WPS8134

Policy Research Working Paper 8134

Land Productivity and Plot Size

Is Measurement Error Driving the Inverse Relationship?

Sam Desiere
Dean Jolliffe

WORLD BANK GROUP
Development Data Group
July 2017

Policy Research Working Paper 8134

Abstract

This paper revisits the decades-old puzzle of the inverse plot-size productivity relationship, which states that land productivity decreases as plot size increases. Existing empirical studies on the inverse plot-size productivity relationship define land productivity or yields as self-reported production divided by plot size. This paper considers an alternative approach to estimating yields based on crop cuts. The crop-cut method entails measuring and harvesting randomly selected subplots by trained technicians, and is recommended by the Food and Agriculture Organization for the accurate measurement of crop production. Using data representative of rural Ethiopia, the analysis indicates that the inverse relationship is strong when based on self-reported production, but disappears when based on crop-cut estimates. The inference from these findings is that the inverse relationship is an artifact of systematic overreporting of production by farmers on small plots, and underreporting on larger plots. The paper also discusses how rejecting the inverse plot-size productivity relationship has significant implications for the inverse farm-size productivity relationship.

This paper is a product of the Development Data Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at at samdesiere@gmail.com or djolliffe@worldbank.org.

Land Productivity and Plot Size:
Is Measurement Error Driving the Inverse Relationship?

Sam Desiere[†a] and Dean Jolliffe[b]

JEL codes: O12, Q12, Q15
Keywords: LSMS-ISA, Ethiopia, Agricultural Productivity, Crop Yields, Plot Size

Acknowledgements: The authors are grateful to the UK Department for International Development Ethiopia and Tim Conway for generous funding assistance. Sam Desiere also benefitted from a travel grant from the FWO Research Organization (V404816N). The authors thank Tassew Woldehanna, Assefa Admassie, Solomon Shiferaw, and Alemayehu Seyoum Taffesse for their generous insight and feedback and Demirew Getachew and Tadele Ferede for their support in the dissemination of this paper. For many comments and useful feedback, the authors also wish to thank Asmelash Haile Tsegay, Alemayehu Ambel, Christian Johannes Meyer, participants at the World Bank – DFID workshop on the Dynamics of Wellbeing in Ethiopia, the 2016 meetings of the African Association of Agricultural Economics, and the 2017 annual meeting of the Center for the Study of African Economies.

[†]Corresponding author (sam.desiere@ugent.be)
[a] Dept. of Agricultural Economics, Ghent University, Coupure Links 653, 9000 Gent, Belgium
[b] Development Research Group, World Bank, Washington, DC, USA

1. Introduction

With an estimated 40 percent of the world's extreme poor living in Sub-Saharan Africa (Ferreira et al., 2016),[1] the vast majority of whom derive their economic livelihood from agricultural activities (Livingston et al., 2001), the productivity of land is central to the global goals of poverty reduction and economic development. More directly demonstrating this point, Irz et al. (2001) estimate that a 10 percent increase in farmland productivity can reduce the number of poor people in Africa by 7 percent. Despite the global importance of land productivity in determining the well-being of the poor, one very basic question of land productivity remains a puzzle: is there, as is commonly accepted, an inverse relationship (IR) between farm size and productivity?

Two empirical regularities between land size and land productivity have puzzled development economists for decades. The first regularity is well-known; that is the farm-size productivity inverse relationship (hereafter referred to as farm-size IR), first noted by Chayanov (1926) in Russia and later rediscovered by Sen (1962) in India. It states that land productivity decreases with farm size. This relationship has been observed in many developing countries and can be considered a stylized fact (Kimhi, 2006; Larson et al., 2014). The farm-size IR has received considerable attention from researchers and policy makers because of its controversial implications for land reform (Collier and Dercon, 2013): if small farms are intrinsically more productive than larger farms, land redistribution would not only shift the distribution of wealth but also increase land productivity generally.

The second empirical regularity has received far less attention. We call it the inverse *plot-size* productivity relationship (and refer to it as plot-size IR). This relationship, observed in many Sub-Saharan African countries, states that smaller plots are more productive than larger plots, even within a farm household (Ali and Deininger, 2015; Assunção and Braido, 2007; Barrett et al., 2010). Although rarely noted, such a relationship implies that land fragmentation—subdividing already small plots into even smaller units—would heighten productivity. But more and smaller plots lead to increased travel time (in moving from plot to plot), more wasted space along borders, and fewer options for mechanization, which typically functions best on larger plots. For these reasons and others, this finding does not seem credible. A common policy recommendation to improve land productivity is not fragmentation but

[1] Estimates from PovcalNet (*iresearch.worldbank.org/PovcalNet*), the World Bank data tool for estimating extreme poverty, indicate that most of the poor in the rest of the world also live in rural areas. For example, in China essentially all the extreme poor (living on less than US$1.90 a day) are estimated to be in rural areas; and in India about 80 percent.

greater land consolidation, and indeed land consolidation programs have been carried out in several countries (Ali and Deininger, 2015; Blarel et al., 1992; Pašakarnis and Maliene, 2010).

The inverse plot-size and farm-size relationships are closely related. The most common, and most intuitive, explanation for the farm-size IR is missing land, labor, credit, or insurance markets; missing markets have been found to explain differences in land productivity between households (Assunção and Ghatak, 2003; Barrett, 1996; Carter and Wiebe, 1990; Eswaran and Kotwal, 1986; Feder, 1985). Households with little land but an abundance of labor apply too much labor to their land because incomplete land markets prevent them from buying or leasing more land, and incomplete labor markets hinder their ability to rent out their labor. Indeed, several studies have shown that the farm-size IR disappears if labor input is valued at the market wage and have shown that on small farms, land is more intensively cultivated than on larger farms (a point noted by many, including Ali and Deininger, 2015; Benjamin and Brandt, 2002; Carter, 1984; Heltberg, 1998; Lamb, 2003). Missing markets cannot, however, explain differences in the productivity of plots owned by the same household. In other words, missing markets can explain the farm-size IR but not the plot-size IR. The existence of the plot-size IR has even led some to reject missing markets as the main explanation for the farm-size IR (Assunção and Braido, 2007; Kagin et a,, 2016).

Two other explanations are often suggested to explain both types of IRs: soil quality, and measurement error in self-reported plot size. Soil quality might explain the IR if smaller plots have on average more fertile soils than larger ones (Benjamin, 1995; Benjamin and Brandt, 2002; Bhalla and Roy, 1988; Chen, Huffman, & Rozelle, 2011), though in one of the few studies to use objective measures of soil quality data, Barrett et al. (2010) show that soil quality contributes only marginally to explaining the IR in Madagascar. They suggest that either differences in labor inputs across plots or measurement error with respect to plot size are the only remaining candidates for explaining the plot-size IR. Lamb (2003) was the first to suggest that measurement error in self-reported land area could explain the inverse farm and plot-size relationship. Carletto, Gourlay and Winters (2015) find that replacing self-reported land size with GPS measurement weakened the farm-size relationship somewhat in several African countries, but the relationship never disappeared and in Uganda was even strengthened (Carletto, Savastano, & Zezza, 2013).

This paper tests a new explanation for the plot-size inverse productivity relationship: measurement error in self-reported production. All recent papers on the IR define yields as self-reported production divided by plot size (Ali & Deininger, 2015; Carletto et al., 2015; Carletto et al., 2013; Larson et al., 2014)[2]— which we refer to throughout the paper as self-reported yields. One reason why systematic measurement error in self-reported production has not yet been considered as an explanation for the IR is that the data requirements to test it are very demanding. Systematic measurement error can only be studied if there is a second, independent measurement of the same concept. In this study, we draw on two waves of data representative of rural Ethiopia which asked farmers to report production for all plots; these plots were also measured with GPS. In addition to the self-reported crop production, crop cuts were conducted on a limited number of randomly selected plots.

First used in the 1950s in India, the crop-cut method is now widely recommended by the Food and Agriculture Organization for accurate measurement of crop production (Fermont and Benson, 2011). This methodology estimates yields by first randomly sampling plots from a master listing, then identifying a random start point within each selected plot, delimiting from that point a small subplot (often a 4m x 4m square), and finally cutting and weighing the harvest from this subplot. Yields are then defined as the harvest in the subplot divided by the area of the subplot.

While crop cuts are often considered to be a gold standard for measuring crop production, this does not mean that they estimate yield without error. As one example, even though mechanical scales are much more precise than self-reports, they still entail rounding errors when reading the weight from the scale. As another example, because crop cutting entails randomly selecting a subplot from a larger plot, there is sampling error if yields within the plot are heterogeneous.

The main assumption in this paper is that whatever error there is in crop-cut estimates, whether from sampling or measurement, the error is independent of the size of the plot.[3] Given that the size of the subplot is fixed, regardless of the size of the entire plot, there is little reason

[2] Most studies about the IR do not discuss how they measured yields, but all the studies cited above used self-reported production. As far as we know, no study has used crop cuts except the very recent study of Gourlay, Kilic, and Lobell (2017).

[3] In this paper we emphasize that measurement error is likely to affect both crop cut and self-reported yield estimates, and while not a necessary condition for our findings, it strikes us as reasonable to assert that the magnitude of error in self-reports is expected to be much larger than that associated with crop cuts. One piece of evidence supporting this assertion is that the coefficient of variation in estimated maize yields from self-reports is about twice as large (in both waves) as the coefficient of variation when based on crop cuts. We also interpret findings from Gourlay, Kilic, and Lobell (2017, Table 2) as aligning with this view. They use as a benchmark production estimates from trained technicians harvesting an entire plot to compare estimates for these same plots when based on self-reports and crop cuts. The average crop cut estimate is within 10 percent of the benchmark, while the self-reported estimates are more than two and a half times greater than the actual harvest.

to expect measurement error in crop cuts to be related to the overall size of the plot from which the subplot is selected.[4]

The key finding in this paper is that when we use crop cuts to measure yields, the plot-size IR disappears. In contrast, and consistent with previous studies, when we use self-reported yield, the relationship is strongly negative. This finding suggests that production is systematically over-reported on small plots and under-reported on larger ones.

In addition to advancing the IR literature, this paper contributes to the less voluminous literature on how systematic measurement error in household surveys can affect statistical inference. This paper adds another example to the literature demonstrating that not all measurement error is white noise (Aguiar and Bils, 2015; Beegle et al., 2012; Gibson and Kim, 2007). As is well-known, but perhaps insufficiently emphasized, this illustrates that the 'Iron Law of Econometrics' – as Hausman (2001) called the observation that measurement error in the independent variables results in attenuation bias in regression coefficients – critically depends on the assumption that measurement error occurs randomly. If measurement error is systematic, the data may produce spurious correlations. Most previous studies focused on systematic measurement error in the independent variables, but this study illustrates that non-random measurement error in the dependent variable can also induce spurious correlations.

In the next section, we discuss the data and provide some descriptive statistics. We then show the conditions in which measurement errors in self-reported production can generate the plot-size IR and discuss the estimation strategy. In the results section, we show that the existence of the plot-size IR depends on the method used to measure yields. In conclusion, section 5 discusses the implications of our findings for the farm-size IR.

2. Data and descriptive statistics

We use data from the first two waves of the Ethiopia Socioeconomic Survey (ESS), which is representative of all rural areas of Ethiopia.[5] This survey is a continuing project to collect panel data in Ethiopia on both household wellbeing and agricultural activities. It is implemented by the Central Statistical Agency of Ethiopia in close collaboration with the LSMS-ISA[6] team of

[4] A caveat to this assertion is the possibility of an 'edge effect'. The concern here is that crop cuts are more likely to be done in a plot's interior, which introduces measurement error correlated with the size of the plot if yields along the edges differ from yields in the interior. This concern is discussed in detail later in the paper.
[5] All data and relevant documentation are available at go.worldbank.org/HWKE6FXHJ0. This includes a manual detailing the crop-cutting procedures the enumerators followed.
[6] LSMS-ISA: Living Standards Measurement Study – Integrated Surveys on Agriculture.

the World Bank, which has a long history of producing high-quality household survey data. The first wave was administered in 2011/2012 to 3,969 rural households, the second in 2013/2014 to 5,262 households – 3,776 panel households and 1,486 new, mainly urban, households. Our data set covers only households that were interviewed in both waves, which are all rural. The survey collected extensive information on household characteristics, consumption, living conditions, and health.

