

Measurement Error and Farm Size: Do Nationally Representative (LSMS-ISA) Surveys Provide Reliable Estimates?

Stein T. Holden, Clifton Makate, and Sarah Tione

School of Economics and Business, NMBU, Ås, Norway

Lilongwe University of Agriculture and Natural Resources, Lilongwe,
Malawi

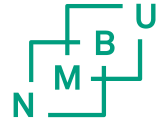
Email: stein.holden@nmbu.no

Homepage: www.steinholden.com

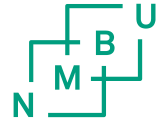
Introduction

- *We assess the reliability of measured farm sizes (ownership holdings) in the Living Standard Measurement Study – Integrated Surveys on Agriculture (LSMS-ISA) in Ethiopia and Malawi based on*
 - *3 survey rounds (2012, 2014, 2016) in Ethiopia*
 - *4 survey rounds (2010, 2013, 2016, 2019) in Malawi*
 - *We utilize a balanced household sample in both countries*
 - *Farm parcels measured by GPS and/or rope/compass*
 - *This gives reliable estimates of reported parcels*
 - *We detect substantial under-reporting of parcels over time within households that largely have been overlooked in previous studies*

Key findings in Ethiopia and Malawi

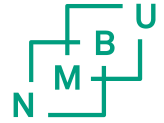


- Parcel attrition leads to
- **Downward bias (23-41%) in estimated average and median farm sizes (ownership holdings) within survey rounds**
- The bias is substantial across both countries, across all survey rounds, and across all regions in both countries
- This non-classical measurement error also contributes to **upward bias in Gini-coefficients** for ownership holding distributions
- We utilized within-household variation in ownership holdings over years to identify and assess the size of the problem
- We tested a variety of models to attempt to correct it



Theoretical framework I

- **Theories to explain real farm size (ownership holding) variation**
 - a. Inheritance and bequeath
 - b. Purchases or sales
 - c. Administrative expropriations and land redistributions
 - d. Private land takings and losses

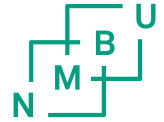


Theoretical framework II

Theories to explain errors in farm size measurement

- a. Farmers' incentives to under-report (reduce burden)
- b. Enumerators also have incentives to reduce the work burden
- c. Surveys tend to focus on the main (large) nearby parcels of a farm and leave out small parcels of less significance and parcels that are located far away
- d. Rented-out parcels are more likely to be left out from the survey as such parcels are not managed by the household included in the survey
- e. Improvements in data collection technologies and methods have improved over time (reduced information asymmetries and transaction costs) and may have improved data quality

Approach



- We are unable to test each of the theoretical propositions explicitly because of data limitations
 - LSMS-ISA data are incomplete wrt parcel transfers and timing of such transfers
 - Moral hazard problems leading to parcel attrition cannot be inferred directly
- In this «second-best world» we test alternative approaches:
 - **Alternative estimators** to correct for real farm size changes over time
 - Indicators for parcel attrition and farm size changes
 - **Use maximum within-household farm size over years as benchmark**

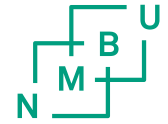
Data: Ethiopian sample

	Year of survey			
Region	2012	2014	2016	Total
Tigray	252	252	252	756
Afar	35	35	35	105
Amhara	498	498	498	1,494
Oromiya	471	471	471	1,413
Somalie	127	127	127	381
Benishangul Gumuz	72	72	72	216
SNNP	665	665	665	1,995
Gambella	50	50	50	150
Harari	73	73	73	219
Dire Dawa	96	96	96	288
Total	2,339	2,339	2,339	7,017

Ownership holdings by year

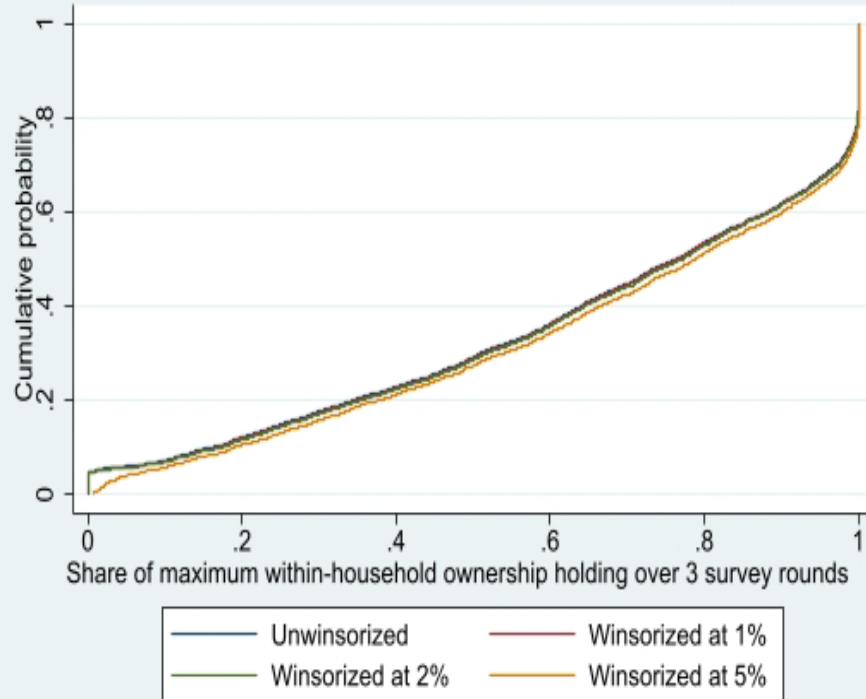
	Unwinsorized Ownership holdings, ha			1% winsorized Ownership holdings, ha			Max ownership holdings 2012- 2016 in ha, winsorized at:		
	2012	2014	2016	2012	2014	2016	1%	2%	5%
Year	2012	2014	2016	2012	2014	2016	1%	2%	5%
Mean	1.140	1.338	1.208	1.066	1.162	1.132	1.572	1.510	1.390
Median	0.702	0.791	0.753	0.702	0.791	0.753	1.117	1.117	1.117
P25	0.278	0.331	0.307	0.278	0.331	0.307	0.560	0.560	0.560
P75	1.408	1.523	1.434	1.408	1.523	1.434	2.009	2.009	2.009
P90	2.397	2.550	2.560	2.397	2.550	2.560	3.344	3.344	3.344
sdev	2.023	3.950	2.128	1.193	1.267	1.271	1.525	1.322	1.033
n	2345	2345	2345	2345	2345	2345	2345	2345	2345
Gini	0.539	0.57	0.545	0.507	0.506	0.515	0.473	0.453	0.413

Within-household ownership holding shares of maximum household ownership holdings in 2012 and 2014



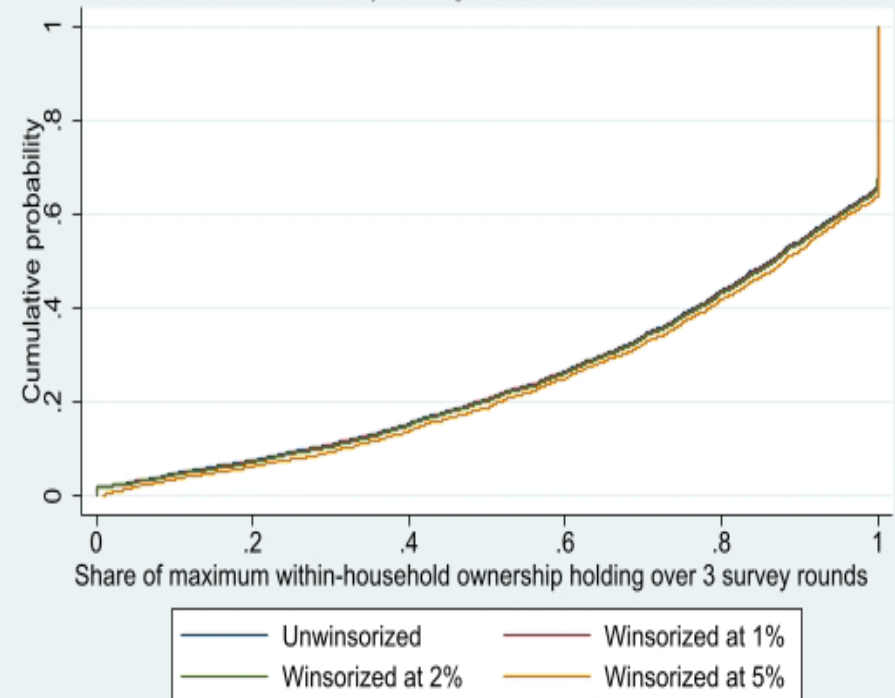
Ethiopia panel 2012: Ownership holding as share of maximum ownership holding

