Norwegian University of Life Sciences (UMB)

Can Area Measurement Error Explain the Inverse Farm Size Productivity Relationship?

Stein T. Holden and Monica Fisher

CLTSUMB

Centre for Land Tenure Studies Working Paper 12/13

Can Area Measurement Error Explain the Inverse Farm Size Productivity Relationship?

Stein T. Holden [a] **and Monica Fisher** [b]

[a] School of Economics and Business/Centre for Land Tenure Studies
Norwegian University of Life Sciences
Ås, Norway
Tel:+47-64965699
Fax:+47-64965701
Email: stein.holden@umb.no

[b] International Maize and Wheat Improvement Center – Ethiopia Office,
c/o ILRI Sholla Campus,
P.O. Box 5689,
Addis Ababa, Ethiopia.
Tel: +251-1-646-2324.
Fax: +251-1-617-2001.
Email: m.fisher@cgiar.org

Can Area Measurement Error Explain the Inverse Farm Size Productivity Relationship?

Abstract

The existence of an inverse relationship (IR) between farm size and productivity in tropical agriculture remains a debated issue with policy relevance. Poor agricultural statistical data, including data on farm sizes and farm plot sizes that typically are self-reported by farmers, can lead to biased results and wrong policy conclusions. This study combines self-reported and GPS-measured farm plot and farm sizes to assess how measurement error affects the IR using three rounds of farm plot and household data from Malawi. The results show that measurement error covers up more than 60% of the IR for the total sample but leads to an upward bias in the IR on farms less than one ha. Land and labor market imperfections in combination with food self-sufficiency motives appear to explain most of the IR and lead to a strong IR on farms below one ha.

JEL classification: O13; J43; Q12.

Key words: Inverse farm size – productivity relationship; Measurement error; Land and labor market imperfections; Land quality; Malawi.

Can Area Measurement Error Explain the Inverse Farm Size Productivity Relationship?

1. Introduction

Economic data for developing countries, including data on agricultural production, are known to be inaccurate, but the degree of inaccuracy is not well understood. The paucity of accurate economic data for developing countries is not just an academic or technical problem; policy decisions based on poor quality data can have a negative impact on the well-being of the population (Jerven, 2013a, b). The relationship between farm size and agricultural land productivity is an example of this problem. The relationship between these two variables has so frequently been reported to be negative (Barrett et al., 2010; Heltberg, 1998; Lamb, 2003) that an inverse relationship (IR) is almost perceived as stylized fact. The assumed IR can then be used by agricultural policymakers to justify the redistribution of land from large to small farmers on the basis of efficiency, as well as equity. However, a few studies have found either no association or a positive relationship between farm size and land productivity (Dorward, 1999; Kevane, 1996; Zaibet and Dunn, 1998). These results suggest that small farms are less efficient, and could lead to agricultural development policies that emphasize large farms. Thus, there is a clear need to accurately determine the relationship between farm size and land productivity to inform the design of agricultural policies that are effective at stimulating agricultural productivity.

The inconsistent reports on the relationship between farm size and productivity might partly reflect site-specific and temporal differences among studies. The inconsistency might also be due to the varying quality of the empirical data used for the analyses. Measurements of farm size are particularly subject to measurement errors in large household surveys, such as the Living

Standards Measurement Study (LSMS), where the standard approach is to ask farmers to estimate the area of their farm plots.[1] This approach is simple and inexpensive, but often inaccurate in developing countries, where farmers have limited understanding of the purposes of surveys and sometimes refuse to respond or provide untrue responses, fearing the information might be used to increase their taxes. Moreover, where land markets are not well developed, farmers are often unfamiliar with area measurement units. Finally, farmers in low-income areas typically have little formal or agricultural education, and rarely possess the measurement and quantitative skills needed to accurately estimate land area (De Groote & Traoré, 2005). For these reasons, estimates of farm size obtained through farmers' estimates in developing countries are considered to be inaccurate.

This paper examines how errors in estimation of farm and plot sizes affect estimates of the IR between farm size and agricultural land productivity. We also examine how errors in area measurement influence tests of alternative hypotheses explaining the IR, such as imperfections in land and labor markets, land quality, price risk, and food security. A three-year panel dataset for Malawi is used to test the following hypotheses:

H1a. Farm size measurement error creates an IR between farm size and net return per unit land, as suggested by Lamb (2003) and Barrett et al. (2010).

H1b. The error is particularly prevalent in models with household fixed effects (Lamb 2003).

H2. Differences in land quality between smaller and larger farms partly explain the IR.

H3. Market imperfections in land and labor markets partly explain the IR.

[1] Exceptions to this approach are the longitudinal Village Level Studies (VLS) conducted by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) in India, where farm plots were measured with tape measure, and the recent LSMS Integrated Surveys on Agriculture (ISA) in Ethiopia, Malawi, Nigeria, Niger, and Tanzania, which used Global Position System (GPS) devices to measure farm and plot sizes.

H4. The IR is driven by covariate risk and food self-sufficiency objectives on small farms (Barrett 1996).

H5. Area measurement error biases the tests for H2-H4.

The present paper provides the most comprehensive assessment to date of the impact of area measurement error on estimation of the IR. The study uses a dataset that includes a household panel with plot observations, self-reported and GPS-measured farm and plot sizes, and a comprehensive set of variables related to land quality, market imperfections, and food self-sufficiency. These data allow us to assess, for the first time, the influences of error in area measurement on tests of the effects of land quality, market imperfections, and small farm size/self-sufficiency objectives on the IR. Results of our study and a few others (Carletto et al., 2011; Tatwangire and Holden, 2013) indicate that farmer-reported farm sizes should not be relied upon in national farm surveys to provide accurate estimates of land productivity.

2. Literature review

The basic inverse relationship (IR) between farm size and land productivity can be summarized by regression equation 1:

$$\frac{y_i}{A_i} = \alpha_1 + \alpha_2 X_i + \alpha_3 A_i + \mu_i, \quad (1)$$

where y is a measure of agricultural productivity, typically yield or net agricultural revenue; A is farm size; X is a vector of control variables that are assumed to influence productivity; and i indexes households. An IR is considered to exist when the estimated α_3 is negative and statistically significant. Results of a large number of studies across many countries support the IR hypothesis (Assunção & Braido, 2007; Barrett, 1996; Barrett et al., 2010; Benjamin, 1995; Berry and Cline, 1979; Bhalla and Roy, 1988; Carletto et al., 2011; Carter, 1984; Cornia, 1985; Eswaran and Kotwal, 1986; Heltberg, 1998; Kimhi, 2006; Lamb, 2003; Tatwangire and Holden, 2013); results of only a few studies did not find the IR to be valid (Dorward, 1999; Kevane, 1996; Zaibet and Dunn, 1998).

Dorward (1999) is the only study we are aware of that examined the IR between farm size and agricultural land productivity in Malawi. The study used data from the 1980s, a time when the Government of Malawi forbade smallholders from growing tobacco, and agricultural policies generally discriminated against the smallholder sector to the benefit of estates. Those policies might explain why Dorward (1999) did not find a significant IR. However, the relationship between farm size and agricultural productivity might have changed with the introduction of pro-poor and smallholder-focused agricultural policies that started in the 1990s.

Explanations for the IR between farm size and agricultural land productivity fall into three main categories: (a) imperfect markets, especially for labor and land; (b) unobserved differences in land quality that are inadequately controlled for in regression equation 1; and (c) measurement error in farm size (Barrett et al., 2010). The IR has most often been explained by imperfect markets. Sen (1966) theorized that when rural labor markets function poorly, the surplus labor of family members is available for farm work at a very low shadow wage. Small farms apply this labor more intensively than large farms, so that the labor-to-land and output-to-land ratios are higher on small farms. Regarding imperfections in the land market, Eswaran and Kotwal (1986) suggested that the optimal land-to-labor ratio is higher for large landowners, because labor is subject to an increasing marginal cost of supervision, leading to decreased output per unit of land cultivated. The policy implication is that any type of land reform that reduces land inequality should positively impact agricultural productivity. Barrett (1996) attributed the IR to imperfect land and insurance markets, and to variable agricultural output prices. Under these conditions, small farmers, who are net buyers of the staple crop, have an incentive to over-supply labor on their own farms in an effort to be self-sufficient and minimize effects of market price fluctuations. In contrast, farmers with large landholdings, who are more likely to be net sellers, have an incentive to under-supply labor on their own farms, to reduce their exposure to price fluctuations when selling to market. As a result, small farms are more productive than large farms.

