

# Land Measurement Bias and Its Empirical Implications

Evidence from a Validation Exercise

*Andrew Dillon*  
*Sydney Gourlay*  
*Kevin McGee*  
*Gbemisola Oseni*



**WORLD BANK GROUP**

Development Data Group  
Surveys and Methods Team  
March 2016

## Abstract

This paper investigates how land size measurements vary across three common land measurement methods (farmer estimated, Global Positioning System (GPS), and compass and rope), and the effect of land size measurement error on the inverse farm size relationship and input demand functions. The analysis utilizes plot-level data from the second wave of the Nigeria General Household Survey Panel, as well as a supplementary land validation survey covering a subsample of General Household Survey Panel plots. Using this data, both GPS and self-reported farmer estimates can be compared with the gold standard

compass and rope measurements on the same plots. The findings indicate that GPS measurements are more reliable than farmer estimates, where self-reported measurement bias leads to over-reporting land sizes of small plots and under-reporting of large plots. The error observed across land measurement methods is nonlinear and results in biased estimates of the inverse land size relationship. Input demand functions that rely on self-reported land measures significantly underestimate the effect of land on input utilization, including fertilizer and household labor.

---

This paper is a product of the Development Data Group, Surveys and Methods Team. 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 [goseni@worldbank.org](mailto:goseni@worldbank.org).

*The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.*

## **Land Measurement Bias and Its Empirical Implications: Evidence from a Validation Exercise**

Andrew Dillon  
Michigan State University

Sydney Gourlay  
Kevin McGee  
Gbemisola Oseni

World Bank

**Acknowledgements:** We appreciate the helpful sampling advice provided by David Megill. The management of the land validation survey data collection process in Nigeria benefited from the collaboration of the Nigerian Bureau of Statistics staff and Colin Williams. All errors are our own. The authors may be contacted at [dillona6@msu.edu](mailto:dillona6@msu.edu), [sgourlay@worldbank.org](mailto:sgourlay@worldbank.org), [kmcgee@worldbank.org](mailto:kmcgee@worldbank.org), and [goseni@worldbank.org](mailto:goseni@worldbank.org).

**JEL Classification:** Q15, 012, Q12, 013

**Keywords:** Land measurement, Agriculture, Nigeria, Survey Methods

## 1 Introduction

Land measurement is critical to empirical development analyses and national agricultural statistical reporting. Land holdings are dependent variables, independent variables or a relative component indicating the scale of production in key empirical relationships estimated in development economics. Aggregate land holdings measure a household's wealth stock in rural areas. Plot sizes are an input in agricultural production functions. Accurate yield measures and input intensities are essential to the estimation of the inverse-land size relationship and input demand functions, and the policy decisions driven by these estimations. A large econometric literature has assessed the effect of measurement error, particularly non-random measurement error on empirical relationships (see for example the reviews in Hausman 2001 and Wooldridge 2008, among others), though specific applications to the above mentioned empirical relationships in development economics remains limited.

The impact of mismeasured land holdings has uncertain, though possibly important, effects on econometric relationships estimated in the literature. Classical measurement error in continuous dependent variables decreases the precision of estimates, but does not necessarily bias parameter estimates of independent variables (Wooldridge 2008). However, classical measurement error in independent continuous variables does potentially bias parameter estimates (Wooldridge 2008). Therefore, measurement error of land size has the potential to distort important and policy relevant econometric relationships especially when land size is included (either directly or indirectly) in the set of independent variables. In their investigation of the commonly found inverse relationship between land size and agricultural productivity (henceforth IR), Carletto, Savastano, and Zezza (2013) and Carletto, Gourlay, and Winters (2015) compared their findings using farmer and GPS estimates of plot size. They find a sizeable difference in the magnitude of the IR when using either land measure. One (or both) of the IR estimates is potentially subject to some degree of measurement error and without a gold standard measure of land size it is difficult to determine which IR estimate is closer to the truth. GPS measures of land area are generally assumed to be more accurate than self-reported land size, though measurement error is non-trivial in many commercial GPS units, particularly as plot size decreases. Another commonly estimated relationship in the literature that can be biased due to land size mismeasurement is the input demand function (Deininger et al 2003; Marennya, Paswel, and Barret, 2007; Erenstein, 2006; Thuo et al, 2011). Input intensification and the estimation of yield gaps related to input intensities are potentially biased when land size is mismeasured.

Despite the relative importance of land measurement, acquiring information on farmers' land area in developing countries can be time consuming, costly, and potentially inaccurate depending on the method employed. In the context of multi-topic household surveys, the costs of correct measurement of one variable

must be weighed against the multiple objectives of the survey. The generally accepted gold standard method of measurement (FAO 1982), the compass and rope (CR) method, is an arduous and time consuming process. Two alternative and less costly methods often used in agricultural surveys involve either (1) asking farmers to estimate the size of their plots (SR or self-reported) or (2) measurement of plot size using a GPS unit. However, both of these methods are potentially subject to greater error than CR through respondent error (farmer estimates) or technological error (GPS). Several studies have found significant land size differences both between GPS and farmer estimates (Goldstein & Udry, 1999; Carletto, Savastano, & Zezza, 2013; Carletto, Gourlay, & Winters, 2015) as well as between CR and GPS estimates (Shoning et al, 2005; Keita & Carfagna 2009). It appears that the potential for land area mismeasurement is significant depending on the measurement method employed, potentially due to either respondent driven or enumerator driven errors. The compass and rope method may also be prone to some degree of measurement error as the dimensions of the plot and compass bearings are determined by the enumerator and not automated. Any measurement with human intervention can be biased, but we treat the CR measurement as the truest measure going forward.

In this paper we aim to investigate (1) how land size measurements vary across measurement methodologies, and (2) the econometric and policy implications of land measurement error in estimation of common agricultural relationships found in the literature. We utilize data from the second wave of the Nigeria General Household Survey Panel (GHS-Panel) as well as a supplementary survey covering a subsample of GHS-Panel plots. As a part of the GHS-Panel, plots were measured using GPS devices and farmers were also asked to estimate the size of their plots. In the supplementary survey several months later, a subset of GHS-Panel plots were again measured according to farmer estimates and GPS but with the addition of a CR measurement. This allows us to compare all three measurement methods for the same plots. Such a comparison has not previously been performed on panel households, with two visits in the same season.

Utilizing the sample design of the GHS-Panel and specially collected validation sample data, we first conduct a descriptive comparison of differences across the three measures followed by an investigation of the household, respondent and plot characteristics that are correlated with farmer-estimated and GPS plot mismeasurement. We then examine two econometric relationships often estimated in the land literature: the inverse farm size-productivity relationship and input demand functions (farm size relationship with utilization of hired labor, fertilizer, and herbicide/pesticide). By estimating these common behavioral equations, we will be able to estimate any bias introduced by using self-reported and GPS plot sizes.

In our comparison of the three different measurement methods, we find self-reported mean land size is not statistically different than mean land size compared to either GPS or CR, but that self-reports overestimate land size on small plots and underestimate land size on large plots relative to either GPS or CR. Over-reporting bias on the smallest tercile of plots is 83%, while self-reports on the largest tercile of plots under-reported by over 20%. The GPS error is relatively consistent ranging from -0.2 percent for large plots to -2.8 percent for medium plots and -2.0 percent for small plots, which is particularly important as this is where the GPS measurement error is most likely to diverge from CR. We also find that plot characteristics such as non-titled land distributed by either the community or family partially explains the SR measurement error. The comparison of the measurement methods overall indicates that GPS estimates provide a relatively accurate estimate of plot size while SR estimates are subject to sizeable error especially for smaller plots.

