University of Toronto
Department of Economics

Working Paper 718

Assessing misallocation in agriculture: plots versus farms

By Fernando Aragon, Diego Restuccia and Juan Pablo Rud

February 04, 2022

Assessing misallocation in agriculture: plots versus farms *

Fernando M. Aragón Diego Restuccia Juan Pablo Rud

February 2022

Abstract

We assess the extent and cost of misallocation in agriculture in less-developed countries comparing the analysis at the plot and farm levels. Using detailed data from Uganda, we show that the plot-level analysis leads to extremely large estimates of reallocation gains, even after adjusting for measurement error and unobserved heterogeneity. These results reflect two empirical limitations of the plot as unit of analysis: excess measurement error and near constant returns to scale production estimates. We find limited evidence of substantial measurement error at the farm level.

JEL classification: O4.

Keywords: Plot, farm, misallocation, measurement error, agriculture, distortions.

*All errors and interpretations are our own. Restuccia gratefully acknowledges the support from the Canada Research Chairs program and the Bank of Canada Fellowship. Aragon gratefully acknowledges financial support from the Social Sciences and Humanities Research Council of Canada, grant 435-2018-0227. Rud gratefully acknowledges the support from the Economic and Social Research Council (UK Research and Innovation), Project ES/S014438/1. The views expressed herein are not necessarily those of the Bank of Canada and are the authors alone. *Contact Information:* Aragón: Department of Economics, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada (e-mail: faragons@sfu.ca); Restuccia: Department of Economics, University of Toronto, 150 St. George Street, Toronto, ON M5S 3G7, Canada (e-mail: diego.restuccia@utoronto.ca); Rud: Department of Economics, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK (e-mail: juan.rud@rhul.ac.uk).

1 Introduction

There is a growing literature documenting substantial dispersion in measures of marginal products of inputs across production units, which has been interpreted as evidence of factor misallocation (Hsieh and Klenow, 2009; Restuccia and Rogerson, 2017). A relevant concern, however, is that the observed dispersion might reflect other factors such as overhead costs, unobserved heterogeneity, or measurement error (Bartelsman et al., 2013; Asker et al., 2014; Foster et al., 2016). There are several emerging approaches to deal with these issues. For instance, an ambitious line of research models specific sources of misallocation to identify their quantitative importance using microdata (Midrigan and Xu, 2014; David and Venkateswaran, 2019; Yang, 2021). Another method exploits the availability of panel data to purge measures of marginal products from time-invariant measurement error and overhead costs (Bils et al., 2017).

A complementary approach involves the use of more granular data. For example, Kehrig and Vincent (2020) study multi-plant firms in the United States and find larger levels of dispersion in marginal revenue products of capital within firms than across firms. The authors show that firms make constraint-efficient reallocation decisions that increase both output and dispersion. In the context of small-scale agriculture, a similar approach emphasizes plot-level variation to address measurement error (Gollin and Udry, 2021; Abay et al., 2019; Desiere and Jolliffe, 2018). The approach exploits a specific feature of agricultural economies whereby the predominant unit of production, the household farm, typically operates more than one plot of land.

The intuition for why the plot-level approach may be useful in this context is simple. If farmers can freely allocate inputs across plots within their farm operation, then the marginal productivity of inputs should be equalized across operated plots. In this case, observed within-farm dispersion in measures of marginal productivity can be attributed to sources other than misallocation, such as unobserved heterogeneity and measurement error. That is, within-household productivity variation across plots represents a measure of mismeasurement that can be subtracted from standard estimates of misallocation.

In this paper, we examine whether the assessment of misallocation in agriculture (and therefore the role of mismeasurement) is affected by the choice of micro data aggregation: plots or farms. Our main finding is that the level of data aggregation does matter, and can lead to quantitatively different conclusions.

Conceptually, the extent of misallocation is measured as the increase in aggregate output that could be obtained from reallocating resources across production units according to an efficient benchmark. The magnitude of these efficiency gains depends on the distribution of total factor productivity and on the parameters of the production function.

Using data from Uganda, we document that efficiency gains at the plot level are extremely high, even after adjusting for measurement error, as in Gollin and Udry (2021). For instance, plot-level estimates suggest efficiency gains of reallocation at the national level of more than 2,200%. Even after adjusting for measurement error efficiency gains are greater than 500%. These gains from factor reallocation are very large compared to similar estimates in agriculture and manufacturing around the developing world. In particular, estimates using the same data but taking the farm as the relevant unit of analysis suggest more modest reallocation gains of around 180% (Aragón et al., 2022).