The second explanation of the IR is the failure in regression analysis to adequately control for unobserved differences in land quality. If large farms have lower quality land than small farms, the omission of land quality variables from equation 1 can bias estimates of α_3 and lead to a spurious IR. Large farms often do have lower quality land than small farms, because large farms rely more on rented land, which tends to be of lower quality, or because high quality land is more often partitioned (Kimhi, 2006).

The third explanation of the IR is measurement error in farm size. Until recently, farm and plot size were typically measured through farmer self-report in large household surveys, such as the Living Standards Measurement Study (LSMS). The difference between farmer self-report and GPS-measured land area can be substantial. Goldstein and Udry (1999) reported a correlation coefficient of only 0.15 between self-reported and measured farm sizes in Ghana. However, Carletto et al. (2011) found a correlation coefficient of 0.77 between self-reported and GPS-measured plot sizes in Uganda, after trimming the right end of the distribution. As discussed by Carletto et al. (2011), error in land area measurement can explain the IR, if small farmers underreport farm size more consistently than large farmers, resulting in artificially inflated yields in the lower part of the distribution. Interestingly, De Groote and Traorè (2005) found a linear negative relationship between measurement error and plot size in southern Mali, implying that farmers tend to overestimate small plots and underestimate large plots.

A number of studies have empirically examined individual or multiple causes of the IR, sometimes implicating error in land area measurement, based on the failure to find any other explanation. Lamb (2003) examined the effects of imperfect factor markets, omitted or imprecisely measured land quality, and error in land size measurement, using data from the longitudinal Village Level Studies (VLS) conducted by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) in India. Land quality was measured at household level, with share of cropland in different soil types, share of irrigated cropland, and the average value of cropland as estimated by a village authority. A gender-specific variable for days unemployed divided by days in the labor market was used as a proxy for labor market imperfections. Land market imperfections were measured as the average share of village land that was sharecropped or rented. Initial analyses found that land quality and land and labor market imperfections explained much, but not all, of the IR. When the role of error in land measurement was examined, Lamb (2003) concluded that this explained much of the IR, because (a) models with household fixed effects gave a much stronger IR than random effects models, and (b) the IR completely disappeared when instrumental variables estimation was used to reduce bias due to error in land area measurement.

Assunção & Braido (2007) also used the ICRISAT Indian VLS panel dataset to examine the three main explanations for the IR, but unlike Lamb (2003), they analyzed the data at the level of the plot, rather than the household. They found the within-household, across-plots IR to be more important than the across-household IR. Based on their results, they rejected household-based explanations for the IR: the slope of the IR remained statistically unchanged when dummy variables for households and periods were used in a regression model of agricultural output per acre. To test whether unobserved land quality attributes or error in cropped area measurement explained the IR, they added variables for the per-acre value of non-labor and labor inputs to the regression model. The rationale was that input choices should partially reveal unobserved plot-level heterogeneity, and should also be correlated with plot size. Inclusion of these variables caused the IR to vanish. The authors concluded that the IR was affected by unobserved plot-specific attributes (e.g., soil quality) or error in cropped area measurement, and that future research should focus on these explanatory factors, rather than market failures.

Barrett et al. (2010) used plot-level regressions of rice yield and land value to assess the effect of omitted land quality and factor market imperfections on the IR. Their cross-sectional dataset, based on 300 farm households in 17 villages in a rice-producing area of Madagascar, included detailed soil quality measurements (e.g., carbon, nitrogen, and pH) and variables for plot characteristics (irrigation and plot value to owner), which provided controls for land quality. Farmers cultivated multiple rice plots during the study year, allowing the use of household fixed effects to test for household-specific factor market imperfections. Regression results suggested that factor market imperfections explained about one-third of the IR, but inter-plot variation in soil quality explained almost none of it. Barrett et al. (2010) concluded that much of the remaining IR was due to error in cropped area measurement.

Carletto et al. (2011) used nationally representative household survey data from the 2005/06 Uganda Living Standards Measurement Study (LSMS) to test the hypothesis that the IR is a statistical artifact related to land measurement error. The study data included both farmer self-reported and GPS-measured plot size information for 65% of their sample, allowing a direct test of the hypothesis. Output data were only available for the farm level, but included plot-level soil quality data as shares of the farm with specific qualities, such as steep slope and irrigation. The regression results strongly rejected the hypothesis that the IR is due to small farmers under-reporting and large farmers over-reporting farm size. In fact, regression analysis of net agricultural revenues on GPS-measured farm size had a stronger IR than on farmer-estimated farm size. This result contradicts the results of Lamb (2003), Assunção & Braido (2007), and Barrett et al. (2010) for India and Madagascar.

Tatwangire and Holden (2013) examined the effect of land markets on the IR, using plot-level panel data for Uganda from 2001, 2003, and 2005. They tested the hypothesis that land market imperfections explain the IR by applying their models separately to freehold, *mailo*, and customary tenure systems. The dataset was longitudinal, allowing specification of fixed and random effects to control for unobserved heterogeneity across households that is time invariant. Nearly all of the plots were GPS-measured, minimizing error in area measurements. The results showed that the IR persisted across the three different tenure systems, but was lower in the freehold system where land markets function better, supporting imperfect land markets as an explanation of the IR. The analyses did not include separate regressions for farmer self-report and GPS-measured land area, so the influence of error in land area measurement on the IR was not examined. Nevertheless, results of this study agree with those of Carletto et al. (2011) for Uganda.

In summary, Lamb (2003) and Assunção & Braido (2007) found indirect evidence that the IR was partly explained by land area measurement error. Barrett et al. (2010) also concluded that land area measurement error contributed significantly to the IR, because factor market imperfections explained only about a third, and land quality explained virtually none of it. However, the studies by Carletto et al. (2011) and Tatwangire and Holden (2013) found that more accurate measurement of land area did not eliminate the IR. In fact, the former study found a stronger IR when GPS-measured land area was used. The present study rigorously tests hypotheses about the three main factors thought to explain the IR, using a detailed, three-year panel dataset with both farmer-reported and GPS-measured land areas. Similarly to Carletto et al. (2011), we assess the effect of error in land area measurement on the IR, but with the advantages that almost all plots were GPS-measured and that our dataset includes three years of data from the same households, which provides better control for unobservable household and farm/plot characteristics. Like Lamb (2003), our study controls for observed and unobserved land quality, using variables for plot attributes (soil type, slope dummy, and plot fertility status) and variables for household fixed effects. The role of factor market imperfections is examined using variables as proxies for the functioning of labor markets.

3. Data

We use farm household and plot level data from a random sample of 450 households from two districts in Central Malawi (Kasungu and Lilongwe) and four districts in Southern Malawi (Chiradzulu, Machinga, Thyolo, and Zomba). These households were surveyed in 2006, 2007, 2009, and 2012. Data on farm sizes and plot sizes were obtained in 2006, 2009, and 2012 by use of handheld GPS as well as by asking farmers to estimate their plot and farm sizes. The data from these three years are used for the analysis. Due to attrition, 378 and 350 of the original 450 households were re-interviewed in 2009 and in 2012, respectively. GPS area measures were taken for 3271 of a total 3354 farm plot observations. It was typically plots that were located far away from the dwelling unit that were less likely to have been measured. This small loss in plot observations, amounting to less than 3%, is not likely to significantly affect our main results. Household attrition is also not expected to significantly influence the key results.

We calculate the net return at plot level as the value of crop output minus the input costs excluding own family labor input, which represents the return to family labor and the land rent. To make the figures comparable across years we inflation-correct the net returns over years by using the consumer price index for Malawi. Two sets of area based net returns are then constructed, one based on the GPS-measured plot sizes and farm sizes and the other based on farmers' own estimates of plot sizes and farm sizes.

4. Estimation strategy

The extent of area measurement error approximated by the difference between farmers' self-reported plot and farm sizes and GPS-measured plot and farm sizes and how it is related to other plot and farm characteristics is assessed parametrically while including quadratic forms for plot and farm sizes.

The model with farmers' estimated plot sizes and farm sizes is first estimated with village fixed effects and household random effects as follows

$$\frac{y_{pit}}{A_{pit}^F} = \alpha_1 + \alpha_2 A_{pit}^F + \alpha_3 FS_{it}^F + \alpha_4 P_{pit} + \alpha_5 T_t + \alpha_6 V + h_i + \varepsilon_{pit} \qquad (2)$$

where y_{pit} is the net agricultural return to plot p for household i in year t, A_{pit}^F is the farmer's estimated plot size on the same plot, FS_{it}^F is the farmer's estimate of own farm size, P_{pit} is a vector of observable plot characteristics (distance from household to plot, soil type, slope class, and soil fertility class), T_t is a vector of year dummies, V is a vector of village dummies, h_i represents household random effects, and ε_{pit} is the error term.