In comparing SR to CR estimates, we find the expected negative relationship between land size and productivity and land size and input utilization. We find considerable bias in SR estimates for plots in the top two thirds of the land distribution leading to overestimation of the negative inverse productivity relationship. We also found considerable differences in the estimate for the input demand functions. SR land measures underestimates the effect of the land size on fertilizer use, pesticide/herbicide use and hired labor.

The remainder of this paper is organized as follows. In the next section, we provide background on land measurement methods as well as briefly review the literature associated with land measurement. In Section 3, we describe the data used in the analysis. We then analyze and attempt to explain the land measurement differences across methodologies in Section 4. In Section 5 we assess the potential impact of mismeasurement on estimates of the IR relationship and input demand functions using the three methodologies. Section 6 concludes.

## **2 Land Measurement: Background and Literature**

The extent of possible land measurement error is highly dependent on the method used to measure land size. Three common measurement methods include compass and rope, plot measurement with GPS devices, and farmer self-reported land size. A longstanding approach to collect land size is the compass and rope (CR) method which uses poles, ropes and a compass to carry out a systematic measure of land size (FAO 1982). It does not rely on advanced technology, only basic geometry and commonly available equipment. The CR method requires only a compass, measuring tape, ranging poles, a programmable calculator, and

two to three persons to measure the area of a plot with significantly more accuracy than by subjective estimates. GPS and CR measures require careful implementation by a field team, though GPS measurements may also be affected by the GPS device's margin of calibration error which can be affected by environmental factors. GPS measurements may be subject to more error than CR since the coordinates measured by the GPS device are not exact.<sup>1</sup> When carefully implemented, CR provides precise estimates, and it can therefore be considered as a benchmark against which to assess the precision of other methods.

Although the CR method is highly accurate when performed properly, it is time consuming and costly due to the necessity of careful training, monitoring, transportation costs to individual farmer plots, and enumerator time spent measuring the plot. The use of GPS, can on average require as little as 28% of the time needed for compass and rope (Keita and Carfagna, 2009; Schoning et al., 2005). Keita and Carfagna (2009) find that on small plots compass and rope ("CR") can take up to 17 times as long as GPS, though GPS measurement still requires survey team relocation to the plot and some survey team time to trace the field's perimeter. Because of this, many surveys rely on a farmer's own estimate of land size which avoids the time and cost of actual measurement. However, self-reported estimates could be subject to greater error given that (1) many farmers in developing countries acquire land through informal means where record keeping (and thus information on plot dimensions) is limited and (2) farmers are more likely to give rounded and inexact size estimates. More recently, the use of GPS units to measure land area has become more common in developing countries due to increased affordability of the technology and portability of the equipment. GPS measurements provide a more exact estimate of land size than farmer estimates and is much less time consuming (and thus less expensive) than measurements using CR.

The spread of GPS measured land size has led to several studies that examine the differences between the GPS and both CR and self-reported measures including the potential impact of measurement error resulting from inaccurate land size measurements. In two examinations of GPS versus CR measures in Sub-Saharan Africa, Schoning et al. (2005) and Keita and Carfagna (2009) find that GPS based measurements are generally lower than CR measures though neither study provides a technical explanation for this difference. Both studies find that the GPS-CR difference is statistically significant for smaller plots but not for larger plots.<sup>2</sup> It appears that for smaller plots, the fixed margin of error associated with determining GPS

---

<sup>1</sup> Standard GPS devices have reported position accuracy of approximately 15 meters depending on the model's satellite calibration algorithm (<http://www8.garmin.com/aboutGPS/>).

<sup>2</sup> Schoning et al. (2005) find a statistically significant difference for plots under 0.5 hectares but not for plots equal to or greater than 0.5 hectares. Keita and Carfagna (2009) separate their sample into 5 clusters based on plot size. They find a statistically significant difference for the clusters with the lowest land size but not for the 2 clusters with larger plots. However, they do not specify the plot size range for each cluster.

coordinates is large relative to the size of the plot and thus results in a greater measurement error. Since most plots in developing countries are relatively small, the potential for land mismeasurement using GPS could be significant. However, rounding of very small plot areas in Schoning et al. (all plots less than 0.01) clouds the reliability of their findings for smaller plots. This paper builds upon these two studies in that we are able to assess the degree of GPS accuracy for even the smallest of areas with the inclusion of CR measures.

Fewer studies have examined the differences between self-reported and GPS measured land sizes. Three studies have found inconsistencies between GPS and self-reported land sizes (Goldstein & Udry 1999; Carletto, Savastano, & Zezza, 2013; Carletto, Gourlay, & Winters, 2015). Although ancillary to their main analysis, Goldstein and Udry (1999) found that the correlation between GPS and self-reported land size was only 0.15 in their dataset from Ghana. The authors do point out that historically, field measurements in the region were based on length and not area and that this could partially explain the lack of a strong correspondence between farmer and GPS estimates. However, Carletto, Savastano, and Zezza (2013) also find a sizeable difference in Uganda though one that varies with plot size. The authors find that the bias is wider for smaller (less than 1.45 acres) and larger (greater than 3.58 acres) plots while self-reported measures are a reasonable estimate for medium size plots. In a similar analysis, Carletto, Gourlay, and Winters (2015) find inconsistencies between GPS and self-reported measurements in Malawi, Uganda, Tanzania, and Niger. They find that in all four countries self-reported plot sizes were higher than GPS for smaller plots but lower for larger plots, suggesting that farmers over-report the area of small plots and under-estimate the area of large plots. Both Carletto, Savastano, and Zezza (2013) and Carletto, Gourlay, and Winters (2015) go further and attempt to explain the difference using a basic econometric model. Carletto, Savastano, and Zezza (2013) find that in addition to plot size, the rounding of self-reported measurements, the age of the household head, and whether the plot was in a dispute with relatives were all positively associated with a greater GPS-self-report difference in land size in Uganda. Carletto, Gourlay, and Winters (2015) similarly find that rounding of farmer estimates are consistently associated with a difference between GPS and self-reported plot size in all four countries.

Following Carletto, Savastano, and Zezza (2013) and Carletto, Gourlay, and Winters (2015), we examine differences between GPS and farmer estimates and attempt to explain those differences, but also include the gold standard CR measure in the analysis. The CR will serve as a better reference measure to accurately assess and explain the measurement error associated with farmer estimates and GPS measurement.



### 3 Data Description

Two data sets are used for the analysis: (1) the second wave of the Nigeria General Household Survey - Panel (GHS-Panel) and (2) data from a subsequent land measurement validation survey (henceforth, validation sample) administered to a subsample of GHS-Panel households interviewed in the second wave.

#### *GHS-Panel Sample*

The GHS-Panel is a detailed multi-topic household and agricultural survey implemented by the Nigeria National Bureau of Statistics (NBS) and supported by the Living Standards Measurement Study – Integrated Surveys on Agriculture (LSMS-ISA) at the World Bank. The GHS-Panel was administered to around 5,000 households and is designed to be nationally and regionally representative. As its name indicates, the GHS-Panel is a panel survey which at present consists of two waves: households were first interviewed in 2010/2011 and then re-interviewed in 2012/2013. While households are followed between waves of the GHS-Panel, plots were not followed and therefore cannot be precisely matched between waves. Therefore, we shall focus our attention on the second wave of the GHS-Panel.