These are striking differences in the assessment of misallocation between the plot and the farm, which also obfuscate the assessment of mismeasurement in micro data. Whereas Gollin and Udry (2021) attribute a large role of measurement error using plot-level data, we find instead a much more modest role for mismeasurement in the same data at the farm level.

Our analysis suggests that focusing on the farm as the relevant unit of analysis provides a more accurate assessment of misallocation in agriculture. This conclusion arises from two important limitations of plot-level analysis that we highlight. First, production function estimates at the plot level imply a much larger contribution of land, and greater returns to scale (close to unity) than estimates using the same or similar data at the farm level (Aragón et al., 2022; Manysheva, 2021). The differences in production function estimates are economically relevant. For instance, efficiency gains from a given plot-productivity dispersion drop from 23.9-fold with plot-level estimates, to 5.3-fold with farm-level estimates. Moreover, whereas the relationship between land input and productivity is negative in plot-level data, it turns positive when assessed at the farm level.

Second, the value of the plot-level approach rests crucially on the reliability of plot-level measurements compared to that at the farm level. In practice, however, measurement error is likely to be substantially worse at the plot level due to the inherent difficulty of attributing all inputs and outputs to a subset of the farm operation, for example as labor and tools are used interchangeably across plots. This issue is particularly relevant in the context of small farmers in developing countries due to the reliance on self-reported values. We provide evidence of excess measurement error in plot-level data by comparing self-reported area of parcels (a unit above the plot but below the farm) to their GPS measures. Not only do we find evidence of substantial and systematic measurement error in land at the parcel level, but also the magnitude of the measurement error at the median observation is more than 4-fold at the parcel than at the farm level.

In addition to the methodological implications of using plot or farm measures, there are important implications for understanding and addressing misallocation. From a policy standpoint, the focus on the farm is relevant given the fact that land institutions in many developing countries allocate land rights at the household level (Restuccia, 2020). Also, in the context of small-scale agriculture, the presence of fixed factors shared across plots makes the household farm the appropriate production unit (De Janvry et al., 1991).

2 Does the level of micro data aggregation matter?

We start by comparing estimates of productivity dispersion and reallocation gains across plots and across farms. In our analysis, a farm is the set of plots cultivated by the household. The data comprises three waves from the Uganda Panel Survey (2009-2014), a household survey collected as part of the World Bank's Integrated Surveys of Agriculture (LSMS-ISA). Gollin and Udry (2021) and Aragón et al. (2022) provide a detailed description of the data.

The concept of misallocation requires a benchmark for comparison. We consider the efficient allocation of resources as the benchmark, namely the allocation of resources that maximizes agricultural output subject to aggregate resources; and calculate the efficiency gain (i.e., the ratio of efficient to actual aggregate output) assuming input reallocation at different geographical levels. The efficiency gain is a standard measure of the extent of misallocation in the literature (Hsieh and Klenow, 2009). We consider an economy comprised of a given set of production units with the following Cobb-Douglas technology,

$$Y_i = s_i (L_i)^{\alpha_L} (X_i)^{\alpha_X}, \tag{1}$$

where L_i and X_i are the amounts of land and labor used in production unit i, and s_i denotes its productivity.

The efficient allocation equates marginal products of land and labor across production units. Denoting $z_i \equiv s_i^{1/(1-\alpha_L-\alpha_X)}$, we can characterize the efficient allocation as:

$$T_i^e = \frac{z_i}{\sum_i z_i} T, \qquad L_i^e = \frac{z_i}{\sum_i z_i} L, \tag{2}$$

where T and L are the aggregate amounts of land and labor. The efficiency gain from reallocation is the ratio $\sum_i Y_i^e / \sum_i Y_i$ where Y_i^e is production-unit output associated with the efficient-input allocation. Note that to calculate the efficiency gain, in addition to actual output, we need estimates of production-unit productivity s_i and production function parameters α_L and α_X.

To assess misallocation at the plot level, we rely on estimates from Gollin and Udry (2021), who use state-of-the-art methods to estimate plot-level productivity and adjust it for measurement error and unobserved heterogeneity. For our analysis we use the two-stage least squares (2SLS) estimates in Gollin and Udry (2021), however, the results are robust to using the alternative instrumental variables correlated random coefficients (IVCRC) estimates.

