University of Toronto
Department of Economics

Working Paper 697

Heterogeneity, Measurement Error, and Misallocation in African Agriculture: A Comment

By Fernando Aragon, Diego Restuccia and Juan Pablo Rud

May 27, 2021

Heterogeneity, Measurement Error, and Misallocation in African Agriculture: A Comment*

Fernando M. Aragón Diego Restuccia Juan Pablo Rud

May 2021

Abstract

Gollin and Udry (2021) estimate the contribution of mismeasurement to productivity dispersion among production units and conclude that previous studies have overestimated the potential efficiency gains from factor reallocation. We show that this conclusion is incorrect based on their own empirical evidence, which instead corroborates the importance of misallocation emphasized in the macro-development literature. We also point out important limitations in the implementation of the plot-level analysis that overstates the importance of mismeasurement in understanding productivity differences.

JEL classification: O4.
Keywords: Plot, farm, misallocation, measurement error, agriculture, land institutions, 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 program. Aragon gratefully acknowledges financial support from the Social Sciences and Humanities Research Council of Canada, grant 435-2018-0227. 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

In a recent article, Gollin and Udry (2021) (henceforth, GU) propose an intuitive method to assess the importance of measurement error in estimates of productivity dispersion across production units using micro data. In particular, the approach exploits a specific feature of agricultural economies whereby the predominant unit of production, the household farm, may operate more than one plot of land. Based on panel data from two African countries, they find that "[l]ate-season production shocks, measurement error, and heterogeneity in inputs together account for as much as 70% of the variance in measured productivity" (p.5). From this evidence they conclude that previous studies of misallocation have overestimated the potential productivity gains from factor reallocation and that the "efficient reallocation of land and other agricultural inputs would not dramatically close the income gaps between African countries and the world's rich economies" (p.5). In this article, we show that i) these conclusions are incorrect based on their own evidence and ii) GU's plot-level analysis overstates the extent of mismeasurement.

The article's empirical evidence in fact corroborates the importance of misallocation in agriculture in the developing world. The efficiency gain associated with GU's preferred measure of productivity, after correcting for mismeasurement, is substantial: a 566% increase from a nationwide reallocation. This magnitude is larger than efficiency gains documented in the related literature and, hence, does not support GU's conclusion that "the gains from a hypothetical reallocation are thus correspondingly overstated by a factor of two or three" (p.48).

What explains seemingly opposing conclusions from the same evidence? GU's approach focuses on the variance of plot-level productivity as a key metric to gauge the extent of mismeasurement. The narrative is that a large variance in plot-level productivity overstates the extent of misallocation as they find a substantial reduction in the dispersion of productivity when adjusting for mismeasurement. For the authors, this suggests that there must be a relatively minor scope for reallocation gains.

We show, however, that the use of plot-level data, as opposed to farm-level data used in the literature, substantially *magnifies* the measure of productivity dispersion. This choice of plot-level analysis means that an assessment of the extent of misallocation based on changes in productivity dispersion is misleading since the starting point is a much higher level of dispersion. Furthermore, we also show that the plot-level specification affects production function estimates in a way that actually *increases* the estimated gains from reallocation.

We also provide evidence that the use of plot-level data exacerbates the problem of measurement error. This can occur, for instance, if farmers round-up plot size or are unable to recall the amount of farm-level inputs allocated to each plot. This measurement error is attenuated when aggregating data at the farm level. We document that measurement error in land input (measured as the difference between self-reported and GPS data) is substantially larger at the parcel level, a unit above the plot, than at the farm level.

As an alternative, we evaluate the extent of measurement error in farm-level data using a similar method that exploits the panel dimension of the data (Bils et al., 2017) instead of suspect plot-level variation. Contrary to GU, we find modest amounts of measurement error using farm-level measures of TFP from GU's data and methods. These results are consistent with findings for farms in Chinese agriculture in Adamopoulos et al. (2017) but different from findings in Bils et al. (2017) for manufacturing plants in India and the United States, reinforcing our emphasis on the importance of farm-level analysis in this context and suggesting caution in making generalizations about the role of measurement error in micro data in different sectors, countries, and applications.