Together, the NBS and the LSMS-ISA aim to improve agricultural data for Nigeria through implementation of the GHS-Panel, particularly with respect to land area measurement. To this end, the GHS-Panel captures plot size using both farmer estimates and GPS measurements. However, extraordinary differences were observed between GPS and self-reported measures of agricultural land areas in the first wave of the GHS-Panel especially when compared to other LSMS-ISA surveys.<sup>3</sup> Table 1 below summarizes the average measurements found in recent LSMS-ISA surveys.<sup>4</sup> The SR-GPS difference observed in Nigeria is considerably larger than in the four other LSMS surveys listed.

<<< TABLE 1 HERE >>>

---

<sup>3</sup> While the differences observed in Nigeria are quite large nationwide, particular states and enumerators see significantly larger divergences. The two states with the highest discrepancy in measurements are Osun and Ondo (10,469% and 11,534%, respectively). Within those two states, three particular interviewers contribute to the majority of the difference suggesting that the problem may be largely the product of human error such as incorrect use of the GPS device or inaccurate recording of self-reported and/or GPS figures.

<sup>4</sup> One important difference between the GHS-Panel and the other LSMS-ISA surveys presented here is that in the GHS-Panel farmers are allowed to use nonstandard area units when estimating plot size while all the other surveys only allow farmers to report in standard units (acres, square meters, or hectares). This may partially explain why the self-reported/GPS difference is so much large in Nigeria.

### *Validation Sample from the GHS-Panel*

In order to further investigate these large differences in Nigeria, a subsample of plots from wave 2 were selected for a subsequent interview where CR, GPS measurements and farmer estimates were taken. This extra information enables comparison of methods and identification of potential measurement error associated with each methodology. The GHS-Panel has a sample of 5,000 households interviewed in the first wave. Out of this, 3,220 are agriculture households with a total of 5,125 plots. Limitations on time and budget prevented implementing compass and rope measurement on all plots owned and/or utilized by households in the GHS-Panel.

Four states were selected for the exercise based on safety and geographical dispersion. The four states are Oyo and Osun from the South West region and Benue and Kogi from the North Central region. Restricting the geographic area to be covered to two zones provided the benefit of limiting the number of enumerators required. Having fewer enumerators lends internal consistency to the measurements and at the same time, minimized training needs.

To ensure the sample covered a range of plot sizes, the plot selection was first stratified by GHS-Wave 2 plot size so we can examine the difference between farmer self-reported, GPS and CR at different plot sizes. Secondly, 100 plots were randomly selected within each stratum.<sup>5</sup> The selection of plots based on the above stratification led to the inclusion of 211 households. Because the greatest expense, both temporal and budgetary, is travel to various households, we then decided to measure all plots within the 211 households for which at least one plot was selected by stratification. This allowed us to increase the sample size without significant increases to the budget. This resulted in a sample size of 518 plots, which is representative of the four states covered in the exercise. Stratification by plot size in the validation sample results in the unequal probability of plot selection within households from the GHS-Wave 2 sample. Sample weights were calculated for the validation sample to make them representative of the same household population sampled in Wave 2.<sup>6</sup> Table 2 contains the distribution of the final validation sample across the four states.

<<< TABLE 2 HERE >>>

---

<sup>5</sup> Strata 1:  $\leq 1000$  sq. meters, 2. 1000-2500 sq. meters, 3. 2500-5000 sq. meters, 4.  $> 5000$  sq. meters.

<sup>6</sup> The final weight for these plots would be equal to the Panel Survey household weight for Wave 2 times the inverse of the subsampling rate for the plots. The subsampling rate for the plots is based on the probability that at least one of Wave 2 household plots is selected. Given the complexity of calculating this probability, it is simpler to determine the probability that none of the plots is selected, and then subtract this from 1 to determine the probability that one or more of the plots of the household are selected.

Stratification by plot size is essential in isolating the population of interest. Because the current literature questions the accuracy of GPS at smaller areas (Keita & Carfagna, 2009, and Schoning et al., 2005, have suggested that GPS measurements on plots less than 0.5 hectares are significantly different than CR assessments), collecting measurement data on plots in each strata is necessary if we are to analyze the accuracy of the GPS and farmer estimates at various plot sizes and, particularly, address the issue of accuracy on small areas. More than 80 percent of plots measured in the Wave 1 GHS-Panel were smaller than 1 hectare. Although compass and rope measurement is more time intensive on larger areas, we did not restrict the sample by employing an upper limit on area. Rather, we consider it critical to measure plots of all sizes in order to enable the analysis of all plot areas. In Nigeria in particular, the farmer-reported areas and GPS areas have been wildly divergent at all plot sizes, not only the smaller plots on which GPS accuracy is called into question. In measuring plots of all sizes, therefore, we aim to gain an understanding of the discrepancy between the two measures at all plot sizes.

Table 3 presents sample averages for the explanatory variables used in the analysis including the validation sample and the wave 2 GHS-Panel sample. In the first three columns, we compare the full wave 2 sample with the validation sample with the difference between the two shown in the third column. There are many significant differences between the validation and wave 2 samples, suggesting the validation sample is not representative of the national GHS-Panel survey. Nigeria is a large and diverse country so this is not surprising given the relatively limited scope of the validation sample. In the middle columns, we limit the wave 2 sample to zones covered by the validation sample (North Central and South West). Comparing the reduced wave 2 sample to the validation sample, there are far fewer significant differences though still a few household characteristics where the samples differ. Restricting the wave 2 sample to the states covered by the validation sample in the far right columns, there are no significant differences between the samples except the land area measurements. The difference between the GPS measurements from the GHS-Panel (from validation states only) is significantly different from the mean GPS measurement from the validation visit, but with limited significance (at 10% level) and low magnitude. There are large differences, both in magnitude and significance level, between the SR estimates collected in the GHP-Panel visit and the validation visit. The change in field staff between the GHS-Panel (professional enumerators) and the validation exercise (enumerated by higher-level staff from NBS headquarters) may partially explain the difference observed in the SR estimates. However, the field staff for the validation exercise were repeatedly instructed to record the farmer's estimate *without* assisting or correcting the farmer. As the validation results may have limited external validity for states outside the validation sample, the paper's analysis is supplemented with the nationally representative wave 2 sample when possible.

&lt;&lt;&lt; TABLE 3 HERE &gt;&gt;&gt;

#### 4 Descriptive Differences between Measurement Methods

In this section we examine differences in plot size measurements across the three land measurement methods using data from the validation exercise<sup>7</sup> as well as the full GHS-Panel wave 2. Since the CR method is considered to be the gold standard measurement method, our analysis assumes that it is the benchmark plot size. Therefore, any deviation from the CR measure will be classified as error or bias relative to this benchmark.

Table 4 presents the differences between SR and CR estimates. As shown in the fourth row of the table, the mean difference between SR and CR is small at only -0.03 acres or -2.1 percent of CR. In order to limit the influence of outliers, in the bottom row of panel A of Table 4, we trim the top and bottom 1 percent of SR and CR measurements. The scaled mean bias between SR and CR is slightly larger at 2.9 percentage points after trimming outliers.<sup>8</sup> The results of an adjusted Wald test are also presented in the far right column of Table 4. The test results comparing overall SR and CR measurements suggest that there is no significant difference between the two. However, this comparison masks much larger and significant differences that vary with plot size.

When we separate the sample into terciles based on CR plot size, both tests are in agreement for small and medium sized plots. We find that the difference between CR and SR is significant for both small (1<sup>st</sup> tercile) and medium (2<sup>nd</sup> tercile) plots. In both cases, the SR estimate is about 0.25 acres larger than the CR measure. Relative to plot size, the difference is much larger for smaller plots with an SR estimate that is on average nearly twice the actual (CR) size. Overall, this seems to indicate that SR measurements are fairly accurate for larger plots, but become less accurate as plot size decreases. This is in partial agreement with Carletto, Savastano, and Zezza (2013) in Uganda who found that SR estimates were a good approximation of GPS measures for medium plots but not small and large plots.

