{f,z}$ (A.2.4), business income $y_{b,z}$ (A.2.3), capital income $y_{k,z}$ (A.2.5), and net transfers $y_{tr,z}$ (A.2.6). All variables are annualized. That is, household-z income is the sum of all income sources:

$$y_z = wh_z + p_a y_{a,z} + p_f y_{f,z} + y_{b,z} + y_{k,z} + y_{tr,z}.$$

A.2.1 Agricultural net production

Household-z agricultural activities are reported separately for nonpermanent crop produced in the rainy (AG:c-i) and dry (simba) (AG:j-o) seasons, $s = \{r, d\}$, tree/permanent crop (AG:p-q), livestock sales (AG:r) and livestock products sales (AG:s).

Nonpermanent crop. Denote by ψ_i the type of nonpermanent crop i.[12] Denote the total quantity of nonpermanent crop-ψ_i harvested by household z per season s by $y_{\psi_i,s,z}$ (AG:g13). The information on harvested crop is available per plot, that is, $y_{\psi_i,s,z} = \sum_d y_{\psi_i,d,s,z}$ where $y_{\psi_i,d,s,z}$ is the amount of crop-ψ_i harvested by household z in season s and plot d, and note that up to 5 types of crop are potentially harvested per plot (AG:d20). These

[11] This information on nondurable nonfood expenditures is collected by item through exclusive 1 week, 1 month, 3 months, and 12 months recalls. For those items with 3 months or less recall period, we convert them to monthly values, deflate, deseasonalize, and annualize each subgroup separately: food from own production, purchased, and received; clothing, utilities, health, and a 'other' category.

[12] Information on all typical nonpermanent crops ϕ_i is available. These crops are: maize (local, composite/OPV, hybrid, hybrid recycled), tobacco (Burley, flue cured, NNDF, SDF, oriental, other), groundnut (chalimbana, CG7, mani-pintar, mawanga, JL24, other), rice (local, faya, pusa, TCG10, IET4094, kilombero, etc.), ground bean, sweet potato, Irish (Malawi) potato, wheat, finger millet (mawere), Sorghum, peral millet (mchewere), beans, soybeans, pigeonpea (nandolo) cotton, sunflower, sugar cane, cabbage, tanaposi, nkhwani, therere/OKRA, tomato, onion, pea paprika, other). For exposition simplicity, we use rainy season variables name codes (AG:b-i) in the Agricultural questionnaire. The procedure for the simba (dry) season name codes (AG:j-o) is analogous, replacing 'c' for 'j' and so on

data are collected for up to 6 plots (R1-R6).[13] Part of the harvested crop is sold, $y^{sold}_{\psi_i,s,z}$ (AG:i02).[14] Household-z monetary revenue from crop-ψ_i sales in a given season is $rev_{\psi_i,s,z}$ (AG:i03). From this revenue and the quantity sold we can infer crop-ψ_i prices-at-the-gate per season s and region r, $p_{\psi_i,s,r}$, as

$$p_{\psi_i,s,r} = q^{50}_{z\in r}\left(\frac{rev_{\psi_i,s,z}}{y^{sold}_{\psi_i,s,z}}\right),$$

where $q^{50}_{z\in r}$ is the median function and $p_{\psi_i,s,r}$ is the median price of crop-ψ_i in region r with positive revenues in season s.[15] For tobacco, which represents 50% of the total value of agricultural production, the price-at-the-gate is also used to value the production that is stored. The reasons for storage tobacco are: wait for the arrival of buyer or sell later at a higher price. A negligible amount is kept for own consumption (AG:i42a). In order to estimate the value of production that is not sold for food items, we use the consumption prices estimated in Section A.1. See section C for a discussion comparing the use of the price-at-the-gate and consumption prices to value unsold produciton.

In crop production, each household z incurs in intermediate input costs per season s associated with $v = \{land,\ hired\,labor,\ transport\ sales,\ fert/pest/herb,\ seed\}$:

1. Rented-in land (per season and plot), $cost^{land}_{s,z} = r_{s,l}l^{rent-in}_{s,z}$ (AG:c07-c09). The associated rental period is either rainy season, full year (rainy and dry seasons), or other (AG:d12). These rental payments take the form of given output (for crossectional households (AG:d08) and for panel households (AG:d10)), or cash and other in-kind payments (for crossectional households (AG:d09) and for panel households (AG:d11)).[16]
2. Hired labor days and wages per day by men, women, and children ($<$ 15 years of age) and/or payment in-kind (crop) (per season and plot), $cost^{hired\ labor}_{s,z}$ (AG:c39);[17]
3. Transportation costs associated with sales, $cost^{transport\ sales}_{\psi_i,s,z}$ (AG:i10). Information is available by crop.
4. Expenditures on organic fertilizers ($l = org$), inorganic fertilizers (up to 4 types) ($l = inorg$), and pesticides/herbicides (up to 4 types) ($l = pest/herb$), $cost^{f/p/h}_{s,z}$;
5. Expenditures on seeds, $cost^{seeds}_{s,z}$.

While the computations of the costs associated with rented-in land, hired labor, and transportation costs of sales are straightforward, the costs of fertilizers/pesticides and seeds require further explanation due to government subsidies:

Fertilizers The expenditure cost of nonsubsidized fertilizers as well as pesticides/herbicides $cost_{\varphi^{nosub}_{l,z}} = exp_{\varphi^{nosub}_{l,z}}$ is available for each $l = \{org,\ inorg,\ pest/herb\}$ (AG:f09/10/18/19/28/29/40). This includes transportation costs. Below we discuss the cost of subsidized fertilizers and pesticides/herbicides.

[13] A plot is defined as a continuous piece of land on which a unique crop (or mixture of crops) is grown under a uniform consistent crop management system. These questions are generally asked to the person that makes the economic decisions on the plot (see AG:d01-d02).

[14] Note that information on sales are available in total per crop, not by plot. In practice, in our computations we do not distinguish whether the quantities and prices of each crop refer to shelled or unshelled product (AG:i02c). This is because whether a crop is sold as shelled or unshelled is crop specific; 98% of maize is reported as shelled. The is not enough variation to allow us to estimate shelled and unshelled prices for a given crop.

[15] We use the same procedure used in Section A.1 to convert all quantities into kg. Reported quantities units include kg, 50 and 90 kg bags, pail, bunch, piece, bale, basket, plate, and others.)

[16] Panel households may not have completed the season when they are interviewed. Question AG:d10 refers to output that will be given, and AG:d11 decomposes payments in cash or kind already paid as well as to be paid in the future.

[17] more than 80% of payment in kind is done with shelled maize (AG:d46j). For this reason we use the price-at-the-gate to estimate the monetary value of the payment in kind.

Seed The cost of nonsubsidized seed is obtained from expenditures $cost_{seed^{nosub}_{i,z}} = exp_{seed^{nosub}_{i,z}}$, and available for different seed (AG:h09/10/18/19/28/29/40). This includes transportation costs. Below we discuss the cost of subsidized seed.

Subsidies The household-z costs on subsidized fertilizers and pesticides/herbicides, and seed (i.e., $cost^{f/p/h}_{\varrho_i,z,s}$ and $cost^{seed}_{\varrho_i,z,s}$) is computed as the sum of the payments of input purchases done by redeeming coupons (AG:e15) plus the transportation cost (AG:e15), the capital gains of trading coupons $purch^{v}_{\varrho_i,z,s} - sales^{v}_{\varrho_i,z,s}$ (AG:e04 and AG:e20), which are negligible.