GU address an important issue, as estimating the quantitative contribution of measurement error to productivity dispersion is policy-relevant, particularly in the context of small-scale farming in low-income countries. While there may be advantages in using plot-level data to identify shocks and 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. Contrary to the article's narrative and conclusions, the empirical evidence in GU corroborates the importance of misallocation in agriculture in developing countries.

2 The evidence

Gollin and Udry (2021) find a substantial reduction in plot-level productivity dispersion after accounting for mismeasurement and, as a result, conclude that the potential gains from reallocation are substantially smaller than those found in the literature. We show, using GU's own estimates, that this conclusion is incorrect, as the change in productivity dispersion does not accurately characterize the extent of remaining misallocation.

Productivity dispersion and misallocation Productivity dispersion across plots in Uganda is extremely large.[1] The top panel in Table 1, columns (1) and (2), reports statistics on the dispersion of unadjusted productivity across plots (TFPA), which correspond to the reported variances in GU (Table A3). We use GU's data and production function estimates to compute the efficient allocations that maximize aggregate output given resources, and the implied efficiency gains at three levels of geographical disaggregation. The efficiency gain, a standard measure of the extent of misallocation in the literature, is simply the average of the ratio of efficient to actual aggregate output across season-years (Hsieh and Klenow, 2009).[2] The results are presented in the bottom panel of Table 1.

[1]For this comment we focus on Uganda because this is the country that features the largest reduction in productivity dispersion due to mismeasurement. Moreover, we are more familiar with the data and institutional context in Uganda (Aragon et al., 2019). Nevertheless, our comments and assessment also apply to Tanzania.

[2]Conceptually, it is not clear what it means to reallocate factors including land across plots as units of production, but we nevertheless for illustration and comparison compute efficiency gains as in GU (see Appendix A6, page 63).

Table 1: Productivity dispersion and efficiency gains in Gollin and Udry's data

	Plot-level TFPA		Plot-level TFPB		Farm-level TFPA	
	2SLS	IVCRC	2SLS	IVCRC	2SLS	IVCRC
	(1)	(2)	(3)	(4)	(5)	(6)
A. Dispersion						
90-10 log difference	2.69	2.67	1.74	1.29	2.07	2.04
90-10 ratio	14.76	14.43	5.67	3.64	7.90	7.68
Variance of log	1.26	1.25	0.53	0.29	0.78	0.77
B. Efficiency gains						
Nationwide	23.92	31.40	6.66	6.66	14.28	17.46
Region	16.38	20.57	5.36	4.69	8.35	10.94
Parish (Village)	4.05	4.46	2.47	2.17	2.11	2.30

Notes: Mean of efficiency gains across season-years. TFPA is unadjusted productivity at the plot level and aggregated to the household level. TFPB is adjusted productivity at the plot level.

Reallocation gains are extremely large. For instance, if the allocations of land and labor were to change to the efficient allocations, agricultural output and the associated aggregate productivity would increase between 23.9 and 31.4-fold, that is an increase of more than 2200%.

The gain from reallocation across plots implied by this data is not a good starting point to challenge the findings of the misallocation literature. These gains are not just exceedingly larger than macro studies of the agricultural sector, but also larger than studies of plant-level reallocation in manufacturing. The closest to plot-level analysis in the literature would be the reallocation across plants in the manufacturing sector of China and India in Hsieh and Klenow (2009). But whereas the variance of log plant-level productivity is similar than those in GU for plots, 1.12 in 1998 China and 1.51 in 2005 India, the reallocation gains are only 115% in China and 128% in India. Clearly, the extent and costs of misallocation are not well characterized by the variance of productivity alone. The extent of misallocation is very different across plots than across plants, but also are the likely measurement issues between the agricultural and manufacturing sectors.

GU's main message is that an adjusted measure of productivity (TFPB) that accounts for late season output shocks, unobserved quality inputs, and measurement error, reduces considerably any apparent misallocation since the variance of adjusted productivity (log TFPB) is 70% lower than the variance of unadjusted productivity (log TFPA).