Table 1 presents efficiency gains and productivity dispersion using different productivity measures. Column 1 uses the baseline plot productivity before adjustment by measurement error and unobserved heterogeneity. This variable is called TFPA in Gollin and Udry (2021). Column 2 uses the measure of plot productivity adjusted for measurement error, called TFPB. Column 3 simply aggregates plot productivity (TFPA) to the farm level by calculating a weighted average. Denoting s_{ij} the unadjusted productivity of plot j in farm i, and given the Cobb-Douglas technology on land and labor, aggregated farm productivity is given by $\sum_j s_{ij}(\phi_{ij}^L)^{\alpha_L}(\phi_{ij}^X)^{\alpha_X}$, where ϕ_{ij}^L and ϕ_{ij}^X are the shares of farm i's total land and labor used in plot j.

We emphasize three relevant observations from Table 1. First, reallocation gains estimated using plot-level data are extremely large (column 1). The estimates imply that if the allocations of land and labor were to change to the efficient allocation at the national level, agricultural output would increase by a factor of 23.9-fold, or more than 2,200%. Reallocation gains remain large even when reallocation is limited to smaller geographical areas: 1,538% within regions and 305% within villages.

The implied reallocation gains are exceedingly larger than those documented in macro studies of the agricultural sector using farm-level data. For example, the estimated reallocation gains (at the national level) in China, Ethiopia and Malawi are 53%, 97% and 259%, respectively (Adamopoulos et al., 2021; Chen et al., 2021; Restuccia and Santaeulalia-Llopis,

Table 1: Efficiency gain and productivity dispersion in plot- and farm-level analysis

	Plot-level data			Farm-level data
	Plot productivity	Plot productivity (adjusted)	Plot productivity aggregated at farm level	Farm productivity
	(1)	(2)	(3)	(4)
A. Efficiency gains				
Nationwide	23.92	6.66	14.28	2.86
Region	16.38	5.36	8.35	2.48
Parish (Village)	4.05	2.47	2.11	1.57
B. Dispersion				
Variance of log	1.26	0.53	0.78	0.84

Notes: Efficiency gain is the ratio of aggregate output in the efficient allocation to actual output averaged over season-years. Columns 1-3 use 2SLS estimates of plot productivity from Gollin and Udry (2021). Column 1 uses the baseline (unadjusted) productivity measure (TFPA), whereas Column 2 uses the adjusted productivity measure (TFPB). Column 3 aggregates TFPA at the farm level by computing a weighed average. Column 4 uses direct estimates of farm productivity from Aragón et al. (2022).

2017). Reallocation gains are comparatively large even within narrower geographical areas. For instance, the within-village reallocation is 305% using plot-level data in Uganda whereas only 24% in China (Adamopoulos et al., 2021).

Second, adjusting for measurement error and unobserved heterogeneity greatly reduces productivity dispersion and efficiency gains. However, the estimated magnitude of misallocation remains very high (column 2). Note that Gollin and Udry (2021)'s adjustment for measurement error reduces log-variance of plot-productivity by almost two thirds from 1.26 to 0.53. This reduction in dispersion is associated with a proportional drop in the implied efficiency gains. Nevertheless, the estimated efficiency gains with adjusted productivity, ranging from 147% to 566%, are sizable and well-above the estimates from other studies emphasizing productivity and reallocation at the farm level.

Third, simply aggregating estimates of plot-productivity to the farm level accomplishes a sizeable reduction in productivity dispersion and efficiency gains (column 3). For instance, the efficiency gains for village-level reallocation are almost half the estimated gains using plot-level productivity and even smaller than when using the plot-productivity adjusted measure in column 2. However, in all cases, the estimated misallocation remains quite large.

The high levels of misallocation across plots in Ugandan agriculture may arise precisely from the analysis at the plot level. To illustrate why this may be the case, we calculate efficiency gains using estimates at the farm level from Aragón et al. (2022) and report them in column 4. We emphasize that these estimates are obtained from the same dataset but aggregated to the farm level.

The estimated gains from reallocation fall substantially when using farm-level data (column 4). The efficiency gains are 186% at the national level and 57% at the village level. These gains are almost half the lowest estimates obtained at the plot level, even after adjusting for measurement error, and closer in magnitude to estimates from macroeconomic studies in other contexts.

The substantial reduction in assessed misallocation at the farm level occurs despite the fact that dispersion of farm-productivity is higher than dispersion of plot-productivity (after adjusting for measurement error). This finding also illustrates the limitation of using productivity dispersion to assess misallocation. In general, the magnitude of efficiency gains is not only a function of productivity dispersion, but also of economies of scale and of the relationship between input allocation and productivity across production units.

The main takeaway is that focusing on the plot as unit of analysis is not a useful starting point for the study of misallocation in agriculture. The large discrepancy in assessed efficiency gains casts doubts on the validity of extrapolating insights obtained using plot-level analysis to results obtained at the farm level.