This way, the total costs of nonpermanent crop production associated with fertilizers/pesticides/herbicides and seed are computed as the sum of nonsubsidized and subsidized costs,

$$
\begin{aligned}
cost^{f/p/h}_{s,z} &= \sum_{l} cost^{f/p/h}_{\varphi^{nosub}_{l,z}} + \sum_{\varrho_i} cost^{f/p/h}_{\varrho_i,z,s} \\
cost^{seed}_{s,z} &= \sum_{i} cost^{seed}_{seed^{nosub}_{i,z}} + \sum_{\varrho_i} cost^{seed}_{\varrho_i,z,s}.
\end{aligned}
$$

Finally, the net product of nonpermanent crop production is,

$$
p_{\psi} y_{\psi,z} = \sum_{s}\sum_{i} rev_{\psi_i,s,z} + \sum_{s}\sum_{i} p_{\psi_i,s,r}(y_{\psi_i,s,z} - y^{sold}_{\psi_i,s,z}) - \sum_{s}\sum_{v} cost^{v}_{s,z}.
$$

for $v = \{land,\ hired\ labor,\ transport\ sales,\ fert/pest/herb,\ seed\}$. The harvested crop-$\psi_i$ that is not sold, i.e., $y_{\psi_i,s,z} - y^{sold}_{\psi_i,s,z}$, is stored $y^{stored}_{\psi_i,s,z}$ (AG:i40) for reasons such as household consumption, seed, and to sell at a later date. There is a question that refer to the loss $y^{lost}_{\psi_i,s,z}$ in the post-harvest period (AG:i36), but this is negligible: less than 40 households report any loss in their maize or tobacco harvest.

Tree/Permanent crop. Denote by ξ_i the type of tree/permanent crop i.[18] The total quantity of permanent crop-ξ_i harvested by household z is $y_{\xi_i,z}$ (AG:p09). Part of the harvested permanent crop is sold, $y^{sold}_{\xi_i,z}$ (AG:q02). Household-z monetary revenue from these sales are $rev_{\xi_i,z}$ (AG:q03). In order to estimate the value of production that is not sold for food items, we use the consumption prices estimated in Section A.1.[19] The product of permanent crop production is,

$$
p_{\xi} y_{\xi,z} = \sum_{i} rev_{\xi_i,z} + \sum_{i} p_{\xi_i,r}(y_{\xi_i,z} - y^{sold}_{\xi_i,z}).
$$

The harvested crop-ξ_i that is not sold, i.e., $y_{\xi_i,z} - y^{sold}_{\xi_i,z}$, is stored $y^{stored}_{\xi_i,z}$ (AG:q39),[20] or lost $y^{lost}_{\xi_i,z}$ in the post-harvest period (AG:q35).[21]

Livestock sales. Denote by ω_i the type of livestock i.[22] Part of the livestock is sold, $k^{sold}_{\omega_i,z}$ (AG:r16), and the value of livestock-ω_i sales in the last 12 months is $rev_{\omega_i,z}$ (AG:r17). Households incur in intermediate input costs associated with livestock, $cost_{\omega,z}$, that are hired labor (to take care of livestock) (AG:r27), animal feed (AG:r28), vaccinations (AG:r29), veterinary services (AG:r30), and expenditures on housing equipment, feeding utensils, and

[18] Typical permanent crops ξ_i are: cassava, tea, coffee, mango, orange, papaya, banana, avocado, guava, lemon, tangerine, peach, custade apple, Mexican apple, masau, pineapple, macadamia, and other

[19] If these are unavailable, we use the price-at-the-gate median sale prices for a given region and season.

[20] The reason is almost exclusively for own consumption (AG:q41a)

[21] Negligible amounts are lost.

[22] Typical livestock ω includes calf, steer/heifer, cow, bull/ox, donkey, mule/horse, goat, sheep, pig, chicken-layer, local hen, chicken-broiler, local-cock, turkey, duck, guinea fowl, beehive, and other (spec).

other inputs (AG:r31). Livestock net product (sales) is

$$p_\omega k_{\omega,z}^{sold} = \sum_i rev_{\omega_i,z} - cost_{\omega,z}.$$

Livestock product. Denote by ζ_i the type of livestock product i.[23] The annual quantity of livestock product-ζ_i produced by household z, $y_{\zeta_i,z}$, is computed as the average amount of production per month (AG:s03) times the months the product was produced (AG:s02). Part of the livestock product is sold, $y_{\zeta_i,z}^{sold}$ (AG:s05, and the reported sales value is $rev_{\zeta_i,z}$ (AG:s06). In order to estimate the value of production that is not sold, we use the consumption prices estimated in Section A.1.[24] The net product of livestock produces for household z is

$$p_\zeta y_{\zeta,z} = \left(\sum_i rev_{\zeta_i,z} + \sum_i p_{\zeta_i,r}(y_{\zeta_i,z} - y_{\zeta_i,z}^{sold}) \right).$$

The livestock product-ζ_i that is not sold, i.e., $y_{\zeta_i,z} - y_{\zeta_i,z}^{sold}$, is used for own consumption, $y_{\zeta_i,z}^{c}$ (AG:s09), or given out as gifts/reimbursements, $y_{\zeta_i,z}^{tr,g}$ (AG:s10-s11).

Renting-in agricultural equipment and structure capital. In the production of permanent and nonpermanent crop, as well as in livestock production, households may rent-in equipment capital (implements and machinery such as hand hoes, axes, ox ploughs, tractors, etc.) and structure capital (e.g. chicken house, storage house, granary, barn, etc.). The value (H:m14) of these rentals in the last 12 months, respectively $r_{k_a^e} k_a^{e,rented-in}$ and $r_{k_a^s} k_a^{s,rented-in}$, is an intermediate cost for agricultural activities that reduces agricultural net production. Therefore, agricultural net production is the sum of nonpermanent crop net production, permanent crop net production, livestock sales, and livestock products net production, minus rentals of agricultural equipment and structure capital, that is,

$$p_a y_{a,z} = p_\psi y_{\psi,z} + p_\xi y_{\xi,z} + p_\omega k_{\omega,z}^{sold} + p_\zeta y_{\zeta,z} - r_{k_a^e} k_a^{e,rented-in} - r_{k_a^s} k_a^{s,rented-in}.$$

A.2.2 Labor market income

Wage income is reported by occupation (main, secondary, and informal). For each activity there is information on the average hours worked per day, average days worked per week, and number of weeks worked per year. This allows for an estimate of yearly labor supply.[25] Wages are reported by activity but potentially in different units of time, mostly on a monthly basis for those with steady labor income (specially in urban areas), and on a weekly or daily basis for those working on *ganyu* or informal activities (e.g., landowners' seasonal labor supply outside their own farm in the lean season).[26] Wage payments include salaries plus additional allowances. These allowances could be in kind (mostly in maize) but are reported in monetary value. By combining the wages and the labor supply in a consistent unit of time we build an estimate for annual labor income for all individuals in the household. By summing individual labor income of all members within households we construct a measure of annual household labor income.

Precisely, household-z annual labor market income, wh_z, aggregates individual cash and in-kind payments/salaries plus allowances/gratuities earned in the market by each and all household members $i \in z$. Individual labor mar-

[23] Livestock products are: milk, eggs, meat, honey, hides, manure and others.

[24] If these are unavailable, we use the price-at-the-gate median sale prices for a given region and season.

[25] Previous LSMS datasets provide information on labor supply with the reference period of "the past 7 days". ISAs complement that information with recalled hours worked per day, week, and month over the past 12 months, which greatly facilitates the determination of the annualized labor supply. In particular, this avoids potential measurement error from labor supply seasonality. See Rosenzweig and Udry (2014) for a discussion on how wages are affected by seasonal weather patterns.