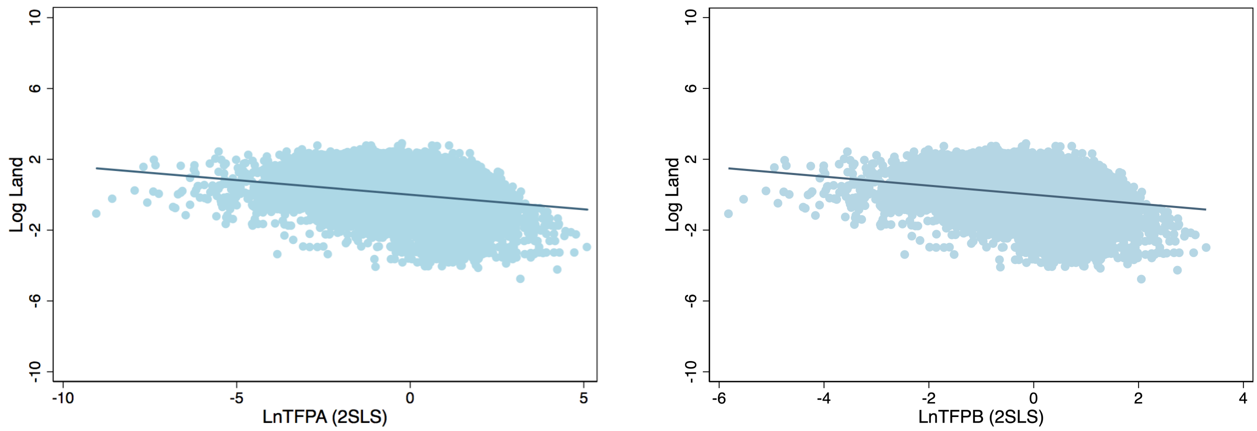
We explore this implication explicitly by analyzing GU's adjusted measure of TFP (TFPB) for efficiency gains at three levels of geographical variation in columns (3) and (4) of Table 1. A nationwide reallocation using the adjusted measure of productivity in GU's data generates an increase in agricultural productivity of 566%. This increase is much larger than those found in the related macro-development literature: 53% in China (Adamopoulos et al., 2017), 97% in Ethiopia (Chen et al., 2021), 186% in Uganda (Aragon et al., 2019), and 259% in Malawi (Restuccia and Santaeulalia-Llopis, 2017). Reallocation gains are large even within narrower geographical areas. For instance, the within-village reallocation is 147% in GU's adjusted data, whereas they are only 24% in China (Adamopoulos et al., 2017).

Given the substantial reallocation gain of 566% in GU's adjusted measure of productivity, the conclusion that "previous estimates of misallocation have probably overestimated the potential productivity losses due to misallocation" (p.46) is not warranted from this evidence.

The pattern of misallocation An important theme in the literature has been the connection between misallocation and land markets, first emphasized by Adamopoulos and Restuccia (2014) and subsequently analyzed using more detailed micro data in different contexts. This literature documents a weak connection between farm inputs and farm productivity, especially arising from frictions in land markets, whereas an efficient allocation of factors would feature a strong positive relationship between inputs and productivity.

We assess whether the adjusted measure of productivity in GU (TFPB) features a different pattern of misallocation. Figure 1 reports two panels, each plotting log land input

Figure 1: Land input and productivity in Gollin and Udry (2021)



(a) Unadjusted TFPA

(b) Adjusted TFPB

(operated land) on log TFP. Panel (a) reports the unadjusted measure (log TFPA) while panel (b) reports the measure adjusted by measurement error (log TFPB). Even though there is much less dispersion in TFPB than in TFPA, in both cases there is a weak negative relation between land input and productivity (fitted line). This is in sharp contrast to the strong positive relationship implied by an efficient allocation.[3] The evidence clearly indicates that land input is not strongly associated with farm TFP and this pattern of misallocation is no different when using GU's adjusted productivity measure.