Importantly, plot-level analysis can lead to the wrong assessments of the extent of misallocation and its possible explanations. For example, a researcher comparing observations of plot-level productivity when adjusting for within-farm variation across plots could wrongly conclude that misallocation in agriculture is unimportant since measurement error seemingly accounts for the bulk of the apparent misallocation. This conclusion, however, hinges on the extremely large initial estimates of efficiency gains and productivity dispersion.

3 What explains plot versus farm-level results?

We highlight two important limitations of plot-level analysis to assess misallocation in agriculture.[1] First, plot-level analysis may provide inadequate production function estimates that obviate the fact that household farms have some fixed factors that are used across plots within the farm, including tools or managerial skills. This leads to drastically different estimates of returns to scale and productivity, key inputs in the assessment of reallocation gains. Second, plot-level analysis inherits excessive measurement error at the plot, inducing excessive productivity dispersion. Together, these issues exacerbate the pattern (and magnitude) of misallocation and overstate the significance of measurement error in agricultural economies.

3.1 Different production function estimates

Production function estimates are key inputs to assess factor misallocation. The main econometric challenge in estimating the production function parameters is the presence of determinants of production, such as productivity shocks, that are unobserved to the econometrician but observed by the farmer. If inputs are chosen by farmers based on these productivity shocks, then there is an endogeneity problem and OLS estimates would be inconsistent.

Panel data offer a way to address this endogeneity problem. For instance, if the unobserved productivity shocks are time-invariant (such as location, soil quality, or farming ability), the production function can be consistently estimated using a panel data model

[1] Conceptually, it is not clear what it means to reallocate resources, including land, across plots since the plot is not an administrative unit of production.

with fixed effects (Ackerberg et al., 2015). In contexts with imperfect input markets, a simplified dynamic panel model can also address auto-regressive, time-varying, productivity shocks (Shenoy, 2017, 2020).[2]

There are several agricultural surveys, such as the World Banks' Integrated Surveys of Agriculture, that already include a panel of households. However, the available plot-level data is mostly cross-sectional.[3] This issue limits the type of methods that can be used to estimate the production function. For instance, Gollin and Udry (2021) are unable to use panel data methods, but rely instead on a clever instrumental variable approach using self-reported productivity shocks on nearby plots. In contrast, Aragón et al. (2022) use the same dataset aggregated at the household level and estimate a panel data model with fixed effects.

Table 2 displays the production function estimates from the cross-sectional plot-level approach and the panel data farm-level specification. The main observation is that there are sizable differences in the contributions of land and labor, and in the implied returns to scale. In particular, the IV estimates using plot-level data suggest a larger contribution of land, and returns to scale closer to unity. This is consistent with the farm being an aggregation of constant-return-to-scales' plots with some fixed factors. Given that productivity is obtained as a residual, these different estimates also generate differences in measures of total factor productivity.

These differences in production function estimates matter for the assessment of misallocation. To illustrate this, we consider the estimates of unadjusted plot-productivity (TFPA) but instead assess efficiency gains using the production function estimates obtained in the household-panel specification in column 2, Table 2. The drop in efficiency gains

[2]Proxy variables methods such as Olley and Pakes (1996), Levinsohn and Petrin (2003) and Ackerberg et al. (2015) also exploit panel data. However, they are of limited use in contexts with suspect factor misallocation because, as shown by Shenoy (2020), their identification assumptions fail when market frictions distort input choices.

[3]Collecting a panel dataset at the plot level does not necessarily address this concern, given the endogeneity of plot formation, which would likely introduce additional biases.

Table 2: Production function estimates at the plot and farm levels

	IV (2SLS)	Panel data with fixed effects
	(1)	(2)
Land contribution (α_L)	0.69	0.37
Labor contribution (α_X)	0.22	0.34
Returns to scale ($\alpha_L + \alpha_X$)	0.91	0.71
Aggregation level	Plot	Household

Notes: Column 1 displays 2SLS estimates reported in Table 9 (column 3) in Gollin and Udry (2021). Column 2 display estimates reported in Table A.1 (column 1) in Aragón et al. (2022).

associated with the change in production parameters is substantial. For instance, simply changing the production function parameters with almost constant returns to scale at the plot level ($\alpha_L + \alpha_X = 0.91$) to the ones with decreasing returns to scale at the farm level ($\alpha_L + \alpha_X = 0.71$) reduces efficiency gains at the national level from 23.9-fold to 5.3-fold.[4] A similar pattern is found for gains at the regional level (from 16.4 to 4.7-fold) and the parish (village) level (from 4.1 to 2.3-fold). The stark difference in measures of misallocation highlights the importance of production function estimates in assessing misallocation.