The unique feature of the survey is its focus on agriculture. To gather detailed and accurate agricultural data at the plot level, households were visited three times during the agricultural year. The first visit, in September-October, collected data on planting activities. During this visit, the area of most plots was measured with GPS.[7] The second visit, in November, implemented the livestock module. The final visit, between January and April, collected data on agricultural production and executed the household questionnaire. The first and final visits include detailed information on labor inputs at the plot level which we also use in our analysis.

In this paper, we exploit the fact that yields from the same plots were measured with two different methods: crop cuts and self-reported production. The crop cuts in the ESS data are a sub-sample of crop cuts that are done for the Government of Ethiopia's (GoE) Annual Agricultural Sample Survey (AgSS) by the same trained technicians, following the same protocols as used by the AgSS. The AgSS data are used by the GoE to measure core agriculture statistics including total volume of crop production, yield of crops and also crop forecasts.

In both survey waves, crop cuts were carried out on 23 major crops: five plots per crop were randomly selected from a list of all plots (stratified on crops) cultivated by the sampled households within an enumeration area. In most cases, the plot selected for crop cutting was mono-cropped, but in those cases where the crop was cultivated on fewer than five plots in pure stand (that is, mono-cropped) within an enumeration area, the crop cut was made on mixed-stand plots.

Once a plot was identified, a rectangular subplot within it was randomly sampled and delineated. In the first wave, the subplot was 2m by 2m square; in the second it was 4m by 4m square. Within the subplot, the crop was harvested by a trained enumerator and weighed. If logistical constraints allowed, crop-cutting occurred simultaneously with the farmer's harvest of the main crop. In both waves, nearly 40 percent of the crop cuts were executed in November

[7] The smallest plots were also measured with tape and compass—considered as a gold standard for measuring small areas.

and more than 90 percent in the October-December period. Both fresh and dry weights were recorded; the correlation between them was over 0.95. We used the dry weights to calculate yields. There were 2,975 crop cuts made in wave 1 and 3,532 in wave 2. We discarded, however, crop cuts for which we had fewer than 150 observations per crop. Ultimately, 5,920 crop cuts remained in the data set, providing yield estimates for 18 crops.[8] For more details on implementation protocols for crop cutting, see Government of Ethiopia, Central Statistical Agency and World Bank (2013).

Farmers reported the harvest per crop and per plot during the third visit. In wave 1 they only reported production for all their plots if a crop cut was implemented on at least one plot, while in wave 2 production was reported for all plots.[9] This explains why there are significantly fewer observations for wave 1 (3,683) than wave 2 (23,638).

Our unit of analysis, plots, cannot be matched across waves. Plots are, however, the smallest unit of a parcel, and parcels can be tracked across waves. Land parcels in Ethiopia are generally small, with mean and median parcel sizes in ESS wave 2 of 0.34 hectares and 0.22 hectares, respectively (and 95 percent of the parcels smaller than 1 hectare). Given the small size of the parcels, we view the parcel fixed-effects estimates as controlling for unobserved soil quality.

By gathering detailed data on labor input per plot, the ESS data also allow us to examine the hypothesis that differences in labor use at the plot level is the main source of the plot-size IR (Barrett et al., 2010). The ESS includes a labor module on planting (first visit) and harvest (third visit) activities. Farmers are asked to report family, hired, and exchange labor by plot, and this information is included in the regression analysis to address the hypothesis that differences in labor inputs across plots is driving the IR. Furthermore, differences in labor input may also explain differences in self-reported yields and crop cuts, a point that will be elaborated in the methodology section.

To examine if crop cuts were indeed random in relation to households and plots, we compare households with at least one plot selected for crop cutting with those without any plot

[8] Barley, maize, millet, sorghum, teff, wheat, chick peas, haricot beans, horse beans, lentils, field peas, vetch, line seed, ground nuts, nueg, rape seed, sesame, fenugreek.

[9] Another difference in questionnaire design between the waves is that only in wave 2 were farmers asked to estimate the share of land devoted to each crop if the plot was cultivated in mixed stand. As a result, yields of plots in mixed stand are more precisely estimated in wave 2 than wave 1. To account for this, we included in the regressions an interaction term between the wave and mixed-stand cropping. Moreover, the results do not change when we restrict the sample to pure-stand or mixed-stand plots. In addition, defining yields on mixed-stand plots in wave 2 as in wave 1, has a limited effect on the IR. The IR for self-reports in the full sample with household fixed effects changes from - 0.397 to -0.402 when estimating yields on mixed-stand plots in wave 2 in the same way as in wave 1.

selected for crop cutting and compare plots with and without crop cuts.[10] Note that the validity of the main finding in this paper does not rest on the random allocation of crop cuts across households since we compare differences in yields based on crop cuts or self-reports for the same plots, but the random allocation does allow us to extend our inferences to the entire rural population of Ethiopia.

Household characteristics of both groups are very similar (Table 1). The only important difference is that households selected for crop cuts owned slightly more land (1.34 hectares versus 1.18 hectares). This is as expected, since households with more land are also more likely to cultivate at least one plot suited for crop cutting.

In terms of plot characteristics, observed differences between plots selected for crop cuts and the other plots (Table 2) are related to the crop-cut selection rule which places a preference on pure-stand plots. By design then, plots selected for crop cutting are more likely to be pure-stand plots (87%) than the plots not selected for crop cutting (57%). Since pure-stand plots are more likely to be further away from the dwelling, crop cuts also tended to be somewhat further away and are slightly larger than the average plot. Differences in fertilizer application (both organic and inorganic), irrigation, labor input, and soil quality are small. The only exception is family labor during the harvest. Plots without crop cuts appear to be more intensively harvested (12.2 days/hectares) than plots with crop cuts (8.1 days/hectares).

To explore whether there is a plot-size IR in our data, we examined maize yields in waves 1 and 2 by quartile of plot size and measurement method (Figure 1), restricting the sample to plots where there had been crop cuts. A first observation is that self-reported maize yields are similar between both waves, while maize yields estimated from crop cuts are substantially lower in wave 2 than wave 1. This observation holds more generally in the data. Crop cuts are on average 40 percent lower in wave 2 than wave 1, while self-reported yields are, on average, 10 percent lower between the waves.[11] A potential explanation is the change in

[10] Sampling weights are not used in the descriptive statistics or the regressions, although the standard errors are corrected for clustering at the enumeration area.

[11] We have no compelling evidence to explain this decline in yields across the waves, but offer an explanation that was suggested by a co-manager of the survey operations. Although most enumerators participated in both waves and were trained intensively beforehand, it was suggested that some enumerators may have been confused and done crop cuts on a 2m x 2m square in wave 2, rather than on the prescribed 4m x 4m. Such errors are unavoidable in large-scale surveys. As a consequence, yields may be underestimated in wave 2, because we always divided the weight of each subplot harvest in wave 2 by 16m². We have some very weak evidence supporting this suggestion – in wave 1, the coefficient of variation (COV) of estimated maize yields was 1.15 for self-reports and 0.53 for crop cuts. In wave 2, dispersion as measured by the scale-independent COV declined slightly for self-reports to 1.11, but increased substantially to 0.64 for crop cuts. The increase in dispersion of the crop

crop-cutting methodology (from 2m by 2m to a 4m by 4m subplot) between the waves. To account for this observation, all regressions include a dummy for the wave and the main results will also be presented separately for each wave.

A second observation is that in both waves whether there is an IR depends on the measurement method. In both waves, self-reported yields decrease with plot size, while crop cuts tend to increase with plot size. Both findings also hold generally in the data. Using all observations, a simple, bivariate regression confirms the existence of the IR in self-reports. Regressing the log of self-reported yields[12] on the log of plot size suggests that yields decrease by 34 percent if plot size doubles (t-statistic=73), while the bivariate regression for crop cuts indicates that yields increase by 9 percent if plot size doubles (t-statistic=10). In the remainder of this paper we will show that these findings continue to hold when accounting for other possible explanations of the IR for self-reports, which suggests that the IR is caused by systematic measurement error in self-reported production.

3. Methodology

We use a simple model to explain how systematic measurement error in self-reported production can generate a plot-size IR. Traditionally, the plot-size IR is estimated by regressing yields on plot size, A, and a set of household and plot characteristics, X:

$$\log(yield_i) = \alpha\log(A_i) + \beta X_i + v_i \qquad (1)$$

where i represents the plot and v_i is a normally distributed error term. The parameter of interest is α, which reflects the association between plot size and yield. An IR can be rejected if $\alpha \geq 0$.[13] When estimating equation 1, household fixed effects are often included to account for unobserved household characteristics, such as wealth. Bevis and Barrett (2016) also include

cut measures is consistent with the suggestion of increased error in wave 2. Again, however, this error is not related to the size of the plot from which the subplot was sampled.

[12] Production was converted to its monetary value using local prices. If we had at least 10 observations, median self-reported prices by crop were calculated at the Woreda level, which is the administrative unit just above the lowest unit (the Kebele). If there were fewer than 10 observations, the median national price was used.

[13] α is typically assumed to be constant. If α depends on A, equation 1 is non-linear, which complicates the estimation procedure (see Verschelde et al., 2013 for an application).

plot fixed effects, thereby controlling for unobservable time invariant plot attributes such as soil quality, and identifying the IR by exploiting differences in plot size across waves.

The dependent variable in the equation is 'observed' yield, $yield_i$. Observed yields differ from 'true' yields, $yield_i^*$, in the presence of measurement error, ε_i. In this study, observed yields are equal to either self-reported production divided by plot size or yields estimated from crop cuts. Both may suffer from measurement error. Errors in self-reports are typically due to recall error, conversion error (from non-standard to standard units), and in the case of self-consumption, the production may never have been measured. Errors in crop-cut estimates can result from imperfect observance of the crop-cut protocols and also from sampling error if yields are heterogeneous within the plot (because the crop cut is a random sub-sample from the plot). In sum, for both methods of yield measurement, the relation between observed and 'true' yields can be expressed as:

$$yield_i \equiv yield_i^* + \varepsilon_i \qquad (2)$$

Measurement error in observed yields can only explain the IR if it is negatively correlated with plot size. To illustrate this, assume that the IR does not exist, or that $corr(yield_i^*, A) = 0$. Consequently, a spurious IR will be observed if, and only if:

$$corr(yield, A) < 0 \quad \Leftrightarrow \quad corr(yield^* + \varepsilon, A) = corr(\varepsilon, A) < 0 \qquad (3)$$

Equation 3 summarizes the point that if measurement error in observed yield is negatively correlated with plot size, observed yield will exhibit plot-size IR even when true yield is independent of plot size. From this expression, it is also clear that "classical" measurement error (i.e., independent of A) will never generate an IR. We argue that measurement error in self-reported yields can be correlated with plot size, but that is unlikely to be true for measurement error in crop cuts. Hence, the IR should disappear when crop cuts rather than self-reported yields are used as the proxy for 'true' yields.

Several types of measurement error in self-reported production satisfy equation 3. Consider, for example the case where farmers always over-report production by a constant positive quantity regardless of plot size. This over-reporting generates an IR because a small, positive measurement error in self-reported production biases self-reported yields more on a small plot than a larger one. Such a constant, positive measurement error could be introduced if farmers always round production upwards to the nearest multiple of 5kg, 10kg, or 25 kg. Similarly, because systematic over-reporting of production on small plots and under-reporting on larger ones also satisfies the condition expressed in equation 3, it too can generate a spurious plot-size IR. Under-reporting on larger plots may occur if farmers fear taxation or do not want to reveal their wealth.