Measurement error correction by winsorizing at 1, 2, and 5% levels in each end of distributions



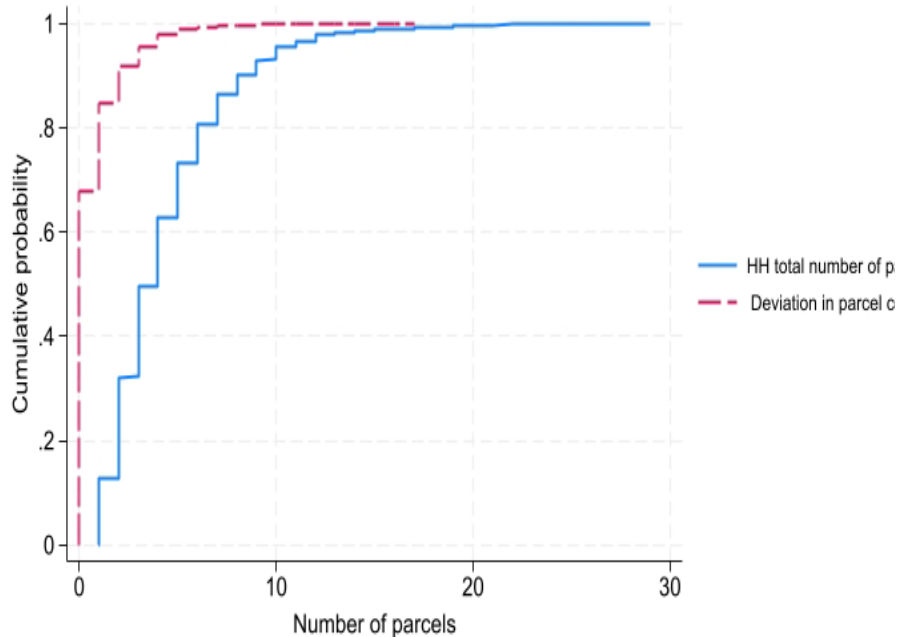
Ethiopia panel 2014: Ownership holding as share of maximum ownership holding

Measurement error corrections by winsorizing at 1, 2, and 5% levels in each end of distributions



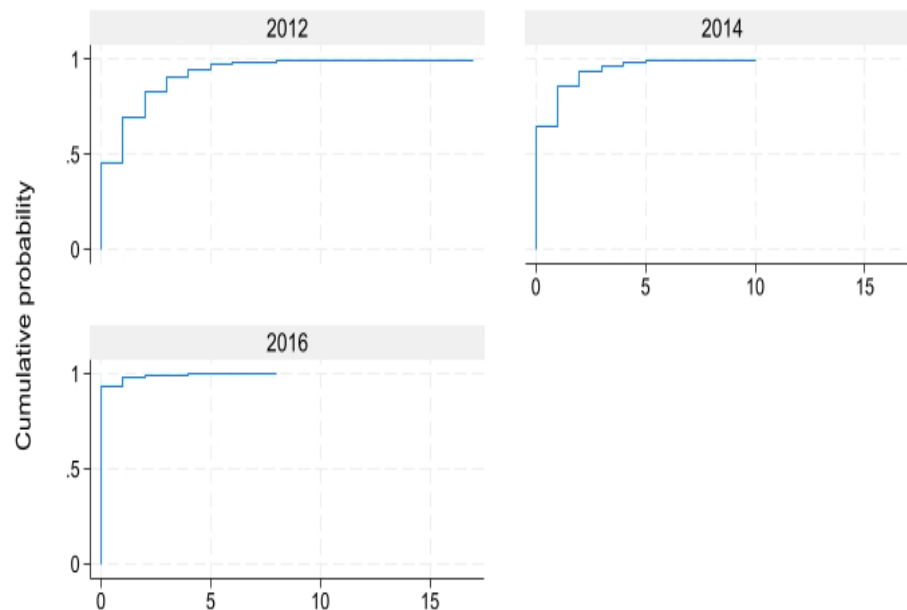
Ethiopia: Parcel count and deviation from max parcel count for households

Ethiopia: Number of parcels and deviation from max number of parcels



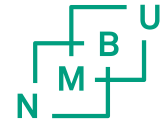
Cumulative distributions full panel sample