&lt;&lt;&lt; TABLE 4 HERE &gt;&gt;&gt;

---

<sup>7</sup> All validation plots with land size information have been included in this analysis. Plots that are missing information on other characteristics are excluded from the regression analysis that follows and hence the number of plot observations is different.

<sup>8</sup> Outliers were trimmed for the pair of SR and CR observations which results in decreases of the SR and CR means.

The CR measurements and GPS measurements are compared in Table 4, panel B. Overall, the difference between the CR and GPS is very small at 0.01 acres but is statistically significant, though only at the 10 percent level. The results remain largely unchanged after trimming outliers. When disaggregating by plot size tercile, the difference for all terciles is very low (never more than 2.8 percent of CR size). The adjusted Wald test results indicate that the difference between CR and GPS is only significant for medium sized plots. These test results are in partial agreement with Schoning et al. (2005) and Keita and Carfagna (2009) who found that GPS was less accurate for smaller plots. However, the difference is so small in magnitude that it may not cause significant bias when GPS is used in econometric specifications. This issue will be further explored in Section 6.

In Table 4, panel C, SR and GPS measured plots are compared. As would be expected, the differences are largely identical to those for SR versus CR measurements. The difference between SR and GPS is slightly higher with farmers reporting plot sizes that are 0.045 acres larger than GPS measurements. Also similar to Table 4, panel A, the difference between SR and GPS is largest (in percentage terms) for small plots, followed by medium plots, and is statistically different for all three plot terciles.

We can expand our comparison of GPS and SR measurements to the full GHS-Panel sample. Table 4, panel D, presents the results for the SR and GPS measurements taken in the wave 2 visit of the GHS-Panel. This highlights the error associated with farmer estimated plot size, relative to the GPS measured plot which was conducted as part of the GHS-Panel. The difference between the two measurements is significantly larger for the GHS-Panel sample than for the validation sample. While the difference is statistically significant for small and medium sized plots, the relative measurement error is largest for the smallest plots (437% of GPS size), followed by medium plots (205%), and then the largest plots (27%, though not significant). These sizeable differences highlight the potential for significant measurement error associated with farmer estimates of plot size.

In order to better examine the differences in the measures across the plot size distribution, we plot the distributions (using kernel methods) for all three validation measures in Figure 1 and the wave 2 measures in Figure 2. Figure 1 clearly illustrates the strong correspondence between the GPS and CR measures. The two distributions are nearly indistinguishable from one another. The SR distribution does largely mimic those for GPS and CR, but there are clear differences at the lower and middle portions of the distribution. The wave 2 GPS and SR distributions in Figure 2 are largely similarly to Figure 1 but they also reflect the larger GPS-SR difference found in wave 2 than in the validation data. In the next section, the reasons for this difference are explored.

<<< FIGURES 1 & 2 HERE >>>

#### 4.1 Determinants of Differences between Methods

In the first econometric specification, we investigate whether the measurement differences observed in the previous section are robust when accounting for the influence of other household and plot characteristics on plot size using the validation sample. In equation 1, land size is measured for each validation sample plot with three observations: one for SR, CR, and GPS. This allows us to directly compare all three measurements in a single regression. In equation 1,  $L_m$  is plot size in acres according to measurement  $m = \{SR, GPS, CR\}$ ,  $SR$  and  $GPS$  are dummy variables for SR and GPS measurement observations,  $Plot$  represents plot fixed effects, and  $\varepsilon$  is the idiosyncratic error term. Because we have multiple observations for each plot, we can estimate equation 1 with plot fixed effects which will capture the influence of any observed or unobserved plot characteristics on plot size for each measurement. In this formulation,  $\delta_1$  and  $\delta_2$  will estimate the mean difference between CR and both SR and GPS after accounting for plot fixed effects.

$$L_m = \delta_0 + \delta_1 SR + \delta_2 GPS + Plot + \varepsilon \quad (1)$$

In a second specification, we estimate the interactions of SR and GPS with an indicator of the CR reported tercile, as the descriptive statistics in Tables 4 indicate substantial variation by tercile in measurement differences.

We can also estimate the effect of self-reported land measures relative to GPS measures using the wave 2 data. As the wave 2 data only includes SR and GPS measurements, we cannot directly compare coefficients across the specifications or control for plot fixed effects. Therefore, we estimate the following equation with the wave 2 data controlling for observable characteristics which may influence self-reported land size:

$$L_m = \beta_0 + \beta_1(SR \times Tercile) + \beta_2 X + \beta_3(SR \times X) + EA + \varepsilon \quad (2)$$

where  $X$  is a matrix of manager, household, and plot characteristics and  $EA$  represents enumeration area fixed effects. In this formulation,  $\beta_1$  captures the inherent difference between SR land size by tercile and the omitted measure (GPS),  $\beta_2$  will capture any characteristics that are correlated with overall plot size

(regardless of measurement method), and the interactions between *SR* and *X* will capture any specific effects of characteristics on SR plot size relative to GPS. The latter will provide insight on what characteristics contribute to the differences between the two measurement methods.

Table 5 presents the results from equations 1 and 2. The results of estimating equation 1 for the validation data with the pooled effect of SR and GPS are presented in Column 1 while the disaggregated effect of SR and GPS by land tercile is presented in Column 2. In the case of the validation data, we are able to directly compare both GPS and SR with the gold standard CR measurements controlling for plot fixed effects. The GPS measurements are significantly smaller than CR in the pooled specification, but the marginal effect (0.012 acres) is not large and the difference is only significant at the 10% level. The difference between SR and CR measurements is insignificant and very close to zero when including the plot fixed effects. When the treatment effects are disaggregated, the effect of SR and GPS relative to CR differs significantly by land tercile. SR land size bias results in an overestimate of 0.2 acres relative to the CR tercile 1 mean of 0.31 acres to a 0.25 acre bias for tercile 2 relative to the tercile 2 mean of 0.83 acres. This suggests that the SR bias is larger for the lowest land tercile though the treatment effect for tercile 1 is lower. The SR bias estimated for the largest land tercile indicates respondents underestimate the land size on larger plots by 0.5 acres relative to the 2.8 acres tercile mean. The disaggregated SR land measurement bias demonstrates this bias is nonlinear across the land distribution with farmers overestimating land size for small plots and underestimating the size of large plots. The disaggregated GPS land measurement bias relative to CR is negative across the land distribution and statistically different than CR in land tercile 2 (-0.02 acres relative to 0.83 acres tercile mean). Hypothesis tests of the equality of the coefficient estimates is rejected at the 5% level for the SR and GPS coefficients.

<<< TABLE 5 HERE >>>

The SR land bias is estimated relative to GPS land measures in the nationally representative Wave 2 sample in Columns 3, 4, and 5. Though CR measures are considered the ‘gold standard’ in land measurement, the validation sample indicates only small differences between CR and GPS measures. The wave 2 results for estimation of equation 1 in column 3 highlights the more striking differences we observed between the GPS and SR measurements obtained in the second wave of the GHS-Panel controlling for plot fixed effects. The SR marginal effect relative to GPS measures in wave 2 is highly significant and much higher than that observed in the validation data (0.7 acres for tercile 1, 1.4 acres for tercile 2, and 1 acres for tercile 3). The SR bias is also nonlinear as in the validation sample.