While we are unable to offer candidate explanations for why measurement error might be negatively correlated with plot size, it is noteworthy that a similar pattern has been observed for self-reported land area. Farmers tend to overestimate the area of small plots but underestimate the area of larger plots (Carletto et al., 2015). Moreover, the order of magnitude of over- and underestimation is substantial. For instance, Carletto et al. (2015) report that farmers overestimate plot size by 103 percent relative to GPS measurement on plots smaller than 0.5 acres (2,000 m²) and underestimate it by 33 percent on plots larger than 5 acres (20,000 m²). If measurement error in self-reported production follows a similar pattern and has a same order of magnitude, then that could explain the IR.

An assumption implicit in the discussion of equation 3 is that crop cuts and self-reported yields are measuring the same underlying concept. But, it might be the case that crop cuts are measuring maximum *potential* yield, while self-reported production measures the actual harvest.[14] While crop cuts are exposed to the same weather shocks and pests as all of the other plots, they are carefully (and completely) harvested. If the IR holds for self-reported yields but disappears for crop cuts, it may simply be that smaller plots are more completely harvested. To control for this possibility, we include in the regressions labor input during planting and harvest at plot level. We also estimate the association between plot size and labor input to see if small plots are indeed more intensively harvested than larger ones.

4. Results

We first establish that the plot-size IR holds for self-reports (Table 3) by testing the conventional explanations for the IR. Next, limiting the sample to plots for which we have both crop cuts and self-reported production (Table 4), we show that the IR disappears if crop cuts are used to estimate yields. We then examine whether these findings continue to hold when accounting for labor inputs (Table 5 and 6) and for differences in crop choice between small

[14] A caveat to viewing crop cuts as maximum attainable yields is that they are exposed to some risk of crop loss from pests and weather in the field. They are not maximum yields as would result in a controlled laboratory.

and large plots (Table 7) and use sensitivity checks to rule out that the IR is driven by differences in crop-cut procedures between the waves, or by influential outliers (Table 8).

Column 1 of Table 3 considers the standard IR regression which includes household fixed effects and plot characteristics.[15] We are able to control for a fairly large set of observable plot characteristics including factors such as: slope, elevation, potential wetness index, distance between plot and dwelling, title status of plot, whether pure stand, irrigated, and several controls for fertilizer type. The household fixed-effects model provides a way to control for all time-invariant household attributes that might be confounding factors affecting the association between yields and plot size. This specification indicates that yields decrease by 40 percent if plot size doubles, which is more strongly negative than indicated by the bivariate association noted earlier. It also suggests that the IR not only holds in Ethiopia but is stronger there than in many other Sub-Saharan African countries (Larson et al., 2014).[16] Column 2 replaces household fixed effects with parcel fixed effects, thereby identifying the IR through variation in plot size within a parcel, both within and across waves. We interpret parcel fixed effects as controls for unobserved plot characteristics, such as soil quality, that are unlikely to change substantially within a parcel. Including parcel fixed effects again does not weaken the IR, showing that differences in soil characteristics between small and large plots do not explain the relationship. The final specification (column 3) includes enumeration area fixed effects and additional controls for household characteristics. The set of household controls is limited to age, gender, literacy of household head and an asset index to proxy wealth. We find no change in the IR when considering this specification. While we view this specification as somewhat less informative than the (geographically much more refined) parcel fixed-effects estimation, we

[15] We also estimate the IR with household-wave fixed effects to address the potential concern that the IR is resulting from shadow prices faced by the household that are changing over the waves in a way that differentially affects productivity across plot size. Across several different samples and specifications, using household-wave fixed effects did not change the findings. For example, in Table 3, column (1) the IR for self-reports is -0.397 with household fixed effects, when controlling for household-wave fixed effects, the IR is essentially the same at -0.401. When restricting the sample to only those plots where we have both crop cuts and self-reports, the IR for self-reports with household fixed effects is -0.287 while the IR with household-wave fixed effects is -0.291. Finally, when estimating the IR for crop cuts with household fixed effects, the coefficient on plot size is 0.119 and the same coefficient when estimated with household-wave fixed effects is 0.118. All the coefficients reported here are statistically significant at the 1% level.

[16] Using a different data source from Ethiopia (Ethiopian Rural Household Survey) based on a purposeful (not probabilistic) sample design from 15 woredas (i.e. districts) from 2004 and 2009, Savastano and Scandizzo (forthcoming) examine heterogeneity in the relationship at different points on the conditional yield distribution. They interpret their analysis as showing both non-linearities in land size - productivity relationship and heterogeneity in these nonlinearities at different points on the conditional yield distribution.

include it here because in the next step of the analyses, the sample will be restricted to plots which were crop cut. In this case, parcel or household fixed effects reduce the sample size substantially, since not all households had at least two plots with a crop cut and even fewer parcels had more than one plot that was crop cut.[17]

In sum, in the full sample, the plot-size IR based on self-reported yields is remarkably stable (Table 3) when examining differences in land productivity between large and small plots within a household (column 1), within the same parcel (column 2), or between households in the same enumeration area (column 3). This shows that the conventional hypotheses, notably missing markets or unobserved plot characteristics, do no explain the plot-size IR.

Table 4 (columns 1 to 3) presents the same regressions as Table 3, but limits the sample to plots that were crop cut. This reduces the sample size from over 25,000 observations to slightly over 5,000. The table (columns 4 to 6) also presents the results of the relation between plot size and crop cuts. The regression results show that the IR for self-reports continues to hold no matter what the level of fixed effects, although the relationship weakens in comparison to the results for the full sample. One explanation of the latter finding is that the IR in wave 2 appeared to be stronger on mixed-stand than on pure-stand plots.[18] Since pure-stand plots were more likely to be crop cut, restricting the sample to plots with crop cuts automatically weakens the IR for self-reports.

The IR disappears, however, if crop cuts are used as the dependent variable (Table 4, columns 4 to 6). This holds for each of the specifications with fixed effects separately for household, parcel and enumeration area. We even observe a positive and highly significant association between plot size and crop cuts in all specifications. Yields, as measured by crop cuts, increase by 10 percent if plot size doubles.

The regressions discussed so far have not included labor inputs, allowing our results to be comparable with previous IR studies, which lack data on labor input at the plot level. Controlling for labor input is, however, important in our study since crop-cut and self-reported

[17] Pooling the waves, 1,405 households had at least two plots selected for a crop cut. Similarly, of the 3,822 parcels with a plot selected for crop cuts, 1,224 had at least two plots selected.

[18] In wave 2, the IR for self-reports with household fixed effects was -0.381 on pure-stand plots, and -0.492 on mixed-stand plots. In wave 1, the IR for self-reports was stronger on pure-stand (-0.308) than on mixed-stand plots (-0.215) but this finding must be interpreted carefully: because the area devoted to each crop in mixed-stand plots was not reported in wave 1, yields on mixed-stand plots were underestimated in that wave. Since we have many more observations on self-reported production in wave 2 than in wave 1, the results for the full sample are primarily driven by wave 2. When estimating the IR for crop cuts separately for pure-stand and mixed-stand plots, in both cases we observed a positive association between plot size and crop cuts. This shows that sample selection issues do not drive the IR for either self-reports or crop cuts.

yields may not necessarily measure the same underlying concept. To some extent, crop cuts may be viewed as measuring maximum attainable yields because they are carefully harvested by trained technicians, while self-reported production is presumably an approximate measure of the actual harvest. Consequently, the observed IR for self-reports would be a real phenomenon if small plots are harvested more completely than large ones. This would not rule out that potential yields, as measured by crop cuts, remain constant across plot size. This hypothesis, which can be tested by including labor input in the regressions, is only partially upheld in the data.

Including labor input in the regressions weakens the IR for self-reports, but the relationship remains negative and statistically significant (Table 5, column 1). Without taking labor input into account, self-reported yields decrease by 30 percent if plot size doubles, and by 15 percent when labor input is accounted for.[19] Although we observe a positive association between yields and all types of labor (family, hired and exchanged) during two agricultural activities (planting and harvest), the positive association is especially pronounced for family labor during the harvest.[20] The positive association between plot size and crop cuts is not sensitive to including labor inputs in the regressions (Table 5, column 2), and we still observe a positive association between labor input and crop cuts.

The observation that the IR is sensitive to controlling for labor input implies an inverse relation between plot size and labor input.[21] In other words, small plots are more intensively cultivated than larger ones. This is confirmed in the full sample. Regressing labor input per hectare on plot size (Table 6), shows that there is a strong, inverse relation between plot size and total labor input per hectare, both for planting and harvesting. Decomposing total labor input in family labor and hired labor, shows a strong negative association between plot size and family labor and a positive, but less robust, relation between plot size and hired labor. Both

[19] Table 5 presents results with household fixed effects only for the restricted sample. The results are similar if enumeration area or parcel fixed effects are used and also hold in the full sample. Results are available upon request.

[20] There is an endogeneity issue here. Yields can be high because more labor is used or more labor is used when production (and thus yields) is high. The positive association between labor input and crop cuts suggest that this second channel is also relevant. We do not attempt to disentangle the causal pathways because our primary concern is to demonstrate that our findings remain robust when labor inputs are accounted for.

[21] We capture information on the type of labor, whether family, exchange, or hired in and the days worked; but we do not capture intensity of effort (nor do we know of any study that has). If for some reason, in addition to investing more labor hours into small plots, the farm households also worked more intensely on small plots, this would further temper the IR. But, this cannot explain why the two yield estimates (based on crop cuts and self-reports) differ.

findings are not novel in the literature and have been reported by Ali and Deininger (2015) in Rwanda and Larson et al. (2014) in several African countries.

Until now we have maintained the questionable assumption that the IR holds across crops. For instance, it may be that more profitable (but perhaps riskier) crops are cultivated on smaller plots, which would explain the IR. To test whether differences in crops grown between small and large plots contribute to the IR, we re-estimated it to include 'crop-wave-enumeration area' fixed effects (Table 7). This specification implies that the IR is identified by exploiting variation in plot size for plots within the same enumeration area on which the same crop is cultivated in the same wave. We also re-estimated the IR for the sample of maize plots.[22] The advantage of the crop-specific regression is that we can express maize yields in quantities (kg/ha) rather than needing to convert quantities into monetary values, and thereby essentially providing an additional sensitivity check. Because all specifications shown in Table 7 include labor input, we can also compare the results with those shown in Table 5. When we do so, the IR for self-reports is even stronger when crop fixed effects are included and for the sample of maize plots, while the impact on the IR for crop cuts is limited. In fact, in most specifications we examined that included crop fixed effects or controls, the IR for self-reports was stronger than in the baseline specifications, which suggests that differences in crop patterns between small and large plots strengthen rather than weaken the IR.

Several robustness checks indicate that our findings hold across different subsamples. First, we estimated the IR separately for wave 1 and wave 2 of the survey (Table 8, columns 1 to 4). This is informative because the size of the crop cut plot changed from a 2m x 2m square in wave 1 to a 4m x 4m square in wave 2. In both waves, we find a negative and significant association between plot size and self-reported yields and a positive association between plot size and crop cuts. As a second robustness check, we assessed the impact of outliers (Table 8, columns 5 and 6). Following the strategy of Larson et al. (2014) we excluded the bottom and top 5 percent of yields. The magnitude of the IR declines, but it is still statistically significant for self-reports. Excluding the tails has little effect on the association between plot size and crop cuts. Perhaps not surprisingly, outliers appear to be more of an issue for self-reported production than for crop cuts.

[22] We also estimated the IR for the sample of teff plots, the second most widespread crop in our data. The IR for self-reported teff yields (including labor input and EA fixed effects) is -0.449, while the association between plot size and crop cuts is positive (0.0454), but not significant at any conventional level.