Measurement error in plot size will create a negative correlation between the dependent variable and the RHS plot size variable and it is possible that the inclusion of the RHS plot size variable can help to reduce the bias in estimate of the relationship between farm size and net return per unit of land. However, with measurement error in the farm size variable as well, a bias in the measurement of the IR is expected to remain. The sign and size of this bias will depend on the nature of the measurement error and how it relates to plot size measurement error across multiple plots of the same household. For example, if the sizes of several plots are overestimated for the same household, it is likely that the farm size also is overestimated, leading to a negative correlation between farm size and plot-level net return per (incorrectly measured) unit land.

Another weakness of the specification above is that there can be unobservable household and plot characteristics that are also correlated with farm size and create a bias in the estimated α_3 parameter. We may eliminate this bias as long as it is time constant, by replacing equation (2) above with equation (3), where village fixed effects and household random effects are replaced by household fixed effects as follows:

$$\frac{y_{pit}}{A^F_{pit}} = \alpha'_1 + \alpha'_2 A^F_{pit} + \alpha'_3 FS^F_{it} + \alpha'_4 P_{pit} + \alpha'_5 T_t + \varphi_i + \varepsilon'_{pit} \qquad (3)$$

where φ_i represents the household fixed effects. The change in the parameter on farm size from equation (2) to equation (3) may say something about the importance of the time invariant unobservables that we are able to control for with our panel data. Fixed effects also control for unobservable plot level characteristics such as unobservable land quality that does not vary over time. Land quality is not likely to change very rapidly as land degradation is a fairly slow process.

The Hausman test has been commonly used to test random effects models against fixed effects models but this test rests on the assumption that the fixed effects model is consistent. Measurement error invalidates this assumption and can cause a larger bias in the fixed effects estimator than in the random effects estimator. The results from the models therefore have to be interpreted with caution (Lamb 2003).

The main remaining problem is measurement errors in plot size and farm size. In order to assess the importance of this measurement error, we measured plot sizes and farm sizes for the same households using handheld GPS devices. Equipped with these additional data we estimate the following equations:

$$\frac{y_{pit}}{A^G_{pit}} = \alpha''_1 + \alpha''_2 A^G_{pit} + \alpha''_3 FS^G_{it} + \alpha''_4 P_{pit} + \alpha''_5 T_t + \alpha''_6 V + h_i + \varepsilon''_{pit} \qquad (4)$$

$$\frac{y_{pit}}{A_{pit}^{G}} = \alpha'''_1 + \alpha'''_2 A_{pit}^{G} + \alpha'''_3 FS_{it}^{G} + \alpha'''_4 P_{pit} + \alpha'''_5 T_t + \varphi_i + \varepsilon'''_{pit} \qquad (5)$$

where A_{pit}^{G} is the GPS-measured plot size, and FS_{it}^{G} is the GPS measured farm size. While equation (4) uses village fixed effects and household random effects, equation (5) uses household fixed effects. By comparison of the estimated parameters of models (2) and (3) with those of models (4) and (5) we can judge the importance of plot and farm size measurement error on the IR and test hypothesis H1a. We can also assess the importance of unobservable time-constant plot and household characteristics by comparing the parameters in equation (4) and (5) after correcting for measurement error. This also allows us to assess whether or not these unobservables have the same effect after versus before correction for measurement error (H1b).

We cannot rule out that GPS-measured plots were measured with some error. By including GPS-measured plot size as well as GPS-measured farm size on the right hand side of equations (4) and (5) we are able to test and control for such measurement error. Significant measurement error in GPS-measured plot sizes should lead to a negative and significant coefficient on the GPS-measured plot size variable.

With an identified IR, we investigate whether or not the IR can be explained by land quality differences (H2) and labor and land market imperfections (H3). If large farms have lower quality land than small farms, and if land quality variables are not included in the analysis, this can lead to a spurious IR. We test for hypothesis H2 by estimating models without and with controls for observable (plot characteristics variables) and unobservable land quality (household fixed effects). It is possible that excess labor crowds more on small farms due to limited access to off-farm employment. Small farms with larger labor endowment may therefore face a lower opportunity cost of labor and apply more labor per unit land on their farm. We include the male and female labor endowment of households to assess this. It is also possible that during peak seasons labor-scarce households employ hired laborers (*ganyu* piecework, agricultural laborers). These labor variables are endogenous time-varying variables. We assess how they are correlated with net return per unit land and how their inclusion affects the IR. We do not have any good instruments to predict them, that is, variables we expect to be uncorrelated with net return per unit land. In a world with imperfect markets, production and consumption decisions of households are interrelated, creating a notorious identification problem. We therefore run models with and without these endogenous variables and observe how the coefficients on the farm size vary. The IR should be reduced when these variables are included; the labor endowment variables should have positive signs and be significant.

To test hypothesis H4 that food insecurity and self-sufficiency motives are part of the explanation of the IR on small farms, we include a dummy variable indicating that maize, Malawi's dominant crop, is grown on the farm plot. We run the models separately for farms less than 1 ha versus farms larger than 1 ha. It is particularly farms less than 1 ha that will need to intensify their production if they aim to be self-sufficient in food crops. Both the supply of family labor and demand for food may increase with the labor endowment of the household such that running models with and without the non-linear effect of the labor endowment can serve as an additional test of the hypothesis.

Finally, the impact of area measurement error on the conclusions regarding hypotheses H2-H4 is simply made by comparing the results of the otherwise equally specified models with high measurement error (models with farmers' self-reported plot and farm sizes) with those of the models with low measurement error (models with GPS-measured plot and farm sizes).

5. Results and discussion

5.1. Plot level measurement error

Figure 1 plots the measurement error estimated as the difference between farmers' estimated plot sizes and the GPS-measured plot sizes versus the GPS-measured plot size. The measurement error tends to be positive for small plot sizes, but turns negative for larger plot sizes. This is consistent with what Carletto et al. (2011) found in Uganda. Figure 2 plots farmers' estimated plot sizes versus GPS-measured plot sizes. This graph makes clear that Malawian farmers estimate their plot sizes as rounded figures, usually reporting areas in acres and tending to round plot sizes to 1 acre, 0.5 acre, 0.25 acre, etc. This is similar to what Carletto et al. (2011) found for their sample of Ugandan farmers. The rounding by farmers of their estimated plots sizes is also demonstrated in Figure 3 with the kdensity graph, which contrasts the plot size distribution with GPS-measured plots which has only one strong peak at about 0.12 ha. However, it is not obvious how these rounded plot size measures translate into farm size measurement error and thus how it affects estimates of farm level agricultural productivity. We explore this in the next section.

*** Figure 1 goes here ***

*** Figure 2 goes here ***

*** Figure 3 goes here ***

5.2. Farm size measurement error

Figure 4 shows that the same tendency exists for farm sizes as for plot sizes, with farmers overestimating the size of small farms and underestimating the size of large farms, consistent with the findings of Carletto et al. (2011) and De Groote and Traorè (2005). Figure 5 shows kernel density distributions for farmers' estimated farm sizes and GPS-measured farm sizes. The rounding to acre-units and upward bias in farmers' estimated farm sizes is very clear.

*** Figure 4 goes here ***

*** Figure 5 goes here ***

Finally, we graph the IR using net return per ha as the measure of land productivity. In Figure 6 this is done using GPS-measured plot sizes and in Figure 7 it is done using farmers' own estimated farm sizes. Both graphs indicate an IR at very small farm sizes, but the IR is weaker although still visible for larger farm sizes with the GPS-measured farm sizes. Considering that there are fewer very small farms when farmers' estimated farm sizes are used, the IR appears to be less important when farmers' estimates are used. The graphical evidence points in the direction of measurement error possibly dampening or eliminating an underlying IR. However, other important factors such as land quality and household and community characteristics must be controlled for before any firm conclusions can be drawn. These graphs have only aimed to provide insights on the extent to which the IR might be a statistical artifact related to land area measurement error. If an IR remains after correcting for most of the measurement error in farm size, then it is relevant to try to explain its existence; that is where hidden land quality and market imperfections have been the key candidate explanations.

*** Figure 6 goes here ***

*** Figure 7 goes here ***

Table 1 shows farmers overestimate small plot sizes and underestimate large plot sizes, suggesting that yields on small plots are underestimated and yields on large plots are overestimated, and this may lead to an underestimate of an IR. It is therefore expected that use of more reliable GPS-estimated plot sizes leads to a stronger IR compared with basing estimation on unreliable estimates by farmers. A key concern with large surveys based on self-reported farm sizes by farmers, is that these data may be unable to detect an IR that exists.