In column 5, we estimate equation 2 using the wave 2 data to determine which characteristics help explain the difference between SR and GPS measurements controlling for EA fixed effects. The specification controls for observable household, manager, plot and interview characteristics in estimating land measurement bias.<sup>9</sup> The results indicate that larger households have larger plots on average. Likewise, whether a plot can be used as collateral and the number of crops grown on a plot are all positively associated with plot size. Larger plots are more likely to be available for use as collateral and have a greater capacity for cultivation of multiple crops.

Household, manager, plot and interview characteristics potentially capture important dimensions of land measurement bias. These characteristics are interacted with the SR indicator in equation 2 to estimate their relative effect to GPS measured land size in wave 2. Observable manager and household characteristics largely do not have statistically significant interactions with self-reported land sizes. Plot characteristics including plot tenure status explain differences between SR and GPS plot sizes in the wave 2 data. SR plot size was higher on plots which were distributed by the community or family compared to purchased plots (the omitted category). This could indicate that farmers are less knowledgeable about the size of plots that did not involve a land market transaction. Farmers are more likely to know the true size of plots acquired through a transaction (whether purchased or rented) since size is important in determining the price of the plot. Interview characteristics did have statistically significant effects on the difference between SR and CR land measurements including the day of the week on which the interview was conducted. Respondents also gave lower SR estimates for interviews that started in the late afternoon compared with those that started in the morning.

One potential concern with our results in column 5 is omitted variable bias. In columns 3 and 4, we control for all observed and unobserved plot characteristics with the plot fixed effects. In column 5, we drop the level of fixed effects to the enumeration area (EA) and include observable characteristics to explain observed plot size. If there were no omitted variables that are important covariates with plot size, we would expect the coefficient estimates on the SR measurement indicators to be the same between columns 4 and 5. However, the estimates in column 5 are nearly twice that of column 4. This could indicate that omitted variable bias is present in that specification. We should be cautious in the interpretation of the effects of observable characteristics interaction with SR land measures.<sup>10</sup> The next section examines the potential implications of SR land measurement bias in the estimation of the inverse farm size and input demand specifications, frequently estimated in empirical applications.

---

<sup>9</sup> These results are available upon request, but not reported in the main table for brevity.

<sup>10</sup> We will continue to incorporate additional characteristics to attenuate any omitted variable bias.



## 5 Inverse Farm Size Relationship and Input Demand Functions

We have shown above that there are significant nonlinear land biases in self-reported land estimates. In this section, we examine two well established agricultural relationships to estimate the effect of these biases on the inverse farm size relationship and input demand functions. First, we propose to estimate the inverse land size-productivity relationship (IR) documented by Carter (1984), Barrett (1997), Assunção and Braido (2008), Carletto, Savastano, and Zezza (2013), and Carletto, Gourlay, and Winters (2015) among others. Using our self-reported plot size and GPS measurements, we propose to estimate equation 3 where

$\ln\left(\frac{y_p}{L_p}\right)$  is the natural logarithm of total output value of  $y_p$  on plot  $p$  divided by plot size ( $L_p$ ), where the observations are stacked such that each plot has an observation using GPS measurement, SR estimation, and CR measurement, with an appropriate dummy variable (as in the analysis of equation 1 and 2):

$$\ln\left(\frac{y_p}{L_p}\right) = \alpha + \beta \ln L_p + EA + \varepsilon_p \quad (3)$$

The standard finding in the literature is a negative estimate of  $\beta$ , or decreasing productivity returns to scale in plot size. We estimate equation 3 using all of the measurements for plots,  $p$ , available for both the validation and full GHS-Panel wave 2 samples. In the validation sample, interactions of the measurement method (GPS or SR) and the CR land tercile variable are also included. All estimates for the validation sample and the wave 2 data include EA fixed effects.

Secondly we estimate input demand functions where previous studies find that a farmer's decision to invest is closely related to the size of the farm as production technologies change with increasing farm size. Deininger et al (2003) find that farmers with more land per capita are less likely to adopt the long term investment of planting trees on their farm. Other researchers have found that farmers that cultivate more land were more likely to invest in improved inputs such as fertilizer, herbicides, pesticides, and improved seed varieties (Marenya, Paswel and Barret, 2007; Erenstein, 2006; Thuo et al, 2011). We estimate the following input demand function using a linear probability model:

$$\text{Prob}(I_p = 1 \mid L_p, X) = \alpha + \delta \ln L_p + \Omega X + EA + \varepsilon_p \quad (4)$$

Where  $I_p$  indicates use of fertilizer, herbicide/pesticide, or hired labor on plot  $p$ . As for equation 3 above, we will estimate equation 4 for each measurement method in both the validation and full samples, including the plot area tercile interactions for the validation sample and EA fixed effects. Differences in the coefficient  $\delta$  can be attributed to the measurement bias in plot size. Using this difference in coefficient estimates, we can estimate the change in predicted probability of agricultural investment (I) due to plot measurement error.

### 5.1 Results: Inverse Farm Size Relationship

Table 6 presents the main IR results. Columns 1 and 2 were estimated using the three validation visit measurements. For all three we find the expected negative relationship. However, there is a large difference between the estimated relationship for SR compared to GPS and CR. The SR estimate is approximately 23 percent above the estimate for the gold standard CR measure. In contrast, the GPS IR estimate is not significantly different from the IR estimated with the CR measurement. This suggests that the mismeasurement bias in econometric results is relatively high for SR but relatively small for GPS. Column 2 disaggregates the plots by land area tercile. While the GPS is not significantly different from the estimate using CR at any tercile, the SR IR estimates do show variation. The SR IR estimates on the smallest plots (tercile 1) are not significantly different from the CR IR estimate of tercile 1, while the 2<sup>nd</sup> and 3<sup>rd</sup> terciles have significant coefficients. The IR relationship estimated with larger plots is significantly underestimated due SR measurement error.

In columns 3 and 4, we estimated the IR relationship using the wave 2 main visit SR and GPS measurements for validation sample plots. The SR IR estimate in column 3 is -0.693, very similar to the estimate in column 1 of -.576. We also find that the wave 2 GPS IR estimate in column 3 is not significantly different than the estimate using SR areas. Column 4 presents the analysis when plots are disaggregated by land area tercile. Unlike the results of the validation data (columns 1 and 2), there appears to be no significant difference between the IR estimates produced with GPS and SR area measures at any tercile. The hypothesis test of the equality of coefficients can not reject the null hypothesis.

<<< TABLE 6 HERE >>>

In columns 5 and 6 we compare results for the wave 2 measurements for the full wave 2 sample. The IR estimates are surprisingly similar for the SR and GPS measures for the full sample. This could indicate that the measurement bias in the full sample is relatively similar for both the wave 2 SR and GPS, though we cannot explicitly measure the bias. While we cannot directly compare these estimates with the validation

sample estimates, it is interesting to note that the IR relationship for wave 2 GPS is almost identical to that of the validation CR specification.

The results have shown that there can be considerable measurement error associated with estimating the IR relationship using the different land measurement methods. It appears that the bias associated with GPS measurements is consistently lower than the bias when using SR. We also find that the SR estimate of the IR is larger in magnitude than for GPS, particularly for larger plots. This is in contrast to Carletto, Savastano, and Zezza (2013) who find the opposite to be the case in Uganda but largely in agreement with the findings of Carletto, Gourlay, and Winters (2015) in Malawi, Uganda, and Niger.

## 5.2 Results: Input Demand Functions

Table 7 presents the results of fertilizer, pesticide/herbicide, and hired labor demand, respectively. The results from the validation sample are provided in columns 1, 2, 5, 6, 9, and 10. For all three we find the expected positive relationship between farm size and use of hired labor, fertilizer and herbicide, albeit very small effects. This is not surprising as the use of hired labor and adoption of non-labor inputs like fertilizer, herbicide and pesticide are low for farm households in developing countries. In the validation data, the interaction terms on GPS and SR with log of plot area are not significantly different than zero in any of the three models, nor when the data is disaggregated by land area tercile.