Having tested and, based on the results, rejected all previous explanations of the IR, we argue that systematic measurement error in self-reported production is the remaining explanation for the IR. The implication is that self-reported production is over-reported on small plots and under-reported on larger ones. Whether over- or under-reported, however, the difference needs to be substantial to generate an IR. A simulation exercise, not reported here but available upon request, suggests that relative to crop cuts, production is likely to be over-reported by up to 50 percent on plots smaller than 80m² and under-reported by 25 percent on plots over 1,500m². Although that measurement error seems large, it is of the same order of magnitude as the error observed in self-reported land area relative to GPS measurement (Carletto et al., 2015). Gourlay et al. (2017) similarly attributes the IR to systematic measurement error in self-reported production, stating that average maize yields are over-reported by at least 85 percent relative to crop cuts.

A possible caveat

The validity of our inference that measurement error in self-reports is the source of the plot-size IR, hinges on the assumption that measurement error in self-reported yields can be correlated with plot size but measurement error in crop cuts is independent of plot size. The latter assumption seems reasonable a priori since crop cuts follow a standardized, objective design that is independent of plot size. The fact that yields are heterogeneous within a plot does not invalidate this assumption because by design enumerators do not systematically select subplots with high yields on small plots and subplots with low yields on larger plots.

There is, however, one phenomenon, the 'edge' effect recently proposed by Bevis and Barret (2016) as an explanation for the IR, which might invalidate this central assumption. The edge effect refers to the observation that yields along the plot's perimeter could be higher than those in the plot's interior because crops along the edges might face less competition for nutrients, water, space and sunlight than crops in the plot's interior. If yields at the outer bounds are indeed sufficiently higher than yields in the interior of a plot, then yields on small plots will be higher than on large plots (because the large plot's perimeter relative to its size is smaller). Bevis and Barrett (2016) show that the IR for maize in Uganda disappears when controlling for plot perimeter relative to plot size.

This edge effect could in principle explain the IR for self-reports, even if this relationship is not observed for crop cuts. To understand this point, it is sufficient to observe that crop cuts, although randomly implemented within plots, are less likely to capture yields along the plot's border because the subplots were selected to avoid the edges[23] (see Appendix B for a more comprehensive discussion). Hence, the central assumption that measurement errors in crop cuts are independent of the plot's size is violated if yields are systematically higher around its edges.

While our data do not allow us to rule out the edge effect as an explanation for the plot-size IR, a recent working paper by Gourlay et al. (2017) shows that the IR for maize in Uganda not only disappears for crop cuts (as our findings confirm), but also disappears when yields are measured by crop-cutting the entire field, rather than just a subplot. The edge effect is completely captured when the entire field is subject to crop cutting, and in this case, cannot explain the IR.[24]

Two additional observations in our data provide some further support to our inference that the plot-size IR is driven primarily by measurement error in self-reported production and not the edge effect. First, some crop cuts do partially capture yields at the plot's exterior. The smaller the plot, the more likely the subplot includes areas near the border. On average, then, it should still be possible to observe an IR for crop cuts if the edge effect is the source of the IR. Our findings provide no evidence at all to support a negative association between plot size and crop cuts and indicate if anything, a positive association. Second, we have simulated the approximate order of magnitude of the edge effect that would be required to induce the empirically observed IR for self-reports. These calculations, available on request, suggest that yields at the exterior need to be 3.5 times higher than yields in the interior if the 'edge effect' extends to 2m from the border and up to 9.5 times if the 'edge effect' only extends to 0.5m from the border.

[23] In practical terms, when the center of a randomly selected subplot (of 4m by 4m) was less than 2m away from the plot's border, a new subplot was randomly selected.

[24] As an additional piece of evidence, if crop cuts were systematically missing edges and edges have significantly higher yields, then the average estimated yield based on crop cuts should be noticeably less than average yields when estimated by harvesting the entire plot (as measured by trained technicians using scales). When examining pure-stand plots, Gourlay et al. (2017) (Table 2) do indeed note the crop cut yield estimates are lower, but the difference is only 2 percent and it is not statistically significant. When considering all plots, pure-stand and mixed-stand, crop cut estimates are about 10 percent greater than estimates based on full-plot harvest.

5. Conclusion

This paper revisits the inverse plot-size productivity relationship and presents evidence to reject it. As in previous studies, the relationship is strong if yield measurement is based on self-reported production—the traditional approach in household surveys to estimate yields. Moreover, conventional explanations, such as unobserved household and plot characteristics, do not actually explain the IR. The relationship disappears and even reverses, however, if yields are measured with crop cuts, which is a standardized, objective method to measure yields. This suggests that the inverse *plot-size* relationship is an artifact of systematic measurement error in self-reported production.

In theory, this result could also be explained by differences in labor input between large and small plots, since crop cuts measure the potential harvest (just before the harvest) and self-reports the actual harvest (after the harvest). If small plots are harvested more completely than large, an IR would be observed only for self-reported yields and not crop cuts. Although there is evidence of an inverse relation between plot size and family labor, the IR for self-reports remained significant even after controlling for labor input.

There is the possibility that self-reported yields include edges and crop cuts do not, and the higher yields at the edges contribute to the plot-size IR. Because we lack data on the shape of the plot, we cannot reject this possibility, although the observation of a positive association between plot size and crop cuts suggests that the edge effect does not drive the results. Our simulations of the required magnitude of the edge effect also lead us to doubt this as the explanation for the IR, at least based on our data from rural Ethiopia. Moreover, recent work on the IR in Uganda, measuring yields with self-reports, crop cuts (on subplots and on the entire field), and remote sensing, corroborates our findings, and also shows that the IR disappears when the entire field, rather than just a subplot, is crop-cut, which rules out the possibility that the edge effect explains the IR (Gourlay et al., 2017).

This leaves systematic measurement error in self-reported production as a remaining explanation for the plot-size IR. Suggesting that self-reported crop production in household surveys is measured with substantial error is not a particularly contested or questioned assertion. Presumably because this error is typically not considered to be systematically related to plot size, it has not been viewed as a causal source of the plot-size IR. But, the plot-size IR can readily be explained by measurement error if the error is negatively correlated with plot size. Our findings suggest that farmers substantially over-report production on small plots and under-report it on larger ones, thereby causing the IR.

Our analyses do not reveal why households systematically misreport production. Perhaps this is related to the finding that farmers seem to consistently overestimate land size on small plots and underestimate it on large ones (Carletto et al., 2015). If farmers use a simple rule of estimating crop production by multiplying their misperception of plot size by some estimate of average yields, then this would result in systematic error in self-reported production as well.[25] This candidate explanation would be consistent with findings from Tversky and Kahneman (1975) that respondents often use simple heuristics when reporting on complex issues.

Another possible explanation for systematic reporting error is heaping (i.e. the tendency to respond in multiples of certain numbers, such as 5 or 10), which often occurs in surveys. Small rounding errors are more of a concern on small plots, where total production is small, than on larger plots, where the error introduced by rounding is small relative to total production. If there is a tendency to round up in heaping, then this asymmetry between small and large plots causes less measurement error in yields, which can generate an IR.

Ultimately understanding the plot-size IR is critical to assessing the effectiveness of different agricultural methods and programs for improving farm productivity. For example, the profitability of fertilizers and new technologies or the impact of development programs directly hinges on correctly measuring land productivity. Further, rejecting the plot-size IR has two important implications for the farm-size IR[26]: (1) The hypothesis of missing markets as the main explanation for the farm-size IR had been discredited because of the presence of the plot-size IR. Missing markets cannot explain differences in the productivity of plots held by the same household (Assunção & Braido, 2007). Attributing the plot-size IR to measurement error thus reinvigorates the more conventional explanation of missing markets for the farm-size IR. From a policy perspective, this implies that reducing frictions in land, labor and credit markets will increase agricultural output. (2) Systematic measurement error in self-reported production might also generate, or at least reinforce, the farm-size IR. Although this could not be tested with our data, attributing the farm-size IR to measurement error in self-reported productivity

[25] Unfortunately, this explanation is not completely satisfactory either. It implies that the IR should disappear if self-reported rather than GPS measured plot size is used as the dependent variable. Many studies have, however, demonstrated a negative association between self-reported plot size and self-reported yields.

[26] The rejection of the plot-size relationship also implies that fragmenting land – subdividing large plots into smaller ones – will not increase production. This policy has never been recommended in practice, although it is an obvious implication of the plot-size IR.

would refute a stylized fact and challenge one of the primary arguments in favor of small-scale agriculture: Small farms might not in fact be more efficient than larger ones.

References

Aguiar, M., & Bils, M. (2015). Has consumption inequality mirrored income inequality? *American Economic Review, 105*(9), 2725-2756. doi: 10.1257/aer.20120599.

Ali, D. A., & Deininger, K. (2015). Is there a farm size–productivity relationship in African agriculture? Evidence from Rwanda. *Land Economics, 91*(2), 317-343.

Assunção, J. J., & Braido, L. H. B. (2007). Testing household-specific explanations for the inverse productivity relationship. *American Journal of Agricultural Economics, 89*(4), 980-990.

Assuncao, J. J., & Ghatak, M. (2003). Can unobserved heterogeneity in farmer ability explain the inverse relationship between farm size and productivity? *Economics Letters, 80*(2), 189-194.

Barrett, C. B. (1996). On price risk and the inverse farm size-productivity relationship. *Journal of Development Economics, 51*(2), 193-215.

Barrett, C. B., Bellemare, M. F., & Hou, J. Y. (2010). Reconsidering conventional explanations of the inverse productivity–size relationship. *World Development, 38*(1), 88-97.

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

Benjamin, D. (1995). Can unobserved land quality explain the inverse productivity relationship? *Journal of Development Economics, 46*(1), 51-84.

Benjamin, D., & Brandt, L. (2002). Property rights, labour markets, and efficiency in a transition economy: The case of rural China. *Canadian Journal of Economics/Revue canadienne d'économique, 35*(4), 689-716.

Bevis, L. E., & Barrett, C. B. (2016). *Close to the Edge: Do Behavioral Explanations Account for the Inverse Productivity Relationship?* Retrieved from http://barrett.dyson.cornell.edu/files/papers/Bevis%20and%20Barrett%20--%20Close%20to%20the%20Edge%20--%20November%202016.pdf

Bhalla, S. S., & Roy, P. (1988). Mis-specification in farm productivity analysis: The role of land quality. *Oxford Economic Papers, 40*(1), 55-73.

Blarel, B., Hazell, P., Place, F., & Quiggin, J. (1992). The economics of farm fragmentation: evidence from Ghana and Rwanda. *The World Bank Economic Review, 6*(2), 233-254.

Carletto, C., Gourlay, S., Winters, P., 2015. From Guesstimates to GPStimates: Land Area Measurement and Implications for Agricultural Analysis. *Journal of African Economies* 24, 593–628. doi:10.1093/jae/ejv011

Carletto, C., Savastano, S., & Zezza, A. (2013). Fact or artifact: The impact of measurement errors on the farm size–productivity relationship. *Journal of Development Economics, 103*(0), 254-261.

Carter, M. R. (1984). Identification of the inverse relationship between farm size and productivity: An empirical analysis of peasant agricultural production. *Oxford Economic Papers*, 131-145.

Carter, M. R., & Wiebe, K. D. (1990). Access to capital and its impact on agrarian structure and productivity in Kenya. *American Journal of Agricultural Economics, 72*(5), 1146-1150.

Chayanov, A. V. (1926). *AV Chayanov on the theory of peasant economy*. Manchester, UK: Manchester University Press.

Chen, Z., Huffman, W. E., & Rozelle, S. (2011). Inverse relationship between productivity and farm size: the case of China. *Contemporary Economic Policy, 29*(4), 580-592.

Collier, P., & Dercon, S. (2013). African agriculture in 50 years: Smallholders in a rapidly changing world? *World Development*.