*** Table 1 goes here ***

Table 1 also shows that measurement errors are relatively smaller at farm level than at plot level but the same pattern of farmers relatively more strongly overestimating small farm sizes remains. Our findings are consistent with those of Carletto et al. (2011) in Uganda and De Groote and Traorè (2005) in Mali who showed that farmers tend to overestimate small plots and underestimate large farms. The same pattern of rounding plot sizes to the nearest acre or half acre was found in Uganda (Carletto et al., 2011) as we find for Malawi.

Table 2 provides parametric regressions with absolute measurement error and relative measurement error as dependent variables. Models with linear and a combination of linear and quadratic forms of plot size and farm size were included with household fixed effects, observable plot characteristics and year dummies. We see that the absolute measurement error decreases linearly with plot size and increases non-linearly (at a decreasing rate) with farm size. The relative measurement error decreases with plot size.

*** Table 2 goes here ***

In the next section we examine the consequences of measurement error for estimation of the IR.

5.3. The Inverse Relationship

Table 3 presents the results from linear panel data models of net returns per unit land with farmers' self-reported farm and plot sizes versus GPS-measured farm and plot sizes. The first two models are based on farmers' self-reported areas. A significant (at 5% level) IR is observed in both models. The plot size variables are also significant and with negative signs, an indication of significant measurement error. The third and the fourth models in Table 3 have the same specifications as models 1 and 2, but are based on GPS-measured farm sizes and plot sizes. Here the IR is highly significant (at 0.1% level) in both models and the IR is two to three times larger than in the first two models. It is also shown that the GPS-measured plot size variable is positive but statistically insignificant in both these models, an indication of very limited measurement error problem in these models. Contrary to the contention of Lamb (2003) and Barrett et al. (2010) that measurement error may explain the IR, but consistent with Carletto et al. (2011), results in Table 3 suggest that measurement error has the effect of dampening rather than increasing the IR. We find that measurement error covers up more than 60% of the IR in our case. The IR is also stronger in the model with household fixed effects than in the model with household random effects when GPS was used for area measurement, while the difference in the IR between fixed effects and random effects models is smaller in the models with large measurement error. This is also contrary to what Lamb (2003) suggested to be the case. We have more confidence in the model with household fixed effects when measurement error is minimized as it should control better for household and farm plot (time-invariant) observables and unobservables.

*** Table 3 goes here ***

Few of the plot characteristics variables are significant in any of the four specifications in Table 3, indicating that the village fixed effects and household fixed effects control well for (time-invariant) observable and unobservable farm and plot characteristics. The year dummy variables account for important variations across years: 2006 and 2009 were years with good rainfall, while severe dry spells were experienced in 2012, although these dry spells had a limited effect on total annual rainfall. The increase in food crop prices from 2006 to 2009 is probably the main explanation for the highly significant and positive coefficients for the 2009 dummy, but the expansion of the Malawi Farm Input Subsidy Program (FISP) in 2005/06 and 2008/09 may also have played a role. The droughts in 2011/12 in combination with high prices may explain the lower coefficients and significance levels for the 2012 dummy, in addition to the fact that the subsidy program was cut back by about 40% from 2010/11 to 2011/12. The 2009 dummy coefficients are also significantly larger in the models with GPS-based areas than in the models with farmers' self-reported areas, another indication of significant bias due to measurement error. The same applies to the estimated constants in the models, which are two to three times as large in the GPS-based models.

To summarize the results thus far, findings in Table 3 lead us to reject hypotheses H1a and H1b. In our empirical case, land area measurement error reduces rather than increases the IR and models with household fixed effects and large measurement error have not exaggerated the IR. To test alternative explanations for the IR (hypotheses H2-H4) and as robustness checks of area measurement error implications for the findings (hypothesis H5), we estimate models without and with observable plot characteristics, without and with male and female labor endowment in linear and quadratic forms, without and with a hired labor dummy variable, without and with a maize plot dummy, and with separate models for farms of less than 1 ha versus farms of greater than 1 ha. These models are all run with farmers' self-reported plot and farm sizes (large measurement error) as well as with GPS-measured areas (small measurement error) and as household random effects and fixed effects models. This also allows us to assess whether area measurement error can bias the findings and conclusions in the hypothesis testing. The IR coefficients and their levels of significance are summarized in columns 2 to 5 of Table 4. The other columns of the table indicate which time-varying variables are included in the models and provide information about the sample. We use this table as the main basis for testing and discussing the remaining hypotheses. We first discuss H2 to H4 one hypothesis at a time before we discuss the last cross-cutting hypothesis H5 about the impact of area measurement error on the outcome of testing the other hypotheses.

*** Table 4 goes here ***

To test hypothesis H2, that land quality explains the IR, we run household random effects and fixed effects models without and with observable land quality and base our judgment on the reported IR coefficients for the models with GPS-measured farm sizes (see the coefficients in columns 2 and 3 of rows 1 and 2 of Table 4). We see that adding observable plot characteristics to either the household random effects or fixed effects models results in only a slight increase in the IR, suggesting a minimal role for land quality in explaining the IR. Another way to test for H2 is by comparing the random effects and fixed effects specifications without and with plot characteristics. The household fixed effects models should control for time-invariant observable and unobservable plot and farm land quality characteristics and should therefore have a lower IR than the household random effects models, if land quality is a primary explanation of the IR. As shown in the table, the IR is higher in the household fixed effects versus random effects models, which indicates that land quality is not the reason for the IR in our data, in agreement with the findings of Barrett et al. (2010). We therefore reject hypothesis H2.

To test hypothesis H3, that market imperfections in land and labor markets explain the IR, we include time-varying male and female labor endowments in linear and combined linear and quadratic forms; a dummy variable for hired labor is also introduced. A perfect land market would wipe out the effect of variations in labor endowments and labor market imperfections. The significance of these labor endowment and labor market participation variables therefore relies on imperfections in both markets. It is also important to keep in mind that these time-varying labor endowment and labor market participation variables are endogenous. Furthermore, market imperfections imply that production decisions are non-separable from consumption and investment decisions, and make it difficult for us to identify credible strong instruments that satisfy the exclusion restrictions. The best we can do therefore is to run models with and without these endogenous variables with alternative model specifications as robustness checks regarding whether omitted variable bias or endogeneity bias represents the larger problem.

We first look at the effect of adding the male and female labor endowment variables in linear form and then in combined linear and quadratic forms in models where observable land quality variables are also included, still focusing on the models with GPS-measured farm sizes (see the coefficients in columns 2 and 3 of rows 3, 4, and 5 of Table 4). We see a substantial reduction in the IR when including the linear labor endowment variables and a further reduction when including the combined linear and quadratic labor endowment variables. In the latter case the IR is reduced to about half of what it is in the models without the labor endowment variables. This lends support to H3 that labor and land market imperfections at least partially explain the IR. The models we have suffer from endogeneity bias related to the included endowment variables, and omitted variable bias related to other time-varying land and labor market imperfections. The IR remains systematically higher in the household fixed effects specifications than the random effects specifications, suggesting that household (time-invariant) unobservables are likely reducing the IR.

Adding another endogenous variable, the hired labor dummy variable, leads to a loss of 504 observations due to missing information. Contrary to results for the labor endowment variables, adding the hired labor dummy variable leads to an increase in the IR (see the coefficients in columns 2 and 3, comparing rows 4 versus 6 and rows 5 versus 7). We cannot rule out that this is caused by attrition or endogeneity bias. However, our attempt to control for labor market participation/non-participation provides no additional explanation for the IR. Our sample dataset also lacks complete data for land market participation and we therefore are unable to investigate the role of this market failure. We tentatively conclude, based on the earlier findings, that labor and land market imperfections appear to be an important explanation for the IR.

Our next hypothesis (H4), that the IR is driven by covariate risk and food self-sufficiency objectives on small farms (Barrett 1996), is tested by including a maize dummy variable in the empirical model. This is an appropriate test given that maize is the dominant food crop in Malawi. About 60% of households are net buyers of maize but many farmers say they want to be self-sufficient in maize, partly because of concerns about food market unreliability during periods of calorie shortfall (Alwang and Siegel, 1999). The vulnerability of households is higher the smaller the farm size, which may create added motivation to intensify food crop production on small versus large farms and lead to a stronger IR on small farms. We divide the sample into farms less than 1 ha and farms more than 1 ha, based on the expectation that households with farm sizes below 1 ha face particularly severe problems in meeting their subsistence needs. Rows 6 to 12 of columns 2 and 3 of Table 4 are devoted to results of models assessing this.