Ethiopia: Deviation in parcel count from max by survey round



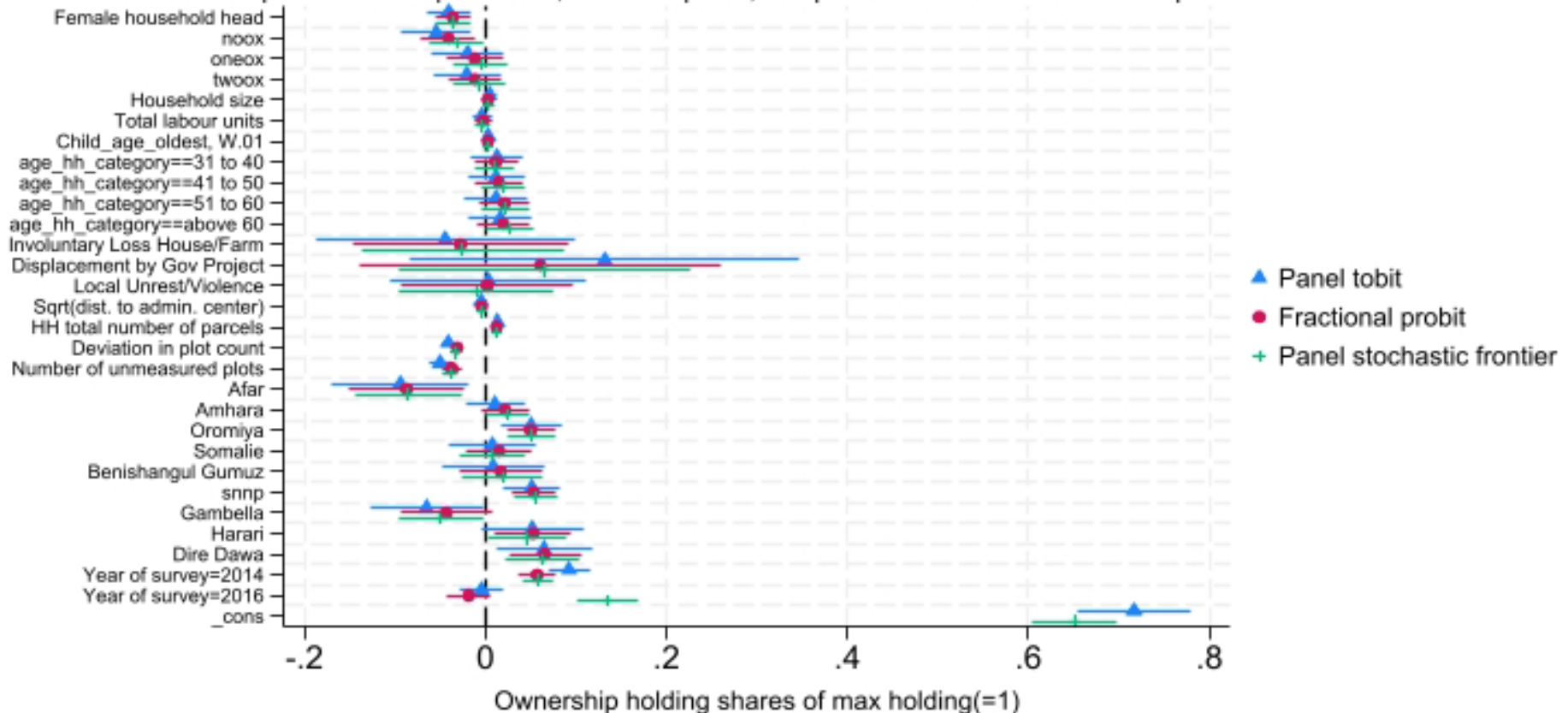
Graphs by Year of survey

Ethiopia Panel Censored Tobit models



Ethiopia: Robustness check: Full panel with alternative estimators

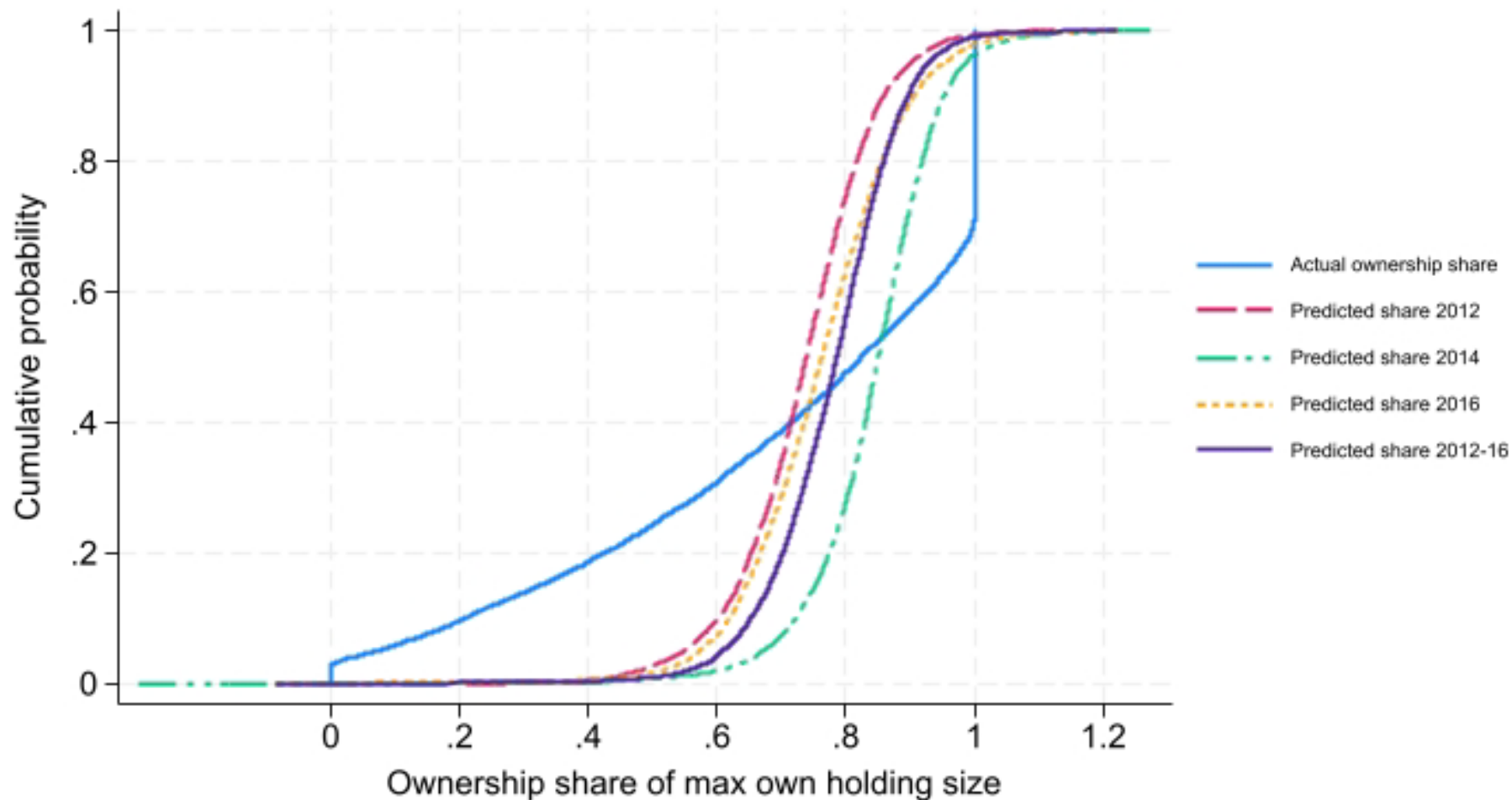
Compare censored panel tobit, fractional probit, and panel stochastic frontier model parameters



All models corrected for attrition with inverse probability weighting, 95% CI

Use 3-round panel 2012-2016. Dependent variable: Ownership holding share of max holding

Ethiopia: Actual versus predicted ownership holding shares



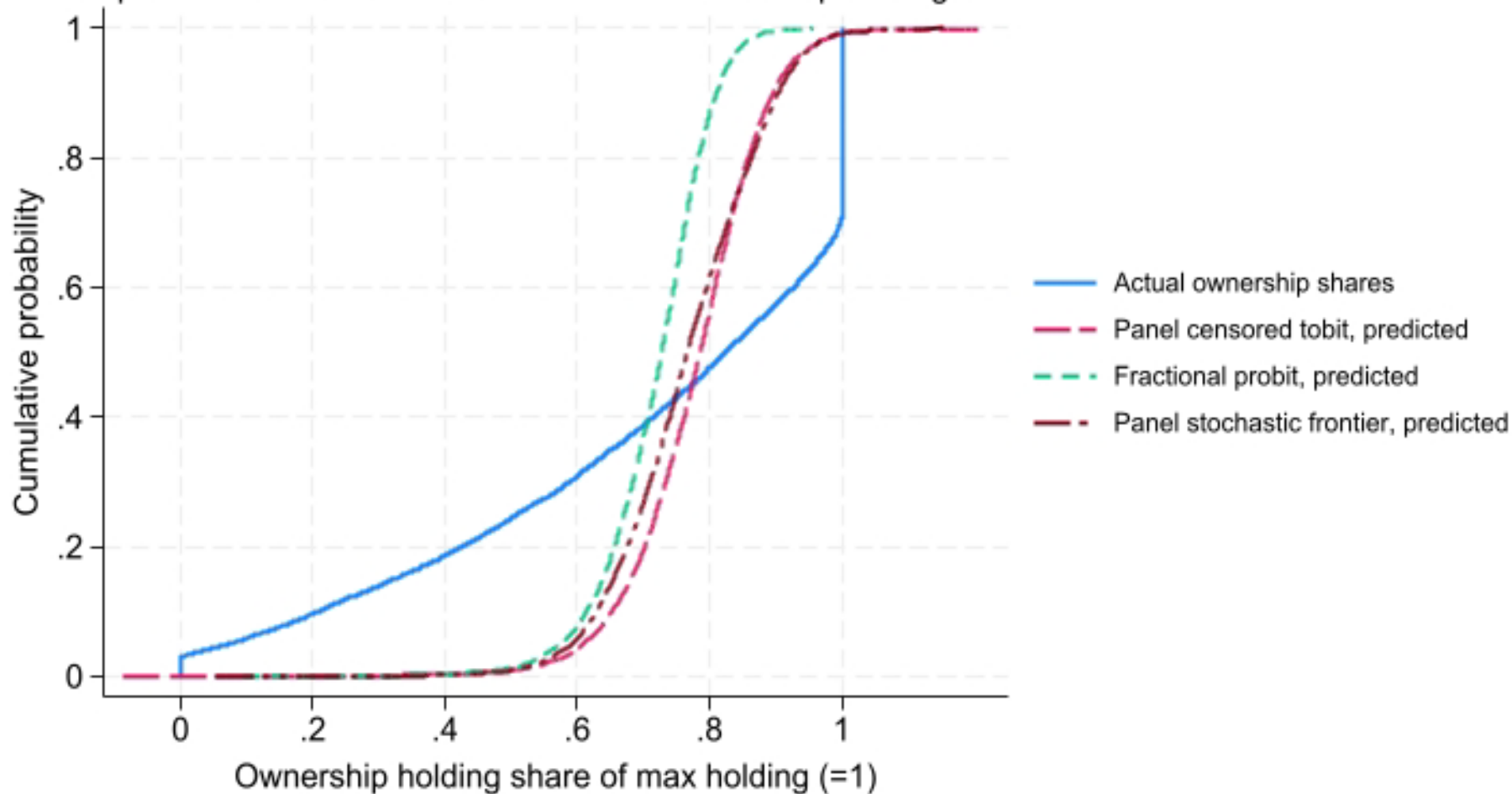
Robustness tests with alternative estimators