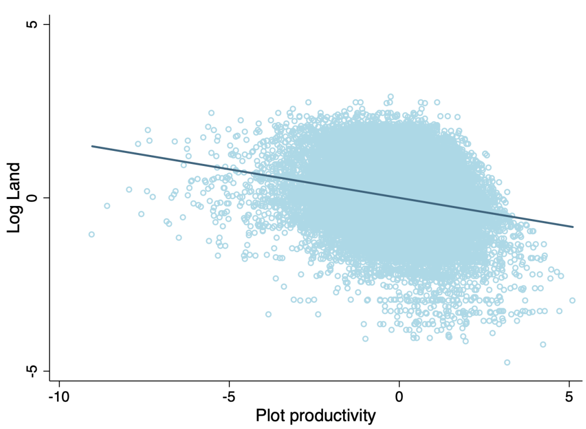
We can also gauge the difference between plot- and farm-level analysis on the extent of misallocation by plotting the relationship between land use and productivity across production units. As shown in equation (2), the efficient allocation of inputs across production units requires a strong positive relationship between productivity and input use. Deviations from this benchmark would be indicative of the extent of misallocation and has been the focus of an expanding literature in development and agricultural economics (Adamopoulos and Restuccia, 2014; Restuccia, 2020).

Figure 1 displays the relationship between land input and productivity from the micro data in Uganda using alternative production units: the plot (Gollin and Udry, 2021) and the

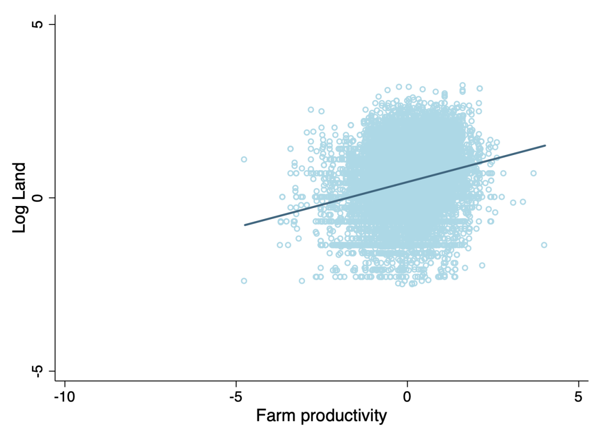
[4]Similar estimates of returns to scale are found in other studies at the farm level such as Shenoy (2017) and Manysheva (2021).

household-farm (Aragón et al., 2022). Based on the estimated returns to scale in Table 2, the slope of the (log) land-productivity relationship in the efficient allocation is 11.1 across plots and 3.4 across farms. In both cases, the main observation from Figure 1 is that the slope of the relationship between productivity and land input is much smaller than that required in the efficient allocation. While the farm estimates (Panel b) show a weak but positive relationship , i.e. a slope of 0.26, the plot-level estimates (Panel a) show a negative relationship, i.e. an elasticity of -0.16. This result implies an even larger deviation from the efficient benchmark in the plot-level analysis, and thus a greater implied factor misallocation. Note that the same pattern between land input and productivity arises when adjusting plot-productivity by the within-farm dispersion across plots as in Gollin and Udry (2021) since the adjustment in this approach amounts to a scalar reduction in variance which is proportional to the productivity variance across plots within farm households.

Figure 1: Land size and productivity across production units



(a) IV estimates (plots)



(b) Panel estimates (farms)

Notes: Both panels display the scatter-plot of size of production unit (measured by area planted) and productivity, and a fitted linear regression. Panel (a) uses plot-level measures of size and productivity from 2SLS estimates in Gollin and Udry (2021). Panel (b) uses farm-level measures of size and productivity from a panel data model in Aragón et al. (2022).

3.2 Excess measurement error in plot-level data

An appealing property of the approach that uses plot-level data is that if the farm household does not face frictions in adjusting inputs across plots, then the within-farm plot-productivity dispersion represents a measure of mismeasurement. However, there are at least two key empirical threats to this approach.

First, there is ample evidence from an established literature emphasizing a variety of frictions within the household, driven by intra-household allocations or variation in plot-level property rights (Shaban, 1987; Goldstein and Udry, 2008; Udry, 1996).[5] These frictions diminish the value of within-farm plot-productivity dispersion as a measure of mismeasurement.