Adding the maize plot dummy variable has no effect on the IR, as shown by comparison of the coefficients for rows 6 and 8 in columns 2 and 3 of Table 4. It appears that the decision to grow maize is not an important explanation for the IR. However, when we do a separate analysis for farms below 1 ha versus above 1 ha, a very strong IR is identified for farms below 1 ha while the IR is much weaker and insignificant on farms above 1 ha (compare the coefficients for rows 9 and 11 in columns 2 and 3). Including the quadratic forms of male and female labor had a significant negative effect on the IR on small farms, demonstrating a non-linear response to labor on small farms such that the IR becomes insignificant. The effect of inclusion of quadratic labor endowments is not found on farms larger than 1 ha. It appears that crowding of labor with limited off-farm employment opportunities but with food needs leads to low shadow wages and labor intensification with increasing labor/land ratios on farms below 1 ha in size but has no effect on larger farms. Our findings therefore lend support to hypothesis H4.

Finally, we test hypothesis H5 that area measurement error biases the tests for hypotheses H2-H4. We already saw that area measurement error affects the IR directly. Hypothesis H5 goes further in examining how area measurement error affects the slope response of the IR to alternative specifications. We examine both the direction and the size of the response in models with large and small area measurement error by comparing models for farmers' estimated farm sizes versus GPS-measured farm sizes.

Inclusion of observable plot characteristics slightly reduces the IR slope in the models with large measurement error while the IR slope slightly increases in the models with small measurement error (see columns 2 to 5 for rows 1 to 4 of Table 4). Inclusion of linear and quadratic labor endowment variables reduces the IR slope substantially in both models with large and small measurement error (see columns 2 to 5 of rows 3, 4, and 5). However, the dampening effect is relatively stronger when the quadratic labor endowment terms are included in the models with large measurement error, with the IR becoming insignificant. Adding the hired labor dummy variable increases the IR in models with large measurement error and small measurement error alike (see columns 2 to 5 for rows 4 to 7). Adding the maize dummy variable has a weak negative and insignificant effect on the IR slope in the models with large measurement error (see columns 4 and 5 of row 6 versus 8), while no effect is found in the models with small measurement error (see columns 2 and 3 of row 6 versus 8).

Finally, the IR is extremely large on farms less than 1 ha in the models with large measurement error, especially when household fixed effects are included (see columns 4 and 5 for rows 8, 9, and 10 of Table 4). While the IR is smaller in all earlier models with large measurement error than in models with small measurement error, the sample with farms less than

1 ha according to self-reported farm sizes has a stronger IR than the sample with farms less than 1 ha according to GPS-measured farm sizes (see columns 2 to 5 for rows 1 to 10). It is important to note that the sample size for self-reported farm sizes below 1 ha is almost 500 observations smaller than the sample size based on GPS-measured farm sizes. In summary, results in Table 4 suggest that land area measurement error leads to some bias in the testing of hypotheses H2 to H4. First, when plot attributes are introduced to the empirical model, the slope of the IR becomes smaller in models with large (area) measurement error suggesting that omitted land quality plays some role, albeit small, in explaining the IR. By contrast, the IR slope slightly increases with the addition of plot attributes in models with small (area) measurement error. Second, while both models with large and small measurement error find support for hypothesis H4, the results in the case of large measurement error are much stronger.

Also of interest is that Table 4 shows for the case of small versus large farms, that the models with large measurement error with household fixed effects yield much higher IRs than the random effects specification (see columns 4 and 5 for rows 9 to 12), while for the models with small measurement error the household fixed effects specification has only a slightly higher IR than the random effects specification (see columns 2 and 3 for rows 9 to 12). These findings for small farms with large measurement error are therefore more in line with the suggestions by Lamb (2003) for the effects of measurement error and use of fixed effects versus random effects specifications.

6. Conclusions

This paper has investigated two main research questions: How do measurement errors in estimation of farm plot and farm sizes affect estimates of the inverse relationship (IR) between farm size and land productivity? And, can area measurement error bias the tests of alternative hypotheses regarding IR explanations, such as those related to market imperfections in land and labor markets, land quality, price risk, and food security? Panel household survey data for Malawi reveals that yields on small farm plots and farms are underestimated and yields on large plots and farms are overestimated, when farmers' estimates of land area are used. The implication is that an eventual IR can be hidden by land area measurement error. We therefore expect and the empirical findings confirm that more reliable GPS-estimated plot and farm sizes lead to estimates of a stronger IR than basing estimation on unreliable self-reported estimates by farmers. This may suggest that some large surveys based on farmers' self-reported farm sizes are unable to detect an IR that exists and the result could be incorrect policy conclusions that overstate problems of inefficiency on small farms and favor large farm development.

Our results show that large area measurement error leads to a substantial underestimation of the IR for the whole sample and a substantial overestimation of the IR on farms smaller than 1 ha, especially when models with household fixed effects are used. The findings and suggestions by Lamb (2003) are therefore opposite of what we find for the total sample but in line with his finding for farms less than 1 ha. Area measurement error may therefore lead to quite unpredictable biases. The measurement error pattern that we find for the case of Malawi, appears to be similar to that found in Uganda by Carletto et al. (2011) and by De Groote and Traorè (2005) in Mali, and indicates that our findings are generalizable to other African countries. An important implication is that collection of more reliable area data, using GPS device or rope and compass, should be part of national farm surveys, or should at least be collected for subsamples if costs are excessive for large samples. Researchers whose analyses use plot and farm sizes should carefully report how land areas were measured, which has often not been done in the existing literature. Land productivity estimates based on unreliable self-reported farm and plot sizes should be treated with caution.

Our study provides support for labor and land market imperfections as important explanations of the IR, in line with other research (Barrett et al., 2010; Heltberg, 1998; Lamb, 2003). Similar to Barrett et al. (2010), we find that land quality is not an important explanation of the IR. The IR is found to be particularly strong on farms with sizes below 1 ha, which is plausibly related to the crowding of family labor (high labor/land ratios and low shadow wages) on small farms combined with subsistence needs as an additional motivation for intensified production in a risky environment with limited off-farm employment opportunities.

Finally, for the case of Malawi we show that land area measurement error influences the results of hypothesis tests of alternative IR explanations. Evidence in support of the hypothesis that covariate risk and food self-sufficiency objectives on small farms drives the IR is considerably stronger in models with large versus small measurement error in farm size. We also find a qualitative difference in test results for the omitted land quality hypothesis: omitted land quality is a partial explanation for the IR, albeit a minor one, in models with large area measurement error, while this explanation is rejected in models with small measurement error.

Acknowledgments

We acknowledge financial support from NORAD through the NOMA program and the "Impact of Fertilizer Subsidies in Malawi" project, while the third survey round was financed by CIMMYT. Special thanks to Rodney Lunduka who was responsible for the first two survey rounds and Julius Mangisoni and Maxwell Mkondiwa who were responsible for the third survey round with valuable assistance also from Nina Bruvik Westberg. Thanks are also due to the students involved in the 2009 survey, all enumerators and data entry staff at Bunda College of Agriculture. Eirik Schmidt Holst helped with data management.

References

Alwang, J., and Siegel, P. B. (1999). Labor Shortages on Small Landholdings in Malawi: Implications for Policy Reforms. *World Development*, 27(8): 1461-75.

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

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., and Hou, J. Y. (2010). Reconsidering conventional explanations of the inverse productivity–size relationship. *World Development*, 38(1), 88-97.

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

Berry, R. A., and Cline, W. R. (1979). Agrarian Structure and Productivity in Developing Countries. Baltimore: Johns Hopkins University Press.

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

Carletto, C., Savastano, S., and Zezza, A. (2011). Fact or artifact. The impact of measurement errors on the farm size-productivity relationship. Policy Research Working Paper 5908. Washington, D.C.: The World Bank.

Carter, M. R. (1984). Identification of the Inverse Relationship Between Farm Size and Productivity: an Empirical Analysis of Peasant Agricultural Production. *Oxford Economic Papers*, 36, 131–145.

Cornia, G. A. (1985). Farm size, land yields and the agricultural production function: an analysis for fifteen developing countries. *World Development*, 13, 513-534.

De Groote, H., and Traoré, O. (2005). The cost of accuracy in crop area estimation. *Agricultural Systems*, 84(1), 21-38.

Dorward, A. (1999). Farm size and productivity in Malawian smallholder agriculture. *The Journal of Development Studies*, 35(5), 141-161.

Eswaran, M., and Kotwal, A. (1986). Access to Capital and Agrarian Production Organization, *Economic Journal*, 96: 482-98.

Goldstein, M., and Udry, C. (1999). Agricultural innovation and resource management in Ghana. Yale University. New Haven, CT. Processed.

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

Jerven, M. (2013a). *Poor Numbers. How we are misled by African Development Statistics and what to do about it. Cornell Studies in Political Economy.* Cornell University Press, Ithaca and London.