The results from the full GHS-Panel Wave 2 sample are provided in columns 3, 4, 7, 8, 11, and 12 of Table 7. Similar to the results from the validation sample, the coefficients on both the SR estimates and the GPS measurements are insignificant. When disaggregated into land area terciles, however, there is evidence of bias with the SR estimate. Fertilizer demand is decreased for the largest plots when using the SR estimates, while SR estimates in the pesticide/herbicide model suggests decreased demand with both medium and large plots. For hired labor, the SR estimate for medium plots is reduced with a coefficient of -0.038, suggesting that SR measurement error may bias the estimate of the demand for hired labor downwards.

Overall, it appears that the effect of land size bias leads to underestimating the probability of use for inputs such as hired labor, fertilizer, and pesticide/pesticide when SR land size are used compared to CR and GPS estimates, particularly for medium and larger plots. Land measurement bias in input demand functions may lead to significant underestimation of the potential effects of land pressure on input intensification.

## 6 Conclusion

Land measurement is a critical component of many statistical analyses centered on agriculture. However, acquiring information on farmers' land area in developing countries can be time consuming, costly, and potentially inaccurate depending on the method employed. Given the relative importance of plot measurements in informing, implementing, and targeting a variety of agricultural interventions, it is vital to have a clear understanding of the potential measurement error associated with each method. In this paper we have investigated how land size measurements vary across measurement methodologies and the econometric and policy implications of land measurement error in estimation of the commonly estimated inverse land size-productivity relationship and selected input demand functions. Using data from the Nigeria GHS-Panel coupled with the validation survey, we were able to compare the three main land measurement methods. Such a comparison using panel data has not been available in previous research, to the best of our knowledge.

In our comparison of the three different measurement methods, we find that measurement error is larger for SR plot size estimates than GPS estimates. On average, farmers overestimate plot size by 2.9 percent while GPS underestimates CR size by only 1 percent. Mean differences in land size hides much of the interesting variation in differences across the land distribution related to measurement method. The error associated with SR is highly variable depending on plot size. The error for the smallest third of plots was over 80 percent compared to 30 percent for medium sized plots and -20 percent for large plots. The GPS error was relatively consistent ranging from -0.2 percent for large plots to -2.8 percent for medium. The comparison of the measurement methods overall indicates that GPS estimates provide a relatively accurate estimate of plot size while SR estimates are subject to sizeable error especially for particularly small plots yielding overestimates and significant underestimates of larger plots.

Lastly, we examine two econometric relationships often estimated in land literature: the inverse farm size-productivity relationship (IR) and input demand functions. By estimating these common behavioral equations, we are able to quantify the effect of land measurement bias. For all three measures, we find the expected negative relationship between land size and productivity. However, there are some considerable differences in the magnitudes of the IR estimate. The bias associated with using SR land size is much higher than when using GPS land size in IR estimates. The much stronger and biased estimated relationship between plot size and agricultural productivity using SR measurements could mislead policy makers to conclude that increasing the size of a farmer's plot will result in a sizeable decrease in productivity when in fact it is only a modest decrease. We also found considerable differences in the estimate for the input demand functions. The SR measurement underestimates the effect of the selected land size on the use of

inputs. Of particular concern is the consistent finding for both the IR and input demand functions that small and larger plots induce different directions of biases due to land measurement error. With GPS and CR measurements, it appears that the link between land size and input use is stronger.

Overall, our findings indicate that GPS estimates are more reliable than SR measurements. Farmer estimated land size was subject to greater measurement error, varied more between the main GHS-Panel wave 2 visit and the validation survey visit, and introduced the greatest bias into estimation of the negative land size-productivity relationship and the input demand functions. With well-trained enumerators, GPS estimates provide an effective approximation of CR measurements at a lower cost.

## References

- Assunção, J. and Braido, L. (2007) “Testing Household-Specific Explanations for the Inverse Productivity Relationship.” *American Journal of Agricultural Economics*, 89 (4): 980-990.
- Barrett, C. (1996). “On price risk and the inverse farm size-productivity relationship.” *Journal of Development Economics*, 51 (2), 193-215.
- Carletto, C., Savastano, S., and Zezza, A., (2013). “Fact or artefact: the impact of measurement errors on the farm size – productivity relationship.” *Journal of Development Economics*, 103: 254-261.
- Carletto, C., Gourlay S., and Winters, P. (2015). “From Guesstimates to GPStimates: Land Area Measurement and Implications for Agricultural Analysis.” *Journal of African Economies*; doi: 10.1093/jae/ejv011
- 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.
- Deininger, K. and Jin, S. (2003). “The Impact of Property Rights on Households’ Investment, Risk Coping, and Policy Preferences: Evidence from China.” *Economic Development and Cultural Change*, 51 (4): 851-882.
- Erenstein, O. (2006). “Intensification or extensification? Factors affecting technology use in peri-urban lowlands along an agro-ecological gradient in West Africa.” *Agricultural Systems*, 90 (1): 132-158.
- FAO (1982). Estimation of crop areas and yields in agricultural statistics, FAO, Rome.
- Goldstein, M. & Udry C. (1999). “Agricultural innovation and risk management in Ghana.” Unpublished, Final report to IFPRI.
- Hausman, J.A. 2001. “Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left.” *Journal of Economic Perspectives*, 15 (4):57–67.
- Keita, N. and E. Carfagna (2009). “Use of modern geo-positioning devices in agricultural censuses and surveys: Use of GPS for crop area measurement.” Bulletin of the International Statistical Institute, the 57th Session, 2009, Proceedings, Special Topics Contributed Paper Meetings (STCPM22), Durban.
- Marenya, Paswel P., and Christopher B. Barrett (2007). "Household-level determinants of adoption of improved natural resources management practices among smallholder farmers in western Kenya." *Food Policy*, 32(4): 515-536.
- Schoning, P., Apuuli, E., and Muwanga-Zake, E. (2005). “Handheld GPS equipment for agricultural statistic surveys: Experiments on area measurements done during fieldwork for the Uganda” Pilot Census of Agriculture, 2003.
- Thuo, Mary, Boris E. Bravo-Ureta, Ibrahima Hathie, and Patrick Obeng-Asiedu (2011). “Adoption of chemical fertilizer by smallholder farmers in the peanut basin of Senegal.” *African Journal of Agricultural and Resource Economics*, 6.
- Wilcoxon, Frank. (1945). “Individual comparisons by ranking methods.” *Biometrics Bulletin*, 1 (6): 80-83.

Wooldridge, Jeffrey. (2008). *Introductory Econometrics: A Modern Approach*. Cengage Learning: Mason, OH.