[26] Informal labor (e.g., *ganyu* in Malawi) while important, is very seasonal and relative small compared to other sources of income (Goldberg, 2016).

ket income information, $wh_{o_m,i}$, is available by occupation: main occupation ($o_m = 1$), secondary occupation ($o_m = 2$), and informal occupations (ganyu) ($o_m = 3$).[27] Individual annual labor income from formal occupations, $wh_{1,i}$ and $wh_{2,i}$, is obtained by multiplying (a) the last payment/salary (H:e25 and H:e39) plus associated allowances/gratuity (H:e27 and H:e41) times (b) the correspondent reference period (hours/weeks/months) for these payments (H:e26, H:e28, H:e40, and H:e42) times (c) the number of hours/weeks/months worked in the last 12 months (H:e22-24 and H:e36-38). Annual individual labor income from informal occupations, $wh_{3,i}$, is computed by multiplying (a) the wage earned per day (H:e59) times (b) the number of hours/weeks/months worked in ganyu in the last 12 months (H:e56-58). To obtain household-z annual labor income we aggregate labor income over individuals $i \in z$ and occupations o_m, that is,

$$wh_z = \sum_{o_m} \sum_{i \in z} wh_{o_m,i}.$$

A.2.3 Business Income

We define annual household business net income using information from all business owned by the household.[28] For each enterprise we compute net income as total annual sales minus costs. In the Malawi ISA, households report the average net income for a bad, standard, and good month. Households are then asked how many of each type of month occurred in the past year. In Uganda, households report both gross income and costs. In Tanzania households report net income directly.[29]

Precisely, household-z privately held businesses (potentially more than one) include owned nonagricultural businesses that process/sell agricultural by-products (e.g. flour, juice, beer, jam, oil, seed, and livestock by-products), sales of forest-based products, street or market trading businesses, taxi or pick up truck drivers, bar/restaurants, professional services (e.g. doctor, accountant, lawyer, and midwife) etc. (H:n09). We define household annual business income, $y_{b,z}$, as the net product of all enterprises owned by the household. For each enterprise b_j we compute net product as the total annual sales minus costs. First, we identify whether the business operation for each and all of the past 12 months is associated with no sales or a low, medium, or high volume of sales (H:n25)—and we have this information separately for each of the household enterprises. Second, we compute the value of total sales per enterprise during the last month in which the household had low, average, and high sales $rev_{b_j,m_q,z}$ with $q = \{low, average, high\}$ (H:n34-39). We combine this information and attribute the same revenue to all months reported to be within the same sales category q. Third, information on variable costs (e.g., raw materials, inventory, freight/transport, fuel/oil, electricity, water, insurance, other) is available for the last month of operation (H:n41a-H:n41h) plus total wages/salaries paid to hired men/women/children (under 16) labor (H:n31). However, we need annual costs. To compute these annual costs we estimate weights that represent the relative differences in sales (if sales are twice as high, costs are twice as high for the base (last month)).[30] This way, we compute household-z business income

$$y_{b,z} = \sum_j \alpha_{b_j,z} y_{b_j,z} = \sum_j \alpha_j \sum_q \left(rev_{b_j,m_q,z} - cost_{b_j,m_q,z}\right)$$

[27] Ganyu is temporary rural daily work in return of cash or kind. Ganyu is an important source of livelihood for most poor households—for some it is even more important than agricultural net production. Anecdotaly, ganyu is also one of the most important coping strategy for most poor households in the crucial hungry period between food stores running out and the next harvest.

[28] Privately held businesses per household (potentially more than one) include owned nonagricultural businesses that process/sell agricultural byproducts (e.g. flour, juice, beer, jam, oil, seed, and livestock by-products), sales of forest-based products, street or market trading businesses, taxi or pickup truck drivers, bar/restaurants, professional services (e.g., doctor, accountant, lawyer, and midwife) etc..

[29] As was the case for agricultural net income, business net income includes the contribution of household labor to household businesses.

[30] To identify what type last month is in terms of volume sales q, we compare (H:n34-39) and (H:n33).

where $\alpha_{b_j,z}$ is the share of profits from business j kept by the household z (H:n14).[31]

A.2.4 Fishery net production

Part of business income is generated from fishing activities. Fishing net income (by fish species) is also collected. These are provided separately for each of the two landing seasons in a year, high and low. We transform the total quantity per species in kilograms depending on the units reported and the form of packaging, which we use to value sold and unsold production. We net fishing income from intermediate input costs such as rented gears (e.g., mosquito nets, beach seine, long/hand line, gillnet, fish traps), rented boats/engines, fuel, oil, and maintenance, hired labor salaries and other in-kind payments.

Household-z fishing activities are provided separately for each of the two landing seasons, $s = \{\text{high,low}\}$.[32] The total quantity of landed fish species ϕ by household z per season s, $y_{f_\phi,s,z}$, is obtained by multiplying the average quantity of landed fish-ϕ per week in season s (F:e04) times the weeks landed per season s (F:e03).[33] Total value of production per season is computed by multiplying total quantitly landed in per season (F:se06) times reported price (F:e08), or imputed median price (if households do not report selling fish).[34]

In fishery production, each season households also incur in intermediate input costs, $cost_{l,s,z}$, that are the sum of: rented gears (per season) (F:d06);[35] rented boats/engines (per season) (F:d12); fuel, oil, and maintenance (per week) (F:d13); hired labor salaries (adults/week and children/week) (F:d14) $\times$ wages (per week) (F:d16, F:d20, and F:d21) plus other payments to hired labor such as in-kind payments (per week) (F:d18), cash payments as share of boat revenue (per week) (F:d20), and other in-kind payments (meals, cigarettes, etc) (per week) (F:d21);[36] and other costs (per week or season) (F:d24).[37] This way, annual household-z net fishery production is,

$$p_f y_{f,z} = \sum_s \sum_\phi rev_{f_\phi,s,z} + \sum_s \sum_\phi p_{f_\phi,s,r}(y_{f_\phi,s,z} - y^{sold}_{f_\phi,s,z}) - \sum_s cost_{f,s,z}$$

where $\sum_s \sum_\phi p_{f_\phi,s,r}(y_{f_\phi,s,z} - y^{sold}_{f_\phi,s,z})$ is the inferred annual monetary value of household-z fish autoconsumption.

A.2.5 Capital Income

Household-z annual capital income, $y_{k,z}$, includes several sources (H:p0a). These sources are savings, interest, or other investment income (code 104), pension income (code 105), rental income from nonagricultural land rental

[31] Not all owners belong to the same household. The percentage of the profits retained by the household is given by H:n14.

[32] For exposition simplicity, we use high season variable name codes (F:c-f) in the Fishery questionnaire. The procedure for the low season name codes (F:g-i) is analogous.

[33] This information is available for the top 5 landed fish species ϕ. We transform the total quantity in kg. depending on the units reported times the form of packaging (piece, dozen/bundle, kg., 5kg. bag, 10kg. bag, 25kg. bag, smalls basket, large basket, other). An additional dimension is the form of fish processing: fresh, sun-dried, smoked, iced, other.

[34] From sales revenue and the quantity sold we can infer fish-ϕ prices per season s and region r, $p_{f_\phi,s,r}$, as

$$p_{f_\phi,s,r} = q^{50}_{z\in r}\left(\frac{rev_{f_\phi,s,z}}{y^{sold}_{f_\phi,s,z}}\right),$$

where $q^{50}_{z\in r}$ is the median price of fish-ϕ in region r with positive revenues.

[35] Typical gears include mosquito nets, beach seine, long/hand line, gillnet, fish traps, cstnet, other.

[36] In practice, only 12 observations are paid in kind, and 2 fully in kind. Given that the species of fish for this in-kind payments is not reported, and hence pricing the value of in-kind is a problem, we decide not to include these 14 observations related to costs.

[37] Per week costs are multiplied by the total number of weeks landing per per season s (F:e03).