This result is not surprising both from an economic and a methodological point of view. First, if misallocation is driven by land institutions that distribute land uniformly and/or prevent reallocation across households, then we should not expect that disaggregating the data at the plot level would address the issue. Second, GU's method simply removes random variation in productivity since the adjustment amounts to a scalar reduction in variance, which is proportional to the productivity variance across plots within households. As a

[3] It is worth noting that GU's estimated plot-level production function is close to constant returns to scale. In this context, low levels of misallocation would essentially require limited TFP dispersion, which is far from what GU's estimates show.

result, the fundamental relationship between land input and productivity is unaffected.[4]

The plot versus the farm The macro-development literature on misallocation focuses on the household farm as the unit of analysis and reallocation instead of the plot, an emphasis that is in part driven by the fact that land institutions in developing countries allocate land rights at the household level (Restuccia, 2020). The distinction between the plot and the farm analysis is important in rationalizing GU's evidence and properly assessing the importance of mismeasurement and misallocation in agricultural economies. In this context, a more pertinent comparison with the misallocation literature is to assess the extent of mismeasurement at the household-farm level.

It is straightforward to use GU's production function to map plot-level productivity to the farm level in a way that is quite comparable to the literature for given estimates of production function parameters. To aid in comparability, we simply use GU's data and estimates to aggregate their measures to the farm level. With a slight change in notation, the plot-level production function (household h, plot j, season-year t) in GU's equation (13) is $Y_{hjt} = e^{\omega_{hjt}}(L_{hjt})^{\alpha_L}(X_{hjt})^{\alpha_X}$. Given n operated plots for household h in period t, we can write the aggregated household-level output in each period as:

$$Y_{ht} = s_{ht}(L_{ht})^{\alpha_L}(X_{ht})^{\alpha_X},$$

where L_h and X_h are the total amounts of land and labor used by the household, and farm TFP is $s_{ht} = \sum_{j=1}^{n} e^{\omega_{hjt}}(\phi^L_{hjt})^{\alpha_L}(\phi^X_{hjt})^{\alpha_X}$, where ϕ^L_{hjt} and ϕ^X_{hjt} are the shares of a household's total inputs (L_h, X_h) used in plot j in period t.

Table 1 reports the variance and efficiency gains for TFPA aggregated at the household-farm level in columns (5) and (6). The results illustrate that a large portion of the adjustment in GU's TFPB measure is already accomplished in the literature by considering the farm as

[4] We find a similar pattern when looking at the allocation of labor across plots.

the unit of analysis. For instance, simply aggregating unadjusted plot TFP to the farm level reduces the variance by about 40% ($1 - 0.78/1.26$). This represents about two-thirds of the reduction in variance emphasized by GU. Similarly, using 2SLS, the nationwide reallocation gain falls from 23.9-fold across plots to 14.3-fold at the household level, whereas within villages falls from 4-fold to 2.1-fold.

GU provide a metric of overstatement of the potential gains from reallocating resources which is based on the difference in the variance of TFP, which is almost 3-fold in the case of Uganda, and conclude that the "extent of misallocation is substantially overstated if the contributions of risk and measurement error to the apparent dispersion of TFP are neglected" (p.45). This conclusion is, however, misleading since in no place in the literature reallocation across plots is the starting point. In particular, when comparing TFP variances, the overstatement relative to TFPB is instead a much more modest 1.28-fold when using farm productivity instead of TFPA. In terms of misallocation, efficiency gains from village-level reallocations may be even smaller across farms than across plots with adjusted productivity.

To summarize, taking GU's adjusted measure of productivity and production function estimates at face value, it is clear from the evidence that factor misallocation remains a critical component of agriculture in the developing world, contrary to GU's conclusions. Moreover, there are reasons to be concerned that GU adjustment over-attributes variation to mismeasurement as we discuss in detail in the next section.

3 Limitations of plot-level analysis

While the evidence indicates a large role for factor misallocation in African agriculture, a proper assessment of mismeasurement remains relevant. However, we show that some aspects of GU's method hinder it. For example, the assumption that the farm household does not face frictions in adjusting inputs across plots is inconsistent with a long literature documenting frictions within the household, driven by intra-household allocations or variation in the property rights of the plots (Shaban, 1987; Goldstein and Udry, 2008; Udry, 1996).

In this section, we focus on the use of plot-level data instead of the standard approach of focusing on household farms. We present evidence that measurement error is larger in parcels, a unit above the plot, compared to farm-level data, biasing upwards the extent of mismeasurement in GU's analysis. Using instead a method that relies on time variation in panel data we find a modest role for measurement error in farm-level data. We also show the implications for productivity dispersion and misallocation when estimating the production function at the farm level.