- In addition to Censored Tobit models by year and Panel Censored Tobit models, robustness tests were done with
- Symmetrically censored least squares estimator (SCLS)
- Fractional probit models
- Panel stochastic frontier models
- They gave similar poor predictions of ownership share holdings of maximum within-household holdings
- We concluded that the maximum within-household ownership holding represents the best predictor of farm size

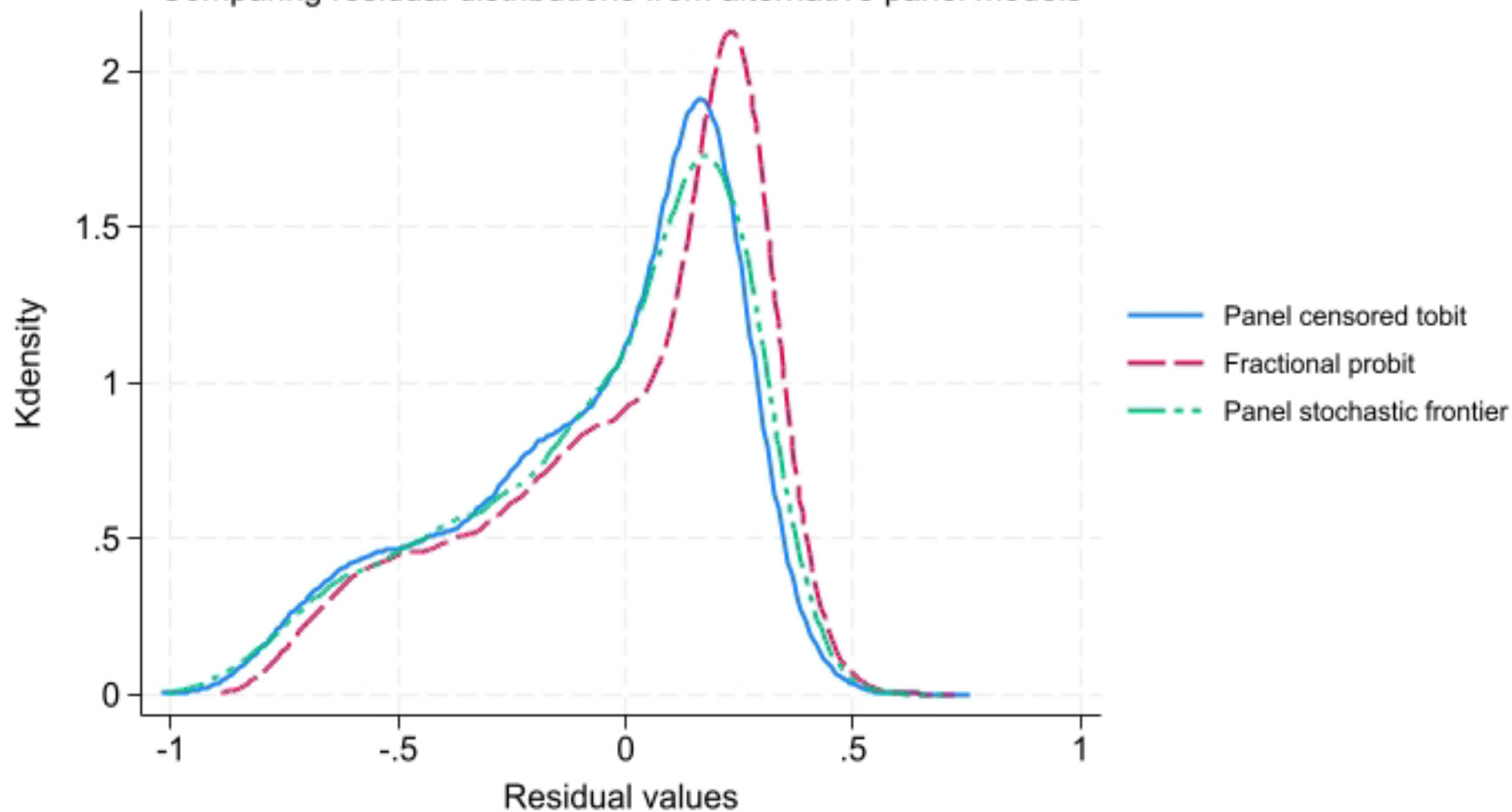
Ethiopia: Robustness checks: Alternative estimators

Full panel models: Predicted distributions of ownership holding shares



Ethiopia: Residual distributions from panel models

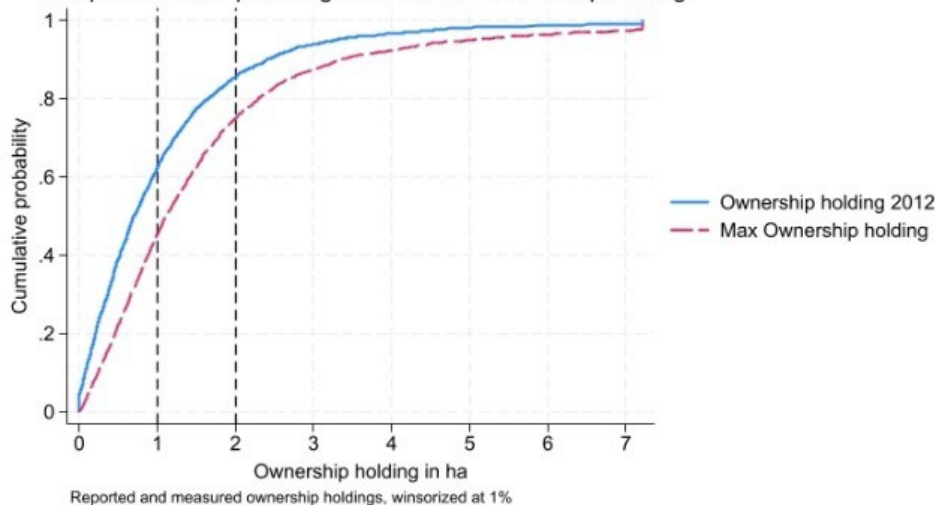
Comparing residual distributions from alternative panel models



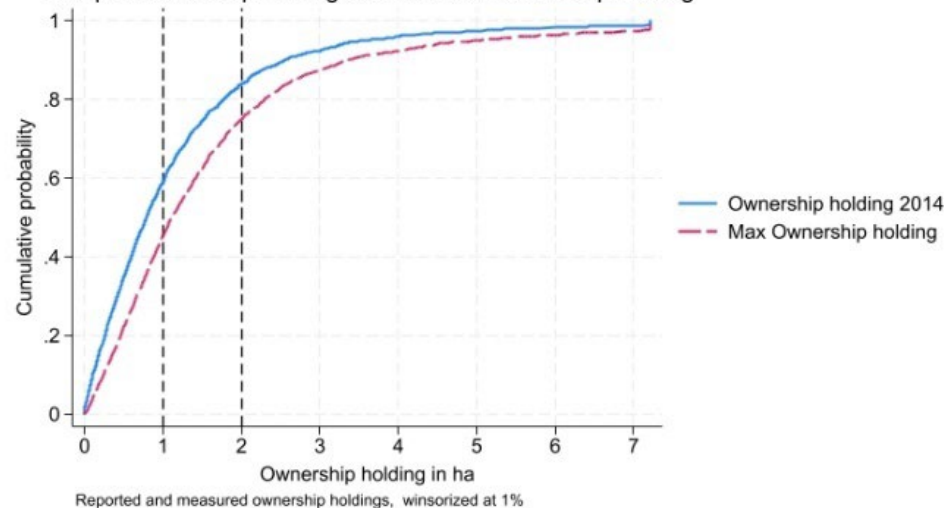
Ethiopia: Farm size distributions by survey round vs. Max within-household holding size