(code 106), apartment, house rental (code 107), shop, store rental (code 108), car, truck, other vehicle rental (code 109), capital gains (including sales) from real estate (code 110), nonagricultural asset sales (code 111), agricultural/fishing asset sales (code 112), and other income such inheritance (code 113), lottery or gambling winnings (code 114), and other income ((spec), code 115).[38] This information is available for the last 12 monhts. Finally, capital income also includes agricultural land rentals (per season), $\sum_s r_{l,s} l_{s,z}^{rent-out}$ (AG:d16-d19), and income from renting out fishery equipment (gears) (per season), $\sum_s \sum_g r_{g,s} k_{g,s,z}^{rent-out}$ (F:e15-e16). This way, household-z capital income is

$$y_{k,z} = \sum_{code} y_{k,z}^{code} + \sum_s r_{l,s} l_{s,z}^{rent-out} + \sum_s r_{g,s} k_{g,s,z}^{rent-out}.$$

A.2.6 Net Transfers

Finally, household annual net transfers are defined as income transfers/gifts received from rural areas/urban areas/other countries minus income transfers/gifts given in the past 12 months. We add the value of received aid (e.g., free maize, other free food, food/cash-for-work programs such as Malawi Social Action Fund or Public Works Program) provided by social safety nets to transfers received. The survey also records remittances in cash received from children 15 years of age or older who no longer live in the household. Neither the Tanzania nor Uganda ISAs have a specific question on in-kind food transfers received by the household. The Malawi survey has such a question, but less than 20% of households report receiving in-kind food transfers in the past 12 months. Nevertheless, in the consumption questionnaire a much higher proportion of households report eating food gifts in the past 7 days. In this 7-day recall data, 62%, 41%, and 25% of households, respectively, in Malawi, Tanzania, and Uganda report consuming food gifts. We therefore include food gifts from the consumption questionnaire (deseasonalized and annualized) in our definition of disposable income. In Malawi, we can compare the contribution of this source of income with other reported transfers. Food gifts represent approximately 6% of total disposable income and dwarfs the 1% contribution of net (cash) transfers in Malawi.

Precisely, household-z annual net transfers are defined as income transfers/gifts received from rural areas/urban areas/other countries $y_{tr,r,z}$ (H:p03) minus income transfers/gifts given out to rural areas/urban areas/other countries $y_{tr,g,z}$ (H:q02) in the last 12 months. These transfers inlcude cash transfers from/to individuals (friends/relatives) (code 101), food transfers (code 102), and nonfood in-kind transfers (code 103). Further, the value of received aid (e.g., free maize, free food (other than maize), food/cash-for-work programs such as MASAF or Public-Works Program (PWP, inputs-for-work program, school feeding program, etc.) provided by social safety nets (social programs) (H:r02) is added to transfers received. Households may also receive remittances from children above 15 years old not living in the household in cash (H:o13-o14) and in-kind (H:o17). Finally, we add as transfers received the annualized and deseasonalized value of household food consumption received from outside the household and estimated in Section A.1.

A.3 Wealth

We measure household-z wealth in terms of net worth, that is, the monetary value of all assets minus liabilities. In Malawi, household wealth largely consists of nonfinancial assets. These assets include houses $p_h h_z$ and other durables $p_d k_{d,z}$ (A.3.1), land $p_l l_z$ (A.3.2), agricultural equipment $p_{k_a^e} k_a^e$ and structures $p_{k_a^s} k_a^s$ (A.3.3), fishery equipment $p_{k_f^e} k_f^e$ (A.3.4), and livestock $p_\omega k_{\omega,z}$ (A.3.5).[39,40] Outstanding debt is given by $debt_z$ (H:s07 and

[38] Capital income is reported by household, not by household member.

[39] Note that the part of agricultural and fishery net production that is stored will show up as wealth one period ahead, but not in the current period.

[40] One cause for potential concern is the lack of a direct question on total savings held at home or at a savings account. We believe this is not an important concern for three reasons. First, there is a specific question about whether the household retrieved savings or received any interest from savings in the last year and only 0.05% of rural households and 3% of urban households replied yes. Second, Beck et al. (2008) report that in Malawi and Uganda checking account fees are more than 20% of household income which makes access close to prohibitive.

H:s09) (A.3.6). That is, household welath (i.e., net worth) is,

$$k_z = p_h h_z + p_d k_{d,z} + p_l l_z + p_{k_a^e} k_a^e + p_{k_a^s} k_a^s + p_{k_f^e} k_f^e + p_\omega k_{\omega,z} - debt_z.$$

A.3.1 Housing (and other durables)

Household-z housing wealth is computed for individuals that own a dwelling (H:f01).[41] The monetary value of housing wealth, $p_h h$, is self reported (H:f02).[42,43] Further, household-z durable goods-d_i (other than housing) typically include furniture, fan, AC, radio, TV, sewing machine, paraffin stove, electric stove, refrigerator, washing machine, bicycle, motorcycle, car, mini-bus, lorry, sattelite dish, solar panel, computer equipment and accessories, generator, etc. The quantity per item of durable good-d_i, k_{d_i} (H:l03), times its self-reported estimated value (H:l05)[44] gives us the monetary value of durable good-d_i owned by household z, $v_{d_i,z}$. The capital value of durable goods (other than houses) for each household-z is

$$p_d k_{d,z} = \sum_i v_{d_i,z}.$$

A.3.2 Land

Household-z land property is composed sum of the value of owned plots of land. A household-z owns a plot if one member of the household does (AG:d04) and ownership is typically acquired by decision of the local leader, inheritance, purchase with title, purchase without a title, or as gift received as bride price (AG:d03).[45] The value of each plot is the estimated selling price $v_{l_{x,z}}$: *"If you were to sell this plot today, how much could you sell it for?"* (AG:d05). The value of land for each household-z is,

$$p_l l_z = \sum_x v_{l_{x,z}}.$$

While the the estimated selling price is preferred to the purchase price, this is not without potential measurement error, particularly for assets such as land, for which the market is largely underdeveloped. For example, in the 2010 survey in Malawi households are asked to provide an estimate of the value of their land and all households do so, but more than four-fifths of households live in areas where no market for land operates (Restuccia and

Third, Brune et al. (2015) find that in a sample of tobacco sellers in Malawi average savings (bank account plus cash held at home) amounted to US$23. Tobacco sellers are a highly selected group as they only represent 13% of rural households and 68% of tobacco sellers are in top quintile of the rural income distribution. The average amount of savings in this selected group represents 1.8% of the average rural household wealth and 0.09% of the average wealth in the top income quintile. The small amount of saving and cash holdings compared to the value of other assets shows that our estimates of wealth inequality would change little by the inclusion of savings and cash holdings.

[41] Alternatives to housing ownership are: employer provides, free (authorized), free (unauthorized), and renting.

[42] "If you sold this dwelling today, how much would you receive for it?"

[43] While accurate information about the size (area) of the house h is not available, making hard to infer the price per housing unit, p_h, housing characteristics such as the number of bedrooms (H:f10), age of the house (H:f05), the quality of the house (e.g. construction material, outer walls, roof, and floor type (H:f06-f09)), source of lighting fuel (H:f11), source of cooking fuel (H:f12), distance to firewood (H:f16), distance to water (H:f38), toilet facility (H:f41), type of garbage disposal (H:f43), etc. are available. That is, it would still be possible to construct a housing price index using these characteristics to infer the housing wealth of those who report ownership but not the value of the house. We do not find this procedure necessary as 97% of housing owners report the value of their dwellings.

[44] "If you wanted to sell on of this [item] today, how much would you receive?"

[45] Otherwise, a land plot cultivated by the household is rented-in and, therefore, is not part of household wealth.

Santaeulàlia-Llopis, 2017). This non-marketed land is either granted by a village chief, inherited, or obtained as bride price. These authors show that the correlation between land quality (at the plot level) and its price is positive and increasing with the amount of marketed land. For instance, for rural households that do not operate any plot obtained through the market, the effects of land quality on land price is a significant 0.116, while for rural households that operated purchased land (with title), this figure increases to 0.503.[46]

A.3.3 Agricultural equipment and structure capital

As in the case of land, the main difference between ISAs and previous LSMS surveys is that in ISAs the quality of capital equipment (and other forms of assets) and its depreciation is measured in a straightforward way by asking: *"How much would you get for this piece of equipment if you sold it now?"* This estimated selling (or resale) price accounts for capital quality and depreciation and avoids the alternative use of the age of the assets (with potential recall error) plus assumptions on the depreciation factor to impute the current value of capital (Deaton, 1997).