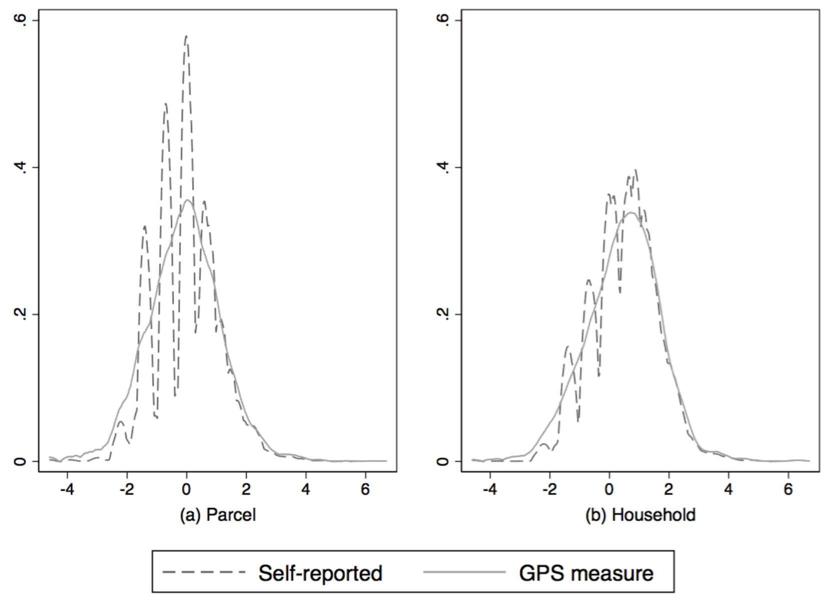
Measurement error in parcel-level data We assess measurement error by comparing self-reported size of land holdings to their GPS measure. GPS data is only available at the parcel level. However, as the parcel is a level of aggregation above the plot, this is a conservative assessment. We use the data for Uganda in Aragon et al. (2019).[5] We obtain parcel size and then aggregate it to the farm level. Figure 2 shows the distribution of self-reported land size and the corresponding GPS measures at the parcel and household levels.

There are three important observations. First, there is bunching in the distribution of self-reported land size. This is consistent with respondents (or surveyors) rounding parcel sizes. Second, deviations from the GPS value, which can be attributed to measurement error, are more pronounced for smaller parcels and farms. In particular, measurement error measured as the ratio of self-reported to GPS land size virtually disappears on the right tail of the distribution. This suggests that measurement error in land is non-classical, something already discussed in the micro-literature on farm size and productivity.[6] Third, measurement

[5] We do this because the GPS data is not provided in Gollin and Udry (2021). We use the Uganda Panel Survey for 2009-10, 2010-11, 2011-12 and 2013-14.

[6] See, for example, the evidence in Abay et al. (2019) on the negative correlation between measurement error and plot size, a pattern we also find in our data at the parcel level. Also note that GU assume classical measurement error.

Figure 2: Self-reported and GPS-measured parcel and farm size



error is larger at the parcel than at the farm level.

Table 2 provides summary statistics of the ratio of self reported to GPS land size, our proxy of measurement error, for parcel data and data aggregated at the household level. We observe that aggregating the data at the household level reduces the magnitude and dispersion of measurement error. For the median observation, measurement error is around 1.7% using parcel-level data, but around 0.04% using household level data. Similarly, the 90-10 ratio drops from 4.2 to 3.5 and the variance of log from 0.54 to 0.45 when aggregating at the household farm level.

Measurement error in farm-level data If plot-level data is problematic, how can we assess the extent of measurement error in measures of misallocation? The intuition behind GU's adjustment is that "observed variation in labor inputs across the plots of a single farmer that is not correlated with either output or other inputs is attributable to measurement error in labor." A similar approach can be applied to variation over time in panel data. This is

Table 2: Measurement error in land

	Measurement error in land	
	Parcel data	Household data
	(1)	(2)
Mean	1.611	1.407
Median	1.017	1.004
p(10)	0.526	0.560
p(90)	2.222	1.961
90-10 ratio	4.222	3.503
Variance (log)	0.536	0.447

Notes: The proxy for measurement error in land is the ratio of self-reported / GPS land size. Estimates using parcel level data (column 1) and aggregated at the household farm level (column 2).

indeed the approach in Bils et al. (2017), exploiting variation in output, relative to variation in inputs, as an alternative measure of revenue productivity (distortions). The extent to which variation over time in inputs is not reflected in variation in output, which varies across levels of revenue productivity, provides a metric of the extent of measurement error in measured distortions.