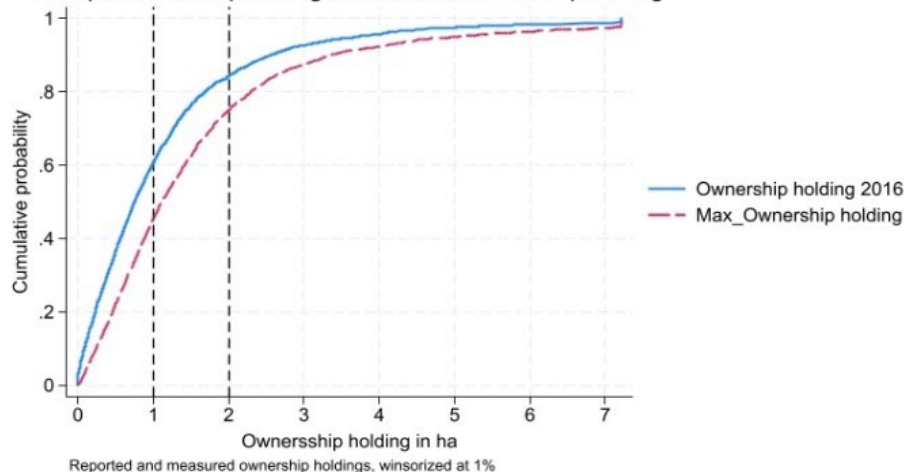
Ethiopia: Ownership holding 2012 vs. Max Ownership holding



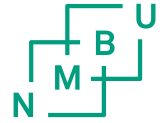
Ethiopia: Ownership holding 2014 vs. Max Ownership holding



Ethiopia: Ownership holding 2016 vs. Max Ownership holding

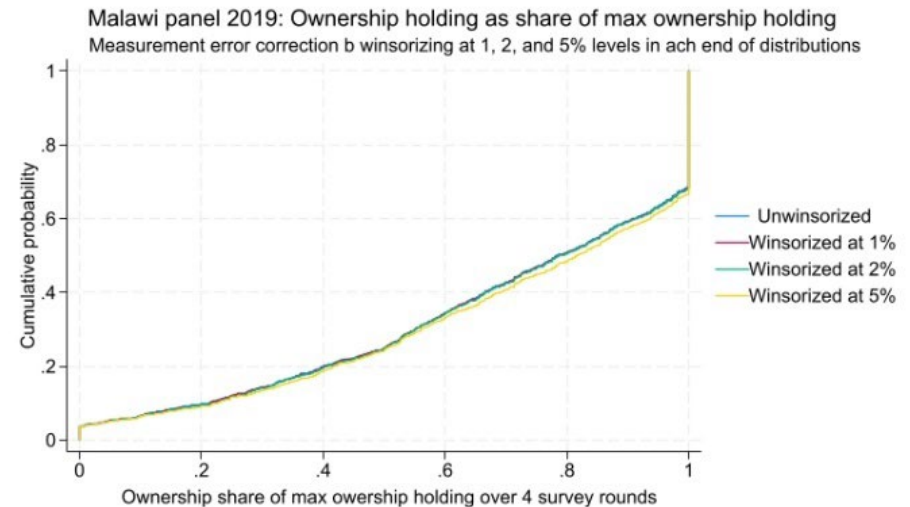
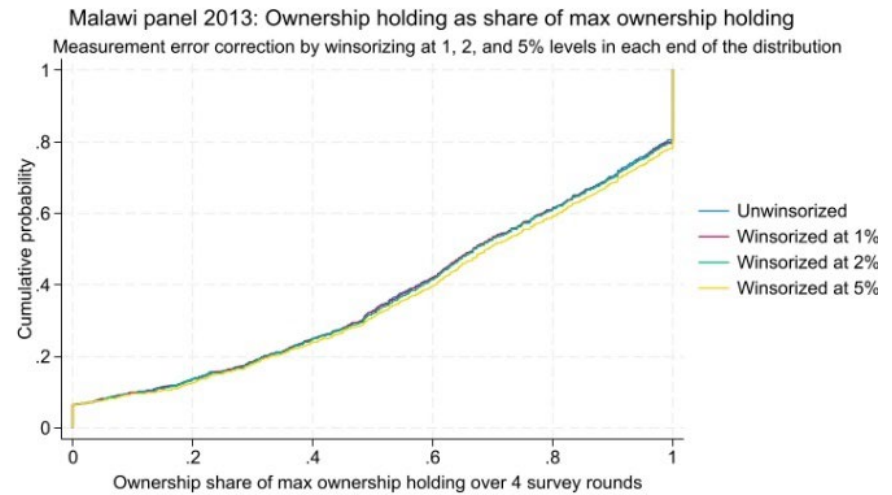
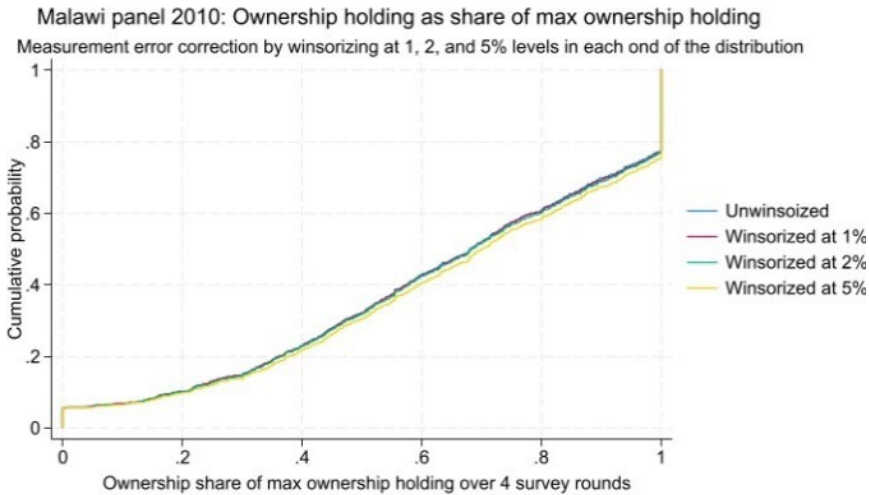
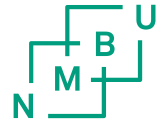