Household-z agricultural equipment (implements and machinery) and structure capital is household wealth that contributes to agricultural net production. Agricultural equipment capital includes implements (e.g. hand hoe, slasher, axe, sprayer, panga knife, sickle, treadle pump, and watering can) and machinery (e.g. ox cart, ox plough, tractor, tractor plough, ridger, cultivator, generator, motorised pump, grain mill, and other (spec.)). Denote $k_a'^e$ the end-of-period owned equipment which is the number of units currently owned (H:m01).Then the value of agricultural equipment, $p_{k_a^e} k_a'^e$, is computed as the number of equipment units times the average value per unit $p_{k_a^e}$ (H:m03).[47] Then the value of structure capital, $p_{k_a^s} k_a'^s$, is computed as the number of units times the average value per unit $p_{k_a^s}$ (H:m03).

A.3.4 Fishery equipment capital

Household-z fishery equipment includes gears (e.g. mosquito nets, beach seine, long/ahnd line, gillnet, fish traps, cstnet, and other) and boats/engines. Denote $k_f'^e$ the end-of-period owned equipment which is the number of units currently owned (F:d03 and F:d09). Then the value of fishery equipment, $p_{k_f^e} k_f'^e h$, is computed as the number of equipment units times the average value per unit $p_{k_f^e}$ (F:d04 and F:d10). I

[46] The valuation of land is available for all three countries, Malawi, Tanzania, and Uganda. We find the price of land varies considerably across countries. In Malawi, the median and mean prices of an acre are, respectively, US$214 and 473. In Tanzania, the price distribution is much wider, the median and the mean are, respectively, US$136 and 1,762. In Uganda, prices are consistently higher; the median and mean are, respectively, US$582 and 1,811. Some of these differences in prices may be driven by differences in land quality. In Uganda and in the north of Tanzania, there are two rain-seasons, the territory is hillier, and there is access to Lake Victoria; the staple crops are also more varied. The average price of land in northern Tanzania, US$2,463 per acre, is similar to Uganda's. In south-west Tanzania the land and climate are similar to those in Malawi (e.g., there is one rain-season and the main crop is maize). The average price of land in southern Tanzania, US$255 per acre, is closer to Malawi's. It is reassuring that land quality and prices are correlated as this suggests that price differences tend to capture genuine variation. The difference in the reported price of land across countries might also be due, in part, to differences in market development. In 1998 the Uganda government enacted the Land Act with the explicit aim of turning dwellers on land held under customary tenure into freeholders. Even if the policy was not fully successful (McAuslan, 2003), the development of a land market in Uganda may help explain the disparity in valuations compared with Malawi. This issue deserves further attention beyond the scope of this paper. Nevertheless, we should keep the issue of land markets (or lack thereof) in mind when comparing the reported monetary value of land, hence, wealth — across poor countries.

[47] "If you wanted to sell one of this [item] today, how much would you receive?"

A.3.5 Livestock

Denote $k'_{\omega_i,z}$ as the end-of-period owned livestock-ω_i (AG:r02). Given $k^{sold}_{\omega_i,z}$ and the value of livestock-ω_i owned by household z, $p_{\omega_i,z}$ (AG:r04)[48]. This gives the estimated value of livestock.

A.3.6 Debt

Outstanding debt is given by $debt_z$ (H:s07 and H:s09). Approximately 16% of households report an outstanding debt. The information on debt from ISA data has the potential caveat that focuses on new loans (taken out in the last 12 months) rather than the total amount of outstanding debt. However, we do not think this is an important concern as the household debt that usually rolls over 12 months, such as mortgages and student loans which account for 90% of household debt in the US, are almost nonexistent in Malawi. This helps explain the low aggregate debt-to-income ratio in Malawi from ISA, 10.9. The debt-to income ratio in the US is roughly 10 times higher, 104, measured by the SCF (Bricker et al., 2014).[49]

[48] "If you sold one of the [livestock] today, how much would you receive from the sale?"

[49] Separately in rural and urban areas of Malawi, we find that the debt-to-income ratio for debtors is respectively, 9.5 and 15.3.

B Robustness of Inequality Measures and Correlations Using Panel Data Averages: Malawi, Uganda and Tanzania

We use the available panel data to compute household-specific averages of consumption, income and wealth and re-compute the inequality measures of Table 3 in the main text with these panel-based measures. This implies measures of consumption, income and wealth that are more permanent in nature. For Malawi we report using the panel waves in 2010/11 and 2013 in Table B-1. First, notice that the dispersion of the panel-based variables is lower than that of the cross-sectional data by roughly one fifth in consumption and wealth, and by one third for income. This reduction might be accounted, at least partly, by genuine idiosyncratic transitory shocks. Perhaps more importantly, our main results with cross-sectional data do not change with this panel-based analysis. First, under the new panel-based variables it is still the case that income inequality is larger than consumption inequality and lower than wealth inequality within rural and urban areas. Second, under the new panel-based variables urban inequality is significantly larger than rural inequality by a similar factor.

The correlations between consumption, income and wealth using these panel-based measures are in Table B-2. Note that the panel-based measures raise the value of the correlations, suggesting a stronger relationship among the permanent component of these variables. More importantly, our main results with cross-sectional data do not change with this panel-based analysis. First, the correlation between income and wealth is larger in urban areas than in rural areas, suggesting a higher transmission from income to wealth in urban areas than in rural areas. Second, it is still the case that the correlation between consumption and income is lower in rural areas than in urban areas, suggesting more insurance in rural areas than in urban areas.

The results for Tanzania using panel waves between 2008 and 201 provide similar insights, see Tables (B-3)-(B-4). The results for Uganda using panel waves for 2009/10, 2010/11 and 2011/12 also offer similar insights, see Tables (B-5)-(B-6)

Table B-1: Inequality in Malawi: Cross-Sectional Vs. Panel Data

(a) Variance of Logs

	Cross-Sectional Data: ISA 2010			Average of Panel Data: ISA 2010-2013		
	Rural	Urban	Full	Rural	Urban	Full
Consumption	0.41	0.55	0.50	0.33	0.43	0.40
Income	0.98	1.56	1.09	0.68	0.93	0.75
Wealth	1.49	4.52	1.96	1.22	3.27	1.61
Inequality Ratios:						
▷ C/I	0.42	0.35	0.46	0.48	0.46	0.53
▷ W/I	1.52	2.90	1.80	1.78	3.51	2.14

(b) Gini

	Cross-Sectional Data: ISA 2010			Average of Panel Data: ISA 2010-2013		
	Rural	Urban	Full	Rural	Urban	Full
Consumption	0.36	0.42	0.41	0.32	0.38	0.36
Income	0.53	0.70	0.60	0.47	0.58	0.51
Wealth	0.60	0.84	0.70	0.58	0.75	0.65
Inequality Ratios:						
▷ C/I	0.68	0.60	0.68	0.68	0.65	0.71
▷ W/I	1.13	1.20	1.17	1.23	1.29	1.27

Notes: The measures of inequality that we study are the variance of logged variables in panel (a) and the Gini index in panel (b). The first set of three columns in each panel shows the inequality measures for consumption, income and wealth using the cross-sectional data from ISA 2010 discussed in the main text of the paper (Section 3). The second set of three columns in each panel shows the inequality measures for consumption, income and wealth computed using as cross-sectional data the household-level averages of the panel data available for Malawi in 2010 and 2013.