Jerven, M. (2013b). The Political Economy of Agricultural Statistics and Input Subsidies: Evidence from India, Nigeria and Malawi. *Journal of Agrarian Change.* Published online: DOI: 10.1111/joac.12025

Kevane, M. (1996). Agrarian Structure and Agricultural Practice. Typology and Application to Western Sudan. *American Journal of Agricultural Economics*, 78(1), 236–245.

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

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

Sen, A. K. (1966). Peasants and dualism with or without surplus labor. *Journal of Political Economy*, 74(5), 425–450.

Tatwangire, A. and Holden, S. T. (2013). Land Market Participation and Farm Size-Productivity Relationship: Implications of Land Tenure Reforms in Uganda. Chapter 8 in Holden, S., Otsuka, K. and Deininger, K. (Eds.) (2013). *Land Tenure Reforms in Asia and Africa: Impacts on Poverty and Natural Resource Management.* Palgrave Macmillan, 187-210.

Zaibet, L.T., and Dunn E.G. (1998). Land tenure, farm size, and rural market participation in developing countries: The case of the Tunisian olive sector. *Economic Development and Cultural Change*, 46(4), 831-848.

Table 1. Summary statistics for the distribution of farmers' estimated versus GPS-measured plot and farm sizes, absolute and relative measurement errors

Distributional statistics	Farmers' estimated plot sizes (ha)	GPS-measured plot sizes (ha)	Measurement error in plot size (ha)	Relative measurement error in plot size, % of GPS-measured plot size	Farmers' estimated farm size (ha)	GPS-measured farm size (ha)	Measurement error in farm size (ha)	Relative measurement error in farm size, % of GPS-measured farm size
Mean	0.436	0.338	1.124	2537	1.146	0.874	0.272	102
p10	0.101	0.061	0.096	24	0.364	0.215	-0.308	-31
p25	0.202	0.125	0.366	117	0.526	0.406	-0.077	-10
p50	0.405	0.262	0.796	340	0.878	0.684	0.125	21
p75	0.607	0.455	1.498	840	1.416	1.115	0.520	91
p90	0.809	0.689	2.520	1972	2.226	1.721	1.067	226
Std. err. (mean)	0.008	0.006	0.025	1308	0.029	0.023	0.023	17
N	2600	2600	2600	2600	1018	1018	1018	1018

Table 2. Factors correlated with absolute and relative area measurement errors

	Absolute plot area measurement error1	Absolute plot area measurement error2	Relative plot area measurement error1	Relative plot area measurement error2
Plot size (ha), GPS-measured	-0.888***	-1.048***	-9441.833**	-2.50e+04*
Farm size (ha), GPS-measured	0.110	0.577***	-102.142	-624.901
Plot size squared, GPS-measured		-0.009		7489.521*
Farm size squared, GPS-measured		-0.008***		8.446
Distance to plot, minutes walk	0.000	0.000	0.138	0.153
Soil type dummy	0.132*	0.111*	2220.002	2128.330
Soil type dummy	0.129	0.100	3082.814	2711.719
Slope dummy	-0.044	-0.031	2849.351	2950.578
Slope dummy	-0.052	-0.039	10.406	949.139
Plot fertility status, dummy	0.011	-0.005	1764.330	2001.044
Plot fertility status, dummy	-0.060	-0.065	-2683.720	-2513.427
Dummy for year 2009	-0.182	-0.089	2526.245	2234.583
Dummy for year 2012	-0.183	-0.252**	-118.543	-573.831
Constant	1.537***	0.992***	2039.328	6272.409**
Prob > F	0.000	0.000	0.028	0.000
Number of observations	3324	3324	3324	3324

Note: Models with household fixed effects. Standard errors are corrected for clustering at household level. Significance levels: *: significant at 10%, **: significant at 5%, ***: significant at 1% or better.

Table 3. Measurement error and the farm size - productivity relationship: Linear farm plot level panel data models with farmers' estimated farm and plot sizes versus GPS-measured farm and plot sizes

	Dependent variable: CPI corrected net return/ha			
	Farmer estimated farm size		GPS-measured farm size	
	With village fixed effects and household random effects	With household fixed effects	With village fixed effects and household random effects	With household fixed effects
	(1)	(2)	(3)	(4)
Farm size, farmer estimate	-4410.693**	-4997.629**		
	(1719.48)	(2094.17)		
Plot size, farmer estimate	-4255.257**	-3327.143*		
	(1733.36)	(1763.78)		
Farm size, GPS measured			-1.07e+04***	-1.39e+04***
			(2715.38)	(3665.18)
Plot size, GPS measured			1969.239	4111.625
			(4554.69)	(4414.38)
Annual rainfall, mm	8.834	11.786	7.711	12.117
	(11.37)	(13.14)	(13.04)	(15.65)
Distance to plot, minutes walk	1.218	1.318	0.356	0.203
	(0.84)	(0.88)	(1.12)	(1.23)
Soil type dummy	-626.064	94.797	1081.81	1625.283
	(2774.73)	(3046.64)	(3995.67)	(4393.11)
Soil type dummy	-1854.833	-1086.005	-902.208	463.413
	(3287.29)	(3467.68)	(3765.84)	(3991.44)
Slope dummy	-1215.437	-558.03	-2656.054	-2500.121
	(2764.47)	(2947.34)	(3773.89)	(3954.99)
Slope dummy	-3687.973	-6.806	-8570.223	-6084.901
	(5691.82)	(6691.06)	(6883.65)	(6968.96)
Plot fertility status, dummy	1035.420	1926.428	76.382	996.112
	(3086.00)	(3327.83)	(3984.82)	(4392.80)
Plot fertility status, dummy	-2468.494	-2372.268	-3792.176	-2685.272
	(4725.09)	(5391.35)	(5524.15)	(6309.73)
Dummy for year 2009	21158.799***	21436.124***	34768.222***	36493.354***
	(3846.84)	(4068.00)	(5742.15)	(6165.43)
Dummy for year 2012	7361.248*	7879.438*	9065.047*	11856.605**
	(4100.40)	(4347.10)	(5291.15)	(5713.61)
Constant	2547.565	20177.866	15692.434	32394.564*
	(18659.62)	(13746.93)	(22248.49)	(16744.15)

Prob > F	0.000	0.000	.	0.000
Number of observations	3284	3284	3271	3271

Note: Standard errors are corrected for clustering at household level. Significance levels: *: significant at 10%, **: significant at 5%, ***: significant at 1% or better.

Table 4. Summary of the Inverse Relationship (IR) with alternative model specifications

Row	GPS-measured farm sizes		Farmers' estimated farm sizes		Alternative time-varying variables included				Sample	
	VFE + HHRE	HHFE	VFE + HHRE	HHFE	Plot attributes	Male and female labor supply	Hired labor dummy	Maize plot dummy	Farm sizes	Sample size
1	-1.06e+04***	-1.38e+04***	-4534.116***	-5217.510**	No	No	No	No	All	3320
2	-8779.08***	-1.16e+04***	-3828.490**	-2992.323	No	Yes, linear	No	No	All	3254
3	-1.07e+04***	-1.39e+04***	-4410.693**	-4997.629**	Yes	No	No	No	All	3278
4	-8957.102***	-1.18e+04***	-3714.844**	-2754.059	Yes	Yes, linear	No	No	All	3354
5	-5662.513*	-7157.095*	-1730.178	-684.931	Yes	Yes, quadratic	No	No	All	3320
6	-1.05e+04***	-1.36e+04***	-3917.324**	-2833.099	Yes	Yes, linear	Yes	No	All	2850
7	-7324.779**	-9222.465**	-2092.852	-857.216	Yes	Yes, quadratic	Yes	No	All	2908
8	-1.05e+04***	-1.36e+04***	-3890.319**	-2783.678	Yes	Yes, linear	Yes	Yes	All	2850
9	-4.49e+04***	-4.69e+04**	-5.93e+04***	-8.02e+04***	Yes	Yes, linear	Yes	Yes	< 1 ha	933/1420
10	-2.27E+04	-1.89E+04	-5.55e+04***	-8.61e+04***	Yes	Yes, quadratic	Yes	Yes	< 1 ha	933/1420
11	-5432.062	-7068.239	-431.444	-312.494	Yes	Yes, linear	Yes	Yes	> 1 ha	1911/1423
12	-5950.964	-8721.926	-29.820	162.413	Yes	Yes, quadratic	Yes	Yes	> 1 ha	1911/1423

Note: VFE=Village fixed effects, HHRE=Household random effects, HHFE=Household fixed effects. Standard errors are corrected for clustering at household level. Significance levels: *: significant at 10%, **: significant at 5%, ***: significant at 1% or better.