Malawi: Sample size: Number of households

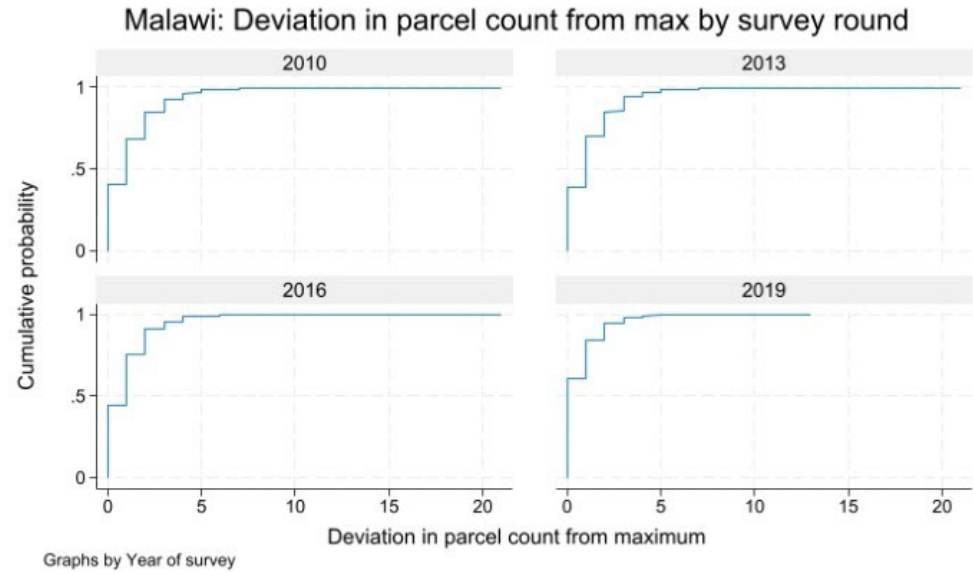
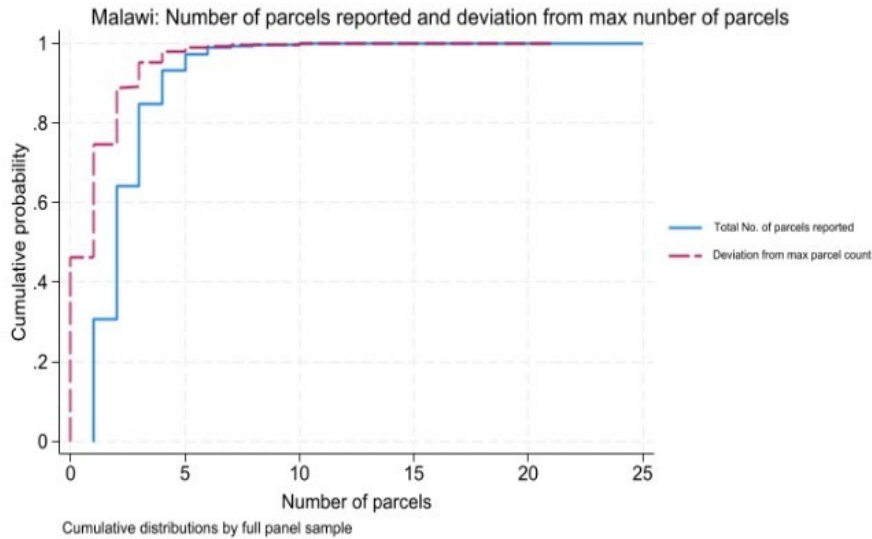
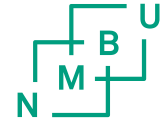


Region	2010	2013	2016	2019	Total
North	86	86	86	86	344
Central	416	416	416	416	1,664
Southern	484	484	484	484	1,936
Total	986	986	986	986	3,944

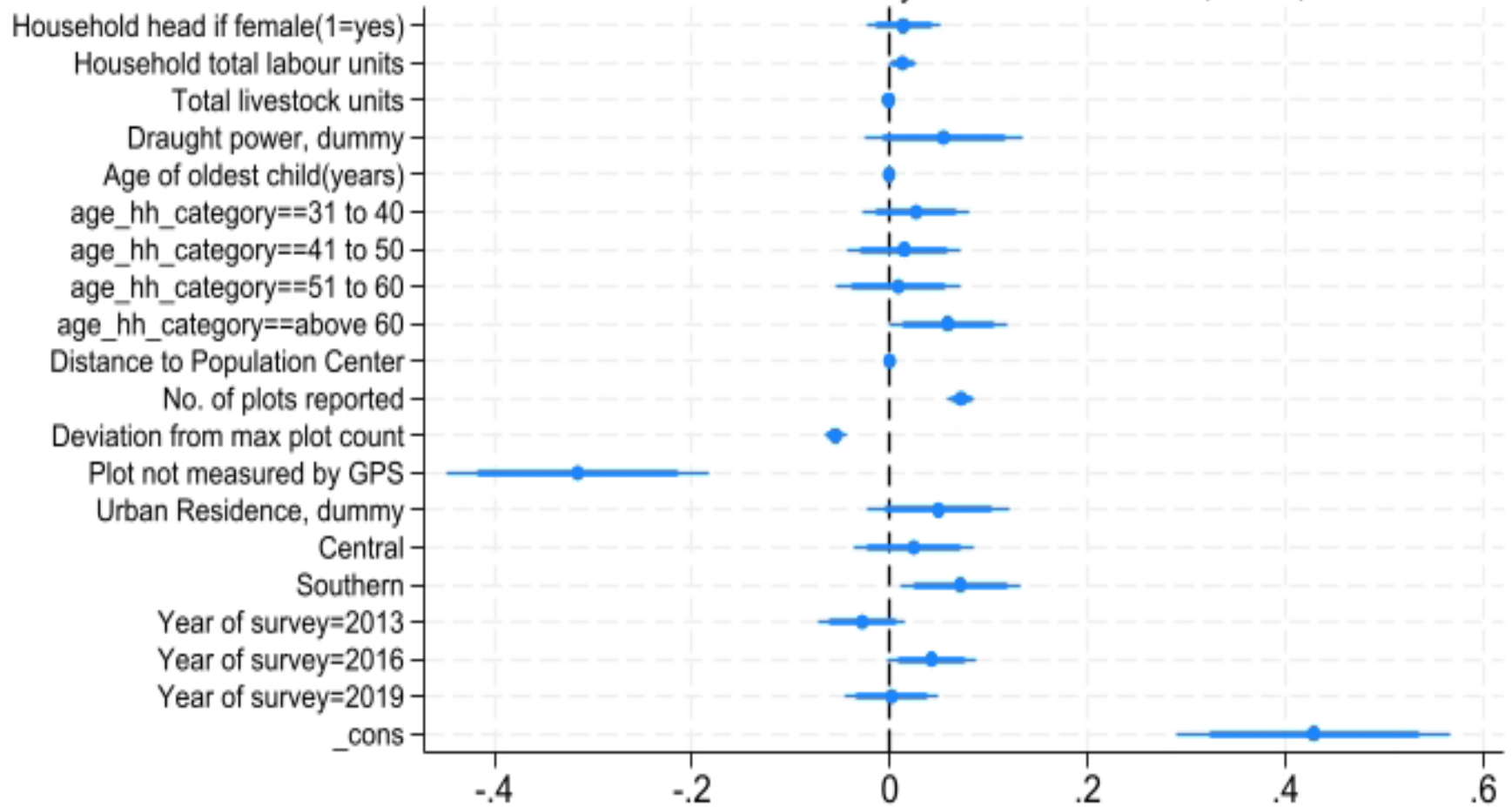
Within-household ownership holding shares of max ownership holding by year