Eswaran, M., & Kotwal, A. (1986). Access to capital and agrarian production organisation. *The Economic Journal, 96*(382), 482-498.

Feder, G. (1985). The relation between farm size and farm productivity: The role of family labor, supervision and credit constraints. *Journal of Development Economics, 18*(2), 297-313.

Fermont, A., & Benson, T. (2011). Estimating yield of food crops grown by smallholder farmers. Washington, DC: International Food Policy Research Institute.

Ferreira, F. H., Chen, S., Dabalen, A., Dikhanov, Y., Hamadeh, N., Jolliffe, D., . . . Sangraula, P. (2016). A global count of the extreme poor in 2012: Data issues, methodology and initial results. *The Journal of Economic Inequality, 14*(2), 141-172.

Gibson, J., & Kim, B. (2007). Measurement error in recall surveys and the relationship between household size and food demand. *American Journal of Agricultural Economics, 89*(2), 473-489. doi:10.2307/4492824.

Gourlay, S., Kilic, T., & Lobell, D. (2017). Could the debate be over? Errors in farmer-reported production and their implications for inverse scale - productivity relationship in Uganda. Working paper presented at the Center for the Study of African Economies (CSAE) Conference 2017, Oxford: United Kingdom.

Government of Ethiopia, Central Statistical Agency, and World Bank (2013). Crop-cutting Manual, Ethiopia Socioeconomic Survey Manuals. Addis Ababa. Retrieved from http://siteresources.worldbank.org/INTLSMS/Resources/3358986-1233781970982/5800988-1367841456879/9170025-1367841502220/ERSS_Crop_Cutting_Manual_English.pdf.

Hausman, J. (2001). Mismeasured variables in econometric analysis: Problems from the right and problems from the left. *The Journal of Economic Perspectives, 15*(4), 57-67.

Heltberg, R. (1998). Rural market imperfections and the farm size— productivity relationship: Evidence from Pakistan. *World Development, 26*(10), 1807-1826.

Irz, X., Lin, L., Thirtle, C., & Wiggins, S. (2001). Agricultural productivity growth and poverty alleviation. *Development Policy Review, 19*(4), 449-466.

Kagin, J., Taylor, J. E., & Yúnez-Naude, A. (2016). Inverse productivity or inverse efficiency? Evidence from Mexico. *The Journal of Development Studies, 52*(3), 396-411.

Kimhi, A. (2006). Plot size and maize productivity in Zambia: Is there an inverse relationship? *Agricultural Economics, 35*(1), 1-9.

Lamb, R. L. (2003). Inverse productivity: land quality, labor markets, and measurement error. *Journal of Development Economics, 71*(1), 71-95.

Livingston, G., Schonberger, S., and Delaney, S. "Sub-Saharan Africa: The state of smallholders in agriculture." Rome: International Fund for Agricultural Development, 2011. https://www.ifad.org/documents/10180/78d97354-8d30-466e-b75c-9406bf47779c.

Larson, D., Otsuka, K., Matsumoto, T., & Kilic, T. (2014). Should African rural development strategies depend on smallholder farms? An exploration of the inverse-productivity hypothesis. *Agricultural Economics, 45*(3), 355-367.

Pašakarnis, G., & Maliene, V. (2010). Towards sustainable rural development in Central and Eastern Europe: Applying land consolidation. *Land Use Policy, 27*(2), 545-549.

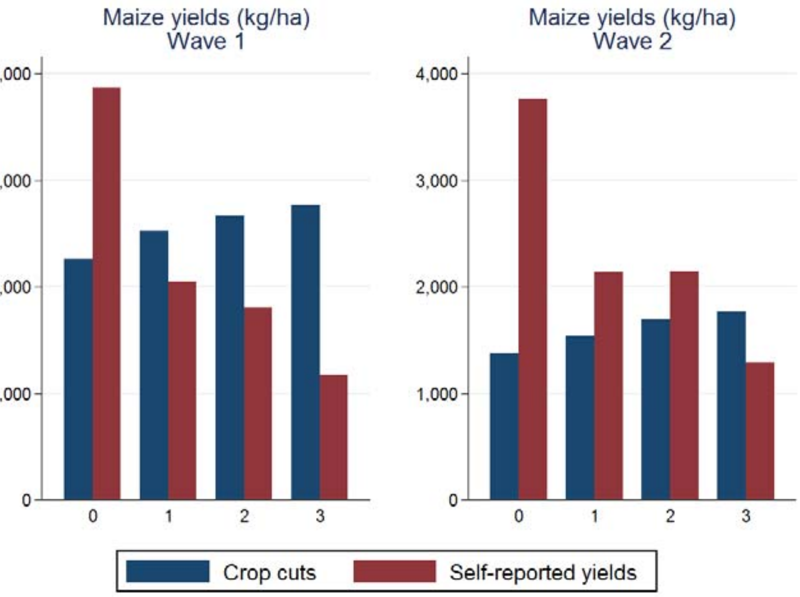
Savastano, Sara and Scandizzo, Pasquale L. (forthcoming). "Farm Size and Productivity: A "Direct-Inverse-Direct" (DID) Relationship:" Manuscript submitted for publication in Policy Research Working Paper. Washington DC: World Bank Group.

Sen, A. K. (1962). An aspect of Indian agriculture. *Economic Weekly, 14*(4-6), 243-246.

Tversky, A., Kahneman, D., 1975. Judgment under uncertainty: Heuristics and biases, in: Wendt, D., Vlek, C. (Eds.), Utility, Probability, and Human Decision Making: Selected Proceedings of an Interdisciplinary Research Conference, Rome, 3–6 September, 1973. Springer Netherlands, Dordrecht, pp. 141–162. doi:10.1007/978-94-010-1834-0_8

Verschelde, M., D'Haese, M., Rayp, G., & Vandamme, E. (2013). Challenging small-scale farming: A non-parametric analysis of the (inverse) relationship between farm productivity and farm size in Burundi. *Journal of Agricultural Economics, 64*(2), 319-342.

Figure 1: Maize-yield estimates, self-reports and crop-cuts for same plot



Notes: Vertical axes are estimated yields in kilograms per hectare, and horizontal axes are land-size quartiles. The quartiles have average plot sizes of 132m², 421m², 1,127m², and 3660m².

Table 1: Household characteristics, with and without crop-cut plots

	No crop cuts	Crop cuts	Difference (t-statistic)	
Landholdings (ha)	1.18	1.34	3.07	***
Applied chemical fertilizer (%)	47	50	1.29	
Asset index	0.16	0.17	2.00	**
Household size	5.75	5.90	1.62	
Age of household head	46.47	45.67	1.42	
Household head can read and write (%)	38	39	0.73	
Female-headed household (%)	20	17	1.50	
N (min/max)	2954/3018	878/890		

Notes: Number of observations differs by variable due to missing variables. T-statistics are based on design-adjusted standard errors corrected for clustering at the enumeration area. ***, **, * denote statistical significance at the 1%, 5% and 10% levels.

Table 2: Plot characteristics, selected and not selected for crop cutting

	No crop cut	Crop cut	Difference (t-statistic)	
Plot size (m²)	1227.30	1501.00	6.60	***
Distance from household dwelling (km)	0.68	0.95	5.96	***
Plot slope (%)	14.55	14.35	0.45	
Plot elevation	1931.35	1968.14	2.10	**
Plot Potential Wetness Index	12.58	12.55	0.52	
Land title certificate (%)	48	50	1.07	
Pure stand (%)	57	87	18.26	***
Applied manure (%)	33	21	8.25	***
Applied compost (%)	5	6	1.05	
Applied organic fertilizer (%)	2	2	0.36	
Irrigation (% of plots)	4	2	3.04	***
Self-reported soil quality (only wave 2)	1.75	1.89	6.45	***
Fertilizer (kg/ha)	44.40	46.21	0.54	
Family labor planting (days/ha)	16.89	16.13	0.99	
Hired labor planting (days/ha)	7.48	7.27	0.18	
Exchange labor planting (days/ha)	15.19	15.08	0.09	
Family labor harvesting (days/ha)	12.23	8.14	7.95	***
Hired labor harvesting (days/ha)	4.81	4.16	1.11	
Exchange labor harvesting (days/ha)	15.33	15.08	0.14	
N (min/max)	20046/21401	3296/5920		

Notes: Number of observations differs by variable due to missing variables. Soil quality was only reported in wave 2. T-statistics are based on design-adjusted standard errors corrected for clustering at the enumeration area. ***, **, * denote statistical significance at the 1%, 5% and 10% levels.

Table 3: Yield - plot size relationship, all plots

	(1)	(2)	(3)
Log plot size (m²)	-0.397***	-0.411***	-0.396***
	(0.0119)	(0.0136)	(0.0110)
Level of fixed effects			
Household	X		
Parcel		X	
Enumeration area			X
Plot inputs, wave dummies	X	x	x
Plot characteristics	X		x
Household characteristics			x
Observations	25,811	27,122	25,004
R-squared	0.194	0.169	0.201

Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% levels. Standard errors in parentheses. Errors clustered at the enumeration area. Plot inputs and wave dummies are included in all specifications, plot characteristics are included in specification (1) and (3). Specification (3) also includes household characteristics. Full results are reported in Appendix A.

Table 4: Yield - plot size relationship, plots with both self-reports and crop cuts

	Self-reports			Crop cuts		
	(1)	(2)	(3)	(4)	(5)	(6)
Log plot size (m²)	-0.287***	-0.293***	-0.303***	0.119***	0.126***	0.104***
	(0.0268)	(0.0289)	(0.0240)	(0.0154)	(0.0164)	(0.0121)
Level of fixed effects						
Household	X			X		
Parcel		X			X	
Enumeration area			X			X
Plot inputs, wave dummies	x	x	x	x	x	x
Plot characteristics	x		x	x		x
Household characteristics			x			x
Observations	5,370	5,869	5,248	5,370	5,869	5,248
R-squared	0.108	0.109	0.110	0.126	0.134	0.120

Notes: The sample is restricted to plots for which there are both self-report and crop-cut yield estimates. ***, **, * denote statistical significance at the 1%, 5% and 10% levels. Standard errors in parentheses. Errors clustered at the enumeration area. Full results are reported in Appendix A.

Table 5: Yield - plot size relationship, controlling for labor inputs

	(1)	(2)
	Self-reports	**Crop cuts**
Log plot size (m²)	-0.153***	0.164***
	(0.0295)	(0.0202)
Labor input (logs)		
Family labor planting (days/ha)	0.131***	0.0734***
	(0.0253)	(0.0182)
Hired labor planting (birr/ha)	0.0469**	0.0347**
	(0.0193)	(0.0162)
Exchange labor planting (days/ha)	0.0127	0.0175*
	(0.0129)	(0.0099)
Family labor harvest (days/ha)	0.282***	0.0714***
	(0.0290)	(0.0246)
Hired labor harvest (birr/ha)	0.117***	0.0393**
	(0.0184)	(0.0155)
Exchange labor harvest (days/ha)	0.111***	0.0306***
	(0.0135)	(0.0086)
Observations	5,181	5,181
R-squared	0.220	0.148

Notes: The sample is restricted to plots for which there are both self-report and crop-cut yield estimates. ***, **, * denote statistical significance at the 1%, 5% and 10% levels. Standard errors in parentheses. Errors clustered at the enumeration area. All specifications include household fixed effects, wave dummies and plot characteristics and inputs. Full results are reported in Appendix A.