Second, even if the farmer can allocate inputs across plots without frictions, to what extent is plot-level measurement comparable to the farm level? The main concern is that measurement error is likely much more pronounced at the plot level than at the farm-household level. For example, a farmer may be able to more accurately report the labor input used at the farm but make errors in attributing it to plots. If this is the case, then productivity dispersion picks up additional measurement error in plot-level data and within-farm productivity dispersion would not be a useful metric in assessing misallocation and mismeasurement. In addition, there are inputs for which measurement in survey data is provided at the household level but not the plot, such as capital or intermediate inputs used in agricultural production, potentially aggravating productivity mismeasurement at the plot relative to the farm. We focus on providing evidence of the extent of measurement error within and across farms.

Most survey data on smallholder agriculture are based on farmers' self-reporting. This feature creates the possibility of misreporting and measurement error: farmers may round-up

[5] Similarly, Kehrig and Vincent (2020) find that productivity dispersion is larger across plants within firms than across firms in the U.S. manufacturing sector.

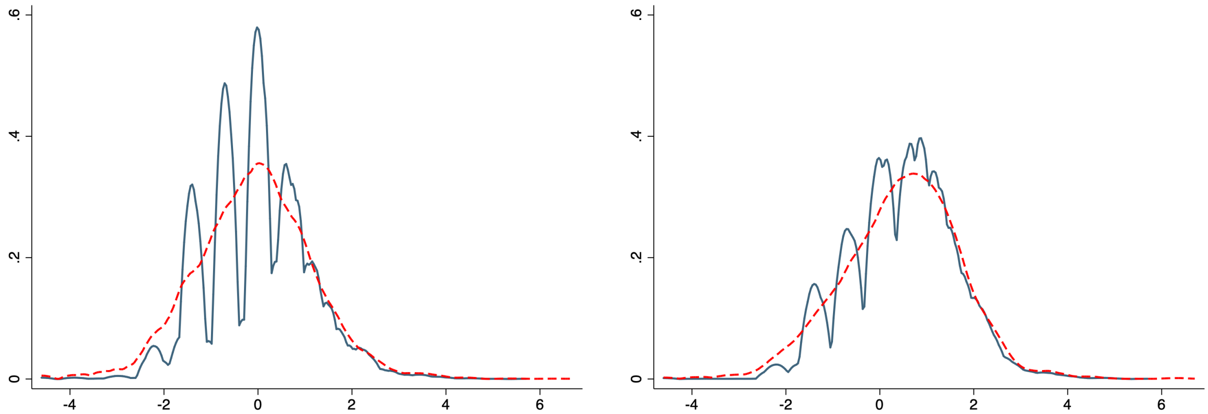
quantities, or simply provide guesstimates instead of actual values. In same cases, measurement error can be introduced when allocating indirect costs (such as capital expenses or management labor) among particular activities or crops. The pervasiveness of measurement error, especially of self-reported land areas, has been documented in several contexts, including the Ugandan case (see, for instance, Judge and Schechter, 2009; Carletto et al., 2015; Gourlay et al., 2019; Abay et al., 2021). This issue is a concern because measurement error can bias estimates of the production function, or be included as part of the residual often attributed to productivity.

We provide direct evidence of measurement error by comparing data at two levels of aggregation within the farm household. We compare the area of land holdings using two measures available in the Ugandan Panel Survey: self-reported by the farmer and GPS measures, which are deemed to be more precise and less subject to farmer's misreporting (Carletto et al., 2017).

Figure 2 displays the distribution of land holdings using both measures. Panel (a) shows the distribution of the original data at the parcel level, while panel (b) shows the distribution of the data aggregated to the farm level. Note that GPS data is not available at the plot level, instead only at the parcel level. However, as the parcel is a level of aggregation above the plot and indeed much closer to the farm size, this represents a conservative assessment of measurement error.

Consistent with measurement error, there are evident discrepancies between both measures. The GPS measure follows a smooth bell-shaped distribution, while the self-reported measures are heaped around certain values. This 'heaping' has been reported in several contexts and is indicative of respondents (or enumerators) rounding the reported size (Abay et al., 2019, 2021; Carletto et al., 2013). The discrepancies between both measures is more pronounced among smaller units, on the left side of the distribution. This evidence suggests

Figure 2: Distribution of landholding size, self-reported and GPS measured



(a) Parcel level (b) Farm level

Notes: Distribution of the log area of landholdings at the parcel level (panel a) and aggregated to the farm level (panel b). Solid lines represent self-reported values, while red-dashed lines represent GPS measures.

that the measurement error is not classical, but correlated to unit size.[6]

The discrepancies in land input, however, are attenuated when using data aggregated at the farm level. For the median observation, the log difference between self-reported and GPS measure (a proxy for measurement error) is 1.9% at the parcel level and drops to 0.45% at the farm level, a 4-fold difference in measurement error at the parcel level on land input alone. Aggregating the data to the farm level also reduces the dispersion of measurement error: the variance of the log of self-reported to GPS land ratio is 0.54 at the parcel level and 0.45 at the farm level.