Malawi: Parcel count and deviation from max parcel count for households



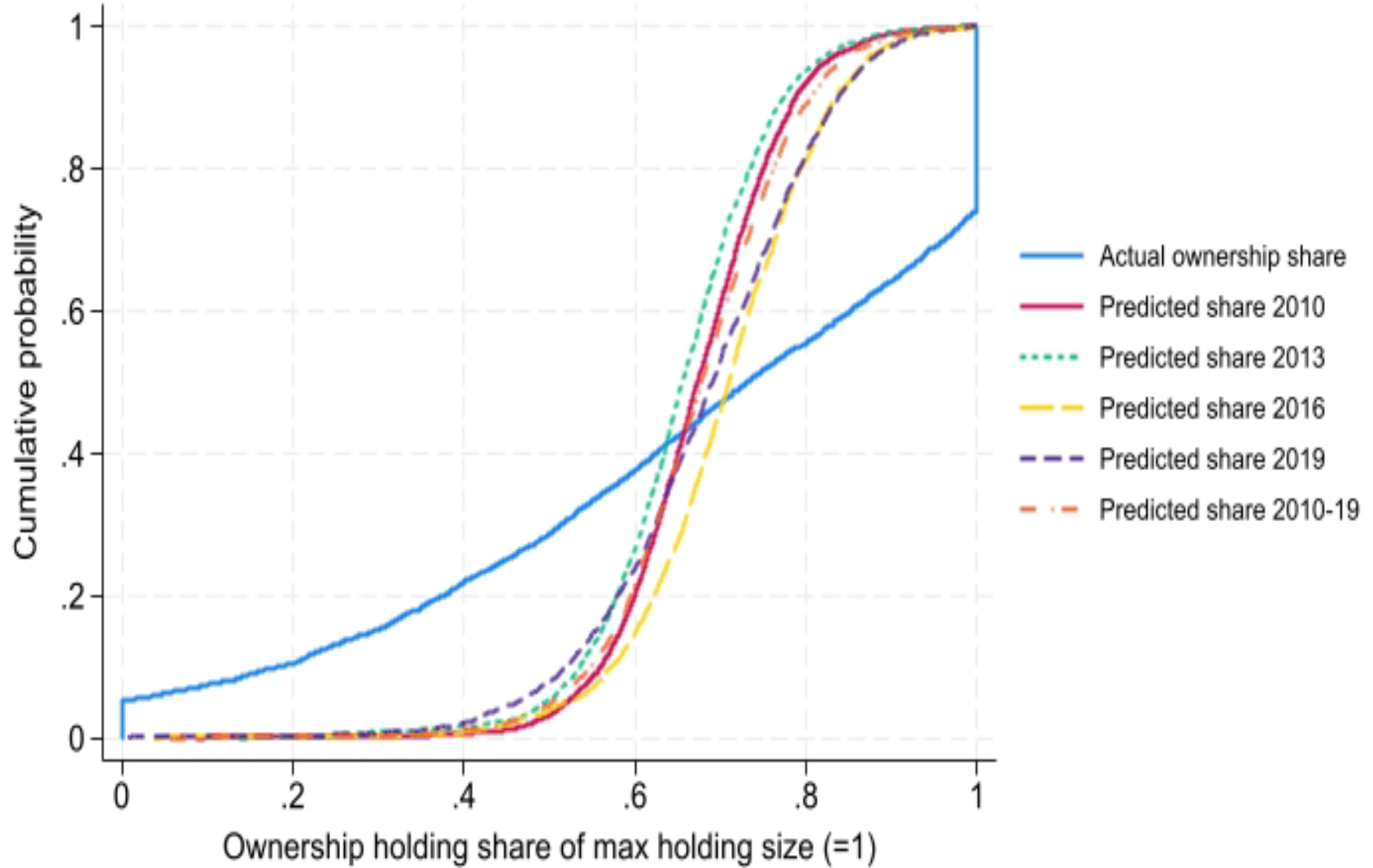
Malawi: Ownership holding share of max holding(=1) Panel Tobit models for survey rounds 2010, 2013, 2016, and 2019



Areas measured with GPS. Estimates with 95 and 99% CIs

Base: Male hh head, <31 years old, with rural residence, in Northern Region, in 2010

Malawi: Actual versus predicted ownership holding shares

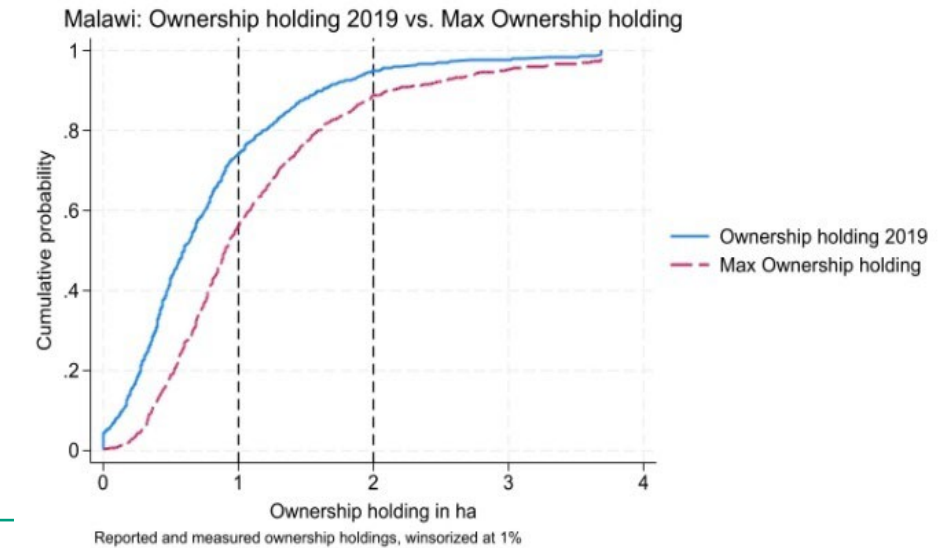
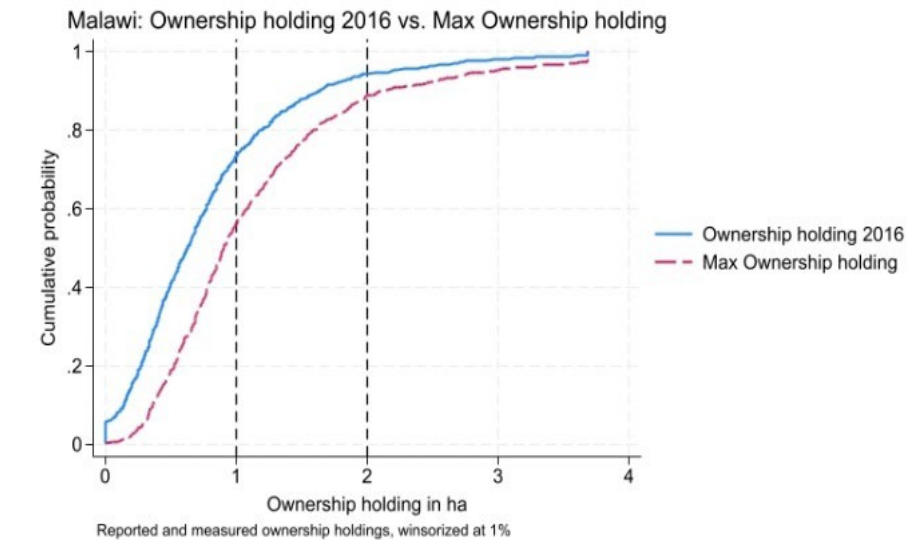
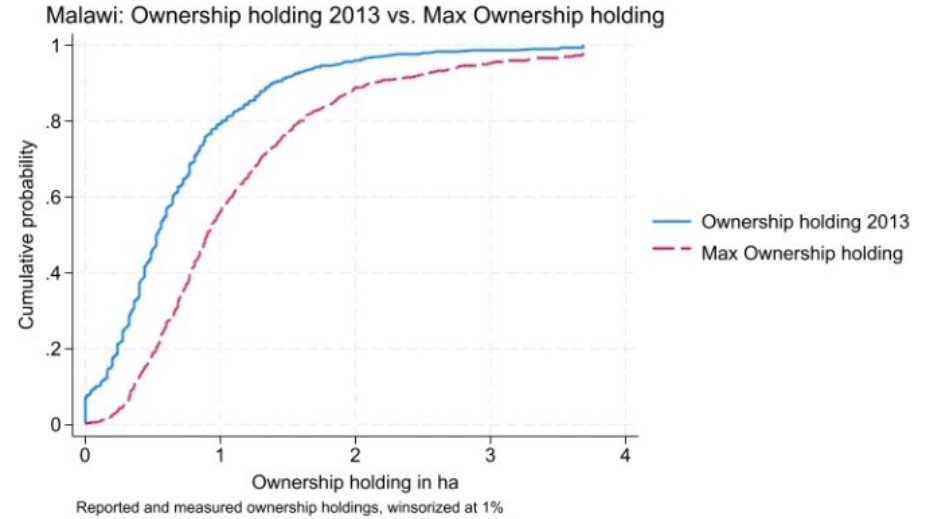
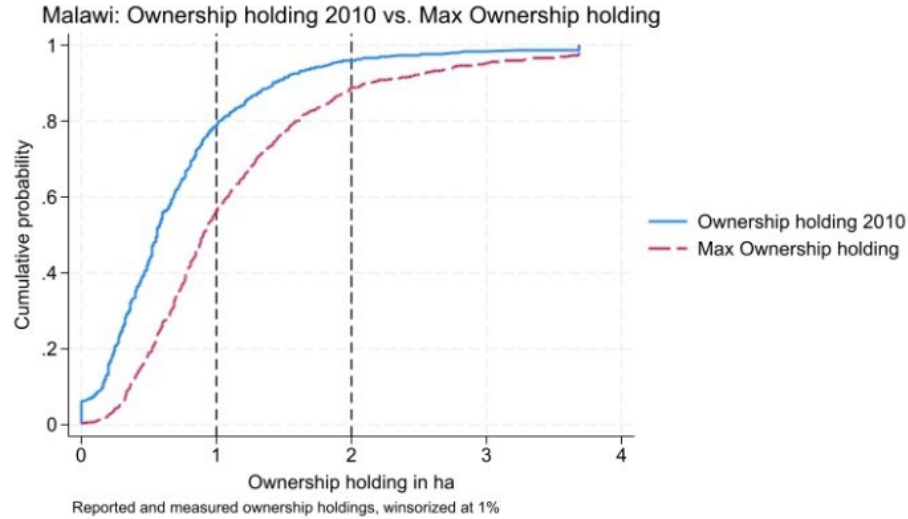
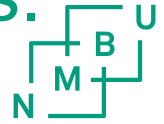