Table 6: Labor input – plot size relationship, by season and labor type

	Planting			**Harvest**		
	Total labor	**Family labor**	**Hired labor**	**Total labor**	**Family labor**	**Hired labor**
Log plot size (m²)	-0.183***	-0.311***	0.0769***	-0.225***	-0.382***	0.0661***
	(0.0176)	(0.0141)	(0.0096)	(0.0143)	(0.0095)	(0.0080)
Observations	25,004	25,004	25,004	24,743	24,743	24,743
R-squared	0.070	0.169	0.035	0.083	0.299	0.038

Notes: The dependent variable is labor input per hectare, the sample includes all plots. ***, **, * denote statistical significance at the 1%, 5% and 10% levels. Standard errors in parentheses. Errors clustered at the enumeration area. All regressions include enumeration area fixed effects, wave dummies, plot characteristics, plot inputs and household characteristics. All dependent variables are expressed in days/m². Full results are reported in Appendix A.

Table 7: Yield - plot size relationship, crop fixed effects

	EA-Wave-Crop fixed effects		Maize	
	Self-reports	Crop cuts	Self-reports	Crop cuts
Log plot size (m²)	-0.323***	0.0778***	-0.439***	0.0595
	(0.0277)	(0.0172)	(0.0601)	(0.0397)
Observations	5,059	5,059	726	722
R-squared	0.265	0.027	0.356	0.182

Notes: The sample is restricted to plots for which there are both self-report and crop-cut yield estimates. ***, **, * denote statistical significance at the 1%, 5% and 10% levels. Standard errors in parentheses. Errors clustered at the enumeration area. All regressions include household and plot characteristics and plot and labor inputs. Specifications for the sample restricted to maize include enumeration area fixed effects. Dependent variables are expressed in Birr/m², except for maize yields, which are expressed in kg/ha. Full results are reported in Appendix A**.**

Table 8: Yield - plot size relationship, by wave and trimmed sample

	Wave 1		Wave 2		Trimmed sample (10% discarded)	
	Self-reports	Crop cuts	Self-reports	Crop cuts	Self-reports	Crop cuts
Log plot size (m²)	-0.160***	0.171***	-0.129***	0.140***	-0.100***	0.110***
	(0.044)	(0.022)	(0.029)	(0.019)	(0.020)	(0.013)
Observations	1860	1860	3,199	3,199	4664	4557
R-squared	0.198	0.101	0.246	0.092	0.168	0.133

Notes: The sample is restricted to plots for which there are both self-report and crop-cut yield estimates. The trimmed sample discards the top and bottom 5 percent of yields. ***, **, * denote statistical significance at the 1%, 5% and 10% levels. Standard errors in parentheses. Errors clustered at the enumeration area. Full results are reported in Appendix A. All regressions include household and plot characteristics, labor input and enumeration area fixed effects. The specifications dealing with outliers also include wave dummies.

Appendix A: Full results for Tables 3 to 8

Table 3: Yield - plot size relationship, all plots (full results)

	(1)	(2)	(3)
Log of plot size (m²)	-0.397***	-0.411***	-0.396***
	(0.0119)	(0.0136)	(0.0110)
Wave	-0.103*	-0.103**	-0.0978*
	(0.0540)	(0.0510)	(0.0550)
Household characteristics			
Asset index			0.682***
			(0.162)
Female headed household (no=1)			0.0801***
			(0.0256)
Age of household head			0.000379
			(0.00337)
Age²			2.78e-06
			(3.32e-05)
Household head can read and write			0.0503**
			(0.0210)
Plot characteristics			
Log distance to dwelling	0.133***		0.110***
	(0.0333)		(0.0257)
Plot slope (percent)	-0.00238		0.000461
	(0.00148)		(0.00120)
Plot elevation (m)	-0.000246*		-0.000198*
	(0.000138)		(0.000108)
Plot Potential Wetness Index	-0.000669		-0.000929
	(0.00591)		(0.00500)
Household has land title (no=1)	0.0105		0.0190
	(0.0407)		(0.0257)
Plot inputs			
Crop in pure stand (no=1)	-0.304***	-0.342***	-0.325***
	(0.0921)	(0.0890)	(0.0974)
Manure applied (no=1)	0.0456	0.112***	0.0191
	(0.0347)	(0.0381)	(0.0281)
Compost applied (no=1)	0.0420	0.0552	0.0196
	(0.0537)	(0.0687)	(0.0410)
Organic fertilizer (no=1)	-0.0266	-0.137	0.0663
	(0.0901)	(0.101)	(0.0819)
Field irrigated (no=1)	-0.272***	-0.222*	-0.202**
	(0.0825)	(0.114)	(0.0938)
Log fertilizer (kg/ha)	0.0838***	0.0753***	0.0779***
	(0.00693)	(0.00662)	(0.00643)
Pure stand (no=1)*Wave 2	0.619***	0.726***	0.619***
	(0.0992)	(0.0956)	(0.105)
Constant	12.64***	12.20***	11.93***
	(0.442)	(0.591)	(0.425)
Observations	25,811	27,122	25,004
R-squared	0.194	0.169	0.201

Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% levels. Standard errors in parentheses. Errors clustered at the enumeration area. Household fixed effects, parcel fixed effects and enumeration area fixed effects are included in specification 1, 2 and 3, respectively.

Table 4: Yield - plot size relationship, plots with both self-reports and crop cuts (full results)

	Self-reports			Crop cuts		
	(1)	(2)	(3)	(4)	(5)	(6)
Log of plot size (m²)	-0.287***	-0.293***	-0.303***	0.119***	0.126***	0.104***
	(0.0268)	(0.0289)	(0.0240)	(0.0154)	(0.0164)	(0.0121)
Wave	-0.135**	-0.0917	-0.101*	-0.444***	-0.475***	-0.427***
	(0.0549)	(0.0734)	(0.0539)	(0.0493)	(0.0522)	(0.0453)
Household characteristics						
Asset index			0.591***			0.319*
			(0.202)			(0.186)
Female headed household (no=1)			0.0533			0.00490
			(0.0520)			(0.0351)
Age of household head			-0.00409			-0.00208
			(0.00632)			(0.00495)
Age²			4.61e-05			3.32e-05
			(5.96e-05)			(4.95e-05)
Household head can read and write (no=1)			0.00557			-0.0270
			(0.0382)			(0.0283)
Plot characteristics						
Log distance to dwelling	0.131*		0.150***	-0.0772*		-0.0188
	(0.0673)		(0.0459)	(0.0444)		(0.0291)
Plot slope (percent)	0.000529		0.00202	-0.00149		-0.000340
	(0.00322)		(0.00230)	(0.00213)		(0.00142)
Plot elevation (m)	-0.000606***		-0.000262	-0.000571***		-0.000328***
	(0.000190)		(0.000160)	(0.000156)		(0.000106)
Plot Potential Wetness Index	0.0183		-0.00625	-0.0134		-0.00882
	(0.0158)		(0.0103)	(0.00898)		(0.00806)
Household has land title (no=1)	-0.00700		0.0193	0.0505		0.0436*
	(0.0676)		(0.0417)	(0.0461)		(0.0258)
Plot inputs						
Crop in pure stand (no=1)	-0.271***	-0.500***	-0.308***	-0.0974	-0.223**	-0.0871
	(0.0919)	(0.154)	(0.0911)	(0.0845)	(0.110)	(0.0797)
Manure applied (no=1)	-0.0444	-0.0413	-0.0129	-0.122***	-0.0882*	-0.0677*
	(0.0618)	(0.0893)	(0.0501)	(0.0456)	(0.0512)	(0.0358)
Compost applied (no=1)	-0.0710	0.0502	-0.116	0.00199	-0.0323	-0.0455
	(0.0931)	(0.107)	(0.0769)	(0.0797)	(0.0978)	(0.0670)
Organic fertilizer (no=1)	-0.234	-0.388**	-0.0841	-0.225	-0.190	-0.254**
	(0.187)	(0.180)	(0.150)	(0.154)	(0.159)	(0.112)
Field irrigated (no=1)	-0.403**	-0.127	-0.290***	0.0201	0.0618	0.0340
	(0.164)	(0.244)	(0.105)	(0.119)	(0.175)	(0.0812)
Log fertilizer (kg/ha)	0.0948***	0.0842***	0.0840***	0.0380***	0.0343***	0.0421***
	(0.0101)	(0.0127)	(0.00844)	(0.00882)	(0.00851)	(0.00679)
Pure stand (no=1)*Wave 2	0.454***	0.606***	0.457***	0.129	0.283*	0.0445
	(0.134)	(0.168)	(0.133)	(0.121)	(0.151)	(0.113)
Constant	13.29***	11.54***	12.34***	10.61***	8.653***	9.700***
	(0.832)	(0.582)	(0.620)	(0.578)	(0.472)	(0.413)
Observations	5,370	5,869	5,248	5,370	5,869	5,248
R-squared	0.108	0.109	0.110	0.126	0.134	0.120

Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% levels. Standard errors in parentheses. Errors clustered at the enumeration area. Specification (1) and (4), (2) and (5), (3) and (6) include household, parcel, and enumeration area fixed effects, respectively.

Table 5: Yield - plot size relationship, controlling for labor inputs (full results)

	(1)	(2)
	Self-reports	Crop cuts
Log of plot size (m²)	-0.153***	0.164***
	(0.0295)	(0.0202)
Wave	-0.128**	-0.421***
	(0.0581)	(0.0518)
Plot characteristics		
Log distance to dwelling	0.0846	-0.0914**
	(0.0585)	(0.0397)
Plot slope (percent)	0.000680	-0.00115
	(0.00285)	(0.00210)
Plot elevation (m)	-0.000318**	-0.000468***
	(0.000156)	(0.000154)
Plot Potential Wetness Index	0.0103	-0.0152*
	(0.0123)	(0.00819)
Household has land title (no=1)	-0.0202	0.0581
	(0.0637)	(0.0466)
Plot inputs		
Crop in pure stand (no=1)	-0.283***	-0.105
	(0.0935)	(0.0872)
Manure applied (no=1)	-0.00388	-0.112**
	(0.0587)	(0.0452)
Compost applied (no=1)	-0.0547	0.0348
	(0.0938)	(0.0802)
Organic fertilizer (no=1)	-0.245	-0.227
	(0.167)	(0.148)
Field irrigated (no=1)	-0.356**	0.0911
	(0.171)	(0.120)
Log fertilizer (kg/ha)	0.0588***	0.0232**
	(0.00971)	(0.00910)
Pure stand (no=1)*Wave 2	0.459***	0.142
	(0.129)	(0.118)
Labor input (logs)		
Family labor planting (days/ha)	0.131***	0.0734***
	(0.0253)	(0.0182)
Hired labor planting (birr/ha)	0.0469**	0.0347**
	(0.0193)	(0.0162)
Exchange labor planting	0.0127	0.0175*
	(0.0129)	(0.00988)
Family labor harvest (days/ha)	0.282***	0.0714***
	(0.0290)	(0.0246)
Hired labor harvest (birr/ha)	0.117***	0.0393**
	(0.0184)	(0.0155)
Exchange labor harvest (days/ha)	0.111***	0.0306***
	(0.0135)	(0.00858)
Constant	10.93***	9.556***
	(0.745)	(0.607)
Observations	5,181	5,181
R-squared	0.220	0.148

Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% levels. Standard errors in parentheses. Errors clustered at the enumeration area. Both specifications include household fixed effects.