Table B-2: Correlation of Consumption, Income and Wealth in Malawi: Cross-Sectional Vs. Panel Data

(a) Rural Malawi

(a1) Cross-Section Data: ISA 2010

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.37	1.00	–
Wealth (W)	0.30	0.17	1.00

(a2) Average of Panel Data: ISA 2010-2013

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.41	1.00	–
Wealth (W)	0.33	0.24	1.00

(b) Urban Malawi

(b1) Cross-Section Data: ISA 2010

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.60	1.00	–
Wealth (W)	0.44	0.40	1.00

(b2) Average of Panel Data: ISA 2010-2013

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.63	1.00	–
Wealth (W)	0.53	0.44	1.00

(c) Malawi

(c1) Cross-Section Data: ISA 2010

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.49	1.00	–
Wealth (W)	0.39	0.31	1.00

(c2) Average of Panel Data: ISA 2010-2013

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.57	1.00	–
Wealth (W)	0.48	0.39	1.00

Notes: The top panel (a) refers to rural Malawi, the center panel (b) refers to urban Malawi, and the bottom panel (c) refers to nationwide Malawi. The first table in each panel shows the correlations of consumption, income and wealth using the cross-sectional data from ISA 2010 discussed in the main text of the paper (Section 3). The second table in each panel shows the correlations of consumption, income and wealth using as cross-sectional data the household-level averages of the panel data available for Malawi in 2010 and 2013.

Table B-3: Inequality in Tanzania: Cross-Sectional Vs. Panel Data

(a) Variance of Logs

	Cross-Sectional Data: ISA 2008			Average of Panel Data: ISA 2008-2010		
	Rural	Urban	Full	Rural	Urban	Full
Consumption	0.49	0.40	0.59	0.41	0.53	0.51
Income	1.44	1.89	1.63	1.37	1.71	1.53
Wealth	2.96	4.45	3.28	2.67	4.61	3.10
Inequality Ratios:						
▷ C/I	0.34	0.21	0.36	0.30	0.31	0.33
▷ W/I	2.06	2.35	2.01	1.95	2.70	2.03

(b) Gini

	Cross-Sectional Data: ISA 2008			Average of Panel Data: ISA 2008-2010		
	Rural	Urban	Full	Rural	Urban	Full
Consumption	0.37	0.60	0.41	0.34	0.38	0.39
Income	0.61	0.66	0.64	0.59	0.61	0.62
Wealth	0.76	0.94	0.81	0.74	0.92	0.79
Inequality Ratios:						
▷ C/I	0.61	0.91	0.64	0.58	0.62	0.63
▷ W/I	1.25	1.42	1.27	1.25	1.51	1.27

Notes: The measures of inequality that we study are the variance of logged variables in panel (a) and the Gini index in panel (b). The first set of three columns in each panel shows the inequality measures for consumption, income and wealth using the cross-sectional data from ISA 2008 discussed in the main text of the paper (Section 3). The second set of three columns in each panel shows the inequality measures for consumption, income and wealth computed using as cross-sectional data the household-level averages of the panel data available for Tanzania in 2008 and 2010.

Table B-4: Correlation of Consumption, Income and Wealth in Tanzania: Cross-Sectional Vs. Panel Data

(a) Rural Tanzania

(a1) Cross-Section Data: ISA 2010

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.41	1.00	–
Wealth (W)	0.31	0.22	1.00

(a2) Average of Panel Data: ISA 2008-2010

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.43	1.00	–
Wealth (W)	0.26	0.16	1.00

(b) Urban Tanzania

(b1) Cross-Section Data: ISA 2010

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.46	1.00	–
Wealth (W)	0.06	0.01	1.00

(b2) Average of Panel Data: ISA 2008-2010

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.42	1.00	–
Wealth (W)	0.06	0.03	1.00

(c) Tanzania

(c1) Cross-Section Data: ISA 2010

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.46	1.00	–
Wealth (W)	0.15	0.10	1.00

(c2) Average of Panel Data: ISA 2008-2010

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.45	1.00	–
Wealth (W)	0.10	0.05	1.00

Notes: The top panel (a) refers to rural Tanzania, the center panel (b) refers to urban Tanzania, and the bottom panel (c) refers to nationwide Tanzania. The first table in each panel shows the correlations of consumption, income and wealth using the cross-sectional data from the ISA 2010 discussed in the main text of the paper (Section 3). The second table in each panel shows the correlations of consumption, income and wealth using as cross-sectional data the household-level averages of the panel data available for Tanzania in 2008 and 2010. Note that the value of dwellings is not reported in the wealth questionnaire. This means that urban households in particular are assigned artificially low wealth.

Table B-5: Inequality in Uganda: Cross-Sectional Vs. Panel Data

(a) Variance of Logs

	Cross-Sectional Data: ISA 2010			Average of Panel Data: ISA 2005-2010		
	Rural	Urban	Full	Rural	Urban	Full
Consumption	0.83	1.04	1.01	0.58	0.68	0.77
Income	1.67	2.02	1.84	2.07	2.36	2.36
Wealth	2.81	4.67	3.15	1.95	3.96	2.40
Inequality Ratios:						
▷ C/I	0.50	0.51	0.55	0.28	0.29	0.33
▷ W/I	1.68	2.31	1.71	0.94	1.68	1.02

(b) Gini

	Cross-Sectional Data: ISA 2010			Average of Panel Data: ISA 2005-2010		
	Rural	Urban	Full	Rural	Urban	Full
Consumption	0.48	0.48	0.53	0.42	0.42	0.48
Income	0.74	0.77	0.76	0.65	0.73	0.72
Wealth	0.75	0.80	0.77	0.70	0.75	0.74
Inequality Ratios:						
▷ C/I	0.65	0.62	0.70	0.65	0.57	0.67
▷ W/I	1.07	1.04	1.01	1.07	1.03	1.03

Notes: The measures of inequality that we study are the variance of logged variables in panel (a) and the Gini index in panel (b). The first set of three columns in each panel shows the inequality measures for consumption, income and wealth using the cross-sectional data from ISA 2009/10 discussed in the main text of the paper (Section 3). The second set of three columns in each panel shows the inequality measures for consumption, income and wealth computed using as cross-sectional data the household-level averages of the panel data available for Uganda in 2005 and 2009/10,.

Table B-6: Correlation of Consumption, Income and Wealth in Uganda: Cross-Sectional Vs. Panel Data

(a) Rural Uganda

(a1) Cross-Section Data: ISA 2010

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.33	1.00	–
Wealth (W)	0.33	0.08	1.00

(a2) Average of Panel Data: ISA 2005-2010

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.45	1.00	–
Wealth (W)	0.54	0.27	1.00

(b) Urban Uganda

(b1) Cross-Section Data: ISA 2010

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.45	1.00	–
Wealth (W)	0.36	0.17	1.00

(b2) Average of Panel Data: ISA 2005-2010

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.36	1.00	–
Wealth (W)	0.54	0.34	1.00

(c) Uganda

(c1) Cross-Section Data: ISA 2010

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.44	1.00	–
Wealth (W)	0.32	0.12	1.00

(c2) Average of Panel Data: ISA 2005-2010

	C	I	W
Consumption (C)	1.00	–	–
Income (I)	0.43	1.00	–
Wealth (W)	0.55	0.34	1.00

Notes: The top panel (a) refers to rural Uganda, the center panel (b) refers to urban Uganda, and the bottom panel (c) refers to nationwide Uganda. The first table in each panel shows the correlations of consumption, income and wealth using the cross-sectional data from ISA 2009/10 discussed in the main text of the paper (Section 3). The second table in each panel shows the correlations of consumption, income and wealth using as cross-sectional data the household-level averages of the panel data available for Uganda in 2005 and 2009/10.

C More Evidence on Income Mobility: Additional Tables and Figures

In Table C-1, we show the income mobility matrices separately for rural and urban Malawi. The patterns of higher persistence of income in urban areas than in rural areas is clear from the diagonal elements of the matrix. The income mobility matrices for Tanzania and Uganda show similar partterns to those of Malawi as described in the main text, Table C-2.