Cumulative distributions, data winsorized at 1% level

Estimated ownership holding sizes in ha

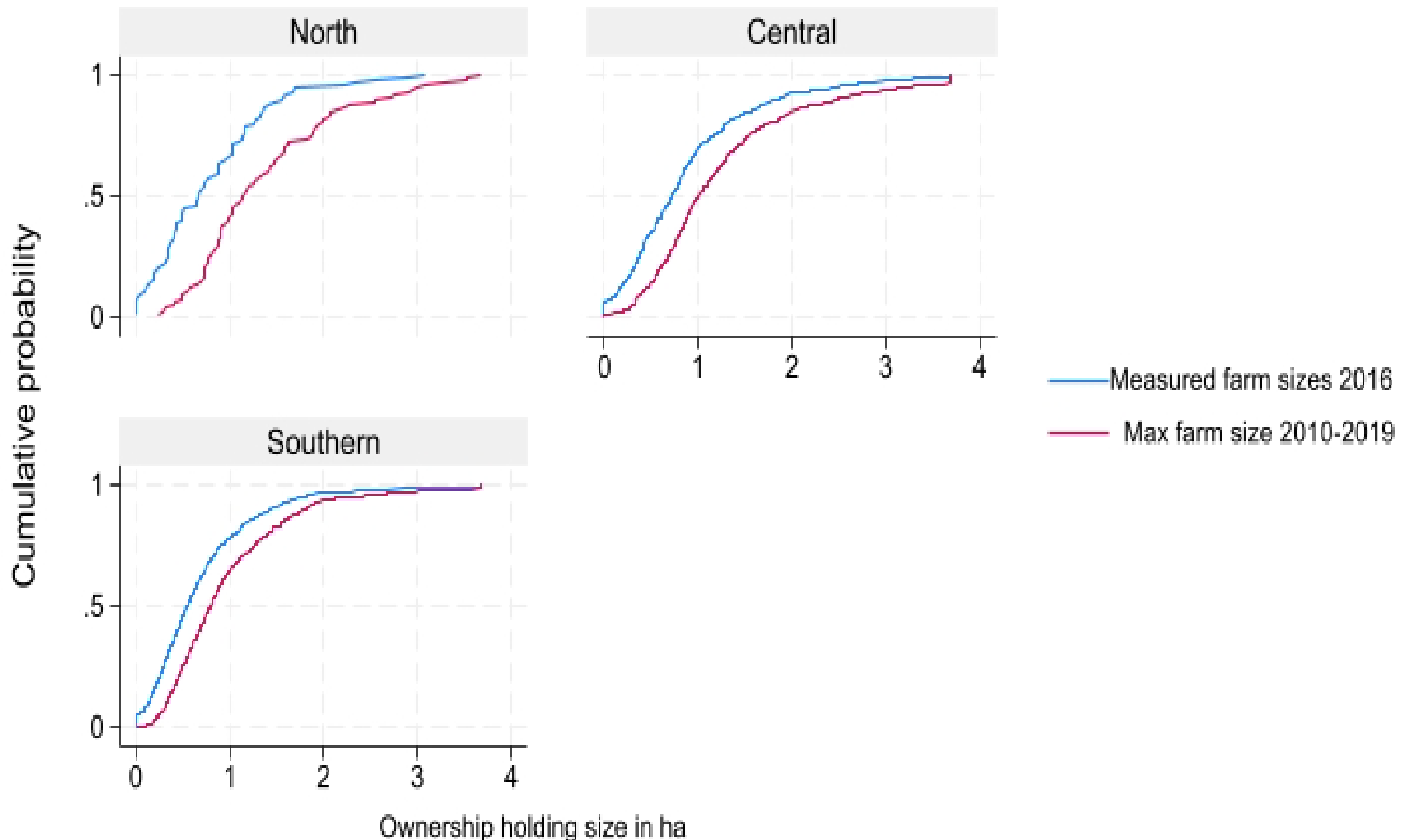
	Ownership holding, unwinsorized				Ownership holding, winsorized 1%				Max ownership holding, winsorized at		
	2010	2013	2016	2019	2010	2013	2016	2019	1%	2%	5%
Year	2010	2013	2016	2019	2010	2013	2016	2019	1%	2%	5%
Mean	0.738	0.697	0.803	0.788	0.713	0.690	0.793	0.781	1.133	1.102	1.025
Median	0.567	0.532	0.631	0.599	0.567	0.532	0.631	0.599	0.902	0.902	0.902
P25	0.308	0.291	0.336	0.332	0.308	0.291	0.336	0.332	0.587	0.587	0.587
P75	0.902	0.890	1.036	1.036	0.902	0.890	1.036	1.036	1.449	1.449	1.449
P90	1.425	1.392	1.639	1.619	1.425	1.392	1.639	1.619	2.125	2.125	1.947
St.dev.	0.837	0.652	0.741	0.718	0.626	0.609	0.688	0.680	0.790	0.702	0.543
St.err.	0.027	0.021	0.024	0.023	0.020	0.019	0.022	0.022	0.025	0.022	0.017
N	986	986	986	986	986	986	986	986	986	986	986
Gini	0.416	0.407	0.415	0.420	0.396	0.401	0.407	0.415	0.358	0.341	0.298

Ownership holding distributions by survey year vs. Maximum within-household ownership holdings



Malawi: Measured farm sizes in 2016 vs. Maximum farm sizes

Measured ownership holdings in 2016, winsorized at 1% versus household Max holding size 2010-2019



Graphs by Regions; 1-North, 2-Central, 3-Southern



Malawi: Gini-coefficients by region

Region	Ownership holdings Unwinsorized	Ownership holdings Winsorized at 1%	Max Ownership holding Winsorized at 1%	Max Ownership holding Winsorized at 2%	Max Ownership holding Winsorized at 5%
North	0.400	0.399	0.313	0.297	0.242
Central	0.409	0.395	0.351	0.330	0.278
Southern	0.413	0.406	0.358	0.345	0.314
Total	0.416	0.406	0.358	0.341	0.298

Discussion

- Further studies needed:
 - Utilize parcel GPS coordinates
 - Match with land registry data in Ethiopia
 - Investigate implications for input use and land productivity estimates based on LSMS data
 - Bias in household production based poverty indicators? (land wealth, total crop production)
 - Other countries?

Conclusions

- Parcel attrition is a severe problem in the LSMS data from both Ethiopia and Malawi
- It leads to substantial under-estimation of average and median farm sizes within survey rounds
- This bias cannot easily be corrected with econometric methods due to data limitations (omitted variables problem)
- The maximum within-household farm size in the balanced panel appears as the most reliable measure of farm size (least likely to suffer from parcel attrition)
- We use it to approximate the size of the biases in farm size estimates