Table 6: Labor input – plot size relationship, by season and labor type (full results)

	Planting			Harvest		
	Total Labor	Family Labor	Hired	Total Labor	Family Labor	Hired Labor
Log of plot size (m²)	-0.183***	-0.311***	0.0769***	-0.225***	-0.382***	0.0661***
	(0.0176)	(0.0141)	(0.00960)	(0.0143)	(0.00950)	(0.00797)
Wave	-0.239***	-0.153**	0.0875**	0.102	0.122*	0.0476
	(0.0799)	(0.0723)	(0.0403)	(0.0698)	(0.0630)	(0.0309)
Plot characteristics						
Log distance to dwelling	0.0676	-0.00727	0.0498	0.184***	0.0550**	0.0358
	(0.0432)	(0.0342)	(0.0313)	(0.0359)	(0.0247)	(0.0273)
Plot slope (percent)	0.000625	0.000395	-0.000670	-0.000481	-0.00185	0.00113
	(0.00185)	(0.00144)	(0.00118)	(0.00154)	(0.00117)	(0.00125)
Plot elevation (m)	-0.000141	-9.34e-05	7.97e-05	3.05e-05	-2.39e-05	0.000112
	(0.000154)	(0.000116)	(0.000102)	(0.000139)	(8.34e-05)	(7.90e-05)
Plot Potential Wetness	0.0106	0.00470	-8.57e-05	-0.000185	-0.000877	0.000700
	(0.00800)	(0.00638)	(0.00607)	(0.00628)	(0.00437)	(0.00517)
Household has land title (no=1)	0.0126	0.0160	-0.0737**	-0.00829	-0.00814	-0.0131
	(0.0394)	(0.0326)	(0.0320)	(0.0339)	(0.0229)	(0.0269)
Plot inputs						
Crop in pure stand (no=1)	0.149	0.311***	0.0426	-0.177	0.00280	0.0314
	(0.113)	(0.115)	(0.0489)	(0.126)	(0.0937)	(0.0581)
Manure applied (no=1)	-0.171***	-0.191***	-0.00558	-0.00402	-0.0582**	0.0319
	(0.0463)	(0.0364)	(0.0303)	(0.0348)	(0.0259)	(0.0201)
Compost applied (no=1)	-0.199***	-0.155***	0.00302	0.0693	0.0689*	-0.0556
	(0.0645)	(0.0568)	(0.0335)	(0.0603)	(0.0359)	(0.0605)
Organic fertilizer (no=1)	0.0820	-0.0430	0.0790	-0.0611	-0.0956	-0.0142
	(0.111)	(0.0856)	(0.101)	(0.0913)	(0.0782)	(0.0602)
Field irrigated (no=1)	-5.31e-06	-0.0444	-0.161*	-0.00700	-0.00977	-0.00528
	(0.0980)	(0.0978)	(0.0820)	(0.0758)	(0.0657)	(0.0369)
Log fertilizer (kg/ha)	0.135***	0.0993***	0.0463***	0.0838***	0.0264***	0.0410***
	(0.00930)	(0.00657)	(0.00887)	(0.00878)	(0.00547)	(0.00727)
Pure stand (no=1)*Wave 2	-0.178	-0.240**	-0.0706	-0.159	-0.165*	-0.0740
	(0.114)	(0.118)	(0.0571)	(0.131)	(0.0970)	(0.0649)
Household characteristics						
Asset index	-0.369	-0.177	0.992***	0.903***	0.126	1.279***
	(0.265)	(0.263)	(0.222)	(0.265)	(0.234)	(0.234)
Female headed household	-0.389***	0.185***	-0.151***	-0.253***	0.118***	-0.103***
	(0.0493)	(0.0484)	(0.0441)	(0.0395)	(0.0266)	(0.0306)
Age of household head	-0.0148***	0.0220***	-0.0181***	-0.0129**	0.0249***	-0.0123***
	(0.00541)	(0.00540)	(0.00493)	(0.00603)	(0.00413)	(0.00445)
Age²	0.000183***	-0.000219***	0.000175***	0.000157**	-0.000238***	0.000121***
	(5.28e-05)	(5.71e-05)	(4.97e-05)	(6.12e-05)	(4.07e-05)	(4.35e-05)
Household head can read and write (no=1)	-0.0646*	0.0402	-0.0465*	0.0391	0.0598***	-0.0565**
	(0.0367)	(0.0309)	(0.0247)	(0.0345)	(0.0226)	(0.0252)
Constant	5.627***	3.873***	0.271	4.259***	3.365***	-0.127
	(0.524)	(0.441)	(0.367)	(0.511)	(0.362)	(0.282)
Observations	25,004	25,004	25,004	24,743	24,743	24,743
R-squared	0.070	0.169	0.035	0.083	0.299	0.038

Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% levels. Standard errors in parentheses. Errors clustered at the enumeration area. Enumeration area fixed effects are included in all regressions.

Table 7: Yield - plot size relationship, crop fixed effects (full results)

	EA-Wave-Crop fixed effects		Maize	
	Self-reports	Crop cuts	Self-reports	Crop cuts
Log of plot size (m²)	-0.323***	0.0778***	-0.439***	0.0595
	(0.0277)	(0.0172)	(0.0601)	(0.0397)
Wave			-0.0609	-0.159
			(0.120)	(0.248)
Household characteristics				
Asset index	0.495**	-0.0481	-0.499	-0.144
	(0.235)	(0.194)	(0.695)	(0.385)
Female headed household (no=1)	0.0833*	0.0376	0.0584	0.0436
	(0.0465)	(0.0292)	(0.109)	(0.0777)
Age household head	-0.00964	0.00187	0.00300	0.00159
	(0.00693)	(0.00451)	(0.0150)	(0.00915)
Age²	9.95e-05	-1.60e-05	2.18e-05	-3.83e-06
	(6.66e-05)	(4.52e-05)	(0.000148)	(8.95e-05)
Household head can read and write (no=1)	0.0196	-0.0217	-0.0569	0.0340
	(0.0364)	(0.0243)	(0.0917)	(0.0684)
Plot characteristics				
Log distance to dwelling	0.0522	0.000392	0.0471	-0.136**
	(0.0423)	(0.0305)	(0.108)	(0.0688)
Plot sSlope (percent)	-0.00177	-0.000639	-0.000374	-0.00299
	(0.00194)	(0.00127)	(0.00577)	(0.00386)
Plot Eelevation (m)	6.51e-05	-0.000116	0.000635**	3.69e-05
	(0.000158)	(0.000134)	(0.000286)	(0.000238)
Plot Potential Wetness Index	-0.0115	-0.00910	-0.0137	0.00202
	(0.00810)	(0.00724)	(0.0247)	(0.0197)
Household has land title (no=1)	0.0647	0.0138	-0.000884	-0.0473
	(0.0478)	(0.0257)	(0.110)	(0.0630)
Plot inputs				
Crop in pure stand (no=1)	-0.0802	-0.192**	-0.415*	-0.0750
	(0.114)	(0.0921)	(0.222)	(0.197)
Manure applied (no=1)	-0.0554	-0.00869	-0.0413	-0.0382
	(0.0442)	(0.0308)	(0.107)	(0.0660)
Compost applied (no=1)	-0.0558	-0.0404	0.247	0.219**
	(0.0825)	(0.0450)	(0.169)	(0.103)
Organic fertilizer (no=1)	-0.0136	-0.196	-0.398**	-0.159
	(0.160)	(0.131)	(0.198)	(0.164)
Field irrigated (no=1)	-0.132	0.0940	-0.291	-0.00751
	(0.139)	(0.0988)	(0.337)	(0.227)
Log fertilizer (kg/ha)	0.0292***	0.00158	0.0335	0.00282
	(0.0108)	(0.00585)	(0.0234)	(0.0141)
Pure stand (no=1)*Wave 2	0.363**	0.148	0.590**	0.0259
	(0.156)	(0.112)	(0.245)	(0.0268)
Labor input (logs)				
Family labor planting (days/ha)	0.103***	0.0192	-0.000233	0.0432***
	(0.0213)	(0.0140)	(0.0324)	(0.0156)
Hired labor planting (birr/ha)	0.0233	0.0183**	-0.0256	-0.0108
	(0.0147)	(0.00899)	(0.0258)	(0.0194)
Exchange labor planting (days/ha)	-0.0117	-0.00311	0.241***	0.0226
	(0.0111)	(0.00720)	(0.0518)	(0.0334)
Family labor harvest (days/ha)	0.204***	0.0333**	0.0547	0.0132
	(0.0237)	(0.0158)	(0.0668)	(0.0294)

Hired labor harvest (birr/ha)	0.0560***	-0.00112	0.0827***	0.0344**
	(0.0158)	(0.00898)	(0.0307)	(0.0167)
Exchange labor harvest (days/ha)	0.0695***	0.0253***	-0.0609	-0.456***
	(0.0109)	(0.00636)	(0.120)	(0.0688)
Constant	10.53***	8.857***	9.812***	7.446***
	(0.626)	(0.448)	(1.049)	(0.829)
Observations	5,059	5,059	726	722
R-squared	0.265	0.027	0.356	0.182

Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% levels. Standard errors in parentheses. Errors clustered at the enumeration area. Specifications (3) and (4) include enumeration area fixed effects.

Table 8: Yield - plot size relationship, by wave and trimmed sample (full results)

	Wave 1		Wave 2		Discarding outliers	
	Self-reports	Crop cuts	Self-reports	Crop cuts	Self-reports	Crop cuts
Log of plot size (m²)	-0.160***	0.171***	-0.129***	0.140***	-0.1000***	0.110***
	(0.0435)	(0.0220)	(0.0294)	(0.0191)	(0.0204)	(0.0133)
Wave					-0.0707	-0.337***
					(0.0460)	(0.0385)
Household characteristics						
Asset index	0.250	0.0880	0.825**	0.0732	0.104	0.0933
	(0.334)	(0.227)	(0.361)	(0.292)	(0.163)	(0.170)
Female headed household (no=1)	-0.0172	-0.00957	0.0650	0.00607	0.0315	0.0175
	(0.0738)	(0.0560)	(0.0537)	(0.0352)	(0.0369)	(0.0241)
Age household head	-0.00419	-0.00150	-0.0130	0.00260	-0.00985*	-0.00450
	(0.00859)	(0.00675)	(0.00809)	(0.00519)	(0.00512)	(0.00370)
Age²	5.83e-05	1.76e-05	0.000113	-1.11e-05	9.32e-05*	4.66e-05
	(7.96e-05)	(6.62e-05)	(7.96e-05)	(5.23e-05)	(4.98e-05)	(3.69e-05)
Household head can read and write (no=1)	0.0331	-0.0477	-0.0340	-0.0344	0.00537	-0.00649
	(0.0536)	(0.0409)	(0.0434)	(0.0313)	(0.0292)	(0.0196)
Plot characteristics						
Log distance to dwelling	0.0390	0.00130	0.0994**	-0.0175	0.0623*	-0.0295
	(0.0680)	(0.0426)	(0.0457)	(0.0355)	(0.0360)	(0.0230)
Plot Slope (percent)	-0.00335	-0.00131	0.00489**	0.000791	0.00220	-4.10e-06
	(0.00361)	(0.00233)	(0.00225)	(0.00150)	(0.00158)	(0.00108)
Plot Elevation (m)	-0.000207	-0.000140	-0.000204	-0.000330**	-2.40e-05	-0.000255***
	(0.000207)	(0.000147)	(0.000171)	(0.000135)	(0.000112)	(8.88e-05)
Plot Potential Wetness Index	-0.00922	-0.0224**	-0.00188	-0.00293	-0.00148	-0.00655
	(0.0131)	(0.0100)	(0.0114)	(0.00982)	(0.00697)	(0.00552)
Household has land title (no=1)	0.0198	0.00973	0.0504	0.0784**	0.00568	0.000353
	(0.0672)	(0.0415)	(0.0513)	(0.0307)	(0.0335)	(0.0208)
Plot inputs						
Crop in pure stand (no=1)	-0.322***	-0.137*	0.165**	-0.0413	-0.169**	-0.0793
	(0.0956)	(0.0699)	(0.0698)	(0.0597)	(0.0698)	(0.0625)
Manure applied (no=1)	0.0933	-0.0719	-0.0487	-0.0131	0.0172	-0.0258
	(0.0783)	(0.0619)	(0.0533)	(0.0384)	(0.0376)	(0.0314)
Compost applied (no=1)	-0.0658	-0.0818	-0.0727	0.0201	-0.125**	-0.0434
	(0.135)	(0.0968)	(0.0856)	(0.0571)	(0.0543)	(0.0601)
Organic fertilizer (no=1)	-0.0433	-0.238**	-0.0748	-0.253	-0.114	-0.182*
	(0.126)	(0.105)	(0.390)	(0.245)	(0.0981)	(0.101)
Field irrigated (no=1)	-0.0136	0.228**	-0.322**	-0.0709	-0.208**	0.0958
	(0.148)	(0.113)	(0.155)	(0.0998)	(0.0857)	(0.0797)
Log fertilizer (kg/ha)	0.0651***	0.0178	0.0326***	0.0324***	0.0375***	0.0190***
	(0.0140)	(0.0125)	(0.00856)	(0.00626)	(0.00676)	(0.00515)
Pure stand (no=1)*Wave 2					0.282***	0.0316
					(0.0996)	(0.0850)
Labor inputs (logs)						
Family labor planting (days/ha)	0.104***	0.0545**	0.176***	0.0925***	0.0965***	0.0659***
	(0.0337)	(0.0218)	(0.0220)	(0.0169)	(0.0170)	(0.0115)
Hired labor planting (birr/ha)	0.0404*	0.0243	0.0460***	0.0234**	0.0240**	0.0176*
	(0.0243)	(0.0201)	(0.0159)	(0.0110)	(0.0109)	(0.00910)