We apply this method using GU data but aggregated to the farm level to prevent additional measurement issues at the plot level. We report, in Table 3, an estimate of λ, the fraction of the dispersion in revenue productivity (TFPR) that is due to true variation in distortions.

We find that, regardless of the specification, the estimate of λ is fairly high (0.90 and 0.83), implying that between 10 and 17% of the variation in misallocation can be ascribed to measurement error. These results are consistent with the findings in Adamopoulos et al. (2017) using Chinese panel data, where this method detects only 4% measurement error in farm-level distortions measures and 10% in cross sectional farm-level data. The extent of farm-level measurement error is substantially smaller than that implied by GU's analysis at the plot level. It is also much smaller than in the manufacturing sector analyzed in Bils et al. (2017) for India and the United States, which suggests caution in making comparisons of

Table 3: Measurement error in distortions

	Using farm TFPA	
	2SLS	IVCRC
	(1)	(2)
$\tilde{\lambda}$	0.90	0.83
S.E.	0.06	0.05
95% C.I.	[0.78 - 1.01]	[0.70 - 0.94]

Notes: Estimates of λ as proposed in Bils et al. (2017), representing the share of the dispersion in TFPR that is due to true variation in distortions. All regressions include year and season fixed effects. Columns (1) and (2) use estimates of TFPA estimated by Gollin and Udry (2021) aggregated to the farm level.

measurement error across sectors, countries, and applications.

Plot versus farm level analysis Taking plot-level measurement issues aside, and the fact that the relevant unit of analysis for the practical purposes of reallocation is the household farm, we now address the differences in insights between the adjusted plot-level estimates in GU and a farm-level analysis. We use estimates of the farm-level production function documented in Aragon et al. (2019). These estimates are obtained using the same dataset as GU and a similar Cobb-Douglas specification. However, we use data aggregated at the household-farm level and exploit within-household variation for identification. In particular, we estimate the following panel data model with fixed effects:[7]

$$\ln Y_{it} = \ln s_i + \alpha_L \ln L_{it} + \alpha_X \ln X_{it} + \delta \text{weather}_{it} + \eta_t + \epsilon_{it},$$

where Y, L and X are the amounts of farm agricultural output, land and labor. weather_{it} and η_t are measures of local weather (temperature and precipitation) and period fixed effects

[7] This approach addresses time-invariant unobserved productivity. Our results are, however, similar when using a dynamic panel approach as in Shenoy (2017). As noted in Ackerberg et al. (2015), this approach can address time-variant unobserved productivity as long as it follows a first-order autoregressive process. We prefer using panel data methods rather than proxy variable methods, such as Olley and Pakes (1996), Levinsohn and Petrin (2003), Ackerberg et al. (2015), or Gandhi et al. (2020), because their key identifying assumptions fail in the presence of input market frictions (Shenoy, 2020).

that capture common productivity shocks, while $\ln s_i$ are household fixed effects. We call this last variable "farm productivity" and use it as a measure of the time-invariant component of household-farm total factor productivity.

Table 4 reports the productivity dispersion and gains from reallocation using plot- and household-level estimates from GU and Aragon et al. (2019). While the dispersion remains larger in the household fixed effects specification than in GU's TFPB (see panel A), the resulting reallocation gains are substantially larger when using GU's estimates (panel B). For example, nationwide reallocation gains are 566% in GU's adjusted measure of productivity compared with 186% in farm-level estimates, whereas village level reallocation is 147% in GU compared with 57% in farm estimates.