An isomorphic representation of the income mobility matrix for Malawi is in the conditional transition probabilities depicted in panel (a) Figure C-1. The vertical axis refers to the quintile of origin in the transition and the horizontal axis shows the conditional transition probability to a destination quintile identified with the colors labeled at the bottom of each figure. We provide analogous mobility matrices for Tanzania (2008-2010) in panel (b) and for Uganda (2009-2011) in panel (c) of Figure C-1. Despite the slight year span difference, the three countries show similar insights: there is larger persistence at the top of the income distribution than at the bottom. Roughly 40% of households in the bottom quintile remain in that quintile after 2-3 years in all countries, and roughly 50% of households in the top quintile remain in the top quintile.[50]

Finally, Figure C-2 shows the predicted future income ranking of households given the current income ranking. The ranking is normalized between position 0 (lowest income) and 100 (highest income). In all panels the horizontal axis shows the income ranking of households today, and the vertical axis the ranking of the same households tomorrow. Each panel household observation is a dot in those graphs. If our observations align with the 45 degree line this would imply that there is absolutely no mobility across these two years. In contrast, if our panel household observations align with a 120 degree line this would imply that the income ranking of households is fully reversed across these two years. The orange line shows the predicted income ranking tomorrow as the outcome of regressing the income ranking tomorrow on a cubic polynomial of the income ranking today. Several observations are in order. Focusing on Malawi (panel (a), Figure C-2), households ranked today at the bottom of the income distribution, position 0, are predicted to rank at roughly at position 35 of the income distribution tomorrow. At the same time, households ranked today at the top of the income distribution, position 100, are predicted to rank roughly at position 80 of the income distribution tomorrow. Tanzania (panel (b)) and Uganda (panel (c)) show similar insights. Instead, the US shows sharp differences, panel (d). In the US, households ranked today at the bottom of the income distribution, position 0, are predicted to rank roughly at position 5 of the income distribution tomorrow. At the same time, households ranked today at the top of the income distribution, position 100, are predicted to rank roughly at position 95 tomorrow. That is, again, there is a higher degree of income mobility in SSA than in the US.

[50] The Malawi ISA 2010-13 panel sample has 2,405 households, the Tanzania ISA 2008-10 panel sample has 2,838 households and the Uganda ISA 2009-2011 panel sample has 1,397 households.

Table C-1: Income Mobility Matrices: Rural and Urban Malawi ISA 2010-2013

(a) Rural Malawi 2010-13

t \ t+1	Q1	Q2	Q3	Q4	Q5	Full
Q1	32.68	26.82	18.72	14.80	6.98	100
Q2	26.52	22.93	23.76	18.51	8.29	100
Q3	20.61	17.27	22.28	24.79	15.04	100
Q4	11.63	20.50	19.39	25.48	22.99	100
Q5	10.39	12.36	15.45	16.29	45.51	100

(b) Urban Malawi 2010-13

t \ t+1	Q1	Q2	Q3	Q4	Q5	Full
Q1	34.55	21.82	21.82	16.36	5.45	100
Q2	26.21	35.92	21.36	14.56	1.94	100
Q3	15.60	24.77	35.78	17.43	6.42	100
Q4	13.21	13.21	18.87	34.91	19.81	100
Q5	4.85	2.91	2.91	19.42	69.90	100

Notes: Income data are divided by quintiles with 1st denoting the poorest quintile and 5th the richest. Our 2010-2013 Malaiwi ISA panel has a size of 2,405 households, with 1,812 in rural areas and 532 in urban areas. The discrepancy between the sum of rural and urban households and the total is driven by migrants.

Table C-2: Income Mobility Matrices in Tanzania and Uganda: From Quintiles to Quintiles

(c) Malawi 2010-13

t \ t+1	Q1	Q2	Q3	Q4	Q5	Full
Q1	33.40	25.68	17.95	17.12	5.85	100
Q2	27.29	24.79	25.21	15.83	6.88	100
Q3	18.71	21.21	22.04	24.53	13.51	100
Q4	12.29	18.75	22.08	24.58	22.29	100
Q5	8.14	9.42	12.85	17.99	51.61	100

(b) Tanzania 2008-10

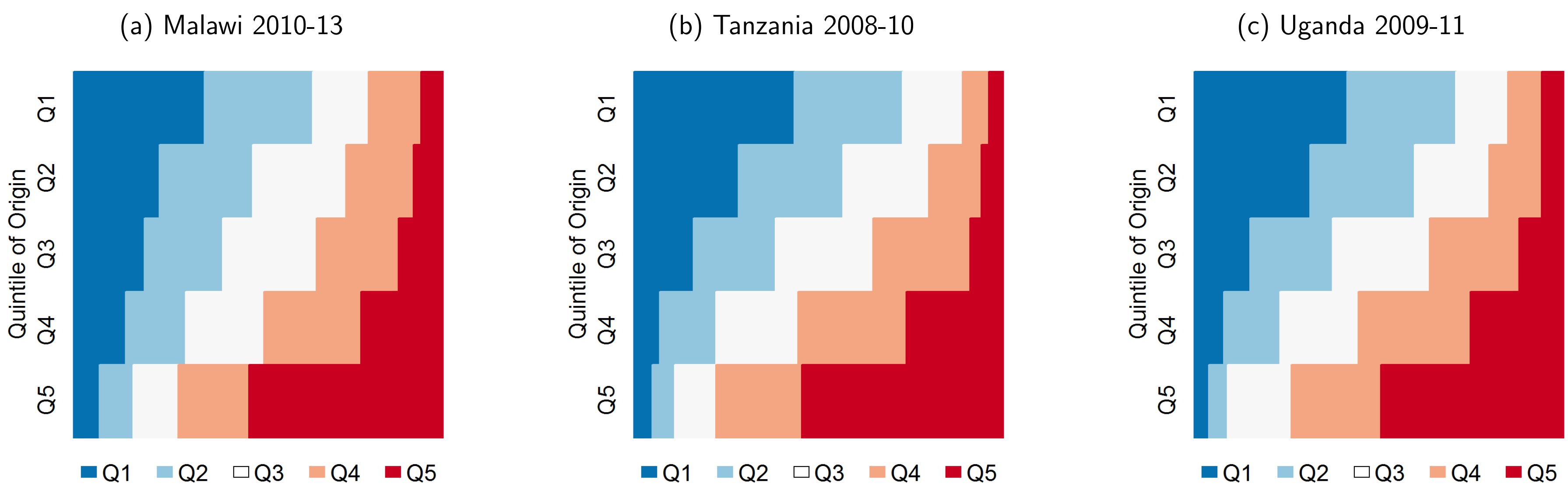
t \ t+1	Q1	Q2	Q3	Q4	Q5	Full
Q1	40.97	26.74	16.49	9.38	6.42	100
Q2	28.84	30.22	21.93	13.64	5.35	100
Q3	17.32	22.13	27.44	24.36	8.75	100
Q4	9.29	14.11	22.55	29.09	24.96	100
Q5	4.20	6.30	11.56	23.82	54.12	100

(c) Uganda 2005-2009/10

t \ t+1	Q1	Q2	Q3	Q4	Q5	Full
Q1	33.33	23.97	20.26	12.64	9.80	100
Q2	25.16	31.95	22.98	15.10	4.81	100
Q3	19.87	25.55	22.05	21.18	11.35	100
Q4	14.04	12.94	21.27	30.04	21.71	100
Q5	7.56	6.00	13.33	21.33	51.78	100