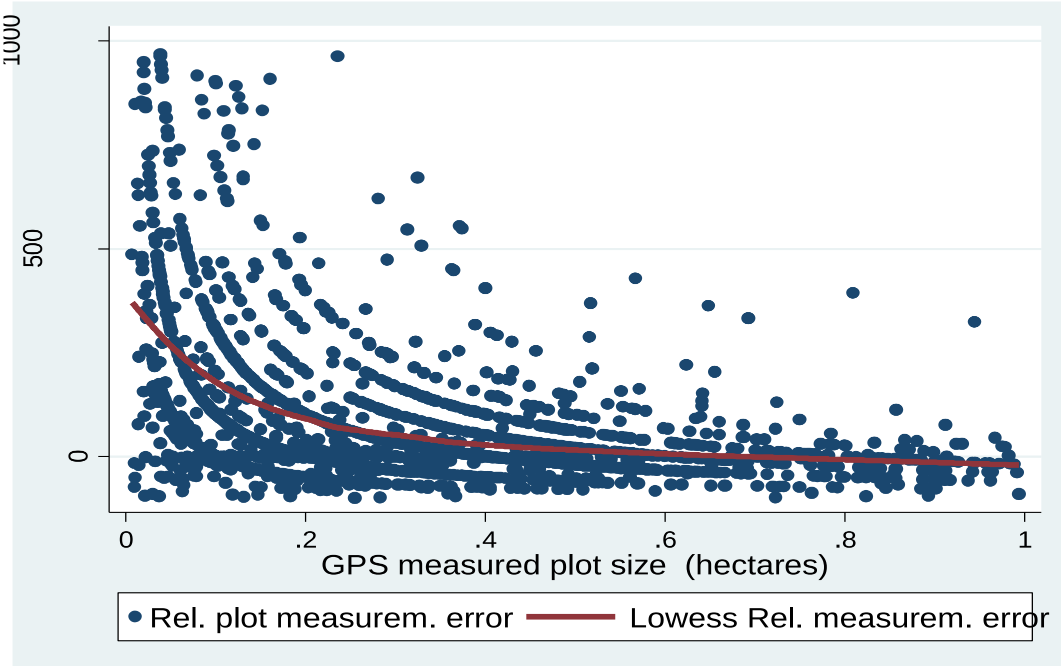

Figure 1. Relative measurement error in farmers' estimated plot size versus GPS-measured plot size

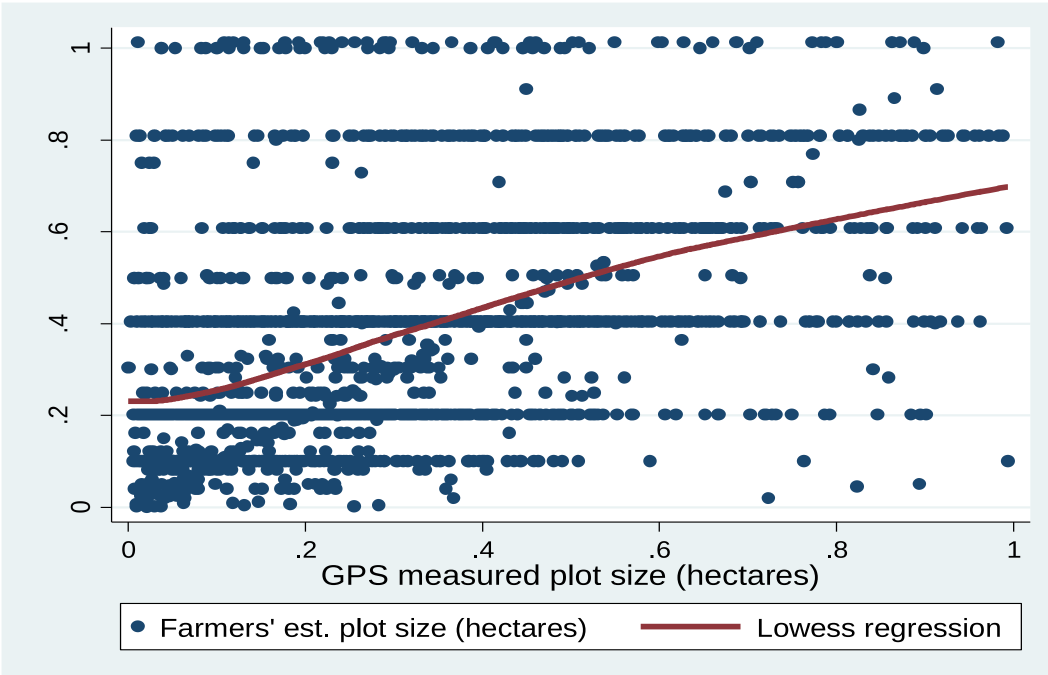

Figure 2. Farmers’ estimated plot size versus GPS-measured plot size

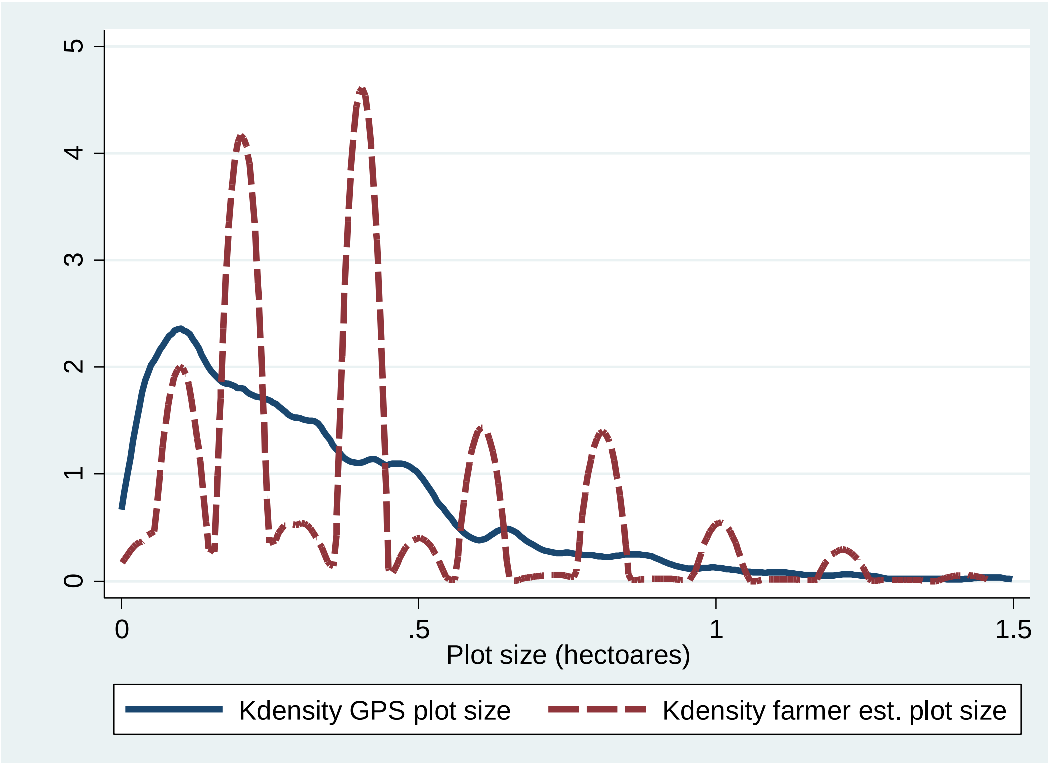

Figure 3. Distribution of GPS-measured and farmers' estimated plot sizes (ha)

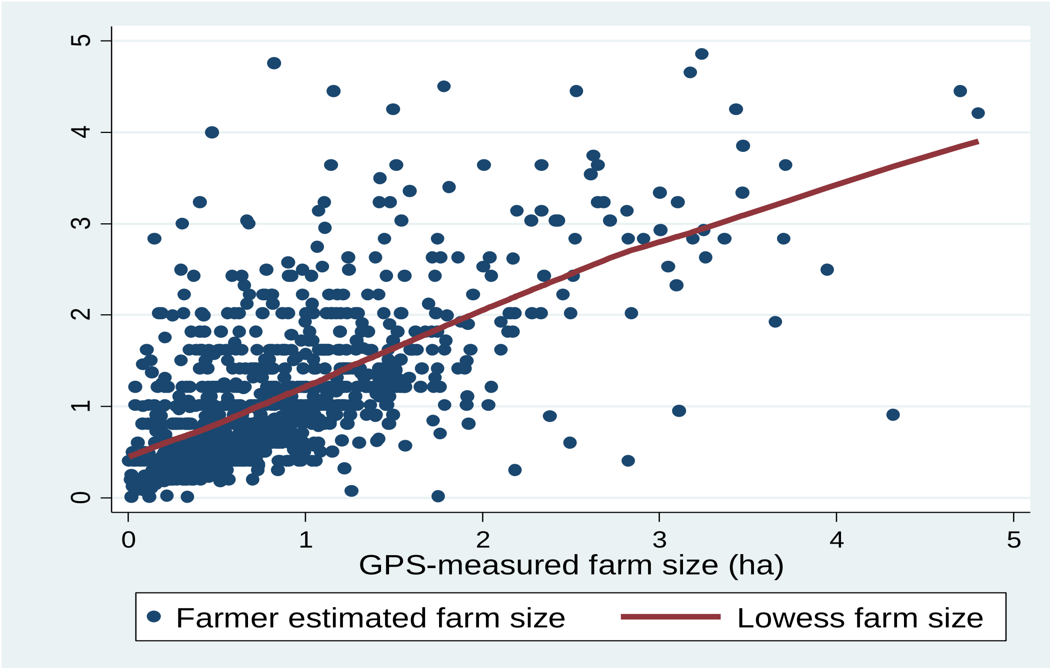

Figure 4. Scatterplot and lowess regression of farmers' estimated farm size versus GPS-measured farm size