Exchange labor planting (days/ha)	0.0293*	0.0200	-0.0127	-0.000202	0.00940	0.0117**
	(0.0152)	(0.0132)	(0.0136)	(0.00921)	(0.00857)	(0.00554)
Family labor harvest (days/ha)	0.268***	0.0769***	0.351***	0.0684***	0.218***	0.0389***
	(0.0432)	(0.0244)	(0.0282)	(0.0183)	(0.0191)	(0.0148)
Hired labor harvest (birr/ha)	0.127***	0.0389**	0.0794***	0.0110	0.0693***	0.0130
	(0.0292)	(0.0178)	(0.0219)	(0.0149)	(0.0138)	(0.0107)
Exchange labor harvest (days/ha)	0.0914***	0.0295***	0.110***	0.0430***	0.0769***	0.0278***
	(0.0160)	(0.00980)	(0.0128)	(0.00807)	(0.00837)	(0.00563)
Constant	9.749***	8.631***	9.550***	8.446***	9.874***	9.363***
	(0.835)	(0.570)	(0.959)	(0.648)	(0.411)	(0.382)
Observations	1,860	1,860	3,199	3,199	4,664	4,557
R-squared	0.198	0.101	0.246	0.092	0.168	0.133

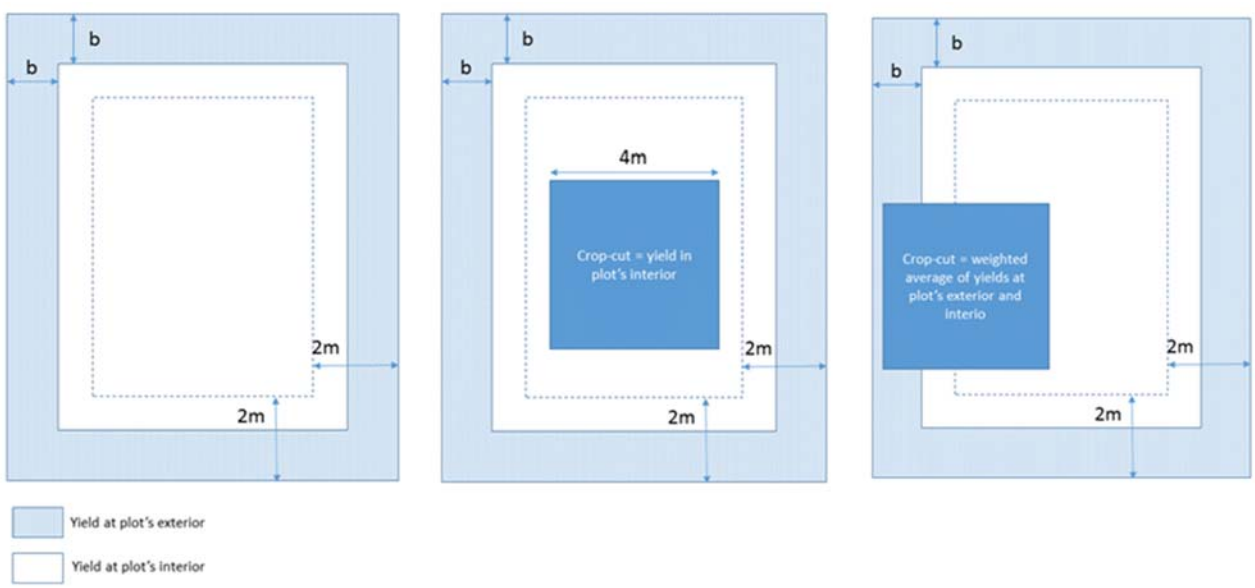
Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% levels. Standard errors in parentheses. Errors clustered at the enumeration area. Enumeration area fixed effects are included in all regressions.

Appendix B: A note on the crop cuts and the edge effect

Crop cuts were implemented in wave 2 of the survey on randomly selected subplots (4m by 4m). The only condition for selecting a subplot was that its center was at least 2m from the border of the plot to avoid subplots parts of which were located outside the border of the plot. This condition is indicated by the dashed line in Figure B.1 (left panel) for a subplot selected in a rectangular plot. The figure also depicts the edge effect as the blue area (with a width of b) along the plot's border. Yields in this area are presumed higher than in the interior.

Because subplots on which crop cuts are implemented always fall inside the border, most randomly selected subplots only measure yields at the interior (Figure B.1, middle panel), although some subplots partially capture yields at the exterior (Figure B.1, right panel).

Figure B.1: Crop cut on a plot with edge effects



Notes: The width, b, delineates the area where the edge effect is important. The center of the subplot (a 4m by 4m square) selected for crop cutting is always situated in the area indicated by the dashed line.

Because subplots are randomly selected, most crop cuts only measure yields at the interior, though some are a weighted average of exterior and interior yields. The average 'expected' yields of the crop cuts can be calculated as follows. The probability of selecting a subplot that only measures yields in the interior equals the ratio of the total area where one can select the center point of a subplot excluding the border area versus the total area where subplots can be crop cut. This area is presented in Figure B.2. Hence, the probability of selecting a subplot in the interior equals:

$$P = P(subplot\ in\ interior = 1) = \frac{(W - 2 * (2 + b))(L - 2 * (2 + b))}{(W - 4)(L - 4)}$$

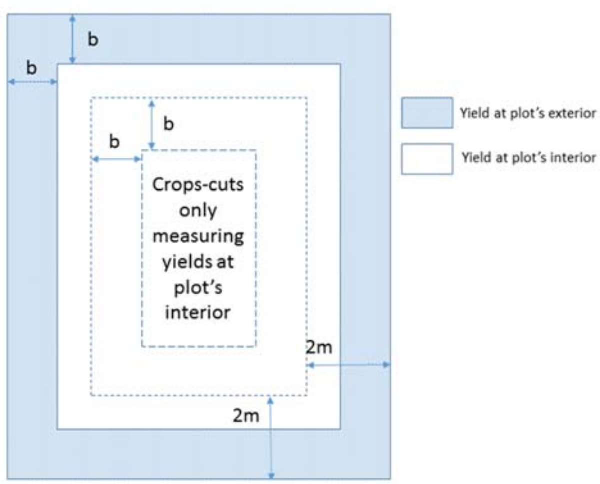
where W is the width and L the length of the plot. The probability of selecting a subplot that partially captures yields around the plot's edges then equals:

$$P(subplot\ in\ interior = 0) = 1 - P = \frac{2b(W + L - 8) - 4b^2}{(W - 4)(L - 4)}$$

Once such a subplot is selected, the share of the area of the subplot that actually measures yields at plot's exterior equals[27]:

$$Share\ of\ subplot\ in\ the\ plot's\ exterior = \frac{1}{16}\int_0^b 2(b - y)dy = \frac{b^2}{16}$$

Figure B.2: Area with crop cuts that do not touch the border



[27] Note that we always assume that the border width, b, is smaller than 2m, i.e., one-half of the subplot's side.

Using these equations, the expected yield measured by a randomly selected crop cut equals the sum of the probability of selecting a subplot in the interior times the yields in the interior and the probability of selecting a subplot that captures the border area times a weighted average of the yields in the interior (Y^I) and exterior (Y^E). Simplifying this expression:

$$E(yields^{CC}) = Y^I + \frac{b^2}{16} * \; P(subplot\ in\ interior = 0)(Y^E - Y^I)$$

This equation reveals two important points: (1) The observed crop cuts are also a weighted averaged of yields at the plot's interior and at the exterior. Since the probability of selecting a subplot that measures yield along the edge decreases with plot size, one also expects an IR between plot size and crop cuts. (2) The IR between plot size and crop cuts is stronger if the border, b, widens.

To get a sense of the magnitude of the IR between plot size and crop cuts caused by edge effects, we developed a simulation model that is based on two assumptions: (1) All plots are squares with sides ranging from 10m to 40m. (2) Yields at the plot's interior are 1,623 kg/ha and at the exterior 5,558 kg/ha—i.e., yields at the exterior are 3.4 times higher than in the interior. We then simulated expected yields for three border widths ($b = 0.5, 1.0,$ and $2.0m$).

Figure B.3 shows the results of the simulations: Edge effects do generate an IR relation between plot size and crop cuts, but the IR weakens as the border width narrows. When the border width is assumed to be 2m, if plot size doubles, yields decrease by 9 percent, while they decrease by less than 1.6 percent if the border width is smaller than 1m (Table B.1).

The theoretical and the simulation models both show that edge effects generate an IR between plot size and crop cuts. The IR is weaker when the border area, where yields are higher, is smaller. In our data, there is no IR between plot size and crop cuts. Rather, we even observed a weakly positive association between plot size and crop cuts. This implies that either yields are not significantly higher around the edges or the area at the plot's border with higher yields is very small (i.e., b is small).

Figure B.3. Simulated IR for crop cuts caused by the 'edge effect', by border width

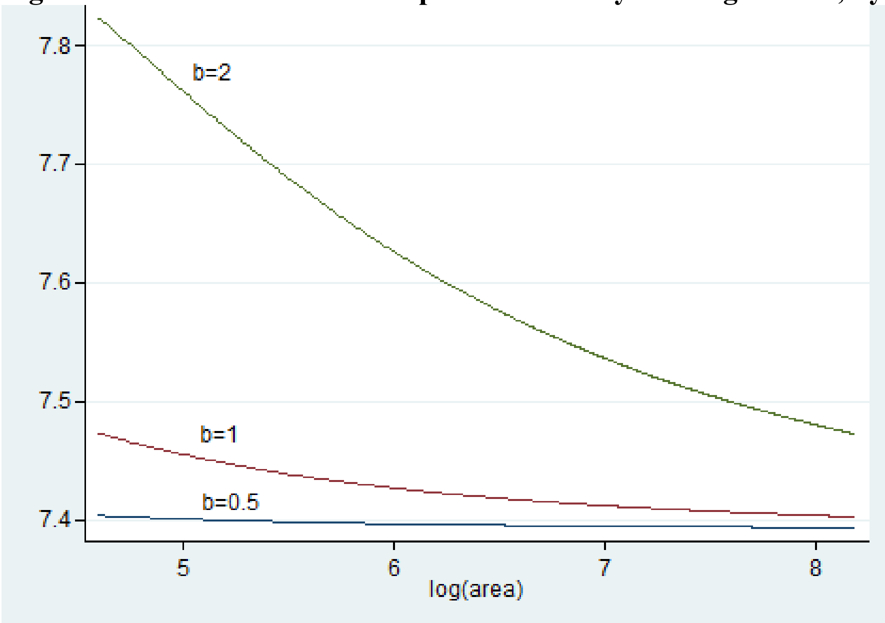


Table B1: Simulated IR for crop cuts, by border width

Border width (m)	IR
0.5	-0.002
1	-0.016
2	-0.089