Table 4: Productivity dispersion and efficiency gains in Gollin and Udry's adjusted data and Aragon et al. (2019)

	Plot-level TFPB		Farm
	2SLS	IVCRC	productivity ($\ln s_i$)
	(1)	(2)	(3)
A. Dispersion			
90-10 log difference	1.74	1.29	2.23
90-10 ratio	5.67	3.64	9.30
Variance of log	0.53	0.29	0.84
B. Efficiency gains			
Nationwide	6.66	6.66	2.86
Region	5.36	4.69	2.48
Parish (Village)	2.47	2.17	1.57
C. Production function estimates			
Land ($\hat{\alpha_L}$)	0.69	0.53	0.37
Labor ($\hat{\alpha_X}$)	0.22	0.43	0.34

Notes: Mean of efficiency gains across season-years. TFPB is plot level productivity adjusted for late season output shocks, unobserved input quality, and measurement error using plot-level variation within household farms. Farm productivity ($\ln s_i$) from Aragon et al. (2019). Estimates in Panel C from Table A2 in Gollin and Udry (2021) and from Aragon et al. (2019).

From this perspective, even though there is a larger variance in farm-level productivity than in GU's adjusted TFPB, the reallocation gains and hence the extent of misallocation are about half in the farm-level analysis than in GU's adjusted data.

The difference arises because the cost of misallocation depends not only on the variance of productivity, but also on the parameters of the production function. Using household-level data, Aragon et al. (2019) finds decreasing returns to scale, while GU estimate a near linear production function at the plot level (with returns to scale ranging from 0.91 to 0.96). Thus, reallocation gains, even from smaller dispersion, are quite large.

There is a vast literature discussing the potential dangers of estimating production functions in the presence of distortions and unobserved variables, both of which may render estimates at the plot level less reliable than at the farm. A clear indication of this limitation is the fact that farm size and productivity are negatively related in GU's plot-level estimates, whereas the correlation is small but positive at the farm level (Aragon et al., 2019).

These considerations suggest that a farm-level analysis, as carried out in the macro development literature, provides a more accurate assessment of the extent of misallocation and relevant measurement error in agriculture than the plot-level analysis in Gollin and Udry (2021).

4 Conclusions

We use the evidence in Gollin and Udry (2021) to argue with the authors' conclusion that the misallocation literature vastly overestimates potential gains from factor reallocation in agriculture in low-income countries. The striking finding is that there is more misallocation associated with GU's adjusted measure of productivity than in the related macro-development literature. We show that GU's plot-level approach induces an increase in raw productivity dispersion and estimates of production function coefficients that generate unusually large reallocation gains. As a consequence, focusing on the reduction in the variance of TFP, as done in GU, is not informative of the extent of misallocation in agriculture.

After accounting for other sources of dispersion, such as measurement error, idiosyncratic shocks and input heterogeneity, GU's own estimates are consistent with substantial gains from factor reallocation, even at the village level. Moreover, GU's adjusted productivity measures are at odds with standard patterns of efficient allocations, as their data shows a negative correlation between productivity and input use.

We show that the standard approach in the macro-development literature of using the household farm as the unit of analysis deals to a large extent with some of the drivers of productivity dispersion between adjusted and unadjusted measures estimated in GU.

We also provide evidence suggesting larger measurement error at the plot level compared with the farm level, implying an upward bias on the extent of mismeasurement in GU's plot-level analysis. We instead assess the extent of measurement error in farm-level data using a similar approach but that leverages on panel data rather than on suspect plot-level data (Bils et al., 2017). We find much smaller scope for measurement error than in GU, and even relative to manufacturing plants in India and the United States, which highlights the importance of farm-level analysis in this context and suggests caution in making comparisons of measurement error across sectors, countries, and applications.

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.

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," Technical Report, National Bureau of Economic Research 2017.

Aragon, Fernando M, Diego Restuccia, and Juan Pablo Rud, "Are small farms really more productive than large farms?," Technical Report, National Bureau of Economic Research 2019.

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

Chen, Chaoran, Diego Restuccia, and Raül Santaeulàlia-Llopis, "The Effects of Land Markets on Resource Allocation and Agricultural Productivity," Technical Report, National Bureau of Economic Research 2021.

Gandhi, Amit, Salvador Navarro, and David A Rivers, "On the identification of gross output production functions," *Journal of Political Economy*, 2020, *128* (8), 2973–3016.

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.

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.

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

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.

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.