## Tables

**Table 1: Difference in Self-reported and GPS Land Size in LSMS-ISA Surveys**

LSMS-ISA Survey*	N	GPS Area (acres)	Self-Reported Area (acres)	Average Bias (SR - GPS)	Bias as % of GPS
Malawi (2010/11) <sup>†</sup>	17767	0.936	0.983	0.047	5.0%
Uganda (2009/10)**	3287	1.73	1.67	-0.06	-3.6%
Tanzania (2010/11) <sup>†</sup>	4583	2.29	2.10	-0.19	-8.4%
Niger (2011) <sup>†</sup>	3001	4.48	4.52	0.04	0.9%
<b>Nigeria (2010/11)</b>	<b>5125</b>	<b>1.18</b>	<b>2.65</b>	<b>1.47</b>	<b>124.1%</b>

*All means are weighted in accordance with the design of the survey.*

*\*Includes plots with both GPS and SR areas (top and bottom 1% of GPS and SR areas trimmed)*

*\*\*Measurements at parcel level, not plot level*

*† Long/rainy season plots only*

**Table 2: Distribution of Validation Sample**

State	# of plots	% of total
Benue	249	48.07
Osun	78	15.06
Oyo	111	21.43
Kogi	80	15.44
<b>Total</b>	<b>518</b>	<b>100</b>



**Table 3: Sample Means for Validation Sample and Wave 2 of the GHS-Panel**

	Full sample			Only validation Zones		Only validation States	
	Wave 2	Validation	Difference	Wave 2	Difference	Wave 2	Difference
<i>Plot size</i>							
Compass and rope (CR)	-	1.305	-	-	-	-	-
GPS	1.180	1.045	0.135*	1.394	0.349***	1.234	0.188*
Self-reported	2.061	2.576	-0.515***	1.737	-0.839***	1.370	-1.206***
<i>Manager characteristics:</i>							
Years of education of manager	4.942	4.800	0.142	5.184	0.383	4.739	-0.061
Manager is male	0.823	0.871	-0.048***	0.878	0.008	0.871	0
Age of manager in years	51.999	54.298	-2.3***	52.945	-1.354*	54.328	0.029
Manager is head of the household	0.911	0.952	-0.04***	0.951	-0.001	0.952	0
<i>Household characteristics:</i>							
Number of household members	7.082	6.403	0.679***	6.581	0.179	6.282	-0.12
Log total value of agricultural capital	8.090	8.522	-0.432***	8.317	-0.205***	8.462	-0.061
<i>Plot characteristics:</i>							
Harvested crops on plot	0.890	0.934	-0.044***	0.914	-0.02	0.917	-0.017
HH can sell or use plot as collateral	0.698	0.622	0.076***	0.654	0.032	0.595	-0.027
Number of crops on plot	2.225	1.836	0.389***	1.844	0.008	1.773	-0.063
Plot has 1 or more tree crops	0.096	0.124	-0.028**	0.157	0.033*	0.128	0.004
Plot Distance in (KMs) to HH	1.516	1.793	-0.277	1.813	0.02	1.690	-0.104
Plot Potential Wetness Index	14.623	13.723	0.9***	13.856	0.134	13.848	0.125
<i>Input use:</i>							
Used fertilizer on plot	0.414	0.075	0.338***	0.178	0.103***	0.078	0.003
Used pesticide or hebicide on plot	0.355	0.430	-0.076***	0.487	0.057**	0.435	0.004
Used hired labor on plot	0.349	0.303	0.047**	0.309	0.006	0.268	-0.034
<i>Plot acquisition:</i>							
Plot purchased outright	0.057	0.038	0.019*	0.048	0.009	0.033	-0.005
Plot rented	0.077	0.170	-0.093***	0.106	-0.064***	0.152	-0.018
Plot distributed by community or family	0.741	0.697	0.044**	0.697	0	0.734	0.037
<i>Observations</i>	4,836	464	-	1,313	-	624	-

Notes: Sample means for GHS wave 2 and validation sample presented. All means are weighted in accordance with the sampling methodology. The difference columns also contain adjusted Wald-test results with significance denoted: \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

**Table 4: Mean Difference in Plot Measurements**

<i>Panel A: Validation: Farmer Self-Reported (SR) &amp; Compass and Rope (CR)</i>								
Plot size tercile (CR)	N	SR		CR		Mean bias (SR - CR)	Mean bias / mean CR	Adjusted Wald-test
		Mean	SD	Mean	SD			
1	165	0.573	1.013	0.313	0.125	0.260	82.9%	0.001
2	166	1.073	0.783	0.826	0.199	0.248	30.0%	0.000
3	164	2.233	1.919	2.818	2.499	-0.584	-20.7%	0.015
Total	494	1.300	1.500	1.328	1.808	-0.027	-2.1%	0.758
Total (trimmed)	472	1.218	1.206	1.183	1.162	0.034	2.9%	0.553
<i>Panel B: Validation: GPS &amp; CR</i>								
Plot size tercile (CR)	N	GPS		CR		Mean bias (GPS - CR)	Mean bias / mean CR	Adjusted Wald-test
		Mean	SD	Mean	SD			
1	165	0.307	0.133	0.313	0.125	-0.006	-2.0%	0.114
2	166	0.803	0.203	0.826	0.199	-0.023	-2.8%	0.000
3	164	2.812	2.514	2.818	2.499	-0.006	-0.2%	0.756
Total	494	1.316	1.817	1.328	1.808	-0.012	-0.9%	0.069
Total (trimmed)	471	1.191	1.178	1.203	1.162	-0.012	-1.0%	0.065
<i>Panel C: Validation: SR &amp; GPS</i>								
Plot size tercile (CR)	N	SR		GPS		Mean bias (SR - GPS)	Mean bias / mean CR	Adjusted Wald-test
		Mean	SD	Mean	SD			
1	165	0.573	1.013	0.307	0.133	0.266	84.9%	0.000
2	166	1.073	0.783	0.803	0.203	0.271	32.8%	0.000
3	164	2.233	1.919	2.812	2.514	-0.579	-20.5%	0.017
Total	495	1.300	1.500	1.316	1.817	-0.016	-1.2%	0.862
Total (trimmed)	472	1.219	1.206	1.174	1.178	0.045	3.7%	0.439
<i>Panel D: Wave 2: SR &amp; GPS</i>								
Plot size tercile (W2 GPS)	N	W2 SR		W2 GPS		Mean bias (SR - GPS)	Mean bias / mean GPS	Adjusted Wald-test
		Mean	SD	Mean	SD			
1	1765	0.870	1.644	0.162	0.088	0.708	436.9%	0.005
2	1766	2.016	2.332	0.661	0.221	1.354	204.8%	0.004
3	1766	3.566	3.222	2.806	2.565	0.759	27.1%	0.150
Total	5297	2.075	2.672	1.129	1.804	0.945	83.7%	0.078
Total (trimmed)	5103	2.061	2.676	1.038	1.337	1.023	98.5%	0.007

Note: All estimates are weighted to account for survey design (with the exception of Signed-rank tests).

**Table 5: Determinants of Differences between Measurement Methods - OLS Results**

	Validation		Wave 2		
	(1)	(2)	(3)	(4)	(5)
<i>Validation measurement method indicator (CR omitted)</i>					
Self-reported	-0.016 (0.090)		0.725*** (0.086)		
GPS	-0.012* (0.007)				
Self-reported × CR tercile 1		0.201*** (0.049)			
Self-reported × CR tercile 2		0.246*** (0.063)			
Self-reported × CR tercile 3		-0.516** (0.258)			
GPS × CR tercile 1		-0.006 (0.004)			
GPS × CR tercile 2		-0.024*** (0.005)			
GPS × CR tercile 3		-0.006 (0.020)			
<i>Wave 2 measurement indicator (GPS omitted)</i>					
Self-reported × GPS tercile 1				0.698*** (0.044)	1.835** (0.740)
Self-reported × GPS tercile 2				1.438*** (0.072)	2.262*** (0.705)
Self-reported × GPS tercile 3				1.008*** (0.123)	3.294*** (0.731)
Constant	1.305*** (0.030)	1.305*** (0.029)	1.228*** (0.034)	1.173*** (0.024)	-0.038 (0.603)
Fixed effects	Plot	Plot	Plot	Plot	EA
Hypothesis Test: SR1=SR2=SR3 (pvalue)		0.017		0.000	0.000
Hypothesis Test: GPS1=GPS2=GPS3 (pvalue)		0.022			
Observations	1392	1392	9672	9376	9376
R-squared	0.000	0.037	0.105	0.105	0.300

*Notes:* OLS estimates with standard errors clustered at the enumeration area in parenthesis. Significance denoted: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Equation 5 includes household, plot manager, plot and interview characteristics as well interactions of these variables with the SR land size. These estimates are available upon request.