We conclude from the evidence that not only measurement error in self-reported land is attenuated at the farm level compared to the parcel, but also focusing on the plot restricts the analysis to self-reported values as opposed to more accurate GPS measures often unavailable at the plot level.

[6]This pattern has been documented in other studies. For instance, Abay et al. (2021) reports a negative correlation between plot size and measurement error in land in four Sub-Saharan African countries.

4 Measurement error and misallocation

If plot-level data is problematic, how can we assess the extent of measurement error in measures of misallocation? A growing literature have instead turned to approaches exploiting panel data. We follow Bils et al. (2017) in exploiting time-variation in the data to address measurement error. The extent to which variation over time in inputs is not reflected in variation in output, which varies across levels of measured distortions, provides a metric of the extent of measurement error.

For comparison, Gollin and Udry (2021) argue that measurement error and other sources of unobserved heterogeneity play a substantial role in accounting for the apparent misallocation in agriculture. Using plot-level data, they show that adjusting for measurement error by the within-farm dispersion in productivity, reduces estimates of misallocation by almost two thirds. As discussed earlier, the implied large reduction in misallocation is due in great part to the exceptionally large estimates of misallocation at the plot level.

This method to identify mismeasurement, however, is not applicable to the farm level since it relies on the assumption of efficient within-farm allocation of resources. An alternative approach, proposed by Bils et al. (2017), exploits panel data to quantify the extent to which misallocation reflects additive measurement error. The starting point is the observation that the ratio of first differences (i.e., the change in revenue divided by the change in inputs) is a measure of marginal product purged from constant measurement error. Based on this insight, they develop a metric, λ, that captures the fraction of the dispersion in revenue productivity (TFPR) that is due to true variation in distortions.

Using the panel household-farm data for Uganda from Aragón et al. (2022), we find that the estimate of λ is fairly high (0.925), implying that only about 7.5% of the variation in misallocation can be ascribed to measurement error. This result is consistent with the findings in Adamopoulos et al. (2021) using Chinese panel data, where this method detects only 4% measurement error in farm-level measures and 10% in cross sectional farm-level data.

The extent of farm-level measurement error is substantially smaller than that implied by an analysis at the plot level. It is also much smaller than that in the manufacturing sector analyzed in Bils et al. (2017) for India and the United States, which suggests caution is needed when making comparisons of measurement error across sectors, countries, and applications.

5 Conclusion

Does exploiting granularity in micro data provide a better assessment of misallocation in developing economies? We address this question using a common dataset for Uganda analyzed at two levels of aggregation: plots versus farms. We show that the plot-level analysis exacerbates the extent of misallocation in agricultural economies and as a consequence overstates the role of mismeasurement on agricultural productivity. In particular, a farm-level analysis provides reasonable measures of the gains from reallocation and a much more limited role for measurement error.

We trace the differential results between the plot and the farm level analyses to differences in the estimates of production function parameters and excess measurement error in plot-level data. While the use of plot-level data may provide some advantages in identifying shocks and certain measurement issues pertaining to productivity in agriculture (and quantify them), this approach is not appropriate when assessing the potential gains from factor reallocation in agriculture, particularly in the context of small-scale farming in low-income countries.

References

Abay, Kibrom A., Gashaw T. Abate, Christopher B. Barrett, and Tanguy Bernard, "Correlated non-classical measurement errors, 'Second best' policy inference, and the inverse size-productivity relationship in agriculture," *Journal of Development Economics*, 2019, *139*, 171 – 184.

Abay, Kibrom A, Leah EM Bevis, and Christopher B Barrett, "Measurement Error Mechanisms Matter: Agricultural intensification with farmer misperceptions and misreporting," *American Journal of Agricultural Economics*, 2021, *103* (2), 498–522.

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer, "Identification properties of recent production function estimators," *Econometrica*, 2015, *83* (6), 2411–2451.

Adamopoulos, Tasso and Diego Restuccia, "The Size Distribution of Farms and International Productivity Differences," *American Economic Review*, June 2014, *104* (6), 1667–97.

_, Loren Brandt, Jessica Leight, and Diego Restuccia, "Misallocation, selection and productivity: A quantitative analysis with panel data from China," *Econometrica*, 2021.

Aragón, Fernando M., Diego Restuccia, and Juan Pablo Rud, "Are small farms really more productive than large farms?," *Food Policy*, 2022, *106*, 102168.