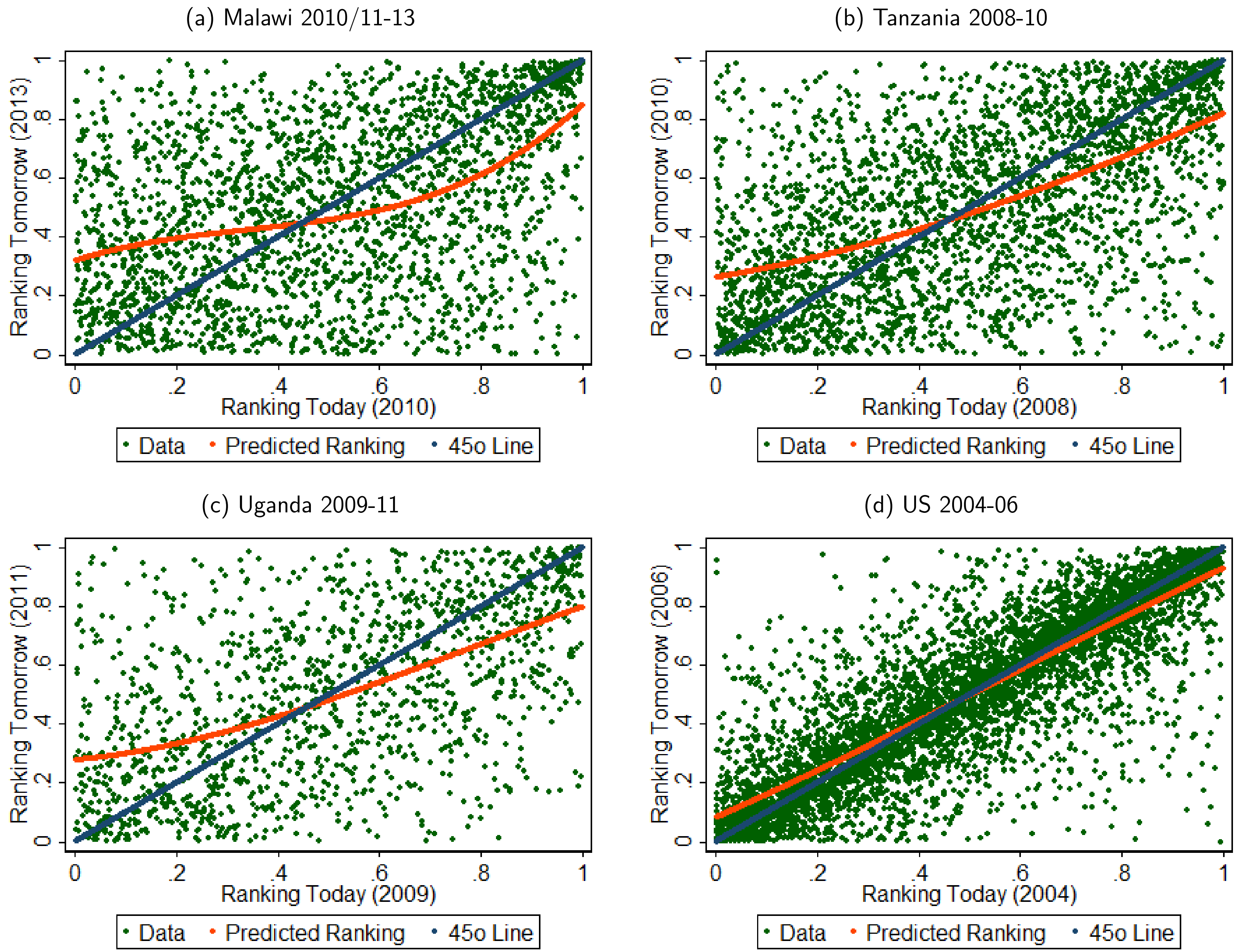
Notes: Income data are divided by quintiles with 1st denoting the poorest quintile and 5th the richest. The 2009-2011 Uganda ISA panel sample is 1,397, and the 2008-10 Tanzania ISA panel sample is 2,838. A discussion of this table is in Section 4.1.

Figure C-1: Income Transition Probability Plots: Malawi, Uganda and Tanzania



Notes: In each panel-year, income data are divided by quintiles with Q1 denoting the poorest quintile and Q5 the richest. In all panels the vertical axis refers to the quintile of origin in the transition and the horizontal axis shows the conditional transition probability to a destination quintile identified with the colors labeled at the bottom of each figure. The Malawi ISA 2010-13 panel sample has 2,405 households, the Uganda ISA 2009-2011 panel sample has 1,397 households, and the Tanzania ISA 2008-10 panel sample has 2,838 households.

Figure C-2: Predicted Income Ranking: From t to $t+2$



Notes: In all panels the horizontal axis shows the income ranking of households today, and the vertical axis the income ranking of the same households tomorrow. That is, each dot is a household observation. In all panels, the blue line denotes the 45 degree line. In all panels the orange line shows the predicted ranking tomorrow as the outcome of a regressing the ranking tomorrow on a cubic polynomial of the ranking today. The Malawi ISA 2010-13 panel sample has 2,405 households (panel (a)), the Tanzania ISA 2008-10 panel sample has 2,838 households (panel (b)) and the Uganda ISA 2009-2011 panel sample has 1,397 households (panel (c)). The 2004-06 PSID sample has 5,649 households (panel (d)).

References

Beck, T., Demirgüç-Kunt, A., and Peria, M. S. M. (2008). Banking services for everyone? barriers to bank access and use around the world. *The World Bank Economic Review*, 22(3):397–430.

Beegle, K., Carletto, C., and Himelein, K. (2012a). Reliability of recall in agricultural data. *Journal of Development Economics*, 98(1):34–41.

Beegle, K., De Weerdt, J., Friedman, J., and Gibson, J. (2012b). Methods of household consumption measurement through surveys: Experimental results from tanzania. *Journal of Development Economics*, 98(1):3–18.

Bricker, J., Dettling, L.-J., Henriques, A., Hsu, J.-W., Moore, K.-B., Sabelhaus, J., Thompson, J., and Windle, R. A. (2014). Changes in U.S. Family Finances from 2010 to 2013: Evidence from the Suvery of Consumer Finances. *Federal Reserve Bulletin*, 100(4).

Brune, L., Giné, X., Goldberg, J., and Yang, D. (2015). Facilitating savings for agriculture: Field experimental evidence from malawi. Technical report, National Bureau of Economic Research.

Carletto, C., Savastano, S., and Zezza, A. (2013). Fact or artifact: the impact of measurement errors on the farm size–productivity relationship. *Journal of Development Economics*, 103:254–261.

Carletto, G., Beegle, K., Himelein, K., Kilic, T., Murray, S., Oseni, M., Scott, K., and Steele, D. (2010). Improving the availability, quality and policy-relevance of agricultural data: The living standards measurement study-integrated surveys on agriculture. In *Third Wye City Group Global Conference on Agricultural and Rural Household Statistic. Washington, DC*, pages 24–25.

Deaton, A. (1997). *The Analysis of Household Surveys. A Microeconometric Approach to Development Policy*. Johns Hopkins University Press.

Deaton, A. and Zaidi, S. (2002). Guidelines for Constructing Consumption Aggregates for Welfare Analysis. Living Standards Measurement Study, Working Paper No. 135, Washington, D.C.: World Bank.

Deininger, K., Carletto, C., Savastano, S., and Muwonge, J. (2012). Can diaries help in improving agricultural production statistics? evidence from uganda. *Journal of Development Economics*, 98(1):42–50.

Goldberg, J. (2016). Kwacha Gonna Do? Experimental Evidence about Labor Supply in Rural Malawi. *American Economic Journal: Applied Economics*, 8(1):129–149.

Grosh, M. and Deaton, A. (2000). *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study*, volume 17, chapter Consumption. Washington, DC: World Bank.

Grosh, M. and Glewwe, P. (2000). *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study, Volume 2*. Washington, DC: World Bank.

McAuslan, P. (2003). A Narrative on Land Law Reform in Uganda. Lincoln Institute of Land Policy - Conference Paper.

MNSO (2012). Malawi Third Integrated Household Survey (IHS3) 2010-2011: Basic Information Document. *MNSO: Malawi National Statistical Office*.

Restuccia, D. and Santaeulàlia-Llopis, R. (2017). Land Misallocation and Productivity. Working Paper 23128, National Bureau of Economic Research.

Rosenzweig, M. R. and Udry, C. (2014). Rainfall Forecasts, Weather and Wages over the Agricultural Production Cycle. NBER Working paper 19808.