**Table 6: Inverse Productivity - OLS Results**

	<i>Independent variable: Log plot yields in Naira/acre</i>					
	Validation		Wave 2 (Validation sample)		Wave 2 (Full sample)	
	(1)	(2)	(3)	(4)	(5)	(6)
Log plot area in acres	-0.576*** (0.144)	-0.577*** (0.146)	-0.693*** (0.138)	-0.694*** (0.139)	-0.580*** (0.045)	-0.577*** (0.044)
<b>GPS</b> × Log plot area in acres	0.006 (0.006)					
<b>SR</b> × Log plot area in acres	-0.131** (0.052)		0.008 (0.030)		-0.002 (0.015)	
<b>GPS</b> × Log plot area in acres × plot area tercile 1		0.009 (0.047)				
<b>GPS</b> × Log plot area in acres × plot area tercile 2		-0.242 (0.387)				
<b>GPS</b> × Log plot area in acres × plot area tercile 3		0.032 (0.079)				
<b>SR</b> × Log plot area in acres × plot area tercile 1		-0.038 (0.058)		0.067 (0.047)		0.013 (0.056)
<b>SR</b> × Log plot area in acres × plot area tercile 2		-0.344** (0.129)		-0.088 (0.068)		-0.163** (0.078)
<b>SR</b> × Log plot area in acres × plot area tercile 3		-0.221* (0.118)		0.041 (0.093)		-0.041 (0.080)
Constant	9.334*** (1.178)	9.469*** (1.239)	10.006*** (1.046)	10.003*** (1.055)	10.605*** (0.310)	10.610*** (0.310)
Hypothesis Test: SR1=SR2=SR3 (pvalue)		0.069		0.287		0.160
Hypothesis Test: GPS1=GPS2=GPS3 (pvalue)		0.794				
Number of observations	1260	1260	840	840	8550	8550
R-Squared	0.302	0.302	0.312	0.312	0.446	0.446

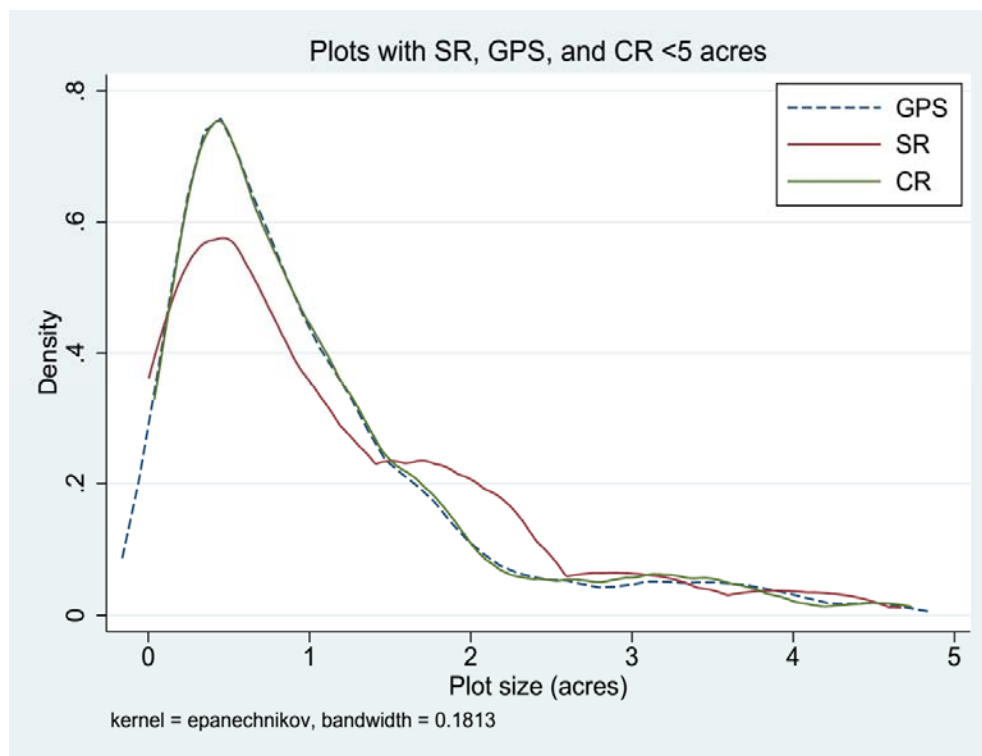
Table 7: Input demand and plot size - results

	Independent variable: used [INPUT] on plot											
	Fertilizer				Pest/Herbicide				Hired Labor			
	Validation		Wave 2		Validation		Wave 2		Validation		Wave 2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log plot size(acres)	0.026** (0.011)	0.026** (0.011)	0.023*** (0.007)	0.025*** (0.006)	0.057** (0.027)	0.057** (0.027)	0.028*** (0.008)	0.030*** (0.007)	0.105*** (0.022)	0.106*** (0.022)	0.031*** (0.008)	0.027*** (0.007)
GPS × Log plot area in acres	0.003 (0.011)				-0.005 (0.029)				-0.007 (0.023)			
SR × Log plot area in acres	0.004 (0.010)		0.001 (0.006)		-0.018 (0.026)		0.001 (0.006)		-0.030 (0.022)		-0.009 (0.006)	
GPS × Log plot area in acres × plot area tercile 1		-0.002 (0.012)				-0.011 (0.033)				0.002 (0.026)		
GPS × Log plot area in acres × plot area tercile 2		0.058 (0.037)				0.011 (0.122)				-0.063 (0.078)		
GPS × Log plot area in acres × plot area tercile 3		0.007 (0.019)				0.005 (0.045)				-0.017 (0.039)		
SR × Log plot area in acres × plot area tercile 1		-0.002 (0.010)		0.014 (0.012)		-0.022 (0.030)		0.048* (0.025)		-0.022 (0.026)		-0.014 (0.021)
SR × Log plot area in acres × plot area tercile 2		0.017 (0.017)		-0.008 (0.013)		-0.028 (0.039)		-0.037* (0.020)		-0.040 (0.029)		-0.038* (0.021)
SR × Log plot area in acres × plot area tercile 3		0.007 (0.017)		-0.046** (0.020)		-0.003 (0.034)		-0.052* (0.027)		-0.042 (0.032)		0.026 (0.024)
Hypothesis Test: SR1=SR2=SR3 (pvalue)		0.495		0.026		0.785		0.007		0.780		0.109
Hypothesis Test: GPS1=GPS2=GPS3 (pvalue)		0.250				0.940				0.646		
Observations	1320	1320	8608	8608	1320	1320	8608	8608	1320	1320	8608	8608
R-Squared	0.440	0.439	0.536	0.536	0.425	0.424	0.390	0.391	0.492	0.491	0.403	0.404

Notes: Linear probability coefficient estimates presented with standard errors in parenthesis. All regressions include enumeration area fixed effects. Indicators for the major crops were omitted to save space. We use the Battese (1997) method to correct for zero values when taking the natural log. The results for Battese dummy variables associated with zero values are omitted to save space. Significance denoted: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

## Figures

**Figure 1: Distributions of Validation Plot Measurements**



**Figure 2: Distributions of Wave 2 Plot Measurements**