Asker, John, Allan Collard-Wexler, and Jan De Loecker, "Dynamic inputs and resource (mis) allocation," *Journal of Political Economy*, 2014, *122* (5), 1013–1063.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta, "Cross-country differences in productivity: The role of allocation and selection," *American Economic Review*, 2013, *103* (1), 305–34.

Bils, Mark, Peter J Klenow, and Cian Ruane, "Misallocation or mismeasurement?," *Working paper No. 599, Center for International Development, Stanford University*, March 2017.

Carletto, Calogero, Sara Savastano, and Alberto Zezza, "Fact or artifact: The impact of measurement errors on the farm size–productivity relationship," *Journal of Development Economics*, 2013, *103*, 254–261.

_, Sydney Gourlay, and Paul Winters, "From guesstimates to GPStimates: Land area measurement and implications for agricultural analysis," *Journal of African Economies*, 2015, *24* (5), 593–628.

_, _, Siobhan Murray, and Alberto Zezza, "Cheaper, faster, and more than good enough: is GPS the new gold standard in land area measurement?," in "Survey Research Methods," Vol. 11 2017, pp. 235–265.

Chen, Chaoran, Diego Restuccia, and Raül Santaeulàlia-Llopis, "The Effects of Land Markets on Resource Allocation and Agricultural Productivity," *Review of Economic Dynamics*, 2021.

David, Joel M and Venky Venkateswaran, "The sources of capital misallocation," *American Economic Review*, 2019, *109* (7), 2531–67.

Desiere, Sam and Dean Jolliffe, "Land productivity and plot size: Is measurement error driving the inverse relationship?," *Journal of Development Economics*, 2018, *130*, 84–98.

Foster, Lucia, Cheryl Grim, John Haltiwanger, and Zoltan Wolf, "Firm-level dispersion in productivity: is the devil in the details?," *American Economic Review*, 2016, *106* (5), 95–98.

Goldstein, Markus and Christopher Udry, "The Profits of Power: Land Rights and Agricultural Investment in Ghana," *Journal of Political Economy*, 2008, *116* (6), 981–1022.

Gollin, Douglas and Christopher Udry, "Heterogeneity, measurement error, and misallocation: Evidence from african agriculture," *Journal of Political Economy*, 2021, *129* (1), 000–000.

Gourlay, Sydney, Talip Kilic, and David B Lobell, "A new spin on an old debate: Errors in farmer-reported production and their implications for inverse scale-Productivity relationship in Uganda," *Journal of Development Economics*, 2019, *141*, 102376.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, November 2009, *124* (4), 1403–1448.

Janvry, Alain De, Marcel Fafchamps, and Elisabeth Sadoulet, "Peasant household behaviour with missing markets: some paradoxes explained," *The Economic Journal*, 1991, *101* (409), 1400–1417.

Judge, George and Laura Schechter, "Detecting problems in survey data using Benfordâs Law," *Journal of Human Resources*, 2009, *44* (1), 1–24.

Kehrig, Matthias and Nicolas Vincent, "Good Dispersion, Bad Dispersion," Technical Report 2020.

Levinsohn, James and Amil Petrin, "Estimating production functions using inputs to control for unobservables," *The Review of Economic Studies*, 2003, *70* (2), 317–341.

Manysheva, Kristina, "Land Property Rights, Financial Frictions, and Resource Allocation in Developing Countries," *mimeo*, 2021.

Midrigan, Virgiliu and Daniel Yi Xu, "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, February 2014, *104* (2), 422–58.

Olley, G Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–1297.

Restuccia, Diego, "The Impact of Land Institutions and Misallocation on Agricultural Productivity," Technical Report, National Bureau of Economic Research March 2020.

_ **and Raul Santaeulalia-Llopis**, "Land misallocation and productivity," Technical Report, National Bureau of Economic Research 2017.

_ **and Richard Rogerson**, "The causes and costs of misallocation," *Journal of Economic Perspectives*, 2017, *31* (3), 151–74.

Shaban, Radwan Ali, "Testing between Competing Models of Sharecropping," *Journal of Political Economy*, 1987, *95* (5), 893–920.

Shenoy, Ajay, "Market failures and misallocation," *Journal of Development Economics*, 2017, *128*, 65–80.

_ , "Estimating the Production Function Under Input Market Frictions," *Review of Economics and Statistics*, 2020, pp. 1–45.

Udry, Christopher, "Gender, Agricultural Production, and the Theory of the Household," *Journal of Political Economy*, October 1996, *104* (5), 1010–46.

Yang, Mu-Jeung, "Micro-level Misallocation and Selection," *American Economic Journal: Macroeconomics*, October 2021, *13* (4), 341–68.