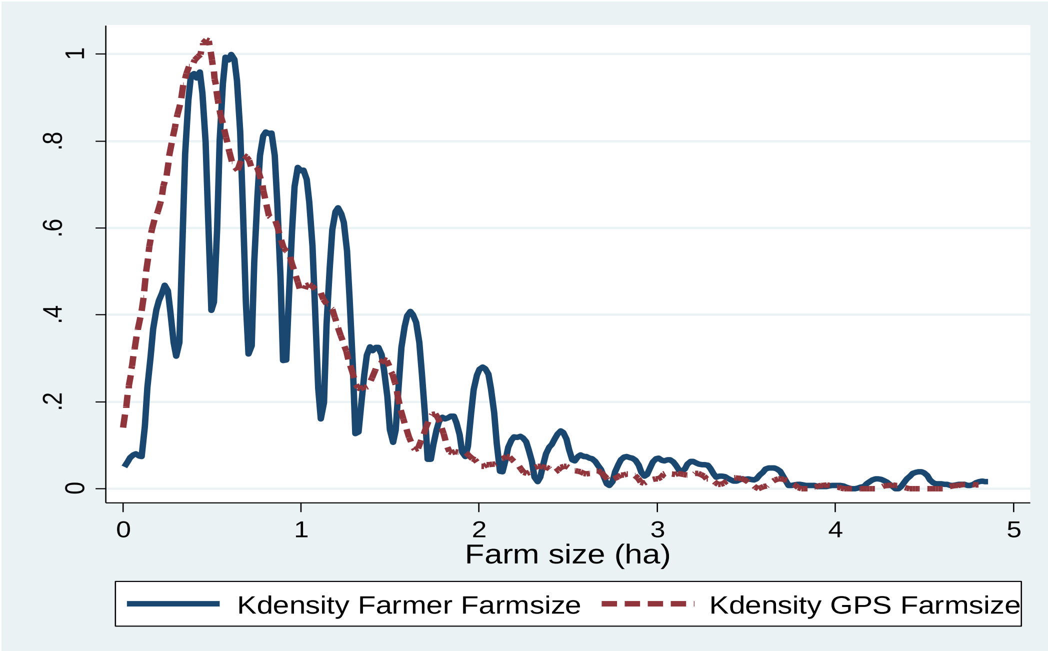

Figure 5. Kernel density distributions for farmers' estimated farm sizes and GPS-measured farm sizes

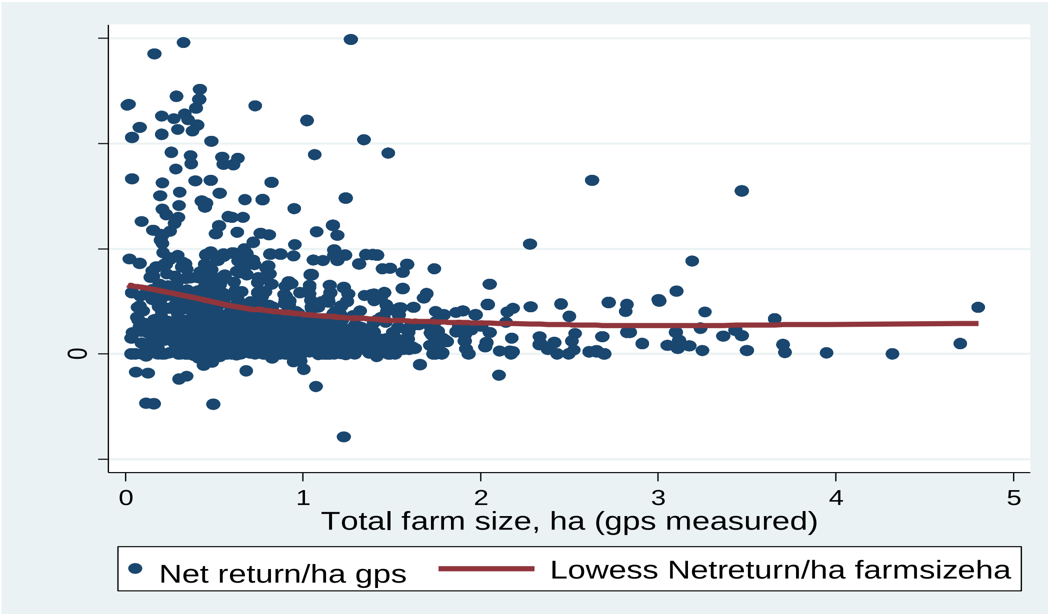

Figure 6. Net return/ha versus farm size (GPS measured)

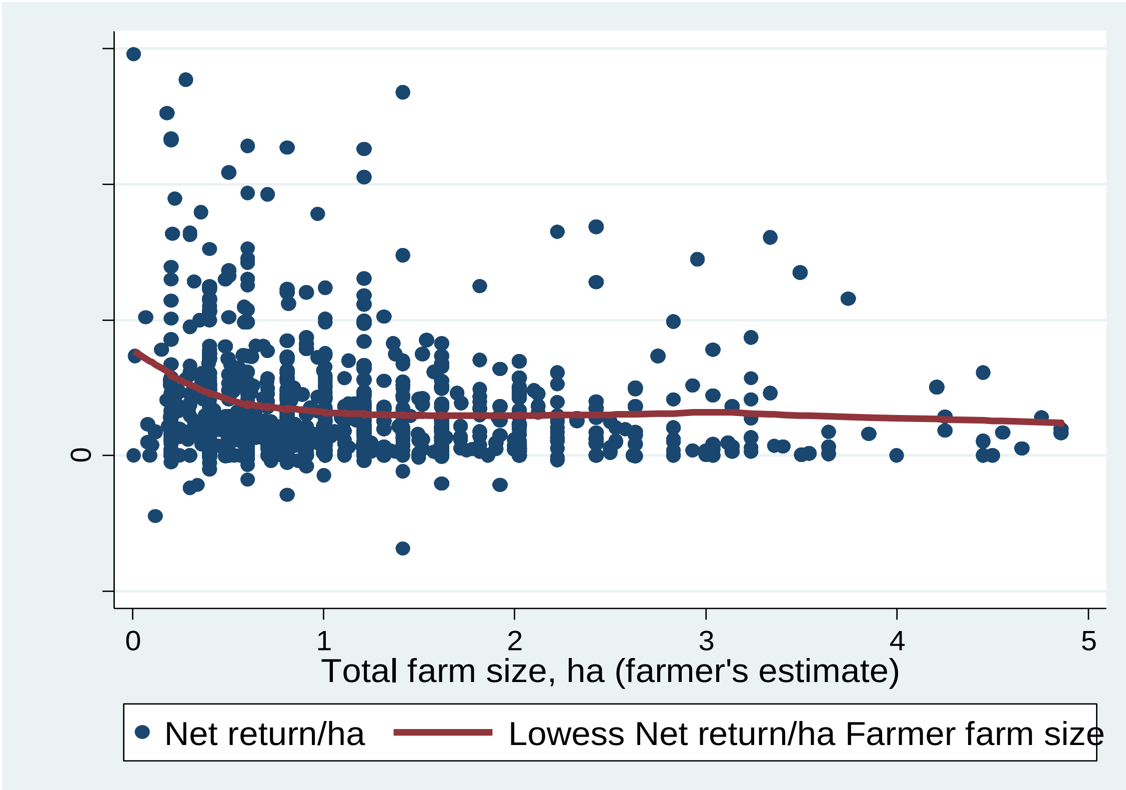

Figure 7. Net return per ha versus farmers' estimated farm sizes